J. Operations Research Soc. of Japan
Vol. 14, Nos. 3 & 4, November 1971

A DICHOTOMOUS SEARCH

SATORU MURAKAMI

Defense Academy

(Received November 12, 1970 and Revised April 2, 1971)

Abstract

An object to be searched is represented by a point lying in an interval with uniform *a priori* probability density. The only available test is to select a point and find out whether the point lies to the left or to the right (with different associated costs) of the test point. In this paper, it is assumed that each test is made at one of n equally spaced points in the whole interval. The problem is that of determining a sequence of test points so as to minimize the expected cost required until the object is located within a unit interval.

By use of the dynamic programming approach, the exact solution is derived, and its asymptotic formula with considerably good accuracy is obtained. A comparison of the approximate solution obtained by Cameron and Narayanamurthy with ours is also given.

1. Introduction

An object to be searched is represented by a point in an interval of length L. Suppose that *a priori* probability density of the location of the point is uniform. It is desired to locate the object. The only available test is to select a point and to find out whether the object lies to its left or its right. The cost of the test is 1 unit if the object lies to the left of the test point and k units if it is on the right. In this search procedure, the interval which contains the object can be unlimitedly diminished by an infinite sequence of test points selected, but we assume it sufficient to locate the object in an interval of an adequate length L/n. We suppose that n is given and each test is made at one of division points. We can locate the object in the minimum interval of length L/n by repeated application of a test under such assumption. The problem is that of determining a sequence of test points so as to minimize the expected cost required and calculating the minimum expected cost.

Several years ago, the model of this type was proposed by S.H. Cameron and S.G. Narayanamurthy [1]. They tried an analysis of a functional equation satisfied by the minimum expected cost and derived the equation $f(n)=p \ln n+C$, where $f(n)$ is the expected cost required by using an optimal policy and p and C are constants. They said that the qualitative and quantitative interpretation of the constant C is an intriguing avenue not yet fully explored. We shall determine the exact solution of the functional equation, *i.e.*, $f(n)$ and the set $\{x^*(n)\}$ of optimal test points for fixed n, where n is an integer. Further, we determine an asymptotic formula of $f(n)$ and give an explicit representation of the constant C mentioned above.

Recently, in dealting with a sorting problem, R. Morris [2] derived the above functional equation with $k=1$, and solved the equation elegantly by use of the convex property of $nf(n)$. In case $k\neq 1$, however, it seems unable to solve the functional equation by the only use of the convex property, and somewhat laborious mathematical manipulation developed in the following sections is necessary.

2. Derivation of Exact Solution

Let $E(n, x)$ be the expected cost of the search procedure required when we select first a test point x and thereafter use an optimal policy.

Using the principle of optimality enunciated by Bellman, we obtain the functional equation governing $f(n)$ for $n\in I^+(2)$,

$$(2.1)\qquad f(n)=\min_{0\le x\le n} E(n,x) = \min_{0\le x\le n}\left[\frac{x}{n}\{1+f(x)\}+\frac{n-x}{n}\{k+f(n-x)\}\right],$$

with $f(1)=0$, $k\in I^{+}(1)$ and $x\in I^{-}(0)$ where $I^{+}(q)$ is the set of integers greater than or equal to q. Defining $E_n(x)=nE(n,x)$ and $f_n=nf(n)$, equation (2.1) is rewritten as follows:

$$(2.2)\qquad f_n=\min_{0\le x\le n} E_n(x) = \min_{0\le x\le n}\{x+k(n-x)+f_x+f_{n-x}\},$$

with $f_0=f_1=0$. From this formula, we see that the functions f_n and $E_n(x)$ are non-negative integers. Therefore, we shall hereafter use f_n which is more convenient for analysis than $f(n)$. We now introduce functions $g(n)$ and $N(i)$ defined by the following equations:

$$(2.3)\qquad N(i)=N(i-1)+N(i-k),\qquad \text{for } i\in I^{+}(2),$$

$$(2.4)\qquad g(n)=g(n-1)+\varphi(n),\qquad \text{for } n\in I^{+}(2),$$

with $g(1)=1+k$, $N(i)=1$ for $i=2-k, 3-k, \cdots, 0, 1$. The function $\varphi(n)$ is 1 if there is an integer j satisfying the equation $n=N(j)$, and 0 otherwise. First, we wish to examine the properties of $N(i)$ and $g(n)$.

