

## Why Unified Statistics Theory by MCMC towards Estimation of Stationary Triple Probabilities of Any Markov Chain?

Dr.Usama H. Abou El-Enien<sup>1</sup>

<sup>1</sup>Administrative Information Systems Department, the High Institute for Tourism, Hotels & Computer, EGYPT

**Abstract:** Unified statistics theory by MCMC is considered. A new proposed algorithm is presented to obtain surely empirical analysis conclusions in order to turn to surely theoretical analysis results about stationary triple probabilities of any genetic or clonal selection algorithm of any general too large dimensional deterministic and probabilistic (the grouping data, both continuous and discrete) linear or nonlinear programming problem in order to introduce a complete framework and to estimate stationary triple probabilities by the proposed algorithm towards the fifth resounding success of unified statistics theory by MCMC.

**Keywords:** Clonal selection algorithms, Genetic algorithms, General linear programming Problem, General nonlinear programming problem, Grouping data, Unified statistics theory by MCMC

### I. INTRODUCTION

Genetic and clonal selection algorithms within the framework of Markov chains and MCMC **and estimation of stationary probabilities** have been studied by El-Nady et al. [1], [2], [3], and [4]. Unified statistics theory by MCMC (UST by MCMC) has been proposed by Abou El-Enien [5]. Unique chromosomes by simple random sampling without replacement theorem for any objective real valued function of  $n$  - variables of the form  $f(x_1, x_2, \dots, x_n)$ , where  $a_i \leq x_i \leq b_i$  for  $i = 1, 2, \dots, n$  within the framework of unified statistics theory by MCMC has been proposed by Abou El-Enien and Khalil [6]. Subset of unique chromosomes by simple random sampling without replacement theorem to solve any too large dimensional deterministic and probabilistic linear or nonlinear programming problems toward two obvious criteria (speed and accuracy) has been proposed by Abou El-Enien [7]. Estimation of stationary transition probabilities of stochastic matrix theorem for any objective real valued function (linear or nonlinear) of  $n$  - variables to estimate **stationary transition probabilities of stochastic matrix of any genetic or clonal selection algorithm** of any general too large dimensional deterministic and probabilistic linear or nonlinear programming problem **and of any Markov Chain** has been proposed by Abou El-Enien [8].

Abou El-Enien estimation of stationary joint probabilities of stochastic matrix theorem for any objective real valued function (linear or nonlinear) of  $n$  - variables to estimate **stationary joint probabilities of stochastic matrix of any genetic or clonal selection algorithm** of any general too large dimensional deterministic and probabilistic linear or nonlinear programming problem **and of any Markov Chain** has been proposed by Abou El-Enien [9].

These works **do not develop estimation of stationary triple probabilities of any genetic or clonal selection algorithm** of any general too large dimensional deterministic and probabilistic (the grouping data, both continuous and discrete) linear or nonlinear programming problem **and do not develop estimation of stationary triple probabilities of any Markov Chain** to provide a general framework. The rest of the paper is organized as follows. In Section 2, we give the formulation of the problem. In Section 3, we state the main result. Then in Section 4, the proof of the main result is given in five steps. In Section 5, we propose the algorithm. In Section 6, we give numerical example. In Section 7, we give some concluding remarks.

### II. FORMULATION OF THE PROBLEM

In this paper, we consider a problem, namely: Why unified statistics theory by MCMC towards estimation of stationary triple probabilities of any Markov Chain?. Throughout this paper, we consider any objective real valued function (linear or nonlinear) of  $n$  - variables  $f(x_1, x_2, \dots, x_n)$ ,

where  $a_i \leq x_i \leq b_i$  for  $i = 1, 2, \dots, n$  are domains of each variable  $x_i$  and  $a_i$  and  $b_i$  are real numbers:

$u \geq 0$  equations:

$$q_i(x_1, x_2, \dots, x_n) = 0 \text{ (linear or nonlinear), } i = 0, \dots, u, \rightarrow \text{ (I)}$$

and  $m - u \geq 0$  inequalities:

$$q_i(x_1, x_2, \dots, x_n) \leq 0 \text{ (linear or nonlinear), } i = u + 1, \dots, m. \rightarrow \text{ (II)}$$

**Proposition 2.1.** We restrict an arbitrary uncountable set  $S = \{a_i \leq x_i \leq b_i \text{ for } i = 1, 2, \dots, n\}$  to be a subset of  $n$ -space  $\mathbf{R}^n$  as a sample space, restrict an arbitrary countable set  $T$  to be set of all  $(x_1, x_2, \dots, x_n)$  in  $S = \{a_i \leq x_i \leq b_i\}$  for which  $P(x_1, x_2, \dots, x_n) > 0$  as a sample space (see [5]).

