ONE-PASS PROCEDURES OF UNEQUAL PROBABILITY SAMPLING

by

Kwok-fai LEE

A

Thesis

Submitted to

(Division of Statistics)

The Graduate School

of

The Chinese University of Hong Kong

In Partial Fulfilment

of the Requirements for the Degree of

Master of Philosophy

(M. Phil.)

May, 1993

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undesigned certify that we have read a thesis, entitled "One-pass procedures of unequal probability sampling" submitted to the Graduate School by Lee Kwok-fai (李國輝) in partial fulfilment of the requirements for the degree of Master of Philosophy in Statistics. We recommend that it be accepted.

Kim-Hung Li

Dr. K.H. Li,
Supervisor

Lee Sik-Yum

Dr. S.Y. Lee

Dr. K.H. Wu

Dr. S.K. Tse,
External Examiner

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

ACKNOWLEDGEMENT

I would like to express my sincere thanks to my supervisor, Dr. Kim-hung Li, for his encouragement and supervision during the course of this research programme. It is also a pleasure to express my gratitude to the entire staff of the Department of Statistics for their kind assistance.

ABSTRACT

Various sampling methods have been proposed for weighted sampling without replacement. For the ease of computer implementation, it is beneficial if the algorithm can draw the sample in one-pass. In this thesis, existing one-pass algorithms will be discussed. A new one-pass algorithm, which is an extension of Chao's (1982) algorithm, is proposed. The new one-pass algorithm generates weighted sample with positive second order inclusion probabilities whenever such a sample exists.

Keywords: Sampling without replacement; Unequal probability sampling; Sequential sampling.

CONTENTS

CHAPTER 1. INTRODUCTION

§1.1 Unequal probabilities sampling schemes without replacement

There are lots of works involving sampling from finite population. We are not only interested in the methods how the sample units are drawn from the population, but also in the inferences that are going to be made from the sample to the population. Given a population of N units, we are interested in estimating the population totals, means or ratios from a sample. The most commonly used sampling procedure to select n ($\le$ N) units with equal probabilities is the simple random sampling method. However, if W_i , $i = 1, \ldots, N$ are known to each population unit and it is believed that W_i is correlated with the interesting variable Y_i, sampling procedure with equal probabilities may not be appropriate. It does not take into account the information available in $\{ W_i \}$. Thus, it arouses the development of the unequal probabilities sampling procedures in which we make use of the auxiliary variable W_i. For simplicity, we assume that W_i is strictly positive and is a measure of the "importance" of Y_i in the sense that Y_i is approximately proportional to W_i. We call W_i the weight of Y_i. A common example of $\{ W_i \}$ in business survey is the company size. The sampling procedures with unequal probabilities mainly consist of two categories. Algorithms in the first category draw units with the inclusion probabilities proportional to their weights while those in second category do not have this requirement. The former is often used in unequal probability sampling survey and the well-known Horivitz-Thompson estimator of the population total is commonly used.

Unequal probabilities sampling procedures offer several advantages. Similar to ratio estimation, it efficiently makes use of the supplementary information. It also resembles more or less a optimally allocated sample drawn from the population stratified by the weights. However, it does not need the stratification. Besides, this technique is also frequently used in selecting the sampling units in the multistage sampling.

Hansen and Hurwitz (1943) demonstrate that the use of unequal selection probabilities frequently can improve the efficiency of our estimator of total. It is also well-known that by assigning varying probabilities of selection to the population units, it is possible to reduce considerably the sampling error of the estimates over those from sampling with equal probabilities (Raj, (1956)). Although it involves mathematical and computational difficulties in the development of the theory in unequal probabilities sampling without replacement, a general theory was first given by Horvitz and Thompson (1952) and various researchers have begun to study on this topic. Hanif and Brewer (1983) list fifty unequal probabilities sampling procedures and Chaudhuri and Vos (1988) also give a review of the sampling procedures from the finite populations with unequal probabilities.

§1.2 Estimation Problems in unequal probabilities sampling scheme without replacement

Denote the population size as N and the sample size as n ($1 < n \le N$). Let Y_i be quantity associated with the i^{th} units and W_i ($W_i > 0$) be the corresponding weight. Besides, denote the probability of inclusion of the i^{th} unit as π_i, $i = 1, \ldots, N$ and the joint probability of inclusion of the i^{th} and j^{th} units as π_{ij}, $i \neq j$, $i,j = 1, \ldots, N$. Horvitz and Thompson (1952) provided an unbiased estimator $\hat{Y}_{HT}$ to estimate $Y = \sum_{i=1}^{N} Y_i$ in the unequal probabilities sampling without replacement. For weighted sampling with probabilities proportional to weight, we have $\pi_i = n\, W_i \,/\, T$ where $T = \sum_{i=1}^{N} W_i$. Let S = { i: Y_i is in the sample of size n } . The estimator is given by

$$\hat{Y}_{HT} = \sum_{i \in S} \frac{Y_i}{\pi_i} .$$

The following is a summary of the general estimation theory for selection with probabilities proportion to weight without replacement (Horvitz and Thompson (1952)). For fixed n, the variance of the unbiased estimator $\hat{Y}_{HT}$, var($\hat{Y}_{HT}$), is

$$\operatorname{var}(\hat{Y}_{HT}) = \sum_{i=1}^{N} Y_i^2 \frac{1-\pi_i}{\pi_i} + \sum_{\substack{i,j=1 \\ i \neq j}}^{N}\sum Y_i Y_j \frac{\pi_{ij} - \pi_i\pi_j}{\pi_i\pi_j}.$$

For $\pi_{ij} > 0$, $1 \leq i \neq j \leq N$, there are two unbiased estimators, v_1, proposed by Horvitz and Thompson (1952), and v_2, proposed by Sen (1953) and Yates and Grundy (1953), for $\operatorname{var}(\hat{Y}_{HT})$, where

$$v_1 = \sum_{i \in s} \frac{Y_i^2}{\pi_i^2}(1 - \pi_i) + \sum_{\substack{i,j \in s \\ i \neq j}}\sum \frac{Y_i\, Y_j}{\pi_i\, \pi_j} \frac{\pi_{ij} - \pi_i\pi_j}{\pi_{ij}} \tag{1.1}$$

and

$$v_2 = \frac{1}{2} \sum_{\substack{i,j \in s \\ i \neq j}}\sum \left\{ \frac{Y_i}{\pi_i} - \frac{Y_j}{\pi_j} \right\}^2 \frac{\pi_i\pi_j - \pi_{ij}}{\pi_{ij}}. \tag{1.2}$$

Both (1.1) and (1.2) can assume negative value, but (1.1) takes a negative value more frequently than (1.2). (1.2) has also performed much better than (1.1) in a number of empirical comparisons, commencing with that in Yates and Grundy's (1953) paper.

If π_{ij} is zero for some $i \neq j$, there is no unbiased variance estimator based on the sample. Thus, in order to have unbiased variance estimator, it is necessary to find a procedure which guarantees that all second order inclusion probabilities are strictly positive.

§1.3 Classification of Unequal probabilities sampling schemes without replacement

Brewer (1983) classifies the sampling procedures in a number of ways and most of the procedures select a sample of size two. One of the classifications is by the manner of selections. It includes five main kinds of procedures.

(i) Systematic procedure

It involves the ordering of the population and the inclusion probabilities are cumulated at each selection of a unit. A random number r ($0 < r \le 1$) is chosen and the j selected unit is the unit whose cumulated value of π_i are the smallest one that is greater than or equal to each of r+j-1.

For example, Madow (1949) proposes an ordered systematic procedure. It arranges the population units with any order and obtains the skip interval I (= T / n), where T is the total weights of the population units and n is the sample size. Also, choose a random start $0 \le s < I$. The first unit selected is that for which the cumulated weight is the smallest one greater than or equal to s. In general, the $(k+1)^{th}$ unit selected is that for which the cumulated weight is the smallest one greater than or equal to s + k I. However, this method may produce a sample with replicates. Consider that the $(k+1)^{th}$ unit is selected and m ($\ge$ 1) intervals are skipped. If the cumulated weight which is the smallest one greater than or equal to s + (k+m) I is the same, the same unit is selected.

This kind of procedure is usually simple. However, since the population units are ordered, π_{ij} may be zero for some pairs. There are difficulties in calculating the joint probabilities of inclusion for the purpose of estimating the variance.

(ii) Draw-by-draw procedure

At each draw, one unit is selected from the population units which are not previously selected. Their probabilities of selections are re-calculated at each draw so that their total probabilities of selection, all over the sample drawn, is proportional to their weights. The idea is direct and simple. But, to draw each unit, it is often required to re-calculate the working probabilities - the probabilities of selection at each successive draw according to the outcomes of the previous draws. The working probabilities are often not simple and the formula for π_i and π_{ij} become rapidly complicated for sample size greater than two.

(iii) Rejective procedure

At each draw, one unit is drawn from the population units with replacement. If any unit is re-selected, all previously selected sample units are abandoned and the process is continued until a sample of size n, with no duplicated selections, is obtained. The procedure can make use of the properties of the sampling with replacement. Nevertheless, similar to draw-by-draw procedure, it requires to calculate the working probabilities, which may not be compact. It also requires much time to select a complete sample as partial sample is discarded during the sampling procedure. The procedure is tedious if there are a large number of the population units.

(iv) Whole sample procedure

Unlike previous procedures, a number of whole sample of size n are given out. The probability for each of the possible sample is chosen so that the probabilities of the selection of each sampling unit in the sample are proportional to weight. Therefore, once all the probabilities of all possible samples have been calculated, the selection procedure is easy. But, it may find difficulties to calculate all these probabilities as the sample size increases.

(v) Others

It includes the procedures that are not classified as one of the above four main methods of selections.

According to the manner of selections, some procedures can draw a complete sample with the data read through only one time and they are called the one-pass algorithms. The characteristics of one-pass algorithm and existing one-pass algorithms will be discussed in Chapter 2.

When the joint inclusion probabilities of all possible combinations of units for two procedures are the same, they are said to belong to the same equivalence class. Each of the procedures discussed in (i), (ii), (iii) and (v) belongs to the same equivalence class with a corresponding whole sample procedure.

CHAPTER 2. ONE-PASS ALGORITHMS

§2.1 Characteristics of one-pass algorithms

Usually, the data are stored in a sequential file. One-pass algorithm makes use of the sequential nature of the file. It draws the sample with the data are read through only one time. Usually, the input and output (I/O) process of the file is time-consuming. If the algorithms require to read the file many times, much of the time is spent on input and output processing. One-pass algorithm, however, saves the I/O time in handling the file. It is also more efficient in file management than other non-one-pass algorithms.

During the sampling procedures, some values are stored up for calculation. One-pass algorithm only requires finite storage to retain relevant information. The storage required remains the same even if there is a large amount of records in the data file. Since the space requirement for the sampling algorithm is known in advance , one would make a better allocation on the use of storage before processing the data.

One-pass algorithms are usually defined by induction and the formulas are written in general form and are less complicated. With this characteristic, one-pass algorithms are most suitable to be implemented on the computer and they are applicable in our daily-life situation.

§2.2 Existing one-pass algorithms

Not all of the procedures can be implemented as one-pass algorithms. There are only a few one-pass algorithms.