It is clear from (2.3) that $N(i)$ is strictly increasing in $i\in I^{+}(1)$ and assumes only integral values. Accordingly, for given n, there is a unique integer i determined by

$$(2.5)\qquad N(i)\le n\le N(i+1)-1.$$

Furthermore, we obtain the following relations

$$(2.6)\qquad N(i)=i=g(i)-k,\qquad \text{for } 1\le i\le k+1,$$

$$(2.7)\qquad N(i+1)>N(i)>i>g(i)-k\ge g(i-1)-k,\qquad \text{for } k+1<i.$$

We can establish

Lemma 1.

(2.8) $$g(n)=i+k, \qquad \text{for } N(i)\leq n\leq N(i+1)-1.$$

Proof. Using (2.4), we have

$$\sum_{j=2}^{n} g(j)=\sum_{j=2}^{n} g(j-1)+\sum_{j=2}^{n} \varphi(j).$$

Therefore, we get

$$g(n)=g(1)+\sum_{j=2}^{n} \varphi(j)=i+k,$$

since $g(1)=1+k$ and $\sum_{j=2}^{j=n} \varphi(j)=i-1$.

Lemma 2. For a function f_n' defined by

$$f_n'=ng(n)-N\{g(n)\},$$

we get

(2.9) $$f'_{n+1}-f_n'=g(n).$$

Proof. From (2.3), (2.4) and Lemma 1, we have

$$\begin{aligned} f'_{n+1}-f_n' &= (n+1)g(n+1)-N\{g(n+1)\}-[ng(n)-N\{g(n)\}] \\ &= g(n)+(n+1)\varphi(n+1)-[N\{g(n+1)\}-N\{g(n)\}] \\ &= g(n)+(n+1)\varphi(n+1)-\varphi(n+1)\{N(i)-N(i-1)\} \\ &= g(n)+\varphi(n+1)\{n+1-N(i-k)\} \\ &= g(n). \end{aligned}$$

Next, we shall prove

Theorem 1.

(2.10) $$f_n=ng(n)-N\{g(n)\}, \qquad \text{for } n\in I^+(1).$$

Proof. The proof is done by induction in n. We observe that (2.10) is true for $n=1, 2$. We assume it is true for all values up to $n-1$. Using this assumption, we shall find $\{x(n)\}$, the set of values which minimize the function $E_n(x)$. Next, substituting an arbitrary element of the set

$\{x(n)\}$ into $E_n(x)$, we shall show that (2.10) is true for n.

It is clear from (2.2) that the set $\{x(n)\}$ is a subset of $I^+(1)\cap I^-(n-1)$ where $I^-(q)$ is the set of integers less than or equal to q. From Lemma 2, we see that

(2.11) $$g(x)=f_{x+1}-f_x$$

holds for $x\in I^+(1)\cap I^-(n-2)$.
Then, defining for $n\in I^+(3)$

(2.12) $$D_n(x)=E_n(x+1)-E_n(x), \qquad \text{for } x\in I^+(1)\cap I^-(n-2),$$

we obtain from (2.2) and (2.11)

(2.13) $$D_n(x)=1-k+g(x)-g(n-x-1), \quad \text{for } x\in I^+(1)\cap I^-(n-2).$$

Here, we see that

(2.14) $$D_n(x)\geq D_n(x-1), \qquad \text{for } x\in I^+(2)\cap I^-(n-2),$$

since the function $g(x)$ is monotonically increasing. We shall assume that for a fixed n there are roots $\beta_s(n)\in I^+(1)$ of the equation

(2.15) $$D_n(x)=0.$$

Let v_n' and V_n' be the minimum and the maximum among the roots, respectively. Also, let $\{\beta(n)\}$ be the set of all $\beta_s(n)$. Then, using the property that $D_n(x)$ is monotonically increasing with x, we have $\{\beta(n)\}=I^+(v_n')\cap I^-(V_n')$ and $\{x(n)\}=I^+(v_n')\cap I^-(V_n'+1)$. We see from (2.13) and (2.14) that for a fixed n there is only one $j_n\in I^+(0)$ satisfying the equation,