**Proposition 2.2.** We divide each interval  $a_i \leq x_i \leq b_i, i = 1, 2, \dots, n$  into  $k$  ( $k$  is a different optional integer number for each interval  $a_i \leq x_i \leq b_i$ ) subintervals  $a_i \leq x_i \leq c_i, c_i \leq x_{i+1} \leq d_i, \dots, w_i \leq x_{k-1} \leq y_i, y_i \leq x_k \leq b_i$  and  $c_i, d_i, \dots, w_i$  and  $y_i$  (the new population) are optional real numbers, list the possible simple random samples without replacement of  $n$  ( $n$  as  $n$  of  $S$ ) subintervals from this new population (see [6]).

**Proposition 2.3.** We restrict any simple random sample without replacement of  $n$  subintervals of Proposition 2.2, get unique chromosomes ( $= 2^k$ ),  $k$  bits  $\{0,1\}$ , substitute in (I) and (II), get subset of unique chromosomes (see [7]), generate all possible combinations of states of subset of unique chromosomes.

### III. MAIN RESULT

In this section, we shall state the main theorem.

**Theorem 3.1.** For any objective real valued function (linear or nonlinear) of  $n$ -variables  $f(x_1, x_2, \dots, x_n)$ , where  $a_i \leq x_i \leq b_i$  for  $i = 1, 2, \dots, n$  are domains of each variable  $x_i$  and  $a_i$  and  $b_i$  are real numbers:

$u \geq 0$  equations:

$$q_i(x_1, x_2, \dots, x_n) = 0 \text{ (linear or nonlinear), } i = 0, \dots, u, \text{ and}$$

$m - u \geq 0$  inequalities:

$$q_i(x_1, x_2, \dots, x_n) \leq 0 \text{ (linear or nonlinear), } i = u + 1, \dots, m, \text{ the following holds:}$$

- (1) A real valued function is one that contains all possible simple random samples without replacement of  $n$  ( $n$  as  $n$  of  $S$ ) subintervals from the new population.
- (2) Every simple random sample without replacement of  $n$  subintervals has unique chromosomes, subset of unique chromosomes and all possible combinations of states of subset of unique chromosomes.
- (3) By applying **genetic or clonal selection algorithms**. We have the probability of each ( $f_{m-1} = s, f_m = t, f_{m+1} = v$ )

$\forall m \geq 1 (P)$ , where

$$P = \frac{\text{number of times } (f_{m-1} = s, f_m = t, f_{m+1} = v) \text{ appeared}}{n - 2}$$

and have **stationary triple probabilities**.

### IV. PROOF OF THE MAIN RESULT

In this section, we prove the main result in Theorem 3.1. We start with a useful theorem.

**Theorem 4.1.** Let  $(S, \beta, P)$  be a probability space and let  $T$  denote the set of all  $x$  in  $S$  for which  $P(x) > 0$ . Then  $T$  is countable (see [10]).

We shall prove Theorem 3.1 in five steps.

**Proof of Theorem 3.1.** Step 1. For the real valued function, we define a probability space  $(S, \beta, P)$  (see [5]).

Step 2. We prove that  $T$  is a countable subset of  $S$ , and define  $n$ -dimensional random variable defined on a sample space  $T$  (see [5]).

Step 3. We divide each interval  $a_i \leq x_i \leq b_i, i = 1, 2, \dots, n$  into  $k$  subintervals, define new population, and list the possible simple random samples without replacement of  $n$  subintervals from this new population.

Step 4. We restrict any simple random sample without replacement of  $n$  subintervals, get unique chromosomes ( $= 2^k$ ),  $k$  bits  $\{0,1\}$ , substitute in (I) and (II), get subset of unique chromosomes and generate all possible combinations of states of subset of unique chromosomes.

Step 5. We pick one state randomly, apply genetic or clonal selection algorithms on the state for  $n$  -iterations, where  $n$  is a large number, count for each  $(f_{m-1} = s, f_m = t, f_{m+1} = v) \forall m \geq 1$  (see [5])

such that  $f_0, f_1, \dots$  are outcome functions

and  $s, t, v$  are any possible sequence of outcomes

the number of times it appeared, calculate the probability of each  $(f_{m-1} = s, f_m = t, f_{m+1} = v)$  ( $P$ ), where

$$P = \frac{\text{number of times } (f_{m-1} = s, f_m = t, f_{m+1} = v) \text{ appeared}}{n - 2}$$

and get the stationary triple probabilities.

On the basis of Steps 1-5, we complete the proof of Theorem 3.1.

