

Unimodality Tests for Global Optimization of Single Variable Functions Using Statistical Methods

Ismail bin Mohd,¹ Siska Candra Ningsih² & Yosza Dasril³

Department of Mathematics, Faculty of Science and Technology,

Universiti Malaysia Terengganu (UMT),

Mengabang Telipot, 21030 Kuala Terengganu, Malaysia

E-mail: ¹ismailmd@kustem.edu.my, ²siskaningsih@yahoo.com,

³yosza@kustem.edu.my

ABSTRACT

This paper discusses the randomness and normality tests to the set of points which are obtained by splitting an interval under consideration into several subintervals for verifying that the optimization problem constitutes a Wiener process. Then, by using the optimization probabilistic algorithm, a best subinterval can be chosen for unimodality test and will be continued with a simple local search method, starting at any point in the such chosen best interval. Furthermore, numerical results are given to illustrate the tests.

Keywords: Global optimization, probability, randomness test, normality test, and unimodality test

1. INTRODUCTION

The task of global optimization of the optimization problem

$$z = f(x) \tag{1.1}$$

subject to

$$a \leq x \leq b \tag{1.2}$$

where $f : R \rightarrow R$, is to find a solution in its feasible solution set for which the objective function achieved its smallest value, the global minimum. Therefore, the global optimization aims in determining not just a local minimum but the smallest local minimum with respect to its feasible solution set which is defined by

$$F = \{x \mid a \leq x \leq b, x \in R\}. \tag{1.3}$$

In (Ismail, 2005) and (Siska, 2006) we have shown how to use the randomness and normality tests to analyze the data obtained by splitting the interval $[a, b] \subseteq X \subset R$ into several subintervals $\Delta_i = [x_i, x_{i+1}] (i = 0, \dots, n-1)$ with $x_0 = a$ and $x_n = b$ for verifying that the optimization problem constitutes a Wiener process.

In this paper, we will describe how to use the unimodality test to the chosen subinterval which is obtained by using the optimization probabilistic algorithm (Ismail, 2005), (Siska, 2006) before the (simple) local search procedure is employed.

2. STOCHASTIC PROCESS

We shall consider the Wiener process (Archetti, 1978) $f(x)$ for $x \in X = [x_0, \bar{x}]$ with

$$f_0 = f(x_0) = \mu$$

and

$$f(x) - f(y) \sim N(0, \sigma^2 |x - y|) \quad \text{for } x, y \in X. \tag{2.1}$$

Given

$$f_i = f(x_i) (i = 1, \dots, n),$$

the distribution of $f(x)$, conditioned by

$$z_n = (x_1, f_1, \dots, x_n, f_n)$$

is normal with some expected value $E[f(x) | z_n] = \mu(x)$ and variance $\text{var}(f(x) | z_n) = \sigma^2(x)$ given by

$$\mu(x) = \begin{cases} f_i \frac{x_{i+1} - x}{x_{i+1} - x_i} + f_{i+1} \frac{x - x_i}{x_{i+1} - x_i} & (f_0 = \mu, x \in \Delta_i, i = (0, \dots, n-1)) \\ f(x_n) & (\Delta_n = [x_n, x], \text{ and } x \in \Delta_n, i = n) \end{cases}$$

and

$$\sigma^2(x) = \begin{cases} \sigma^2_i \frac{(x - x_i)(x_{i+1} - x)}{x_{i+1} - x_i} & (f_0 = \mu, x \in \Delta_i, i = (0, \dots, n-1)) \\ \sigma^2(x - x_n) & (\Delta_n = [x_n, x], \text{ and } x \in \Delta_n, i = n) \end{cases}$$

respectively.

Theorem 2.1 (Archetti, 1980)

If $f(x), x \in [a, b]$, is a Wiener process such that $f(a) = f_a$, and by considering the distribution

$$f(z) = P\{\min_{a \leq t \leq b} f(t) \leq z \mid f(b) = h\},$$

then

$$f(z) = \begin{cases} 1 & (z \geq \min(f_a, f_b)) \\ \exp\left(-2 \frac{(f_a - z)(f_b - z)}{\sigma^2(b - a)}\right) & (z < \min(f_a, f_b)) \end{cases} \blacklozenge$$

3. STATISTICAL TESTS

Given $f : S \subseteq R^1 \rightarrow R^1$ and the points $x_i (i = 1, \dots, n)$ split the search interval $X \subset S$ in equal parts. Let

$$z_i = f_i - f_{i-1} \quad (i = 1, \dots, n) \tag{3.1}$$

The goodness of fit of the Wiener process for f can be checked by testing that whether $z_i (i = 1, \dots, n)$ is a random sample drawn from a normal distribution where the details are as follows.