(2.16) $$g\{\beta_s(n)\}=j_n+k.$$

Further, from (2.13) we obtain that $1-k+g\{\beta_s(n)\}-g\{n-\beta_s(n)-1\}=0$, therefore, $g\{n-\beta_s(n)-1\}=1+j_n$. From equation (2.16), and Lemma 1, we obtain

(2.17) $$N(j_n)\leq\beta_s(n)\leq N(j_n+1)-1,$$

(2.18) $$N(1-k+j_n)\leq n-\beta_s(n)-1\leq N(2-k+j_n)-1.$$

Sum up each sides of (2.17) and (2.18). Noticing (2.3), we have

$$N(j_n+1)+1\leq n\leq N(j_n+2)-1.$$

Here, defining $i_n=g(n)-k$, we see from Lemma 1 that $j_n=i_n-1$. Then, we obtain

$$(2.19)\qquad N(i_n)+1\leq n\leq N(i_n+1)-1.$$

Also, we see that equation (2.19) does not hold if $n=N(i_n)$.

We shall obtain $\{x(n)\}$ according to the following two cases:

(i) The case $n\neq N(i_n)$.

Substituting $j_n=i_n-1$ into (2.17) and (2.18) we have

$$(2.20)\qquad n-N(i_n+1-k)\leq\beta_s(n)\leq n-1-N(i_n-k),$$

$$(2.11)\qquad N(i_n-1)\leq\beta_s(n)\leq N(i_n)-1.$$

Also, we see from (2.19) that for arbitrary $n\neq N(i_n)$ we have

$I^+(v_n)\cap I^-(V_n)=\phi$ where

$$(2.22)\qquad v_n=\max\{N(i_n-1),\ n-N(i_n+1-k)\},$$

$$(2.23)\qquad V_n=\min\{N(i_n)-1,\ n-1-N(i_n-k)\}.$$

Since both equations (2.20) and (2.21) must be satisfied by $\beta_s(n)$, we obtain $\{\beta(n)\}\subset I^+(v_n)\cap I^-(V_n)$. Next, we see easily that $D_n(v_n)=0$ and $D_n(V_n)=0$. Accordingly, from (2.14) we have $\{\beta(n)\}\supset I^+(v_n)\cap I^-(V_n)$, therefore, $\{\beta(n)\}=I^+(v_n)\cap I^-(V_n)$, moreover, $v_n=v_n'$ and $V_n=V_n'$. Consequently, we obtain

$$(2.24)\qquad \{x(n)\}=I^+(v_n)\cap I^-(V_n+1).$$

(ii) The case $n=N(i_n)$.

In this case, the equation (2.15) has no root since (2.19) does not hold. Then, we shall determine $\{x(n)\}$ by examining whether $D_n(x)$ is positive or negative.

Substituting $x=N(i_n-1)$ or $N(i_n-1)-1$ into (2.13),
and using $g\{N(i_n-1)\}=i_n-1+k$, $g\{N(i_n-1)-1\}=i_n-2+k$,
we obtain $D_n\{N(i_n-1)\}=1>0$ and $D_n\{N(i_n-1)-1\}=-1<0$.
Thus, we see from (2.12) and (2.14) that the function $E_n(x)$ is minimum only when $x=N(i_n-1)$. Therefore, we have

$$\{x(n)\}=I^+\{N(i_n-1)\}\cap I^-\{N(i_n-1)\}=N(i_n-1). \tag{2.25}$$

Next, substituting $n=N(i_n)$ into (2.22) and (2.23), we have
$N(i_n)-N(i_n+1-k)<N(i_n-1)$ and $N(i_n)-1-N(i_n-k)<N(i_n)-1$.
Therefore, we have $v_n=N(i_n-1)$ and $V_n=N(i_n-1)-1$.
Thus, equation (2.25) can be rewritten as $\{x(n)\}=I^+(v_n)\cap I^-(V_n+1)$.
From the fact mentioned above, we see that $\{x(n)\}$ can be represented by (2.24) for all $n\in I^+(2)$. That is, for arbitrary n we have