§2.2.1 Chao's algorithm

Chao (1982) proposes a sampling algorithm which selects n items from $\{ Y_1, \dots , Y_N \}$ without replacement and the population size N does not have to be known in advance. Let $W_1, W_2, \dots , W_N$ (> 0) be the corresponding weights of the units. The algorithm is constructed via induction. Let $S_k = \{ Y_1, \dots, Y_k \}$ and a sample of size n is selected from S_k without replacement. Furthermore, let $\pi(k; i)$ be the probability that Y_i is selected, that is, the first order inclusion probability with respect to S_k. Then

$$\sum_{i=1}^{k} \pi(k; i) = n. \tag{2.1}$$

If it is required that

$$\pi(k; i) \propto W_i \qquad \text{for } 1 \le i \le k, \tag{2.2}$$

some $\pi(k; i) \left(= \dfrac{n W_i}{\sum_{j=1}^{k} W_j} \right)$ may be greater than one. In this case, Y_i with the largest $\pi(k; i)$ is sampled out with probability one. This unit is called self-selective unit. Reconsider (2.1) and (2.2) with both n and k reduced by one until all $\pi(k; i)$'s are less than or equal to one. Given the triangular array $\{ \pi(k; i);\ i = 1, \dots, k;\ k \ge n \}$, the following four sequences of sets ($k \ge n$) are defined:

$$E_k = \{ i: \pi(k; i) = \pi(k+1; i) = 1; i \le k \},$$

$$F_k = \{ i: \pi(k; i) = 1, \pi(k+1; i) < 1; i \le k \},$$

$$G_k = \{ i: \pi(k; i) < 1, \pi(k+1; i) < 1; i \le k \},$$

$$H_k = \{ k+1 \}.$$

Let I_k be the number of units in $E_k \cup F_k$. The I_k units in $E_k \cup F_k$ should be sampled with probability one at stage k. Let $U_k = \pi(k+1; k+1)$. For j in F_k, define $T_{kj} = \{ 1 - \pi(k+1; j) \} / U_k$. Let $T_k = \sum_{j \in F_k} T_{kj}$, with the convention that $T_k = 0$ if F_k is empty. Also define

$$R_{kj} = \begin{cases} 0 & (j \in E_k) \\ T_{kj} & (j \in F_k) \\ (1 - T_k)/(n - I_k) & (j \in G_k) \end{cases}$$

It is shown that for $1 \le i \le k$ and $k \ge n$,

$$\pi(k+1; i) = (1 - U_k R_{ki}) \pi(k; i), \qquad \pi(k+1; k+1) = U_k,$$

where R_{ki} is the conditional probability that Y_i $(1 \le i \le k)$ is removed, given that Y_{k+1} is selected. It is also shown that for all $m \le n$ and all $1 \le i_1 < \dots < i_m \le k$, $k \ge n$,

$$\pi(k+1; i_1, \dots, i_m) = (1 - U_k \sum_{j=1}^{m} R_{k i_j}) \pi(k; i_1, \dots , i_m)$$

and for $1 \le i_1 < \dots < i_{m-1} \le k$,

$$\pi(k+1; i_1, \dots, i_{m-1}, k+1) = U_k (1 - \sum_{j=1}^{m-1} R_{k i_j}) \pi(k; i_1, \dots , i_{m-1}).$$

For m=2, we have $\pi(k; i,j) \le \pi(k; i) \pi(k; j)$ for $1 \le i < j$. The procedure is:

(i) Select Y_{k+1} with probability U_k.

(ii) If Y_{k+1} is not selected, retain the sample at stage k, which is of sample size n.

(iii) If Y_{k+1} is selected, replace an unit in the sample, say K, by Y_{k+1} to form the new sample at stage k+1, which is still of sample size n. The unit K is selected as follow. With probability T_{ki}, select the i^{th} unit from F_k; if no unit from F_k is selected, select one unit at random from those n - I_k remaining units in the sample.

§2.2.2 Other algorithms

Richardson (1989) proposes a one-pass weighted-selection procedure. For m > n, a sample cannot be drawn from (Y_1, ... , Y_m) unless

$$n W_i / T_m < 1 \text{ for } i = 1, \ldots, m \qquad (2.3)$$

where $T_i = \sum_{j=1}^{i} W_j$. The method first sets up a pool of n units as follow.

(i) Read a chosen number, say m,(m > n) of units from the list and check that (2.3) holds. If it does not, increase m until it does.

(ii) Using any standard method to draw a sample of size n proportional to weights without replacement out of the list in (i).

After setting up the pool, the $(k+1)^{th}$ unit is examined. Evaluate the probability $n W_{k+1} / T_{k+1}$. If this value is greater than one, the unit is put on one side for re-examination until all units in the main list has been exhausted. If not, admit the $(k+1)^{th}$ unit to the pool with the evaluated probability and choose one of the current pool unit with probability 1/n and delete it when the $(k+1)^{th}$ unit is admitted.

This method may not work for some combinations of weights. For example, assume that a sample of size two is drawn from four population units with weights 1, 2, 3 and 4. The condition (2.1) is satisfied when m = 4. However, it provides no method how the initial pool of sample size n is set. Besides, as some units are put on one side for re-examinations, the storage requirement is not known in advance.

Schultz (1990) proposes an ordered random selection with probabilities proportional to weight without replacement. It assumes that $W_i \le T_N / n$, where T_N is the total weights of N population units and n is the sample size. Then

(i) Set i = 0

(ii) Set i = i+1 and generate z, a uniform random variate on (0,1).

(iii) If $z \le n\, W_i / T_N$, select unit i. Return to step (ii).

Each unit is selected with inclusion probability proportional to its weight and the expected sample size is n. Besides this method, Schultz also proposes a modified one-pass method such that a sample proportional to weight is drawn out with an exact sample size n with a very high probability. Schultz finds that the probability of not reaching a sample of size n depends on the ordering, and differences in weights of successive population units, but, it can be expected to be low. Again, the total of all weights in the population, T_N, should be known in advance.

Li (1992) introduces a computer implementation of the Yates-Grundy draw-by-draw procedure. It is a sequential method that the total weight is not known in advance and the sample, of which the inclusion probabilities are not proportional to weights, can be drawn in one-pass. When the sample size n is large and n/N is small, it is shown that Li's algorithm is superior to the original Yates-Grundy algorithm in comparing the efficiency of the algorithms.

§2.3 Second order inclusion probabilities

Bethlehem and Schuerhoff (1984) provides the following necessary and sufficient conditions for the strictly positive second inclusion probabilities of the sample drawn by Chao's algorithm.

Theorem 2.1 (Bethlehem and Schuerhoff (1984))

For Chao's algorithm, all the second order inclusion probabilities are strictly positive if, and only if at each stage k ($n < k \le N$), the sample contain at most n-2 self-selective units.

Chao's method does not always guarantee that all the second order inclusion probabilities are positive. Here is an example.

Example:

Suppose the population consists of four units, Y_1, Y_2, Y_3 and Y_4, with weights $W_1 = 8$, $W_2 = 5$, $W_3 = 2$ and $W_4 = 2$. A sample of size two is selected. Since $\pi_i = \frac{n W_i}{\sum_{j=1}^{N} W_j}$, we have $\pi_1 = \frac{4}{17}$, $\pi_2 = \frac{4}{17}$,

$\pi_3 = \frac{10}{17}$ and $\pi_4 = \frac{16}{17}$. It can be verified that sample with the following second order inclusion probabilities exists.

$\pi_{1,2} = \frac{28}{51}$, $\pi_{1,3} = \frac{10}{51}$, $\pi_{1,4} = \frac{10}{51}$, $\pi_{2,3} = \frac{1}{51}$, $\pi_{2,4} = \frac{1}{51}$ and

$\pi_{3,4} = \frac{1}{51}$.

All the second order inclusion probabilities are strictly positive. Nevertheless, it is impossible to use Chao's method to draw a sample with positive second order inclusion probabilities. Consider the first three units. $\pi(3;\ 1) = \frac{16}{15} > 1$ and thus Y_1 is self-selective unit. Thus, unit 2 and unit 3 cannot be included in the sample simultaneously. Based on Chao's algorithm, no sample with positive second order inclusion probabilities is drawn out. There is also no unbiased variance estimation.

We have examined the characteristics of one-pass algorithms and some existing algorithms. Clearly, one-pass algorithm is useful when it handles a sequential file and is implemented on the computer. Though not many of the existing procedures have computer one-pass implementation, there are still much rooms for development. Besides, the algorithm proposed by Chao has certain advantages. If the existing algorithm can be extended such that it can handle the problem of the strictly positive second order inclusion probabilities, it will be more helpful in estimating the variance. In this thesis, we propose an algorithm such that the problem of strictly positive second order inclusion probabilities can be handled.

CHAPTER 3. A NEW ONE-PASS ALGORITHM

§3.1 Introduction

For simplicity, we use the term "weighted sample" to stand for weighted sample without replacement and with inclusion probability proportional to weight. Suppose the population is stored in a sequential file. There are totally N records from which a weighted sample of size n ($1 < n \le N$) is selected. Let R_i be the i^{th} record and W_i be its corresponding weight. Each W_i will be divided into two parts, $W_i^{(1)}$ and $W_i^{(2)}$ so that $W_i = W_i^{(1)} + W_i^{(2)}$, $W_i^{(1)} > 0$ and $W_i^{(2)} \ge 0$. We will create two sets. As the contents of the sets change as we pass through the file, we denote them as $S_{1,k}$ and $S_{2,k}$, where $S_{1,k} = \{(R_i, W_i^{(1)}): i \in L_k\}$ and $S_{2,k} = \{(R_i, W_i^{(2)}): i \in L_k\}$ where $L_k \subseteq \{1,\dots,k\}$, $1 \le k \le N$. We refer k as the stage number. Denote the maximum weights in each of these two sets as $M_k^{(j)} = \max\{ W_i^{(j)}: i \in L_k \}$, $j = 1,2$ and the corresponding maximum weight in the set L_k as $M_k = \max\{W_i: i \in L_k\}$. Also, denote the total weight in each of these two sets as $T_k^{(j)} = \sum_{i \in L_k} W_i^{(j)}$, $j = 1,2$, and the total weights in these two sets as $T_k = T_k^{(1)} + T_k^{(2)}$.

To make the algorithm to be one-pass, both $S_{1,k}$ and $S_{2,k}$ are not needed to be actually stored. Chao's algorithm is applied to $S_{1,k}$ and $S_{2,k}$ separately and thus after each stage only a sample of size n from $S_{1,k}$ and another sample of size n from $S_{2,k}$ are stored. We require $S_{1,N}$ to be such that a weighted sample of size n with weights $\{W_i^{(1)}\}$ and with strictly positive second order inclusion probability can be drawn from $S_{1,N}$. In addition, a weighted sample of size n from $S_{2,N}$ with weights $\{W_i^{(2)}\}$ exists. To achieve these goals, some criteria are required to form $S_{1,k}$'s and $S_{2,k}$'s. Sometimes a record with 'large' weight is read in and its weight cannot be divided into two parts without violating the above requirement on $S_{1,k}$'s and $S_{2,k}$'s. The record will be stored in a new set $S_{3,k} = \{ R_i : 1 \le i \le k \text{ and } i \notin L_k \}$. Let J_k be the size of $S_{3,k}$. In the algorithm, for all $R_j \in S_{3,k}$, $W_j = V_k \equiv \max_{1 \le i \le k} \{W_i\}$. At each stage k, we are interested in keeping the least total weights that have to be read in order to ensure that the sample exists. We define $C_k = n M_k - T_k$ and $C_k^* = n V_k - T_k$. It follows that $C_k^* = C_k + n (V_k - M_k) \ge C_k$. Clearly, when $J_k = 0$, we have $V_k = M_k$ and $C_k^* = C_k$. If $C_N \le 0$, the weighted sample drawn from the population exists. With the similar argument for $S_{1,k}$ and $S_{2,k}$, we define $A_k = n M_k^{(1)} - T_k^{(1)}$ and $B_k = n M_k^{(2)} - T_k^{(2)}$.

Therefore, at the end of stage k, the i^{th} record ($i \le k$) will either have its weight divided into two parts $W_i^{(1)}$ and $W_i^{(2)}$ and $(R_i, W_i^{(1)})$ and $(R_i, W_i^{(2)})$ be included in $S_{1,k}$ and $S_{2,k}$ respectively or have R_i added to $S_{3,k}$. The requirements to be fulfilled at the end of the stage k , where L_k contains at least n+1 elements, are listed as follows. There are only two possible cases at the end of stage k.

CASE I.

I1. $J_k = 0$.

I2. $M_k = M_k^{(1)} + M_k^{(2)}$.

I3. $A_k < (n-2)\, M_k^{(1)}$.

I4. Suppose $(\alpha-1)\, M_k \le C_k < \alpha\, M_k$, for an integer α.

If $C_k \ge 0$,

then $(\alpha-1)\, M_k^{(1)} \le A_k < \alpha\, M_k^{(1)}$,

$(\alpha-1)\, M_k^{(2)} \le B_k < \alpha\, M_k^{(2)}$.

else (i.e. $C_k < 0$)

$A_k < 0$, $B_k < 0$.

I5. Further, if $(\alpha-1)\, M_k < C_k < \alpha\, M_k$ and $C_k > 0$, we additionally require $(\alpha-1)\, M_k^{(1)} < A_k$.

CASE II.

II1. $0 < J_k < n$.

II2. $M_k = M_k^{(1)} + M_k^{(2)}$.

II3. $A_k < (n-2)\, M_k^{(1)}$.

II4. $(n-1)\, V_k \le C_k^* < n\, V_k$.

II5. $(n-1)\, M_k^{(2)} - (V_k - M_k) \le B_k < n\, M_k^{(2)}$.

II6. $(n-1)\, M_k^{(1)} - (V_k - M_k) < A_k < n\, M_k^{(1)}$.

II7. For all $j \in S_{3,k}$, $W_j = V_k$.

§3.2 Examination of all possible cases

Suppose the above requirements are satisfied at the end of stage k, where the size of L_k is at least n+1. When W_{k+1} is read in, we have the following fifteen possible cases. We will show that for each case there is a way to ensure that the above requirements remain to be satisfied at the end of stage k+1.

CASE 1 : $J_k = 0$, $C_k < W_{k+1} \le M_k$.