Randomness Test

The randomness of $z_i (i = 1, \dots, n)$ can be checked by the simple correlation coefficient

$$R = \frac{\frac{1}{n} \sum_{i=1}^n iz_i - \bar{z}\bar{K}}{S_z S_1} \tag{3.2}$$

where

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i, \quad S_z^2 = \frac{1}{n} \sum_{i=1}^n (z_i - \bar{z})^2, \quad \bar{K} = \frac{n+1}{2}, \quad \text{and} \quad S_1 = \frac{n^2 - 1}{12}$$

If $z_i (i = 1, \dots, n)$ are independent on i , then the random variable

$$T_n = R \sqrt{\frac{(n-p)}{(1-R^2)}} \tag{3.3}$$

follows a t-distribution with $v = n - p$ degrees of freedom.

The tabulated t-distribution can be used to construct the statistical tests for different significance levels α and sample sizes n . The observations $z_i (i = 1, \dots, n)$ are independent if $|T_n| < T_{\alpha, v}$.

Normality Test

In order to examine the normality of $z_i (i = 1, \dots, n)$, we firstly compute:

$$\sigma_i = \max \left\{ F_0(z_{(i)}; \bar{z}; S_z) - \frac{i-1}{n}, \frac{i}{n}, \frac{i}{n} - F_0(z_{(i)}; \bar{z}; S_z) \right\} \tag{3.4}$$

where $z_{(i)} (i = 1, \dots, n)$ are values $z_i (i = 1, \dots, n)$ rearranged in increasing order, and $F_0(x; \bar{x}; s)$ is the standard normal distribution function given by

$$F_0(x; \bar{x}; s) = \int_{-\infty}^{(x-\bar{x})/s} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du, \tag{3.5}$$

and secondly, we compute

$$D_n = \max_{i=1, \dots, n} (\sigma_i) \tag{3.6}$$

The variables are drawn from a normal distribution if $D_n < D_{\alpha, n}$.

4. THE PARAMETERS

The following parameters are needed in our algorithm.

$$\mu = f(x_0), \Delta = \frac{\bar{x} - x_0}{n-1}, x_i = x_0 + i\Delta \quad (i = 1, \dots, n-1) \tag{4.1}$$

and $\hat{\sigma}$, the maximum likelihood estimate of σ , is computed as

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n-1} \frac{(f_i - f_{i-1})^2}{x_i - x_{i-1}} \tag{4.2}$$

5. STOPPING CRITERION

The stopping criterion for chosen the best subinterval is provided by Theorem 2.1. Let L_n be the sum of the intervals of local inadequacy of the model, after n function evaluations, performed in $x_i (i = 0, \dots, n-1)$ with $x_i \leq x_{i+1} (i = 0, \dots, n-2)$. At the next step, $f(x)$ is assumed unimodal in any of the intervals of L_n and a sample path of the Wiener process for $x \in X - L_n$.

Define the following probability

$$P_i = P\left\{ \min_{x \in [x_i, x_{i+1}]} f(x) \leq f_n^* - \varepsilon \mid f_i, f_{i+1} \right\}. \tag{5.1}$$

We have

$$P_i = \begin{cases} \exp\left(-2 \frac{(f_n^* - \varepsilon - f_{i+1})(f_n^* - \varepsilon - f_i)}{\sigma^2 (x_{i+1} - x_i)}\right) & ([x_i, x_{i+1}] \not\subset L_n) \\ 0 & ([x_i, x_{i+1}] \subset L_n) \end{cases} \tag{5.2}$$

where $f_n^* = \min\{f_i, i = 0, \dots, n-1\}$ and ε is the prefixed accuracy level. Finally, we compute the probability

$$PF_n = P\left\{ \min_{x \in X} f(x) > f_n^* - \varepsilon \mid z_n \right\} = \prod_{i=0}^{n-2} (1 - P_i). \tag{5.3}$$

The algorithm will terminate when $PF_n > PT$ where PT is a given probability level, and as example $PT = 0.99$. If $PF_n \leq PT$, then change the value of n .