$$\{x(n)\}=I^+(v_n)\cap I^-(V_n+1). \tag{2.26}$$

We shall now show that the functions f_n and $\{x(n)\}$ given by the assumption and (2.26) satisfy equation (2.2). We have already shown that $E_n(\omega)=\min_{0\leq x\leq n} E_n(x)$ for arbitrary element ω of $\{x(n)\}$. Accordingly, it is sufficient to show that for specific element $x=v_n$,

$$f_n=E_n(x)=ng(n)-N\{g(n)\}.$$

Noting that $v_n=\max\{N(i_n-1),\ n-N(i_n+1-k)\}$, we shall examine two cases:

(i) The case $N(i_n-1)>n-N(i_n+1-k)$.

We have $x=N(i_n-1)$, $N(i_n-k)\leq n-x<N(i_n+1-k)$, therefore,

$$g(x)=i_n-1+k,\quad g(n-x)=i_n.$$

(ii) The case $N(i_n-1)\leq n-N(i_n+1-k)$.

We have $x=n-N(i_n+1-k)$, $N(i_n-1)\leq x<N(i_n)$, therefore,

$$g(x)=i_n-1+k,\quad g(n-x)=i_n+1.$$

Consequently, summarizing two cases (i) and (ii) yields

(2.27) $g(x)=i_n-1+k,$

(2.28) $g(n-x)=i_n+\phi(n-x),$

where $\phi(n-x)$ is 1 if $n-x=N(i_n+1-k)$, and 0 otherwise.

Accordingly, using the assumption, (2.27) and (2.28) result in

$$f_x=(i_n-1+k)x-N(i_n-1+k),$$

$$f_{n-x}=\{i_n+\phi(n-x)\}(n-x)-N\{i_n+\phi(n-x)\}.$$

Therefore, we have

$$\begin{aligned} f_n &= E_n(x) \\ &= x+k(n-x)+f_x+f_{n-x} \\ &= (n-x)\phi(n-x)+N(i_n)-N\{i_n+\phi(n-x)\}+(i_n+k)n-N(i_n+k) \\ &= (n-x)\phi(n-x)+\{N(i_n)-N(i_n+1)\}\phi(n-x)+(i_n+k)n-N(i_n+k) \\ &= \{n-x-N(i_n+1-k)\}\phi(n-x)+(i_n+k)n-N(i_n+k) \\ &= (i_n+k)n-N(i_n+k) \\ &= ng(n)-N\{g(n)\}. \end{aligned}$$

This completes the proof.

Corollary 1. The set $\{x^*(n)\}$ of the solutions of equation (2.2) is given by

$$\{x^*(n)\}=I^+(v_n)\cap I^-(V_n+1),$$

where

$$v_n=\max\{N(i_n-1),\ n-N(i_n+1-k)\},$$

$$V_n=\min\{N(i_n)-1,\ n-1-N(i_n-k)\}.$$

Corollary 2. $g(n)=f_{n+1}-f_n,$ for $n\in I^+(1)$.

Corollary 3. The function f_n is strictly increasing and convex.

It follows from (2.4) and Corollary 2 that $g(n)=f_{n+1}-f_n>0$ and

$f_{n+1}+f_{n-1}-2f_n=g(n)-g(n-1)\geq 0.$

Corollary 4. The function f_n is linear in any interval

$$N(i)\leq n\leq N(i+1)-1.$$

The result follows immediately from Lemma 1 and Theorem 1.