CASE 2 : $J_k = 0$, $(\alpha-1)\, M_k + W_{k+1} \le C_k < \alpha\, M_k$, for a positive integer α.

CASE 3 : $J_k = 0$, $W_{k+1} \le M_k$, $\alpha\, M_k \le C_k < \alpha\, M_k + W_{k+1}$, for a positive integer α.

CASE 4 : $J_k = 0$, $M_k < W_{k+1} < T_k / (n-1)$.

CASE 5 : $J_k = 0$, $M_k < T_k / (n-\alpha) \le W_{k+1} < T_k / (n-\alpha-1)$, for a positive integer $\alpha \le n-2$.

CASE 6 : $J_k = 0$, $T_k / (n-\alpha) \le M_k < W_{k+1} < T_k / (n-\alpha-1)$, for a positive integer $\alpha \le n-2$.

CASE 7 : $J_k = 0$, $W_{k+1} > M_k$, $W_{k+1} \ge T_k$.

CASE 8 : $0 < J_k < n-1$, $W_{k+1} = V_k$.

CASE 9 : $J_k = n-1$, $W_{k+1} = V_k$.

CASE 10 : $0 < J_k \le n-1$, $W_{k+1} \le M_k$, $W_{k+1} \le V_k - T_k$.

CASE 11 : $0 < J_k \le n-1$, $V_k - T_k < W_{k+1} \le M_k$.

CASE 12 : $0 < J_k \le n-1$, $M_k < W_{k+1} \le V_k - T_k$.

CASE 13 : $0 < J_k \le n-1$, $M_k < W_{k+1}$, $V_k - T_k < W_{k+1} < V_k$.

CASE 14 : $0 < J_k \le n-1$, $M_k < V_k < W_{k+1}$, $W_{k+1} \ge T_k + J_k V_k$.

CASE 15 : $0 < J_k \le n-1$, $M_k < V_k < W_{k+1} < T_k + J_k V_k$.

Let us examine each case.

CASE 1 : $J_k = 0,\ C_k < W_{k+1} \le M_k$.

Choose $W_{k+1}^{(1)}$ such that

$$\max\left(A_k + \delta_1 ,\ \delta_2 ,\ W_{k+1} - M_k^{(2)} \right) \le W_{k+1}^{(1)} \le \min\left(W_{k+1} - B_k - \delta_3 ,\ M_k^{(1)},\ W_{k+1} \right) \qquad (3.1)$$

where

$$0 < \delta_1 < \min\left((W_{k+1} - C_k)/2,\ M_k^{(1)} - A_k,\ W_{k+1} - A_k \right),$$

$$0 < \delta_2 < \min\left(M_k^{(1)},\ W_{k+1},\ (W_{k+1} - B_k)/2 \right)$$

and

$$0 < \delta_3 < \min\left((W_{k+1} - C_k)/2,\ M_k^{(2)} - B_k,\ (W_{k+1} - B_k)/2 \right).$$

To prove that δ_1, δ_2 and δ_3 exist, we show that each component in the right hand side is strictly greater than zero.

(1) $(W_{k+1} - C_k)/2 > 0$

(2) $M_k^{(1)} - A_k > 0$

(3) $W_{k+1} - A_k > 0$ because if $C_k < 0$, then $A_k < 0$ and hence, $W_{k+1} - A_k > 0$; and if $C_k \ge 0$, then $A_k \ge 0$ and $B_k \ge 0$, and thus $W_{k+1} - A_k > C_k - A_k = B_k \ge 0$.

(4) $M_k^{(1)} > 0$

(5) $W_{k+1} > 0$

(6) $(W_{k+1} - B_k) / 2 > 0$ because if $C_k < 0$, then $B_k < 0$ and hence, $W_{k+1} - B_k > 0$; and if $C_k \ge 0$, then $A_k \ge 0$ and $B_k \ge 0$, and thus $W_{k+1} - B_k > C_k - B_k = A_k \ge 0$.

(7) $M_k^{(2)} - B_k > 0$

Now we show the existence of $W^{(1)}_{k+1}$. Again we need only to prove that each component in the right hand side of (3.1) is larger than or equal to each component in the left hand side.

(1) $W_{k+1} - B_k - \delta_3 - (A_k + \delta_1) = W_{k+1} - C_k - (\delta_1 + \delta_3) > 0$

(2) $W_{k+1} - B_k - \delta_3 - \delta_2 > 0$

(3) $W_{k+1} - B_k - \delta_3 - (W_{k+1} - M^{(2)}_k) > 0$

(4) $M^{(1)}_k - (A_k + \delta_1) > 0$

(5) $M^{(1)}_k - \delta_2 > 0$

(6) $M^{(1)}_k - (W_{k+1} - M^{(2)}_k) = M_k - W_{k+1} \geq 0$

(7) $W_{k+1} - (A_k + \delta_1) > 0$

(8) $W_{k+1} - \delta_2 > 0$

(9) $W_{k+1} - (W_{k+1} - M^{(2)}_k) = M^{(2)}_k > 0$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1} , W^{(1)}_{k+1})\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1} , W^{(2)}_{k+1})\}$ where $W^{(2)}_{k+1} = W_{k+1} - W^{(1)}_{k+1}$ and $S_{3,k+1} = \emptyset$. It can be easily proved that conditions I1 - I5 hold at the end of stage k+1.

CASE 2 : $J_k = 0$, $(\alpha-1) M_k + W_{k+1} \leq C_k < \alpha M_k$, for a positive integer α.

It implies that $M_k > W_{k+1}$, $(\alpha-1) M_k < C_k < \alpha M_k$ and $(\alpha-1) M_k \leq C_k - W_{k+1} < \alpha M_k$. Since $(\alpha-1) M_k < C_k < \alpha M_k$, we have $(\alpha-1) M^{(1)}_k < A_k < \alpha M^{(1)}_k$, and $(\alpha-1) M^{(2)}_k \leq B_k < \alpha M^{(2)}_k$. Also, $A_k < (n-2) M^{(1)}_k$. Choose $W^{(1)}_{k+1}$ such that

$$\max (W_{k+1} - B_k + (\alpha-1) M_k^{(2)} , \delta_1) \le W_{k+1}^{(1)} \le \min (A_k - (\alpha-1) M_k^{(1)} - \delta_2, W_{k+1}) \quad (3.2)$$

where

$$0 < \delta_1 < \min ((A_k - (\alpha-1) M_k^{(1)})/2 , W_{k+1})$$

and

$$0 \le \delta_2 \le \min ((A_k - (\alpha-1) M_k^{(1)})/2 , C_k - W_{k+1} - (\alpha-1) M_k),$$

δ_2 can be zero only when $(\alpha-1) M_k = C_k - W_{k+1}$.

Each component in the right hand side of δ_1 and δ_2 is strictly positive except $C_k - W_{k+1} - (\alpha-1) M_k$ may be zero. We then check the non-emptiness of the interval in (3.2).

(1) $A_k - (\alpha-1) M_k^{(1)} - \delta_2 - (W_{k+1} - B_k + (\alpha-1) M_k^{(2)})$

$= C_k - (\alpha-1) M_k - W_{k+1} - \delta_2$

≥ 0

(2) $A_k - (\alpha-1) M_k^{(1)} - \delta_2 - \delta_1 > 0$

(3) $W_{k+1} - (W_{k+1} - B_k + (\alpha-1) M_k^{(2)}) = B_k - (\alpha-1) M_k^{(2)} \ge 0$

(4) $W_{k+1} - \delta_1 > 0.$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1} , W_{k+1}^{(1)})\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1} , W_{k+1}^{(2)})\}$ where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$, and $S_{3,k+1} = \emptyset$. Thus, $C_{k+1} = C_k - W_{k+1}$. Conditions I1 - I5 hold at the end of stage k+1 as

$$A_{k+1} - (\alpha-1) M_k^{(1)} = A_k - W_{k+1}^{(1)} - (\alpha-1) M_k^{(1)} \ge \delta_2 \ge 0,$$

$$B_{k+1} - (\alpha-1) M_k^{(2)} = B_k - W_{k+1} + W_{k+1}^{(1)} - (\alpha-1) M_k^{(2)} \ge 0,$$

$$M_k^{(2)} - W_{k+1}^{(2)} = M_k^{(2)} - W_{k+1} + W_{k+1}^{(1)}$$
$$\geq M_k^{(2)} - W_{k+1} + W_{k+1} - B_k + (\alpha-1)\, M_k^{(2)}$$
$$> 0$$

and

$$M_k^{(1)} - W_{k+1}^{(1)} \geq M_k^{(1)} - A_k + (\alpha-1)\, M_k^{(1)} + \delta_2 > 0.$$

CASE 3 : $J_k = 0$, $W_{k+1} \leq M_k$, $\alpha\, M_k \leq C_k < \alpha\, M_k + W_{k+1}$, for a positive integer α.

It implies that $W_{k+1} \leq M_k$, $\alpha\, M_k \leq C_k < (\alpha+1)\, M_k$ and $(\alpha-1)\, M_k \leq C_k - W_{k+1} < \alpha\, M_k$. Since $\alpha\, M_k \leq C_k < (\alpha+1)\, M_k$, we have $\alpha\, M_k^{(1)} \leq A_k < (\alpha+1)\, M_k^{(1)}$ and $\alpha\, M_k^{(2)} \leq B_k < (\alpha+1)\, M_k^{(2)}$. Also, $A_k < (n-2)\, M_k^{(1)}$. Choose $W_{k+1}^{(1)}$ such that

$$\max\,(\, A_k - \alpha\, M_k^{(1)} + \delta_1\,,\ \delta_2\,,\ W_{k+1} - M_k^{(2)}\,) \leq W_{k+1}^{(1)} \leq$$
$$\min\,(\, W_{k+1} - B_k + \alpha\, M_k^{(2)} - \delta_3\,,\ M_k^{(1)}) \qquad (3.3)$$

where

$$0 < \delta_1 < \min\,(\,(\alpha\, M_k - C_k + W_{k+1})/2,\ (\alpha+1)\, M_k^{(1)} - A_k),$$
$$0 < \delta_2 < \min\,(\,(W_{k+1} - B_k + \alpha\, M_k^{(2)})/2,\ M_k^{(1)}\,)$$

and

$$0 < \delta_3 < \min\,(\,(W_{k+1} - B_k + \alpha\, M_k^{(2)})/2,\ (\alpha+1)\, M_k^{(2)} - B_k,$$
$$(\alpha\, M_k - C_k + W_{k+1})/2\,).$$

It is easy to show that each component in the right hand side of δ_1, δ_2 and δ_3 is strictly positive. We then check the existence of $W_{k+1}^{(1)}$ in (3.3).

(1) $W_{k+1} - B_k + \alpha M_k^{(2)} - \delta_3 - (A_k - \alpha M_k^{(1)} + \delta_1)$

$= W_{k+1} - C_k + \alpha M_k - (\delta_1 + \delta_3)$

> 0

(2) $W_{k+1} - B_k + \alpha M_k^{(2)} - \delta_3 - \delta_2 > 0$

(3) $W_{k+1} - B_k + \alpha M_k^{(2)} - \delta_3 - (W_{k+1} - M_k^{(2)}) = (\alpha+1) M_k^{(2)} - B_k - \delta_3 > 0$

(4) $M_k^{(1)} - (A_k - \alpha M_k^{(1)} + \delta_1) = (\alpha+1) M_k^{(1)} - A_k - \delta_1 > 0$

(5) $M_k^{(1)} - \delta_2 > 0$

(6) $M_k^{(1)} - (W_{k+1} - M_k^{(2)}) \geq 0$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1}, W_{k+1}^{(1)})\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1}, W_{k+1}^{(2)})\}$ where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$, and $S_{3,k+1} = \emptyset$. Thus, $C_{k+1} = C_k - W_{k+1}$. Conditions I1 - I4 hold at the end of stage k+1 as

$$\alpha M_k^{(1)} - A_{k+1} = \alpha M_k^{(1)} - A_k + W_{k+1}^{(1)} \geq \delta_1 > 0,$$

$$\alpha M_k^{(2)} - B_{k+1} = \alpha M_k^{(2)} - B_k + W_{k+1} - W_{k+1}^{(1)} \geq \delta_3 > 0,$$

$$M_k^{(1)} - W_{k+1}^{(1)} \geq 0,$$

$$M_k^{(2)} - W_{k+1}^{(2)} = M_k^{(2)} - W_{k+1} + W_{k+1}^{(1)} \geq 0$$

and

$$W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)} \geq B_k - \alpha M_k^{(2)} + \delta_3 \geq \delta_3 > 0.$$

If, in addition to $\alpha M_k < C_k < \alpha M_k + W_{k+1}$, it implies that $\alpha M_k < C_k < (\alpha+1) M_k$ and $(\alpha-1) M_k < C_k - W_{k+1} < \alpha M_k$. Since $\alpha M_k < C_k < (\alpha+1) M_k$, we additionally have $\alpha M_k^{(1)} < A_k$. The previous choice of $W_{k+1}^{(1)}$ can still guarantee that the condition I5 is satisfied because

$$A_{k+1} - (\alpha-1) M_k^{(1)} = A_k - W_{k+1}^{(1)} - (\alpha-1) M_k^{(1)} \geq A_k - M_k^{(1)} - (\alpha-1) M_k^{(1)} > 0.$$

CASE 4 : $J_k = 0$, $M_k < W_{k+1} < T_k / (n-1)$.