6. THE CONTROL OF PARAMETER

The $\{\gamma_j\}$ sequence of parameters for controlling the convergency of the algorithm, can be determined by Theorem 2.1 given by the following probability.

$$P_{\gamma_j} = P\left\{\min_{x \in X} f(x) > f_n^* - \gamma_j \mid z_n\right\}. \tag{6.1}$$

The parameter γ_j is kept constant if $P_{\gamma_j} \leq P_T$ where P_T is a prefixed probability level. If $P_{\gamma_j} > P_T$, then compute

$$\gamma_{j+1} = 2 / \gamma_j. \tag{6.2}$$

7. THE CHOSEN SUBINTERVAL

A subinterval Δ_p is chosen such that the probability

$$\begin{aligned} P\left\{\min_{x \in \Delta_p} f(x) < f_{n-1}^* - \gamma \mid z_{n-1}\right\} &= \max_i P\left\{\min_{x \in \Delta_i} f(x) < f_{n-1}^* - \gamma \mid z_{n-1}\right\} \\ &= \max_i P\left\{\min_{x \in \Delta_i} f(x) < f_{n-1}^* - \gamma \mid f_i, f_{i+1}\right\} \\ &= \max_i \exp\left(-2 \frac{(f_{n-1}^* - \gamma - f_{i+1})(f_{n-1}^* - \gamma - f_i)}{\sigma^2(x_{i+1} - x_i)}\right) \\ &= \text{size}(\Delta_p) \end{aligned} \tag{7.1}$$

where $f_{n-1}^* = \min(f_0, f_1, \dots, f_{n-1})$, and γ is some positive value.

Now, divide $\text{size}(X)$ by $\text{size}(\Delta_p)$ to obtain subinterval $\Delta_j = [x_j, x_{j+1}] (j = 0, \dots, m)$ where

$$m = \left\lceil \frac{\bar{x} - x_0}{\text{size}(\Delta_p)} \right\rceil. \tag{7.2}$$

Suppose that we have

$$f_t = f(x_t) = \min(f_0, f_1, \dots, f_m). \tag{7.3}$$

Therefore, the (best) chosen subinterval is an interval which contains x_t and can be expressed by

$$I = [x_{t-1}, x_{t+1}]. \tag{7.4}$$

8. UNIMODALITY TEST

Definition 8.1

If $x^* \in (a, b)$ is a minimizer of f over (a, b) , $f(x) \geq f(x^*)$ ($x < x^*$), and $f(x^*) \leq f(x)$ ($x^* < x$) for $x \in [a, b]$, then f is said to be unimodal in $[a, b]$. ♦

Suppose that f unimodal in $I = [x_{t-1}, x_{t+1}]$ where the subinterval $I = [x_{t-1}, x_{t+1}]$ is chosen as described in Section 7. Denote $x_1^{(0)} = x_{t-1}$ and $x_2^{(0)} = x_{s+1}$ and compute $f_3^{(0)} = f(x_3^{(0)})$ and $f_4^{(0)} = f(x_4^{(0)})$ for any pair $x_3^{(0)}, x_4^{(0)}$ in $(x_1^{(0)}, x_2^{(0)})$ which satisfy $x_1^{(0)} < x_3^{(0)} < x_4^{(0)} < x_2^{(0)}$. Based on the values of $f_3^{(0)}$ and $f_4^{(0)}$ (Ismail, 1989), the width of $I = [x_{s-1}, x_{s+1}]$ can be reduced according to the following algorithm.

Algorithm (Unimodality Test)

Data: $x_1^{(0)}, x_2^{(0)}, x_1^{(0)} < x_3^{(0)} < x_4^{(0)} < x_2^{(0)}, f_3^{(0)} = f(x_3^{(0)})$, and $f_4^{(0)} = f(x_4^{(0)})$.