Taking $k=6$ as an example, the values of $f(n)$ and $x\in\{x^*(n)\}$ are shown in Fig. 1 and Fig. 2, respectively. As mentioned above, we have a single optimal test point if $n=N(i_n)$ holds, otherwise we have multiple optimal test points. For example, if the initial n is 34, the optimal test point is uniquely determined and is 27. If the object is to the right of the test point 27, we are given the information that the object lies within the interval (27, 34), *i.e.*, an interval of length 7. We have, then, the new value of n, 7 for the second trial, and as Fig. 2 shows, the next test point should be 6. (With regard to the original interval of length 34 the second test point is 33). In the similar way we obtain the optimal policy, *i.e.*, the sequence of the optimal test points, 27, 33,

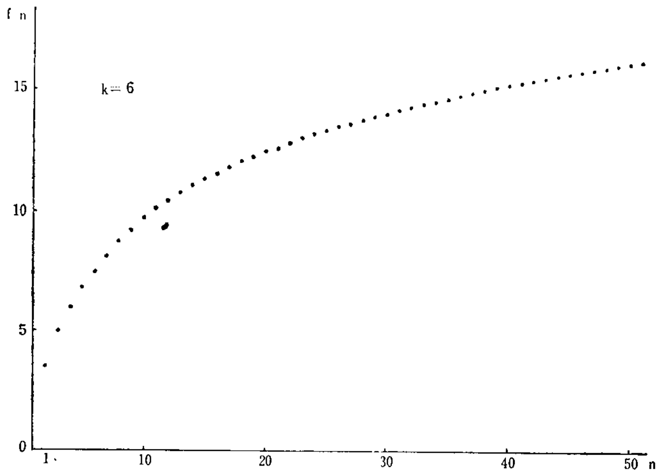

Fig. 1. Optimal value function, $f(n)$.

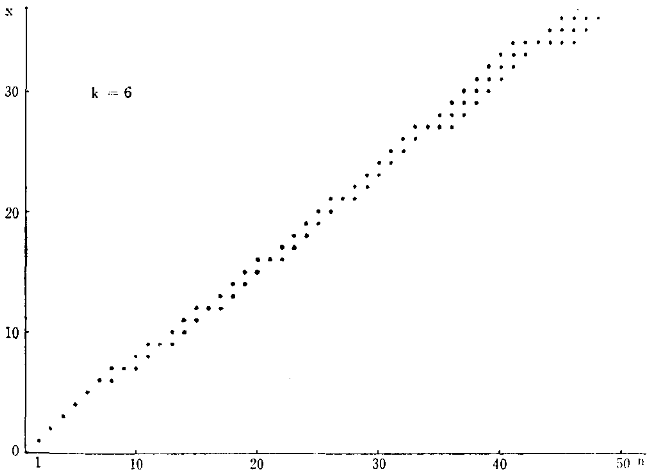

Fig. 2. Optimal test point; $x \in \{x^*(n)\}$.

32, 31, 30, 29, 28.

If the object is to the left of the test point 27, we are given the information that the object lies within the interval (0, 27). We have, then, the new value of n, 27 for the second trial, and as Fig. 2 shows, the next test point should be 21. In this way, the optimal test point is always uniquely determined. Contrary to the above case, if the initial value of n is 36, for example, the optimal test point is 27, 28 or 29 and we are given a number of optimal policies.

3. Asymptotic Formula of the Solution

In the preceding section, the exact expression of f_n was derived, but the calculation of its value is rather cumbersome especially for large n. Therefore, some approximate expression is sought in this section. First, we shall find an asymptotic formula of $N(i_n)$ as $n \to \infty$. The equation satisfied by $N(i_n)$,

$$N(i_n+k) = N(i_n+k-1) + N(i_n) \tag{3.1}$$

is a k th order recurrence equation and its characteristic equation is given by

(3.2) $$z^k = z^{k-1} + 1.$$

Let z_j $(j=0, 1, \cdots, k-1)$ be the roots of this equation. We see easily that z_j are all simple roots. Therefore the general solution of (3.1) is written as follows:

(3.3) $$N(i_n) = \sum_{j=0}^{k-1} C_j z_j^{i_n},$$

where C_j are constants. Let us examine the distribution of z_j's on the complex plane. Putting $z_j = r_j \exp i\theta_j$ where $i=\sqrt{-1}$, and substituting z_j into (3.2), we have

$$r_j^{2k} = r_j^{2k-2} + 2r_j^{k-1} \cos (k-1)\theta_j + 1,$$

therefore,

(3.4) $$r_j^k \leq r_j^{k-1} + 1,$$

with equality only if $\cos(k-1)\theta_j = 1$, that is, z_j is a positive root. However, we see easily that equation (3.2) has only one positive root $z_0 = r_0$. Hence, we have

$$r_0^k = r_0^{k-1} + 1,$$

$$r_j^k < r_j^{k-1} + 1, \qquad \text{for } j \neq 0.$$

We see from this that $r_j < r_0$ for all $j \neq 0$. Therefore, we have

(3.5) $$|z_j| < r_0, \qquad \text{for } j \neq 0.$$

Using (3.3) and (3.5), we obtain an asymptotic formula of $N(i_n)$

$$N(i_n) = C_0 r_0^{i_n} + \sum_{j=1}^{k-1} C_j z_j^{i_n},$$

$$= (C_0 + O_n) r_0^{i_n},$$

where O_n is an infinitesimal converging to 0 as $n \to \infty$. Here, letting

$r_0=1/\alpha$ we obtain

(3.6) $$N(i_n)=(C_0+O_n)\,\alpha^{-i_n},$$

(3.7) $$1-\alpha=\alpha^k.$$

Therefore,

(3.8) $$i_n=p\ln N(i_n)/C_0+O_n,$$

where

(3.9) $$p=-1/\ln\alpha.$$

Nevertheless, C_0 cannot be determined by using this method. Then, in order to determine the expression of C_0, we will seek two different asymptotic formulae of $f(n)$ in case $n=N(i_n)$. Hereafter, we put $i_n=i$ and $N(i_n)=N_i$ for simplicity. First, from Theorem 1, we have

$$f(N_i)=i+k-N_{i+k}/N_i\,.$$

Substituting (3.6) and (3.7) into this equation, we have

(3.10) $$f(N_i)=i+k-\frac{1}{1-\alpha}+O_i\,.$$

Furthermore, substituting (3.8) into this equation, we have

(3.11) $$f(N_i)=k-\frac{1}{1-\alpha}+p\ln N_i/C_0+O_i.$$

Next, we find another expression of $f(N_i)$. Applying Corollary 1, we have only one solution $x=N_{i-1}$ for $n=N_i$. Therefore, substituting $n=N_i$, $x=N_{i-1}$ and $N_i=N_{i-1}-N_{i-k}$ into (2.1), we obtain directly

(3.12) $$f(N_i)=\frac{N_{i-1}}{N_i}\cdot\{1+f(N_{i-1})\}+\frac{N_{i-k}}{N_i}\{k+f(N_{i-k})\}.$$

After several calculations, we see that

(3.13) $$\frac{N_{i-1}}{N_i}f(N_{i-1})=\alpha f(N_{i-1})+O_i.$$

In a similar way, we derive

(3.14) $$\frac{N_{i-k}}{N_i} f(N_{i-k}) = \alpha^k f(N_{i-k}) + O_i.$$

Also, we obtain easily

(3.15) $$\frac{N_{i-j}}{N_i} = \alpha^j + O_i.$$

Then, substituting (3.13), (3.14) and (3.15) into (3.12), we have

(3.16) $$f(N_i) = \tau + \alpha f(N_{i-1}) + (1-\alpha) f(N_{i-k}) + O_i,$$

where

(3.17) $$\tau = \alpha + k(1-\alpha).$$

We shall now introduce a function F_i satisfying the equation

(3.18) $$F_i = \tau + \alpha F_{i-1} + (1-\alpha) F_{i-k},$$

with $F_i = 0$ for $i \in I^+(1-k) \cap I^-(0)$.