As $W_{k+1} < T_k / (n-1)$, we have $W_{k+1} < (n M_k - C_k) / (n-1)$. Since $M_k < W_{k+1}$, it follows that $C_k < M_k$. We have $A_k < M_k^{(1)}$, $B_k < M_k^{(2)}$ and $A_k < (n-2) M_k^{(1)}$. Choose $W_{k+1}^{(1)}$ such that

$$\max (W_{k+1} - (n M_k^{(2)} - B_k)/(n-1) + \delta_1 , M_k^{(1)}) \le W_{k+1}^{(1)} \le$$
$$\min ((n M_k^{(1)} - A_k)/(n-1) - \delta_2, n M_k^{(1)} - A_k - \delta_3 , W_{k+1} - M_k^{(2)}) \quad (3.4)$$

where

$$0 < \delta_1 < \min (((n M_k - C_k)/(n-1) - W_{k+1})/ 2, ((n M_k^{(1)} - A_k) + (n M_k^{(2)} - B_k)/(n-1) - W_{k+1}) / 2, (M_k^{(2)} - B_k)/(n-1)),$$

$$0 < \delta_2 < \min (((n M_k - C_k)/(n-1) - W_{k+1})/ 2, (M_k^{(1)} - A_k)/(n-1))$$

and

$$0 < \delta_3 < \min (((n M_k^{(1)} - A_k) + (n M_k^{(2)} - B_k)/(n-1) - W_{k+1}) / 2, (n-1) M_k^{(1)} - A_k).$$

To prove that δ_1, δ_2 and δ_3 exist, we show that each component in the right hand side is strictly greater than zero.

(1) $((n M_k - C_k)/(n-1) - W_{k+1})/ 2 > 0$

(2) $((n M_k^{(1)} - A_k) + (n M_k^{(2)} - B_k)/(n-1) - W_{k+1}) / 2$

$\ge ((n M_k - C_k)/(n-1) - W_{k+1})/2$

> 0

(3) $(M_k^{(2)} - B_k)/(n-1) > 0$

(4) $(M_k^{(1)} - A_k)/(n-1) > 0$

(5) $(n-1) M_k^{(1)} - A_k > 0$

To show the existence of such a $W_{k+1}^{(1)}$, we check the non-emptiness of the interval in (3.4).

(1) $(n M_k^{(1)} - A_k)/(n-1) - \delta_2 - (W_{k+1} - (n M_k^{(2)} - B_k)/(n-1) + \delta_1) > 0$

(2) $(n M_k^{(1)} - A_k)/(n-1) - \delta_2 - M_k^{(1)} = (M_k^{(1)} - A_k)/(n-1) - \delta_2 > 0$

(3) $n M_k^{(1)} - A_k - \delta_3 - (W_{k+1} - (n M_k^{(2)} - B_k)/(n-1) + \delta_1) > 0$

(4) $n M_k^{(1)} - A_k - \delta_3 - M_k^{(1)} = (n-1) M_k^{(1)} - A_k - \delta_3 > 0$

(5) $W_{k+1} - M_k^{(2)} - (W_{k+1} - (n M_k^{(2)} - B_k)/(n-1) + \delta_1)$

$= (M_k^{(2)} - B_k)/(n-1) - \delta_1$

> 0

(6) $W_{k+1} - M_k^{(2)} - M_k^{(1)} > 0$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1}, W_{k+1}^{(1)})\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1}, W_{k+1}^{(2)})\}$ where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$, and $S_{3,k+1} = \emptyset$. Thus, $C_{k+1} = (n-1) W_{k+1} - (n M_k - C_k)$. Conditions I1 - I5 hold at the end of stage k+1. That is

$$A_{k+1} < 0,$$

$$B_{k+1} < 0,$$

$$A_{k+1} < (n-2) W_{k+1}^{(1)}$$

and

$$M_k^{(1)} < W_{k+1}^{(1)} \le W_{k+1} - M_k^{(2)}.$$

where

$$A_{k+1} = (n-1) W_{k+1}^{(1)} - (n M_k^{(1)} - A_k),$$

$$B_{k+1} = (n-1) W_{k+1}^{(2)} - (n M_k^{(2)} - B_k).$$

CASE 5 : $J_k = 0,\ M_k < T_k / (n-\alpha) \le W_{k+1} < T_k / (n-\alpha-1)$, for a positive integer $\alpha \le n-2$.

It implies that $M_k < W_{k+1}$, $C_k < \alpha M_k$ and $(\alpha-1) W_{k+1} \le (n-1) W_{k+1} - T_k < \alpha W_{k+1}$ for a positive integer $\alpha \le n-2$. Since $C_k < \alpha M_k$, we have $A_k < \alpha M_k^{(1)}$ and $B_k < \alpha M_k^{(2)}$. Also, $A_k < (n-2) M_k^{(1)}$. Choose $W_{k+1}^{(1)}$ such that

$$\max\left((n M_k^{(1)} - A_k)/(n-\alpha) + \delta_1 ,\ W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha-1) + \delta_2 \right)$$
$$\le W_{k+1}^{(1)} \le \min\left((n M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3,\ W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha) \right) \qquad (3.5)$$

where

$$0 \le \delta_1 \le \min\left((n M_k^{(1)} - A_k)/2(n-\alpha)(n-\alpha-1),\ W_{k+1} - (n M_k - C_k)/(n-\alpha) \right),$$
$$0 < \delta_2 < \min\left(((n M_k - C_k)/(n-\alpha-1) - W_{k+1})/2 ,\ (n M_k^{(2)} - B_k)/(n-\alpha)(n-\alpha-1) \right)$$

and

$$0 < \delta_3 < \min\left(((n M_k - C_k)/(n-\alpha-1) - W_{k+1})/2 ,\ (n M_k^{(1)} - A_k)/2(n-\alpha)(n-\alpha-1) \right).$$

δ_1 can be zero only when $(\alpha-1) W_{k+1} = (n-1) W_{k+1} - T_k$.

Each component in the right hand side of δ_1, δ_2 and δ_3 is strictly positive because

(1) $(n M_k^{(1)} - A_k)/2(n-\alpha)(n-\alpha-1) > 0$

(2) $W_{k+1} - (n M_k - C_k)/(n-\alpha) \ge 0$ as

$$(n-1) W_{k+1} - T_k \ge (\alpha-1) W_{k+1} ,$$
$$(n-1) W_{k+1} - (n M_k - C_k) \ge (\alpha-1) W_{k+1} ,$$

$$(n-\alpha)W_{k+1} \geq n M_k - C_k ,$$

$$W_{k+1} \geq (n M_k - C_k)/(n-\alpha) .$$

(3) $((n M_k - C_k)/(n-\alpha-1) - W_{k+1})/2 > 0$ as

$$(n-1) W_{k+1} - T_k < \alpha W_{k+1} ,$$

$$(n-1) W_{k+1} - (n M_k - C_k) < \alpha W_{k+1} ,$$

$$(n-\alpha-1)W_{k+1} < n M_k - C_k ,$$

$$W_{k+1} < (n M_k - C_k)/(n-\alpha-1).$$

(4) $(n M_k^{(2)} - B_k)/(n-\alpha)(n-\alpha-1) > 0.$

To show the existence of $W_{k+1}^{(1)}$, we check the interval in (3.5) is not empty.

(1) $(n M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3 - ((n M_k^{(1)} - A_k)/(n-\alpha) + \delta_1)$

$= (n M_k^{(1)} - A_k)/(n-\alpha)(n-\alpha-1) - (\delta_1 + \delta_3)$

> 0

(2) $(n M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3 - (W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha-1) + \delta_2)$

$= (n M_k - C_k)/(n-\alpha-1) - W_{k+1} - (\delta_2 + \delta_3)$

> 0

(3) $W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha) - ((n M_k^{(1)} - A_k)/(n-\alpha) + \delta_1)$

$= W_{k+1} - (n M_k - C_k)/(n-\alpha) - \delta_1$

≥ 0

(4) $W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha) - (W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha-1) + \delta_2)$

$= (n M_k^{(2)} - B_k)/(n-\alpha)(n-\alpha-1) - \delta_2$

$> 0.$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1}, W^{(1)}_{k+1})\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1}, W^{(2)}_{k+1})\}$ where $W^{(2)}_{k+1} = W_{k+1} - W^{(1)}_{k+1}$, and $S_{3,k+1} = \emptyset$. Thus, $C_{k+1} = (n-1)\ W_{k+1} - T_k$. Conditions I1 - I5 hold at the end of stage k+1 as

$$\begin{aligned}
A_{k+1} - (\alpha-1)\ W^{(1)}_{k+1} &= (n-1)\ W^{(1)}_{k+1} - (n\ M^{(1)}_k - A_k) - (\alpha-1)\ W^{(1)}_{k+1} \\
&= (n-\alpha)\ W^{(1)}_{k+1} - (\ n\ M^{(1)}_k - A_k\ ) \\
&\geq (n-\alpha)\ \delta_1 \\
&\geq 0,
\end{aligned}$$

$$\begin{aligned}
\alpha\ W^{(1)}_{k+1} - A_{k+1} &= \alpha\ W^{(1)}_{k+1} - (n-1)\ W^{(1)}_{k+1} + (n\ M^{(1)}_k - A_k) \\
&= (n\ M^{(1)}_k - A_k) - (n-\alpha-1)\ W^{(1)}_{k+1} \\
&\geq (n-\alpha-1)\ \delta_3 \\
&> 0,
\end{aligned}$$

$$\begin{aligned}
B_{k+1} - (\alpha-1)\ W^{(2)}_{k+1} &= (n-1)\ W^{(2)}_{k+1} - (n\ M^{(2)}_k - B_k) - (\alpha-1)\ W^{(2)}_{k+1} \\
&= (n-\alpha)\ W_{k+1} - (n-\alpha)\ W^{(1)}_{k+1} - (\ n\ M^{(2)}_k - B_k\ ) \\
&\geq 0
\end{aligned}$$

and

$$\begin{aligned}
\alpha\ W^{(2)}_{k+1} - B_{k+1} &= (n\ M^{(2)}_k - B_k) - (n-\alpha-1)\ W_{k+1} + (n-\alpha-1)\ W^{(1)}_{k+1} \\
&\geq (n-\alpha-1)\ \delta_2 \\
&> 0.
\end{aligned}$$

CASE 6 : $J_k = 0$, $T_k / (n-\alpha) \leq M_k < W_{k+1} < T_k / (n-\alpha-1)$, for a positive integer $\alpha \leq n-2$.

It implies that $M_k < W_{k+1}$, $\alpha\ M_k \leq C_k < (\alpha+1)\ M_k$ and $(\alpha-1)\ W_{k+1} < (n-1)\ W_{k+1} - T_k < \alpha\ W_{k+1}$. Since $\alpha\ M_k \leq C_k < (\alpha+1)\ M_k$, we have $\alpha\ M^{(1)}_k \leq$

$A_k < (\alpha+1)\ M_k^{(1)}$ and $\alpha\ M_k^{(2)} \le B_k < (\alpha+1)\ M_k^{(2)}$. Also $A_k < (n-2)\ M_k^{(1)}$. Choose $W_{k+1}^{(1)}$ such that

$$\max\left((n\ M_k - A_k)/(n-\alpha) + \delta_1,\ W_{k+1} - (n\ M_k^{(2)} - B_k)/(n-\alpha-1) + \delta_2,\ M_k^{(1)} \right)$$
$$\le W_{k+1}^{(1)} \le \min\left((n\ M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3,\ W_{k+1} - M_k^{(2)} \right) \qquad (3.6)$$

where

$$0 < \delta_1 < \min\left((n\ M_k^{(1)} - A_k)/2(n-\alpha)(n-\alpha-1),\ W_{k+1} - M_k^{(2)} - (n\ M_k^{(1)} - A_k)/(n-\alpha) \right),$$
$$0 < \delta_2 < \min\left(((n\ M_k - C_k)/(n-\alpha-1) - W_{k+1})/2,\ ((\alpha+1)\ M_k^{(2)} - B_k)/(n-\alpha-1) \right)$$

and

$$0 < \delta_3 < \min\left(((n\ M_k - C_k)/(n-\alpha-1) - W_{k+1})/2,\ ((\alpha+1)\ M_k^{(1)} - A_k)/(n-\alpha-1),\ (n\ M_k^{(1)} - A_k)/\ 2(n-\alpha)(n-\alpha-1) \right).$$

Each component in the right hand side of δ_1, δ_2 and δ_3 is strictly positive. We then check that the interval in (3.6) is not empty.