1. case true of

- 1.1. $f_3^{(0)} < f_4^{(0)} : [x_1^{(1)}, x_2^{(1)}] = [x_1^{(0)}, x_4^{(0)}]$ // The minimizer, $x^* \in [x_1^{(0)}, x_4^{(0)}]$
- 1.2. $f_3^{(0)} > f_4^{(0)} : [x_1^{(1)}, x_2^{(1)}] = [x_3^{(0)}, x_2^{(0)}]$ // The minimizer, $x^* \in [x_3^{(0)}, x_2^{(0)}]$
- 1.3. $f_3^{(0)} = f_4^{(0)} : [x_1^{(1)}, x_2^{(1)}] = [x_3^{(0)}, x_4^{(0)}]$ // The minimizer, $x^* \in [x_3^{(0)}, x_4^{(0)}]$
- 1.4. **default:** {} ! An error

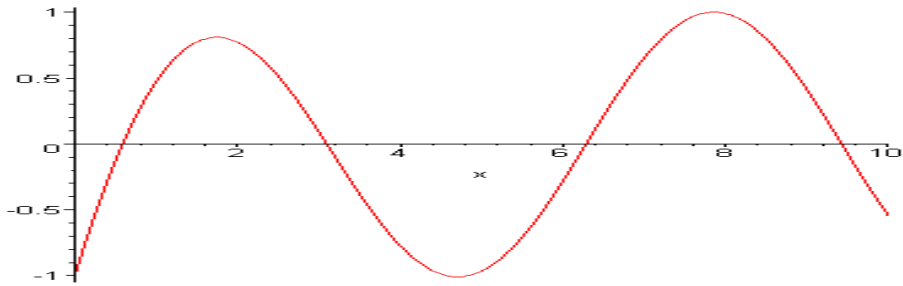
2. return. ♦

However, we do not know exactly about the unimodality property of f in the subinterval $[x_1^{(0)}, x_2^{(0)}] = [x_{s-1}, x_{s+1}]$ even though this chosen interval produced by our algorithm is assumed has fulfilled the unimodality property. Therefore, in order to obtain more evidence, during implementation, if for the sequences of points $x_1^{(0)}, x_1^{(1)}, x_1^{(2)}, \dots$ or $x_2^{(0)}, x_2^{(1)}, x_2^{(2)}, \dots$ we can get $f_1^{(i)} > f_1^{(i+1)}$ or $f_2^{(i)} > f_2^{(i+1)}$ respectively, then it can we assumed that f is uimodal in that subinterval where $f_j^{(i)} = f(x_j^{(i)})$ ($i = 0, 1, 2, \dots; j = 1, 2$). Statistically, this criteria can be achieved since the subinterval $[x_1^{(0)}, x_2^{(0)}] = [x_{t-1}, x_{t+1}]$ has been tested. This statistical tests have been done before the unimodality test.

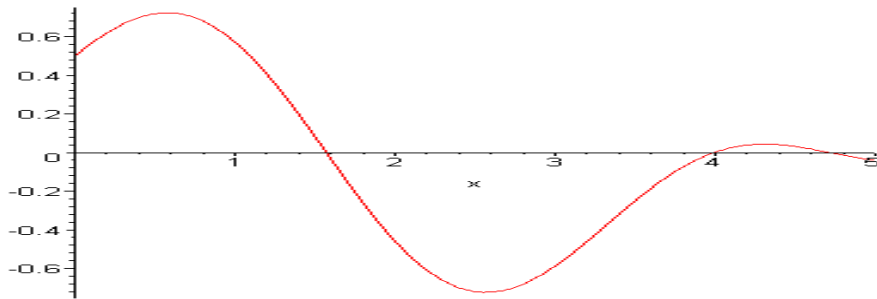
9. TEST EXAMPLES

The statistical and unimodality tests have been tested on five functions as follows:

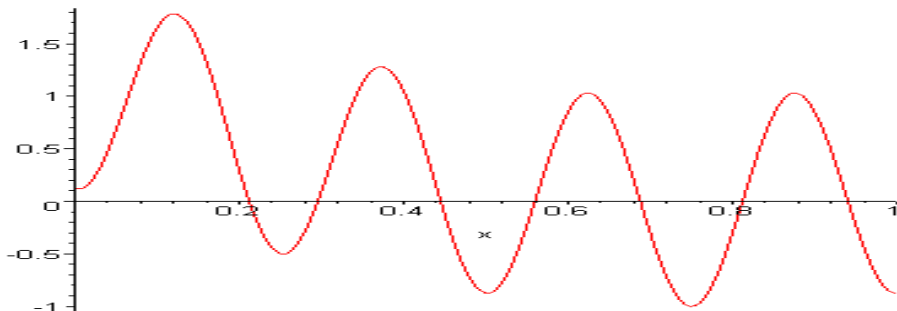
Example 1: Minimize $f(x) = \sin(x) - e^{-x}$ subject to $0 \leq x \leq 10$ where its graph is given by



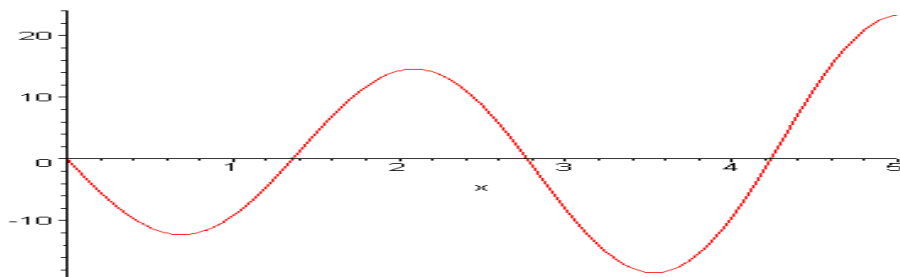
Example 2: Minimize $f(x) = 1/2 \cos(x) + 1/3 \sin(2x)$ subject to $0 \leq x \leq 5$ where its graph is given by



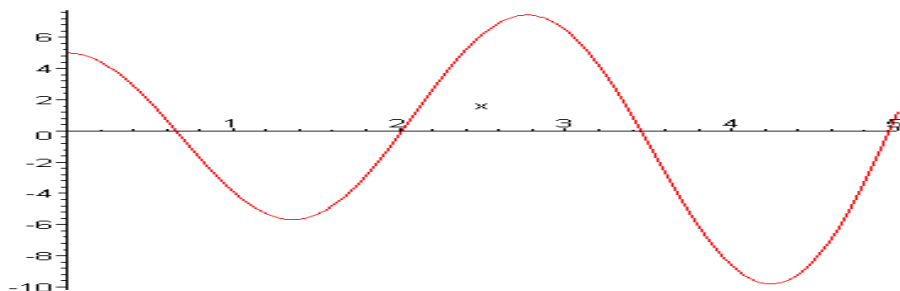
Example 3: Minimize $f(x) = 2(x - 0.75)^2 + \sin\left(8\pi x - \frac{\pi}{2}\right)$ subject to $0 \leq x \leq 1$ where its graph is given by



Example 4: Minimize $f(x) = -12\sin(2x) - 4x\cos(2x)$ subject to $0 \leq x \leq 5$ where its graph is given by



Example 5: Minimize $f(x) = x\cos(x) - x^2\sin(x)$ subject to $0 \leq x \leq 5$ where its graph is given by



10. NUMERICAL RESULTS

The results of the statistical and unimodality tests for the considered examples are given in the Tables 1.1, 1.2, 2.1, 2.2, ..., 5.2. In Tables 1.1, 2.1, 3.1, ..., and 5.1, n denotes the number of samples, T_n and D_n are computed using equations (3.3) and (3.6) respectively. The values of $T_{0.05}$, $T_{0.01}$, $D_{0.05,n}$, and $D_{0.01,n}$ are obtained from the t-distribution table and the table of Critical Values for the Liliefors test for Normality (Sheskin, 2000). Tables 1.2,

TABLE 1.1
Results of randomness and normality tests for Example 1

n	T_n	$T_{0.05}$	$T_{0.01}$	D_n	$D_{0.05,n}$	$D_{0.01,n}$
10	0.365	2.306	3.355	0.185	0.258	0.294
20	0.268	2.101	2.878	0.152	0.190	0.231
30	0.222	2.048	2.763	0.126	0.161	0.187
40	0.194	2.024	2.711	0.120	0.140	0.163
50	0.174	2.011	2.682	0.130	0.125	0.146
100	0.124	1.980	2.682	0.115	0.089	0.103

2.2, 3.2, ..., 5.2 contain the results of the unimodality test up to four iterations for each example.