Then, we obtain

(3.19) $$f(N_i) = F_i + O_i.$$

Next, solving (3.18) yields

$$F_i = \tau \sum_{j=0}^{i-1} \alpha^j N_{j+2-k}.$$

Thus, we have

(3.20) $$f(N_i) = \tau \sum_{j=0}^{i-1} \alpha^j N_{j+2-k} + O_i.$$

From (3.10), we see easily

(3.21) $$f(N_{i+1}) - f(N_i) = 1 + O_i.$$

Therefore, from (3.20) and (3.21) we have

(3.22) $$N_i = \tau^{-1} \alpha^{2-k-i} (1 + O_i).$$

Also, equation (3.6) can be written as $N_i = C_0 \alpha^{-i}(1+O_i)$. Comparing these equations it follows that

$$C_0 = \tau^{-1} \alpha^{2-k}.$$

Substituting this value into (3.8) and (3.11) we obtain

$$i = p \ln N_i + 2 - k + p \ln \tau + O_i, \tag{3.23}$$

$$f(N_i) = p \ln N_i + 2 - \frac{1}{1-\alpha} + p \ln \tau + O_i. \tag{3.24}$$

Similarly we have the following theorem.

Theorem 2. An asymptotic formula of $f(n)$ can be written as

$$f(n) = p \ln N_i + 2 - \frac{1}{1-\alpha} \cdot \frac{N_i}{n} + p \ln \tau + O_n. \tag{3.25}$$

Comparing this with the following asymptotic formula given by Cameron and Narayanamurthy

$$f(n) = p \ln n + C + O_n, \tag{3.26}$$

we see that (3.25) coincides with (3.26) only if $n = N_i$ and we obtain the following corollary.

Corollary 5.

$$C = 2 - \frac{1}{1-\alpha} + p \ln \tau. \tag{3.27}$$

For $k=6$, equations (3.7) and (3.9) give $\alpha = 0.778$ and $p = 3.984$. Therefore, from (3.17) and (3.27), we obtain $\tau = 2.109$ and $C = 0.467$, respectively.

Formula (3.24) is shown in Fig. 3. We see from this graph that the asymptotic formula rapidly approaches the exact values with increasing N_i.

One area applying the search procedure described in this paper is the following. Suppose we are given an expensive body the destruction

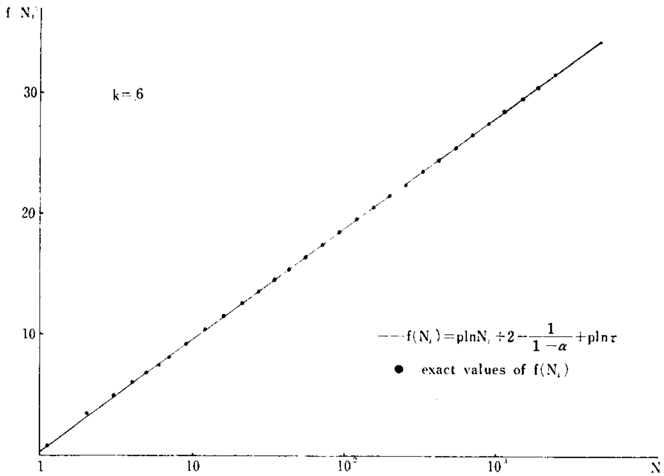

Fig. 3. Exact and approximate values of $f(N_i)$.

point of which is unknown, but to be determined by some experiments. The destruction point may be a critical load or voltage and the like such that the body will destruct if overloaded. Let us assume the destruction point lies within some interval with a uniform *a priori* probability density. We wish to locate the destruction point with a desired precision by repeated application of a load/voltage test. If the test point happens to be chosen smaller than (to the right of) the destruction point, destruction does not occur and we are given the information that the destruction point is larger than (to the left of) the test point. In this case the body can be used for the next test. On the other hand, if the test point happens to be larger than the destruction point the body destructs, giving the information that the destruction point lies below the test point. In this case the next test requires a new test body and will cost $k(>1)$ times the former case. The proper choice of the test points is then the problem discussed above.

Acknowledgement

The author is deeply indebted to Assistant Prof. T. Kisi for many helpful suggestions and advices. He would also like to express his sincere appreciation to Prof. M. Ogi for his helpful comments; and to a referee for pointing out the Morris' paper.

References

[1] Cameron, S.H., and Narayanamurthy, S.G., "A Search Problem," Operations Research, **12**, 623-629 (1964).
[2] Morris, R., "Some Theorems On Sorting," SIAM J. Appl. Math., **17**, 1-6 (1969).