(1) $(n\ M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3 - \left((n\ M_k - A_k)/(n-\alpha) + \delta_1 \right) > 0$

(2) $(n\ M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3 - (W_{k+1} - (n\ M_k^{(2)} - B_k)/(n-\alpha-1) + \delta_2)$

$= (n\ M_k - C_k)/(n-\alpha-1) - W_{k+1} - (\delta_2 + \delta_3)$

> 0

(3) $(n\ M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3 - M_k^{(1)}$

$= ((\alpha+1)\ M_k^{(1)} - A_k)/(n-\alpha-1) - \delta_3$

> 0

(4) $W_{k+1} - M_k^{(2)} - \left((n\ M_k - A_k)/(n-\alpha) + \delta_1 \right) > 0$

(5) $W_{k+1} - M_k^{(2)} - (W_{k+1} - (n M_k^{(2)} - B_k)/(n-\alpha-1) + \delta_2)$

$= ((\alpha+1) M_k^{(2)} - B_k)/(n-\alpha-1) - \delta_2$

> 0

(6) $W_{k+1} - M_k^{(2)} - M_k^{(1)} > 0$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1}, W_{k+1}^{(1)})\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1}, W_{k+1}^{(2)})\}$ where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$, and $S_{3,k+1} = \emptyset$. Thus, $C_{k+1} = (n-1) W_{k+1} - T_k$. Conditions I1 - I5 hold at the end of stage k+1 as

$$\begin{aligned}
\alpha W_{k+1}^{(1)} - A_{k+1} &= \alpha W_{k+1}^{(1)} - ((n-1) W_{k+1}^{(1)} - (n M_k^{(1)} - A_k)) \\
&= n M_k^{(1)} - A_k - (n-\alpha-1) W_{k+1}^{(1)} \\
&\geq (n-\alpha-1) \delta_3 \\
&> 0 \\
\alpha W_{k+1}^{(2)} - B_{k+1} &= n M_k^{(2)} - B_k - (n-\alpha-1) W_{k+1} + (n-\alpha-1) W_{k+1}^{(1)} \\
&\geq (n-\alpha-1) \delta_2 \\
&> 0 \\
A_{k+1} - (\alpha-1) W_{k+1}^{(1)} &= (n-1) W_{k+1}^{(1)} - (n M_k^{(1)} - A_k) - (\alpha-1) W_{k+1}^{(1)} \\
&= (n-\alpha) W_{k+1}^{(1)} - (n M_k^{(1)} - A_k) \\
&\geq (n-\alpha) \delta_1 \\
&> 0.
\end{aligned}$$

CASE 7 : $J_k = 0$, $W_{k+1} > M_k$, $W_{k+1} \geq T_k$.

Set $S_{1,k+1} = S_{1,k}$, $S_{2,k+1} = S_{2,k}$ and $S_{3,k+1} = \{ R_{k+1} \}$. Also, $J_{k+1} = 1$. Clearly, $T_{k+1} = T_k$. Thus, $V_{k+1} = W_{k+1}$. Since

$$0 < T_k \le W_{k+1},$$

$$0 < T_{k+1} \le W_{k+1},$$

$$n\, W_{k+1} > n\, W_{k+1} - T_{k+1} \ge (n-1)\, W_{k+1},$$

$$(n-1)\, V_{k+1} \le C^*_{k+1} < n\, V_{k+1},$$

Obviously, $A_{k+1} = A_k < n\, M^{(1)}_k = n\, M^{(1)}_{k+1}$, $B_{k+1} < n\, M^{(2)}_{k+1}$ and $A_{k+1} = A_k < (n-2)\, M^{(1)}_k = (n-2)\, M^{(1)}_{k+1}$. Since $W_{k+1} \ge T_k$, we have $W_{k+1} - M_k \ge (n-1)\, M_k - C_k$.

Consider

$$A_{k+1} - (n-1)\, M^{(1)}_{k+1} + (\, V_{k+1} - M_{k+1}\,)$$

$$= A_k - (n-1)\, M^{(1)}_k + (\, W_{k+1} - M_k\,)$$

$$\ge A_k - (n-1)\, M^{(1)}_k + (n-1)\, M_k - C_k$$

$$= (n-1)\, M^{(2)}_k - B_k$$

> 0 as $B_k = n\, M^{(2)}_k - T^{(2)}_k = (n-1)\, M^{(2)}_k + (\, M^{(2)}_k - T^{(2)}_k\,) < (n-1)\, M^{(2)}_k$.

Hence, $A_{k+1} > (n-1)\, M^{(1)}_{k+1} - (\, V_{k+1} - M_{k+1}\,)$. Similarly, $B_{k+1} \ge (n-1)\, M^{(2)}_{k+1} - (\, V_{k+1} - M_{k+1}\,)$. Thus, all conditions II1 - II7 are satisfied.

CASE 8 : $0 < J_k < n-1$, $W_{k+1} = V_k$.

We simply set $S_{1,k+1} = S_{1,k}$, $S_{2,k+1} = S_{2,k}$ and $S_{3,k+1} = S_{3,k} \cup \{\, R_{k+1}\,\}$.

CASE 9 : $J_k = n-1$, $W_{k+1} = V_k$.

From (II4), (II6) and (II5), we have $(n-1)\, V_k \le C^*_k < n\, V_k$, $(n-1)\, M^{(1)}_k - (V_k - M_k) < A_k < n\, M^{(1)}_k$ and $(n-1)\, M^{(2)}_k - (V_k - M_k) \le B_k < n\, M^{(2)}_k$. Also, $A_k < (n-2)\, M^{(1)}_k$. Choose L such that

$$M^{(1)}_k \le L \le V_k - M^{(2)}_k \tag{3.7}$$

Obviously, such a L in (3.7) exists. Set $W_j^{(1)} = L$ for $j = k+1$ or $R_j \in S_{3,k}$. Update $S_{1,k+1} = S_{1,k} \cup \{(R_j, W_j^{(1)}) : j = k+1 \text{ or } R_j \in S_{3,k}\}$, $S_{2,k+1} = S_{2,k} \cup \{(R_j, W_j^{(2)}) : W_j^{(2)} = W_j - W_j^{(1)}, \text{ and } j = k+1 \text{ or } R_j \in S_{3,k}\}$ and $S_{3,k+1} = \emptyset$. It can be proved that the conditions I1 - I5 hold at the end of stage k+1. That is,

$$A_{k+1} < 0,$$

$$B_{k+1} < 0,$$

$$A_{k+1} < (n-2)\, W_{k+1}^{(1)},$$

$$M_k^{(1)} \le W_{k+1}^{(1)} \le V_k - M_k^{(2)}.$$

It is because $A_{k+1} = n\, W_{k+1}^{(1)} - T_{k+1}^{(1)} = n\, L - (n\, L + T_k^{(1)}) = - T_k^{(1)} < 0$ and hence $A_{k+1} < (n-2)\, W_{k+1}^{(1)}$; $B_{k+1} = n\, W_{k+1}^{(2)} - T_{k+1}^{(2)} = n\, W_{k+1}^{(2)} - (n\, W_{k+1}^{(2)} + T_k^{(2)}) = - T_k^{(2)} < 0$.

CASE 10 : $0 < J_k \le n-1$, $W_{k+1} \le M_k$, $W_{k+1} \le V_k - T_k$.

From (II4), $W_{k+1} \le M_k$ and $W_{k+1} \le V_k - T_k$, it implies $(n-1)\, V_k \le C_k^* - W_{k+1} < n\, V_k$. From (II6) and (II5), $(n-1)\, M_k^{(1)} - (V_k - M_k) < A_k < n\, M_k^{(1)}$ and $(n-1)\, M_k^{(2)} - (V_k - M_k) \le B_k < n\, M_k^{(2)}$. Also, $A_k < (n-2)\, M_k^{(1)}$. Choose $W_{k+1}^{(1)}$ such that

$$\max\left(W_{k+1} - B_k + (n-1)\, M_k^{(2)} - (V_k - M_k),\ \delta_1,\ W_{k+1} - M_k^{(2)} \right)$$

$$\le W_{k+1}^{(1)} \le \min\left(A_k - (n-1)\, M_k^{(1)} + (V_k - M_k) - \delta_2,\ M_k^{(1)},\ W_{k+1}\right) \tag{3.8}$$

where

$$0 < \delta_1 < \min\left((A_k - (n-1) M_k^{(1)} + (V_k - M_k))/2 \ , \ M_k^{(1)} \ , \ W_{k+1} \right)$$

and

$$\begin{aligned} 0 < \delta_2 < \min (& C_k - (n-1) M_k + 2 (V_k - M_k) - W_{k+1} \ , \\ & (A_k - (n-1) M_k^{(1)} + (V_k - M_k))/2 \ , \\ & n M_k^{(2)} + n (V_k - M_k) - (n-1) V_k + A_k - W_{k+1}). \end{aligned}$$

We have to show that each component in the right hand side of δ_1 and δ_2 is strictly positive.

(1) $(A_k - (n-1) M_k^{(1)} + (V_k - M_k))/2 > 0$ as

$$\begin{aligned} & C_k^* \geq (n-1) V_k + W_{k+1}, \\ & n V_k - T_k \geq (n-1) V_k + W_{k+1}, \\ & \quad V_k > T_k, \\ & \text{since } T_k^{(2)} > M_k^{(2)}, \\ & \quad T_k > T_k^{(1)} + M_k^{(2)}, \\ & \quad V_k > T_k^{(1)} + M_k^{(2)} \ , \\ & \quad V_k > n M_k^{(1)} - A_k + M_k^{(2)} \ , \\ & \text{hence, } A_k - (n-1) M_k^{(1)} + (V_k - M_k) > 0. \end{aligned}$$

(2) $M_k^{(1)} > 0$

(3) $W_{k+1} > 0$

(4) $C_k - (n-1) M_k + 2 (V_k - M_k) - W_{k+1}$

$$\begin{aligned} & = C_k^* - W_{k+1} - (n-1) V_k + V_k - M_k \\ & \geq V_k - M_k \\ & > 0 \end{aligned}$$

(5) $n\, M_k^{(2)} + n\,(V_k - M_k) - (n-1)\, V_k + A_k - W_{k+1} \geq n\, M_k^{(2)} - B_k > 0$ as

$$(n-1)\, V_k \leq C_k + n\,(V_k - M_k) - W_{k+1},$$

$$A_k + B_k \geq (n-1)\, V_k - n\,(V_k - M_k) + W_{k+1},$$

$$- B_k \leq n\,(V_k - M_k) - (n-1)\, V_k + A_k - W_{k+1}.$$

We then check the non-emptiness of the interval in (3.8).

(1) $A_k - (n-1)\, M_k^{(1)} + (V_k - M_k) - \delta_2$

$- (\, W_{k+1} - B_k + (n-1)\, M_k^{(2)} - (V_k - M_k)\,)$

$= C_k - (n-1)\, M_k + 2\,(V_k - M_k) - W_{k+1} - \delta_2$

> 0

(2) $A_k - (n-1)\, M_k^{(1)} + (V_k - M_k) - \delta_2 - \delta_1 > 0$

(3) $A_k - (n-1)\, M_k^{(1)} + (V_k - M_k) - \delta_2 - (\, W_{k+1} - M_k^{(2)}\,)$

$= M_k^{(2)} - (n-1)\, M_k^{(1)} + (V_k - M_k) + A_k - W_{k+1} - \delta_2$

$= n\, M_k^{(2)} + n\,(V_k - M_k) - (n-1)\, V_k + A_k - W_{k+1} - \delta_2$

> 0

(4) $M_k^{(1)} - (\, W_{k+1} - B_k + (n-1)\, M_k^{(2)} - (V_k - M_k)\,)$

$= M_k^{(1)} - (n-1)\, M_k^{(2)} + (V_k - M_k) + B_k - W_{k+1}$

> 0

(5) $M_k^{(1)} - \delta_1 > 0$

(6) $M_k^{(1)} - (\, W_{k+1} - M_k^{(2)}\,) \geq 0$

(7) $W_{k+1} - (\, W_{k+1} - B_k + (n-1)\, M_k^{(2)} - (V_k - M_k)\,)$

$= B_k - (n-1)\, M_k^{(2)} + (V_k - M_k)$

≥ 0

(8) $W_{k+1} - \delta_1 > 0$

(9) $W_{k+1} - (W_{k+1} - M_k^{(2)}) > 0$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1}, W_{k+1}^{(1)})\}$ and $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1}, W_{k+1}^{(2)})\}$ where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$ and $S_{3,k+1} = S_{3,k}$. Thus, $C_{k+1}^{*} = C_k^{*} - W_{k+1}$. It can be proved that conditions II1 - II7 hold at the end of stage k+1 as

$$(n-1)\, M_k^{(1)} - (V_k - M_k) < A_{k+1} < n\, M_k^{(1)},$$
$$(n-1)\, M_k^{(2)} - (V_k - M_k) \le B_{k+1} < n\, M_k^{(2)},$$
$$A_{k+1} < (n-2)\, M_k^{(1)},$$
$$0 < W_{k+1}^{(1)} \le M_k^{(1)},$$
$$0 \le W_{k+1}^{(2)} \le M_k^{(2)}$$

where

$$A_{k+1} = A_k - W_{k+1}^{(1)},$$
$$B_{k+1} = B_k - W_{k+1}^{(2)}.$$

It is because

$$\begin{aligned} & A_{k+1} - (n-1)\, M_k^{(1)} + (V_k - M_k) \\ = & A_k - W_{k+1}^{(1)} - (n-1)\, M_k^{(1)} + (V_k - M_k) \\ \ge & \delta_2 \\ > & 0 \end{aligned}$$

and

$$\begin{aligned} & B_{k+1} - (n-1)\, M_k^{(2)} + (V_k - M_k) \\ = & B_k - W_{k+1} + W_{k+1}^{(1)} - (n-1)\, M_k^{(2)} + (V_k - M_k) \\ \ge & 0. \end{aligned}$$

CASE 11 : $0 < J_k \le n-1,\ V_k - T_k < W_{k+1} \le M_k$.