TABLE 1.2
Results of unimodality test for Example 1

Iteration	0		1		2		3	
j	$x_j^{(0)}$	$f(x_j^{(0)})$	$x_j^{(1)}$	$f(x_j^{(1)})$	$x_j^{(2)}$	$f(x_j^{(2)})$	$x_j^{(3)}$	$f(x_j^{(3)})$
1	4.3178	-0.9396	4.4977	-0.9882	4.6177	-1.0054	4.6177	-1.0054
2	4.8576	-0.9973	4.8576	-0.9973	4.8576	-0.9973	4.7776	-1.0063
3	4.4977	-0.9882	4.6177	-1.0054	4.6976	-1.0090	4.6710	-1.0085
4	4.6777	-1.0087	4.7376	-1.0084	4.7776	-1.0063	4.7243	-1.0088

TABLE 2.1
Results of randomness and normality tests for Example 2

n	T_n	$T_{0.05}$	$T_{0.01}$	D_n	$D_{0.05,n}$	$D_{0.01,n}$
10	0.318	2.306	3.355	0.139	0.258	0.294
20	0.228	2.101	2.878	0.118	0.190	0.231
30	0.188	2.048	2.763	0.102	0.161	0.187
40	0.164	2.024	2.711	0.103	0.140	0.163
50	0.147	2.011	2.682	0.104	0.125	0.146
100	0.105	1.980	2.682	0.110	0.089	0.103

TABLE 2.2
Results of unimodality test for Example 2

Iteration	0		1		2		3	
j	$x_j^{(0)}$	$f(x_j^{(0)})$	$x_j^{(1)}$	$f(x_j^{(1)})$	$x_j^{(2)}$	$f(x_j^{(2)})$	$x_j^{(3)}$	$f(x_j^{(3)})$
1	2.3970	-0.6999	2.3970	-0.6999	2.3970	-0.6999	2.4985	-0.7200
2	3.0819	-0.5388	2.8536	-0.6610	2.7014	-0.7093	2.7014	-0.7093
3	2.6253	-0.7210	2.5492	-0.7236	2.4985	-0.7200	2.5661	-0.7239
4	2.8536	-0.6610	2.7014	-0.7093	2.5999	-0.7229	2.6338	-0.7202

TABLE 3.1
Results of randomness and normality tests for Example 3

n	T_n	$T_{0.05}$	$T_{0.01}$	D_n	$D_{0.05,n}$	$D_{0.01,n}$
10	0.174	2.306	3.355	0.188	0.258	0.294
20	0.104	2.101	2.878	0.167	0.190	0.231
30	0.083	2.048	2.763	0.124	0.161	0.187
40	0.714	2.024	2.711	0.153	0.140	0.163
50	0.064	2.011	2.682	0.113	0.125	0.146
100	0.045	1.980	2.682	0.103	0.089	0.103

TABLE 3.2
Results of unimodality test for Example 3

Iteration	0		1		2		3	
j	$x_j^{(0)}$	$f(x_j^{(0)})$	$x_j^{(1)}$	$f(x_j^{(1)})$	$x_j^{(2)}$	$f(x_j^{(2)})$	$x_j^{(3)}$	$f(x_j^{(3)})$
1	0.7390	-0.9619	0.7390	-0.9619	0.7455	-0.9934	0.7455	-0.9934
2	0.7680	-0.8988	0.7583	-0.9780	0.7583	-0.9780	0.7540	-0.9948
3	0.7487	-0.9994	0.7455	-0.9934	0.7497	-1.000	0.7483	-0.9991
4	0.7583	-0.9780	0.7519	-0.9989	0.7540	-0.9948	0.7512	-0.9996