## V. PROPOSED ALGORITHM

We prepared programs by using MATLAB 7.5. We named the proposed algorithm **About El-Enien Estimation of Stationary Triple Probabilities (About El-Enien ESTP)**, the basic steps of **About El-Enien ESTP** algorithm are as follows:

1. Divide each interval  $a_i \leq x_i \leq b_i, i = 1, 2, \dots, n$  into  $k$  subintervals, define the new population.
2. List the possible simple random samples without replacement of  $n$  subintervals from this new population. For any restricted simple random sample without replacement of  $n$  subintervals, do the following:
  - a. Input number of bits  $k$ .
  - b. Get unique chromosomes =  $2^k$ .
  - c. Substitute in (I) and (II).
  - d. Get subset of unique chromosomes.
  - e. Generate all possible combinations of states of subset of unique chromosomes and give each state a number.
  - f. Pick one state randomly.
  - g. Apply genetic or clonal selection algorithms on the state for  $n$  -iterations, where  $n$  is a large number.
  - h. Count for each  $(f_{m-1} = s, f_m = t, f_{m+1} = v) \forall m \geq 1$

such that  $f_0, f_1, \dots$  are outcome functions

and  $s, t, v$  are any possible sequence of outcomes

the number of times it appeared.

- i. Calculate the probability of each  $(f_{m-1} = s, f_m = t, f_{m+1} = v)$  ( $P$ ), where

$$P = \frac{\text{number of times } (f_{m-1} = s, f_m = t, f_{m+1} = v) \text{ appeared}}{n - 2}$$

- j. Get the stationary triple probabilities.

## VI. NUMERICAL RESULTS AND DISCUSSION

Consider the following function:

$$f(x) = x \cdot \sin(10 \pi \cdot x) + 1, x \in [-1, 2]$$

if  $k = 5$  bits,  $m = 2$  chromosomes, probability of crossover = 0.6, probability of mutation = 0.9 then

**unique chromosomes** = { 00000 = -1.000000, 00001 = -0.903226, ....., 11101 = 1.806452, 11110 = 1.903226, 11111 = 2.000000 } and globally optimum value = 1.225806.

All possible combinations of states of unique chromosomes =  $2^{k \times m} = \{ 0: (00000, 00000), 1: (00000, 00001) \dots, 1023: (11111, 11111) \}$

Pick one state randomly (state 1006), apply GAs with bit mutation on the state for  $n = 100000$  and then get the stationary triple probabilities.

1006, 17, 1006, 17, 972, ....., 63, 960, 30, 977, 55.

One of the stationary triple probabilities

$P(1006, 17, 1006) = 0.00001$ .

## VII. DISCUSSION

In this paper, the main result is About El-Enien estimation of stationary triple probabilities for any objective real valued function (linear or nonlinear) of  $n$  - variables. Using this, we propose About El-Enien estimation of stationary triple probabilities method to estimate **stationary triple probabilities of any genetic or clonal selection algorithm** of any general too large dimensional deterministic and probabilistic linear or nonlinear programming problem **and of any Markov Chain**.

## REFERENCES

- [1] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are genetic algorithms Markov?. *InterStat*, *InterStat*, (2010), no.4.
- [2] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are clonal selection algorithms Markov?. *InterStat*, *InterStat*, (2010), no.5.
- [3] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are clonal selection algorithms MCMC?. *International Journal of Open Problems in Computer Science and Mathematics*, ICSRS Publication, **4**(2011), no.3, pp. 37-55.
- [4] Kh. El-Nady, U. Abou El-Enien, A. Badr, Why are genetic algorithms MCMC2 Markov Chain Monte Carlo. *Advances in Modelling*, B *Signal Processing and Pattern Recognition*, **54**(2011), no 1-2.
- [5] U. H. Abou El-Enien, *A New Unified MCMC Methods Toward Unified Statistics Theory by MCMC*, LAP, Germany, 2012.
- [6] U. Abou El-Enien, M. Khalil, Why unified statistics theory by MCMC towards resoundingsuccess ?. *Indian Journal of Computer Science: Theory and Practical.*, GBS Publishers & Distributors (I), **1**(2013), no.1, pp. 13-16.
- [7] U. Abou El-Enien, Why unified statistics theory by MCMC towards linear and nonlinear programming problems?, *International Journal of Research in Business and Technology*, Techmind Research, **4** (2014), no. 3, pp. 531-533.
- [8] U. Abou El-Enien, Why unified statistics theory by MCMC towards estimation of stationary transition probabilities of stochastic matrix?, *Journal of Computer and Mathematical Sciences*, **6**(2015), no. 7, pp. 411-415.
- [9] U. Abou El-Enien, Why Unified Statistics theory by MCMC towards Estimation of Stationary Joint Probabilities of Stochastic Matrix?, *International Journal of Engineering Research And Management*, **3**(2016),no. 3, pp. 100-102.
- [10] T. Apostol, *Calculus*, vol.II, John Wiley, Inc., 1969.

## Author



**Dr. Usama H. M. Abou El-Enien** received the Ph.D. degree from Alexandria University in 2010. He is currently assistant professor in the department of Administrative Information Systems at The High Institute for Tourism, Hotels & Computer. His major research interests in Markov Chains, MCMC and **Unified statistics theory by MCMC**.