From (II4) and $V_k - T_k < W_{k+1} \le M_k$, it implies that $(n-2)\ V_k < C^*_k - W_{k+1} < (n-1)\ V_k$.

We first consider W_{k+1}. From (II4), (II6) and (II5), we have $(n-1)\ V_k \le C^*_k < n\ V_k$, $(n-1)\ M^{(1)}_k - (V_k - M_k) < A_k < n\ M^{(1)}_k$ and $(n-1)\ M^{(2)}_k - (V_k - M_k) \le B_k < n\ M^{(2)}_k$. Also, $A_k < (n-2)\ M^{(1)}_k$. Choose $W^{(1)}_{k+1}$ such that

$$\max\left(W_{k+1} - B_k + (n-2)\ M^{(2)}_k - 2\ (V_k - M_k)\ ,\ \delta_1\ ,\ W_{k+1} - M^{(2)}_k \right)$$
$$\le W^{(1)}_{k+1} \le \min\left(A_k - (n-2)\ M^{(1)}_k + 2\ (V_k - M_k) - \delta_2\ ,\right.$$
$$\left. W_{k+1} - B_k + (n-1)\ M^{(2)}_k - \delta_3,\ M^{(1)}_k\ ,\ W_{k+1} \right) \qquad (3.9)$$

where

$$0 < \delta_1 < \min\left(\left(A_k - (n-2)\ M^{(1)}_k + 2\ (V_k - M_k) \right)/2\ ,\right.$$
$$\left. (W_{k+1} - B_k + (n-1)\ M^{(2)}_k)/2\ ,\ M^{(1)}_k\ ,\ W_{k+1} \right),$$
$$0 < \delta_2 < \min\left(C_k - (n-2)\ M_k + 4(V_k - M_k) - W_{k+1}\ ,\right.$$
$$\left(A_k - (n-2)\ M^{(1)}_k + 2\ (V_k - M_k) \right)/2\ ,$$
$$\left. A_k - (n-2)\ M^{(1)}_k + 2\ (V_k - M_k) - W_{k+1} + M^{(2)}_k \right)$$

and

$$0 < \delta_3 < \min\left(M^{(2)}_k + 2\ (V_k - M_k)\ ,\right.$$
$$\left. (W_{k+1} - B_k + (n-1)\ M^{(2)}_k)/2\ ,\ n\ M^{(2)}_k - B_k \right).$$

To prove that δ_1, δ_2 and δ_3 exist, we need to show that each component in the right hand side is strictly greater than zero.

(1) $\left(A_k - (n-2)\ M^{(1)}_k + 2\ (V_k - M_k) \right)/2$

$= \left(A_k - (n-1)\ M^{(1)}_k + (V_k - M_k) + M^{(1)}_k + (V_k - M_k) \right)/2$

> 0

(2) $(W_{k+1} - B_k + (n-1)\, M_k^{(2)})/2 > 0$

as $C_k^* - W_{k+1} < (n-1)\, V_k$, which implies $A_k + B_k + n\,(V_k - M_k) - W_{k+1} < (n-1)\, V_k$, and thus, $W_{k+1} - B_k + (n-1)\, M_k^{(2)} > A_k + (V_k - M_k) - (n-1)\, M_k^{(1)} > 0$.

(3) $M_k^{(1)} > 0$

(4) $W_{k+1} > 0$

(5) $C_k - (n-2)\, M_k + 4\,(V_k - M_k) - W_{k+1}$

$$= C_k^* - W_{k+1} - 2\, M_k - (n-4)\, V_k$$

$$> (n-2)\, V_k + 2\,(V_k - M_k) - (n-2)\, V_k$$

$$> 0$$

(6) $A_k - (n-2)\, M_k^{(1)} + 2\,(V_k - M_k) - W_{k+1} + M_k^{(2)}$

$$= A_k - (n-1)\, M_k^{(1)} + (V_k - M_k) + M_k^{(1)} + (V_k - M_k) - W_{k+1} + M_k^{(2)}$$

$$> V_k - W_{k+1}$$

$$> 0$$

(7) $M_k^{(2)} + 2\,(V_k - M_k) > 0$

(8) $n\, M_k^{(2)} - B_k > 0$

To show the existence of $W_{k+1}^{(1)}$, we check the non-emptiness of the interval in (3.9).

(1) $A_k - (n-2)\, M_k^{(1)} + 2\,(V_k - M_k) - \delta_2$

$$- (\, W_{k+1} - B_k + (n-2)\, M_k^{(2)} - 2\,(V_k - M_k)\,)$$

$$= C_k - (n-2)\, M_k + 4(V_k - M_k) - W_{k+1} - \delta_2$$

$$> 0$$

(2) $A_k - (n-2)\, M_k^{(1)} + 2\,(V_k - M_k) - \delta_2 - \delta_1 > 0$

(3) $A_k - (n-2)\, M_k^{(1)} + 2\, (V_k - M_k) - \delta_2 - (\, W_{k+1} - M_k^{(2)}\,) > 0$

(4) $W_{k+1} - B_k + (n-1)\, M_k^{(2)} - \delta_3$

$- (\, W_{k+1} - B_k + (n-2)\, M_k^{(2)} - 2\, (V_k - M_k)\,)$

$= M_k^{(2)} + 2\, (V_k - M_k) - \delta_3$

> 0

(5) $W_{k+1} - B_k + (n-1)\, M_k^{(2)} - \delta_3 - \delta_1 > 0$

(6) $W_{k+1} - B_k + (n-1)\, M_k^{(2)} - \delta_3 - (\, W_{k+1} - M_k^{(2)}\,)$

$= n\, M_k^{(2)} - B_k - \delta_3$

> 0

(7) $M_k^{(1)} - (\, W_{k+1} - B_k + (n-2)\, M_k^{(2)} - 2\, (V_k - M_k)\,)$

$= B_k - (n-1)\, M_k^{(2)} + (V_k - M_k) + M_k^{(2)} + (V_k - M_k) - W_{k+1} + M_k^{(1)}$

$\geq V_k - W_{k+1}$

> 0

(8) $M_k^{(1)} - \delta_1 > 0$

(9) $M_k^{(1)} - (\, W_{k+1} - M_k^{(2)}\,) = M_k - W_{k+1} \geq 0$

(10) $W_{k+1} - (\, W_{k+1} - B_k + (n-2)\, M_k^{(2)} - 2\, (V_k - M_k)\,)$

$= B_k - (n-1)\, M_k^{(2)} + (V_k - M_k) + M_k^{(2)} + (V_k - M_k)$

> 0

(11) $W_{k+1} - \delta_1 > 0$

(12) $W_{k+1} - (\, W_{k+1} - M_k^{(2)}\,) > 0$

It can be proved that conditions II1 - II7 hold and $C^*_{k+1} = C^*_k - W_{k+1}$ up to this point . That is,

$$(n-2)\, M_k^{(1)} - 2\, (V_k - M_k) < A_{k+1} < (n-1)\, M_k^{(1)},$$

$$(n-2)\, M_k^{(2)} - 2\, (V_k - M_k) \leq B_{k+1} < (n-1)\, M_k^{(2)},$$

$$A_{k+1} < (n-2)\, M_k^{(1)},$$
$$0 < W_{k+1}^{(1)} \le M_k^{(1)},$$
$$0 \le W_{k+1}^{(2)} \le M_k^{(2)}.$$

where

$$A_{k+1} = A_k - W_{k+1}^{(1)},$$
$$B_{k+1} = B_k - W_{k+1}^{(2)}.$$

It is because

$$\begin{aligned} A_{k+1} - (n-2)\, M_k^{(1)} + 2\,(V_k - M_k) &= A_k - W_{k+1}^{(1)} - (n-2)\, M_k^{(1)} + 2\,(V_k - M_k) \\ &\ge \delta_2 \\ &> 0 \end{aligned}$$

$$\begin{aligned} & B_{k+1} - (n-2)\, M_k^{(2)} + 2\,(V_k - M_k) \\ &= B_k - W_{k+1} + W_{k+1}^{(1)} - (n-2)\, M_k^{(2)} + 2(V_k - M_k) \\ &\ge 0 \end{aligned}$$

$$(n-1)\, M_k^{(2)} - B_{k+1} = (n-1)\, M_k^{(2)} - B_k + W_{k+1} - W_{k+1}^{(1)} \ge \delta_3 > 0$$

After considering W_{k+1}, we reconsider the J_k elements in $S_{3,k}$. Since $(n-2)\, V_k < C_k^* - W_{k+1} < (n-1)\, V_k$, (or equivalently, $(n-2)\, V_k < C_{k+1}^* < (n-1)\, V_k$), we re-label the index as $(n-2)\, V_k < C_k^* < (n-1)\, V_k$ for convenience and there are J_k elements in $S_{3,k}$ and A_k, B_k and C_k are the updated value.

From the previous intermediate step, we have $(n-2)\, V_k < C_k^* < (n-1)\, V_k$, $(n-2)\, M_k^{(1)} - 2\,(V_k - M_k) < A_k < (n-1)\, M_k^{(1)}$ and $(n-2)\, M_k^{(2)} - 2\,(V_k - M_k) \le B_k < (n-1)\, M_k^{(2)}$. Also, $A_k < (n-2)\, M_k^{(1)}$. When the J_k elements in $S_{3,k}$ are considered, we let $V_k = W_{k+1}^{'(1)} + W_{k+1}^{'(2)}$ and choose $W_{k+1}^{'(1)}$ such that

$$\max\left((n M_k^{(1)} - A_k)/2 + \delta_1 ,\ V_k - (n M_k^{(2)} - B_k) + \delta_2 ,\ M_k^{(1)} \right) \le W'^{(1)}_{k+1}$$

$$\le \min\left(n M_k^{(1)} - A_k - \delta_3 ,\ V_k - (n M_k^{(2)} - B_k)/2 ,\ V_k - M_k^{(2)} \right) \qquad (3.10)$$

where

$$0 < \delta_1 < \min\left((n M_k^{(1)} - A_k)/4 ,\ V_k - (n M_k - C_k)/2 ,\ (2(V_k - M_k) - (n-2) M_k^{(1)} + A_k)/2 \right) ,$$

$$0 < \delta_2 < \min\left((n M_k - C_k - V_k)/2 ,\ (n M_k^{(2)} - B_k)/2 ,\ (n-1) M_k^{(2)} - B_k \right)$$

and

$$0 < \delta_3 < \min\left((n M_k^{(1)} - A_k)/4 ,\ (n M_k - C_k - V_k)/2 ,\ (n-1) M_k^{(1)} - A_k \right).$$

We have to prove that each component in the right hand side of δ_1, δ_2 and δ_3 is strictly positive.

(1) $(n M_k^{(1)} - A_k)/4 > 0$

(2) $V_k - (n M_k - C_k)/2 > 0$ as

$$C_k^* > (n-2) V_k,$$

$$C_k + n V_k - n M_k > n V_k - 2 V_k ,$$

$$C_k - n M_k + 2 V_k > 0$$

(3) $(2(V_k - M_k) - (n-2) M_k^{(1)} + A_k)/2 > 0$

(4) $(n M_k - C_k - V_k)/2 > 0$ as

$$C_k^* < (n-1) V_k,$$

$$C_k + n V_k - n M_k < n V_k - V_k ,$$

$$n M_k - C_k - V_k > 0$$

(5) $(n M_k^{(2)} - B_k)/2 > 0$

(6) $(n-1)\ M_k^{(2)} - B_k > 0$

(7) $(n-1)\ M_k^{(1)} - A_k > 0$

To show the existence of such a $W_{k+1}^{(1)}$, we check the non-emptiness of the interval in (3.10).