TABLE 4.1
Results of randomness and normality tests for Example 4

n	T_n	$T_{0.05}$	$T_{0.01}$	D_n	$D_{0.05,n}$	$D_{0.01,n}$
10	0.354	2.306	3.355	0.153	0.258	0.294
20	0.254	2.101	2.878	0.135	0.190	0.231
30	0.209	2.048	2.763	0.119	0.161	0.187
40	0.182	2.024	2.711	0.110	0.140	0.163
50	0.164	2.011	2.682	0.111	0.125	0.146
100	0.117	1.980	2.682	0.101	0.089	0.103

TABLE 4.2
Results of unimodality test for Example 4

Iteration	0		1		2		3	
j	$x_j^{(0)}$	$f(x_j^{(0)})$	$x_j^{(1)}$	$f(x_j^{(1)})$	$x_j^{(2)}$	$f(x_j^{(2)})$	$x_j^{(3)}$	$f(x_j^{(3)})$
1	3.4119	-17.8772	3.4734	-18.3364	3.4743	-18.3364	3.5007	-18.4403
2	3.5963	-18.3052	3.5963	-18.3052	3.5553	-18.4588	3.5553	-18.4588
3	3.4743	-18.3364	3.5143	-18.4687	3.5007	-18.4403	3.5189	-18.4747
4	3.5348	-18.4816	3.5553	-18.4588	3.5280	-18.4813	3.5371	-18.4809

TABLE 5.1
Results of randomness and normality tests for Example 5

n	T_n	$T_{0.05}$	$T_{0.01}$	D_n	$D_{0.05,n}$	$D_{0.01,n}$
10	0.299	2.306	3.355	0.198	0.258	0.294
20	0.220	2.101	2.878	0.195	0.190	0.231
30	0.183	2.048	2.763	0.208	0.161	0.187
40	0.159	2.024	2.711	0.191	0.140	0.163
50	0.143	2.011	2.682	0.181	0.125	0.146
100	0.102	1.980	2.682	0.181	0.089	0.103

TABLE 5.2
Results of unimodality test for Example 5

Iteration	0		1		2		3	
j	$x_j^{(0)}$	$f(x_j^{(0)})$	$x_j^{(1)}$	$f(x_j^{(1)})$	$x_j^{(2)}$	$f(x_j^{(2)})$	$x_j^{(3)}$	$f(x_j^{(3)})$
1	8.0562	-65.1982	8.0562	-65.1982	8.1480	-65.9024	8.1480	-65.9024
2	8.4694	-63.4606	8.3316	-65.4767	8.3316	-65.4767	8.2704	-65.8994
3	8.1939	-66.0303	8.1480	-65.9024	8.2092	-66.0389	8.1888	-66.0237
4	8.3316	-65.4767	8.2398	-66.0041	8.2704	-65.8994	8.2296	-66.0235

11. CONCLUSION

From the Tables 1.1, 2.1, 3.1, ..., 5.1, for the given functions, we can say that variables are random variables since $|T_n| < T_{\alpha, \nu}$, and the values of $D < D_{\alpha, n}$ suggests that the variables are drawn from a normal distribution.

The Tables 1.2, 2.2, 3.2, ..., 5.2 and the given functions confirm that the best subinterval determined by the suggested probabilistic algorithm, is unimodal.

12. ACKNOWLEDGEMENT

The authors wish to express their appreciation to Universiti Malaysia Terengganu and the Government of Malaysia for their financial support in this project.

REFERENCES

ARCHETTI, F. and B. BETRO. 1978. *A Probabilistic Algorithm For Global Optimization*.

ARCHETTI, F. 1980. *Analysis of Stochastic Strategies For The Numerical Solution of The Global Optimization Problem*. Working paper in *Numerical Techniques For Stochastic System* North-Holland.

ISMAIL, B.M., D. YOSZA, and C.N. SISKA. 2005. *Probabilistic Algorithm for Optimization Using Newton's Method*, Refereed Conference Proceedings of Australian Operations Research Society and 11th. Australian Optimization Day, 191-234.

ISMAIL, B.M., and H.H. MALIK. 1989. *Introduction to Optimization Methods*, Universiti Pertanian Malaysia, Serdang.

SISKA, C.N., D. YOSZA, and B.M. ISMAIL. 2006. *Probabilistic Algorithm in Optimization*, MATEMATIKA, 22(1).

SHEKIN, J. D. 2000. *Handbook of Parametric and Nonparametric Statistical Procedures*. London.