(1) $n\ M_k^{(1)} - A_k - \delta_3 - (\ (n\ M_k^{(1)} - A_k)/2 + \delta_1\ )$

$= (n\ M_k^{(1)} - A_k)/2 - (\delta_1 + \delta_3)$

> 0

(2) $n\ M_k^{(1)} - A_k - \delta_3 - (\ V_k - (n\ M_k^{(2)} - B_k) + \delta_2\ ) > 0$

(3) $n\ M_k^{(1)} - A_k - \delta_3 - M_k^{(1)} = (n-1)\ M_k^{(1)} - A_k - \delta_3 > 0$

(4) $V_k - (n\ M_k^{(2)} - B_k)/2 - (\ (n\ M_k^{(1)} - A_k)/2 + \delta_1\ ) > 0$

(5) $V_k - (n\ M_k^{(2)} - B_k)/2 - (\ V_k - (n\ M_k^{(2)} - B_k) + \delta_2\ ) > 0$

(6) $V_k - (n\ M_k^{(2)} - B_k)/2 - M_k^{(1)}$

$= (\ 2(V_k - M_k) - (n-2)\ M_k^{(2)} + B_k\ )/2$

≥ 0

(7) $V_k - M_k^{(2)} - (\ (n\ M_k^{(1)} - A_k)/2 + \delta_1\ )$

$= (\ 2(V_k - M_k) - (n-2)\ M_k^{(1)} + A_k\ )/2 - \delta_1$

> 0

(8) $V_k - M_k^{(2)} - (\ V_k - (n\ M_k^{(2)} - B_k) + \delta_2\ ) = (n-1)\ M_k^{(2)} - B_k - \delta_2 > 0$

(9) $V_k - M_k^{(2)} - M_k^{(1)} > 0$

Set $S_{1,k+1} = S_{1,k} \cup \{(\ R_{k+1},\ W_{k+1}^{(1)}\ )\} \cup \{\ (\ R_j,\ W_{k+1}^{'(1)})\ :\ R_j \in S_{3,k}\ \}$, $S_{2,k+1} = S_{2,k} \cup \{(\ R_{k+1},\ W_{k+1}^{(2)}\ )\} \cup \{\ (\ R_j,\ W_{k+1}^{'(2)}\ )\ :\ W_{k+1}^{'(2)} = V_k - W_{k+1}^{'(1)}$, $R_j \in S_{3,k}\ \}$ and $S_{3,k+1} = \emptyset$, where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$. Thus, $C_{k+1} = (\ n-J_k\ )\ V_k - (\ n\ M_k - C_k\ )$. It can be proved that the conditions I1 - I5 hold at the end of stage k+1. That is,

$$(n-J_k-2)\ W'^{(1)}_{k+1} < A_{k+1} < (n-J_k-1)\ W'^{(1)}_{k+1},$$
$$(n-J_k-2)\ W'^{(2)}_{k+1} \le B_{k+1} < (n-J_k-1)\ W'^{(2)}_{k+1},$$
$$A_{k+1} < (n-2)\ W'^{(1)}_{k+1},$$
$$M^{(1)}_k \le W'^{(1)}_{k+1} \le V_k - M^{(2)}_k.$$

where

$$A_{k+1} = (n-J_k)\ W'^{(1)}_{k+1} - (\ n\ M^{(1)}_k - A_k),$$
$$B_{k+1} = (n-J_k)\ W'^{(2)}_{k+1} - (\ n\ M^{(2)}_k - B_k).$$

It is because

$$A_{k+1} - (n-J_k-2)\ W'^{(1)}_{k+1} = 2\ W'^{(1)}_{k+1} - (\ n\ M^{(1)}_k - A_k\ ) \ge 2\ \delta_1 > 0,$$
$$(n-J_k-1)\ W'^{(1)}_{k+1} - A_{k+1} = (\ n\ M^{(1)}_k - A_k\ ) - W'^{(1)}_{k+1} \ge \delta_3 > 0,$$
$$B_{k+1} - (n-J_k-2)\ W'^{(2)}_{k+1} = 2\ V_k - 2\ W'^{(1)}_{k+1} - (\ n\ M^{(2)}_k - B_k\ ) \ge 0,$$
$$(n-J_k-1)\ W'^{(2)}_{k+1} - B_{k+1} = (\ n\ M^{(2)}_k - B_k\ ) - V_k + W'^{(1)}_{k+1} \ge \delta_2 > 0$$

and

$$\begin{aligned}(n-2)\ W'^{(1)}_{k+1} - A_{k+1} &= (J_k - 2)\ W'^{(1)}_{k+1} + (\ n\ M^{(1)}_k - A_k\ ) \\ &\ge (\ n\ M^{(1)}_k - A_k\ ) - W'^{(1)}_{k+1} \\ &\ge \delta_3 \\ &> 0\end{aligned}$$

CASE 12 : $0 < J_k \le n-1,\ M_k < W_{k+1} \le V_k - T_k$.

From (II4) and $0 < W_{k+1} \le V_k - T_k$, it implies that $(n-1)\ V_k \le C^*_k - W_{k+1} < n\ V_k$. Since $(n-1)\ V_k \le C^*_k < n\ V_k$, we have $(n-1)\ M^{(1)}_k - (V_k - M_k) < A_k < n\ M^{(1)}_k$ and $(n-1)\ M^{(2)}_k - (V_k - M_k) \le B_k < n\ M^{(2)}_k$. Also, $A_k < (n-2)\ M^{(1)}_k$. Choose $W^{(1)}_{k+1}$ such that

$$M^{(1)}_k \le W^{(1)}_{k+1} \le \min\ (n\ M^{(1)}_k - A_k - \delta_1\ ,\ W_{k+1} - M^{(2)}_k\ ) \qquad (3.11)$$

where

$$0 < \delta_1 < (n-1)\ M_k^{(1)} - A_k.$$

Obviously, $(n-1)\ M_k^{(1)} - A_k$ is greater than zero and the interval of $W_{k+1}^{(1)}$ in (3.11) is non-empty. Set $S_{1,k+1} = S_{1,k} \cup \{ (R_{k+1}, W_{k+1}^{(1)}) \}$ and $S_{2,k+1} = S_{2,k} \cup \{ (R_{k+1}, W_{k+1}^{(2)}) \}$ where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$ and $S_{3,k+1} = S_{3,k}$. It can be proved that the conditions II1 - II7 hold at the end of stage k+1. That is,

$$(n-1)\ W_{k+1}^{(1)} - (V_k - W_{k+1}) < A_{k+1} < n\ W_{k+1}^{(1)},$$
$$(n-1)\ W_{k+1}^{(2)} - (V_k - W_{k+1}) \le B_{k+1} < n\ W_{k+1}^{(2)},$$
$$A_{k+1} < (n-2)\ W_{k+1}^{(1)},$$
$$M_k^{(1)} \le W_{k+1}^{(1)} \le W_{k+1} - M_k^{(2)}$$

where

$$A_{k+1} = (n-1)\ W_{k+1}^{(1)} - (\ n\ M_k^{(1)} - A_k\ ),$$
$$B_{k+1} = (n-1)\ W_{k+1}^{(2)} - (\ n\ M_k^{(2)} - B_k\ ).$$

It is because

$$\begin{aligned} & A_{k+1} - (n-1)\ W_{k+1}^{(1)} + (V_k - W_{k+1}) \\ &= (n-1)\ W_{k+1}^{(1)} - (\ n\ M_k^{(1)} - A_k\ ) - (n-1)\ W_{k+1}^{(1)} + (\ V_k - W_{k+1}\ ) \\ &= V_k - W_{k+1} - (\ n\ M_k^{(1)} - A_k\ ) \\ &= V_k - W_{k+1} - T_k^{(1)} \\ &> 0 \quad \text{as} \end{aligned}$$

$$(n-1)\ V_k \le C_k^* - W_{k+1} = n\ V_k - T_k - W_{k+1},$$
$$0 \le -\ T_k + V_k - W_{k+1},$$
$$T_k \le V_k - W_{k+1},$$
$$T_k^{(1)} < T_k \le V_k - W_{k+1}.$$

Therefore, $A_{k+1} > (n-1)\ W_{k+1}^{(1)} - (V_k - W_{k+1})$ will hold. Similarly for $B_{k+1} \ge (n-1)\ W_{k+1}^{(2)} - (V_k - W_{k+1})$.

CASE 13 : $0 < J_k \le n-1,\ M_k < W_{k+1},\ V_k - T_k < W_{k+1} < V_k$.

From (II4), $M_k < W_{k+1}$ and $V_k - T_k < W_{k+1} < V_k$, it implies that $(n-2)\ V_k < C_k^* - W_{k+1} < (n-1)\ V_k$. We first consider W_{k+1}. Since $(n-1)\ V_k \le C_k^* < n\ V_k$, we have $(n-1)\ M_k^{(1)} - (V_k - M_k) < A_k < n\ M_k^{(1)}$ and $(n-1)\ M_k^{(2)} - (V_k - M_k) \le B_k < n\ M_k^{(2)}$. Also, $A_k < (n-2)\ M_k^{(1)}$. Choose $W_{k+1}^{(1)}$ such that

$$\max\left(n\ M_k^{(1)} - A_k - 2\ (V_k - W_{k+1}) + \delta_1\ ,\ M_k^{(1)} \right) \le W_{k+1}^{(1)}$$
$$\le \min\left(W_{k+1} - (\ n\ M_k^{(2)} - B_k) + 2\ (\ V_k - W_{k+1}\ ),\ n\ M_k^{(1)} - A_k - \delta_2,\ W_{k+1} - M_k^{(2)} \right) \quad (3.12)$$

where

$$0 < \delta_1 < \min\left(C_k^* - W_{k+1} - (n-4)\ V_k - 2\ W_{k+1}\ ,\ W_{k+1} - M_k^{(2)} - n\ M_k^{(1)} + A_k + 2\ (V_k - W_{k+1})\ ,\ V_k - W_{k+1} \right)$$

and

$$0 < \delta_2 < \min\left((n-1)\ M_k^{(1)} - A_k\ ,\ V_k - W_{k+1} \right).$$

To prove that δ_1 and δ_2 exist, we show that each component in the right hand side is strictly greater than zero.

(1) $C_k^* - W_{k+1} - (n-4)\ V_k - 2\ W_{k+1}$
$= C_k^* - W_{k+1} - (n-2)\ V_k + 2\ (V_k - W_{k+1})$
> 0

(2) $W_{k+1} - M_k^{(2)} - n\ M_k^{(1)} + A_k + 2\ (V_k - W_{k+1})$
$= A_k - (n-1)\ M_k^{(1)} + (V_k - M_k) + (V_k - W_{k+1})$
> 0

(3) $V_k - W_{k+1} > 0$

(4) $(n-1)\ M_k^{(1)} - A_k > 0$

To show the existence of $W_{k+1}^{(1)}$ in (3.12), we then check the non-emptiness of the interval.

$$\begin{aligned}
(1)\quad & W_{k+1} - (n M_k^{(2)} - B_k) + 2 (V_k - W_{k+1}) \\
& - (n M_k^{(1)} - A_k - 2 (V_k - W_{k+1}) + \delta_1) \\
= \; & W_{k+1} - (n M_k - C_k) + 4 (V_k - W_{k+1}) - \delta_1 \\
= \; & C_k^* - W_{k+1} - (n-4) V_k - 2 W_{k+1} - \delta_1 \\
> \; & 0
\end{aligned}$$

$$\begin{aligned}
(2)\quad & W_{k+1} - (n M_k^{(2)} - B_k) + 2 (V_k - W_{k+1}) - M_k^{(1)} \\
= \; & B_k - (n-1) M_k^{(2)} + (V_k - M_k) + (V_k - W_{k+1}) \\
> \; & 0
\end{aligned}$$

$$\begin{aligned}
(3)\quad & n M_k^{(1)} - A_k - \delta_2 - (n M_k^{(1)} - A_k - 2 (V_k - W_{k+1}) + \delta_1) \\
= \; & 2 (V_k - W_{k+1}) - (\delta_1 + \delta_2) \\
> \; & 0
\end{aligned}$$

$$(4)\quad n M_k^{(1)} - A_k - \delta_2 - M_k^{(1)} = (n-1) M_k^{(1)} - A_k - \delta_2 > 0$$

$$(5)\quad W_{k+1} - M_k^{(2)} - (n M_k^{(1)} - A_k - 2 (V_k - W_{k+1}) + \delta_1) > 0$$

$$(6)\quad W_{k+1} - M_k^{(2)} - M_k^{(1)} > 0$$

It can also be proved that the conditions II1 - II7 hold and $C_{k+1}^* = C_k^* - W_{k+1}$ up to this point. That is,

$$(n-2) W_{k+1}^{(1)} - 2 (V_k - W_{k+1}) < A_{k+1} < (n-1) W_{k+1}^{(1)},$$

$$(n-2) W_{k+1}^{(2)} - 2 (V_k - W_{k+1}) \le B_{k+1} < (n-1) W_{k+1}^{(2)},$$

$$A_{k+1} < (n-2) W_{k+1}^{(1)},$$

$$M_k^{(1)} \le W_{k+1}^{(1)} \le W_{k+1} - M_k^{(2)}$$

where

$$A_{k+1} = (n-1)\, W^{(1)}_{k+1} - (\, n\, M^{(1)}_{k} - A_{k}\,),$$
$$B_{k+1} = (n-1)\, W^{(2)}_{k+1} - (\, n\, M^{(2)}_{k} - B_{k}\,).$$

It is because

$$\begin{aligned}
& A_{k+1} - (n-2)\, W^{(1)}_{k+1} + 2\, (V_{k} - W_{k+1}) \\
&= W^{(1)}_{k+1} - (\, n\, M^{(1)}_{k} - A_{k}\,) + 2\, (V_{k} - W_{k+1}) \\
&\geq \delta_{1} \\
&> 0,
\end{aligned}$$

$$\begin{aligned}
& B_{k+1} - (n-2)\, W^{(2)}_{k+1} + 2\, (V_{k} - W_{k+1}) \\
&= W_{k+1} - W^{(1)}_{k+1} - (\, n\, M^{(2)}_{k} - B_{k}\,) + 2\, (V_{k} - W_{k+1}) \\
&\geq 0
\end{aligned}$$

and

$$(n-2)\, W^{(1)}_{k+1} - A_{k+1} = (\, n\, M^{(1)}_{k} - A_{k}\,) - W^{(1)}_{k+1} \geq \delta_{2} > 0.$$

After considering W_{k+1}, we reconsider the J_k elements in $S_{3,k}$. Since $(n-2)\, V_k < C^*_k - W_{k+1} < (n-1)\, V_k$, (or equivalently, $(n-2)\, V_k < C^*_{k+1} < (n-1)\, V_k$,), we re-label the index as $(n-2)\, V_k < C^*_k < (n-1)\, V_k$ for convenience and there are J_k elements in $S_{3,k}$. From the previous intermediate step, we have $(n-2)\, V_k < C^*_k < (n-1)\, V_k$, $(n-2)\, M'^{(1)}_k - 2\, (V_k - M'_k) < A'_k < (n-1)\, M'^{(1)}_k$ and $(n-2)\, M'^{(2)}_k - 2\, (V_k - M'_k) \leq B'_k < (n-1)\, M'^{(2)}_k$. Also, $A'_k < (n-2)\, M'^{(1)}_k$. Note that A'_k, B'_k, C'_k, $M'^{(1)}_k$, $M'^{(2)}_k$ and M'_k are updated after the previous intermediate step. Similar in CASE 11, when the J_k elements in $S_{3,k}$ are considered, we let $V_k = W'^{(1)}_{k+1} + W'^{(2)}_{k+1}$ and choose $W'^{(1)}_{k+1}$ such that

$$\max\left((n M_k'^{(1)} - A_k')/2 + \delta_1 ,\ V_k - (n M_k'^{(2)} - B_k') + \delta_2 ,\ M_k'^{(1)} \right) \le W_{k+1}'^{(1)} \le$$

$$\min\left(n M_k'^{(1)} - A_k' - \delta_3 ,\ V_k - (n M_k'^{(2)} - B_k')/2 ,\ V_k - M_k'^{(2)} \right) \qquad (3.13)$$

where

$$0 < \delta_1 < \min\left((n M_k'^{(1)} - A_k')/4 ,\ V_k - (n M_k' - C_k')/2 ,\ \left(2(V_k - M_k') - (n-2) M_k'^{(1)} + A_k' \right)/2 \right),$$

$$0 < \delta_2 < \min\left((n M_k' - C_k' - V_k)/2 ,\ (n M_k'^{(2)} - B_k')/2 ,\ (n-1) M_k'^{(2)} - B_k' \right)$$

and

$$0 < \delta_3 < \min\left((n M_k'^{(1)} - A_k')/4 ,\ (n M_k' - C_k' - V_k)/2 ,\ (n-1) M_k'^{(1)} - A_k' \right)$$

Set $S_{1,k+1} = S_{1,k} \cup \{(R_{k+1}, W_{k+1}^{(1)})\} \cup \{ (R_j, W_{k+1}'^{(1)}) : R_j \in S_{3,k} \}$, $S_{2,k+1} = S_{2,k} \cup \{(R_{k+1}, W_{k+1}^{(2)})\} \cup \{ (R_j, W_{k+1}'^{(2)}) : W_{k+1}'^{(2)} = V_k - W_{k+1}'^{(1)}, R_j \in S_{3,k} \}$ and $S_{3,k+1} = \emptyset$, where $W_{k+1}^{(2)} = W_{k+1} - W_{k+1}^{(1)}$. Thus, $C_{k+1} = (n-J_k) V_k - (n M_k' - C_k')$. It can be proved that the conditions I1 - I5 hold at the end of stage k+1.

CASE 14 : $0 < J_k \le n-1$, $M_k < V_k < W_{k+1}$, $W_{k+1} \ge T_k + J_k V_k$.

From (II4) and $W_{k+1} \ge T_k + J_k V_k$, it implies that $(n-1) W_{k+1} \le C_k^* + n W_{k+1} - (n+J_k) V_k < n W_{k+1}$. From (II4), (II6) and (II5), we have $(n-1) V_k \le C_k^* < n V_k$, $(n-1) M_k^{(1)} - (V_k - M_k) < A_k < n M_k^{(1)}$ and $(n-1) M_k^{(2)} - (V_k - M_k) \le B_k < n M_k^{(2)}$. Also $A_k < (n-2) M_k^{(1)}$. We examine the J_k elements in $S_{3,k}$. Let $V_k = W_{k+1}^{*(1)} + W_{k+1}^{*(2)}$, where $W_{k+1}^{*(1)}$, $W_{k+1}^{*(2)}$ are weights included in $S_{1,k+1}$ and $S_{2,k+1}$ in this stage respectively.

(i) $J_k = 1$.

Choose $W^{*(1)}_{k+1}$ such that

$$M^{(1)}_k \le W^{*(1)}_{k+1} \le \min (n M^{(1)}_k - A_k - \delta_1 , V_k - M^{(2)}_k) \qquad (3.14)$$

where

$$0 < \delta_1 < (n-1) M^{(1)}_k - A_k .$$

Obviously, $(n-1) M^{(1)}_k - A_k$ is greater than zero and the interval of $W^{*(1)}_{k+1}$ in (3.14) is non-empty. Set $S_{1,k+1} = S_{1,k} \cup \{ (R_j, W^{*(1)}_{k+1}) : R_j \in S_{3,k} \}$, $S_{2,k+1} = S_{2,k} \cup \{ (R_j, W^{*(2)}_{k+1}) : W^{*(2)}_{k+1} = V_k - W^{*(1)}_{k+1}, R_j \in S_{3,k} \}$ and $S_{3,k+1} = \{ R_{k+1} \}$. Thus, $C^*_{k+1} = C^*_k + n W_{k+1} - (n+1) V_k$. It can be proved that the conditions II1 - II7 hold at the end of stage k+1. That is,

$$(n-1) W^{*(1)}_{k+1} - (W_{k+1} - V_k) < A_{k+1} < n W^{*(1)}_{k+1},$$
$$(n-1) W^{*(2)}_{k+1} - (W_{k+1} - V_k) \le B_{k+1} < n W^{*(2)}_{k+1},$$
$$A_{k+1} < (n-2) W^{*(1)}_{k+1},$$
$$M^{(1)}_k \le W^{*(1)}_{k+1} \le V_k - M^{(2)}_k .$$

where

$$A_{k+1} = (n-1) W^{*(1)}_{k+1} - (n M^{(1)}_k - A_k),$$
$$B_{k+1} = (n-1) W^{*(2)}_{k+1} - (n M^{(2)}_k - B_k).$$

It is because

$$A_{k+1} - (n-1) W^{*(1)}_{k+1} + (W_{k+1} - V_k)$$
$$= (n-1) W^{*(1)}_{k+1} - (n M^{(1)}_k - A_k) - (n-1) W^{*(1)}_{k+1} + (W_{k+1} - V_k)$$
$$= W_{k+1} - V_k - (n M^{(1)}_k - A_k)$$

$= W_{k+1} - V_k - T_k^{(1)}$

> 0 as

$$(n-1)\, W_{k+1} \le C_k^* + n\, W_{k+1} - (n+1)\, V_k,$$

$$(n-1)\, W_{k+1} \le n\, V_k - T_k + n\, W_{k+1} - (n+1)\, V_k,$$

$$T_k \le W_{k+1} - V_k,$$

since $$T_k^{(1)} < T_k,\ T_k^{(1)} < W_{k+1} - V_k.$$

Therefore, $A_{k+1} > (n-1)\, W_{k+1}^{*(1)} - (W_{k+1} - V_k)$ will hold. Similarly for $B_{k+1} \ge (n-1)\, W_{k+1}^{(2)} - (W_{k+1} - V_k)$.

(ii) $1 < J_k \le n-1$.

Choose $W_{k+1}^{*(1)}$ such that

$$\max\,(\ V_k - [\ W_{k+1} - V_k - (n\, M_k^{(2)} - B_k)\ ] \,/\, (J_k-1),\ \ M_k^{(1)}\ ) \le W_{k+1}^{*(1)} \le$$
$$\min\,(\ [\ W_{k+1} - V_k - (n\, M_k^{(1)} - A_k)\ ] \,/\, (J_k-1) - \delta_1,\ \ n\, M_k^{(1)} - A_k - \delta_2,$$
$$V_k - M_k^{(2)}\ ) \qquad (3.15)$$

where

$$0 < \delta_1 < \min\,(\ [\ 2(W_{k+1} - V_k) - (n\, M_k - C_k) - (J_k-1)\, V_k\ ] \,/\, (J_k-1),$$
$$[\ W_{k+1} - V_k - (n\, M_k^{(1)} - A_k) - (J_k-1)\, M_k^{(1)}\ ] \,/\, (J_k-1)\ )$$

and

$$0 < \delta_2 < \min\,(\ (n\, M_k^{(1)} - A_k) - V_k + [\ W_{k+1} - V_k - (n\, M_k^{(2)} - B_k)\ ]$$
$$/\, (J_k-1)\ ,\ (n-1)\, M_k^{(1)} - A_k\ ).$$

To prove that δ_1 and δ_2 exist, we show that each component in the right hand side is strictly greater than zero.

(1) $[\ 2(W_{k+1} - V_k) - (n\, M_k - C_k) - (J_k-1)\, V_k\ ] \,/\, (J_k-1)$

$= [\ 2(W_{k+1} - V_k) - T_k - J_k\, V_k + V_k\ ] \,/\, (J_k-1)$

$$\geq [\ 2(W_{k+1} - V_k) - W_{k+1} + V_k\ ] \ /\ (J_k-1)$$

$$= (W_{k+1} - V_k) \ /\ (J_k-1)$$

$$> 0$$

(2) $[\ W_{k+1} - V_k - (n\, M_k^{(1)} - A_k) - (J_k-1)\, M_k^{(1)}\ ] \ /\ (J_k-1)$

$$= [\ W_{k+1} - V_k - T_k + (n\, M_k^{(2)} - B_k) - (J_k-1)\, V_k + (J_k-1)(V_k - M_k^{(1)})\ ] \ /(J_k-1)$$

$$\geq [\ (n\, M_k^{(2)} - B_k) + (J_k-1)(\ V_k - M_k^{(1)})\ ] \ /\ (J_k-1)$$

$$> 0$$

(3) $(n\, M_k^{(1)} - A_k) - V_k + [\ W_{k+1} - V_k - (n\, M_k^{(2)} - B_k)\ ] \ /\ (J_k-1)$

$$= [\ (J_k-1)(n\, M_k^{(1)} - A_k) - (J_k-1)V_k + W_{k+1} - V_k - (n\, M_k^{(2)} - B_k)\ ] \ /\ (J_k-1)$$

$$= [\ J_k(n\, M_k^{(1)} - A_k) - T_k - (J_k-1)V_k + W_{k+1} - V_k\ ] \ /\ (J_k-1)$$

$$= [\ J_k(n\, M_k^{(1)} - A_k) - T_k + W_{k+1} - J_k V_k\ ] \ /\ (J_k-1)$$

$$\geq J_k(n\, M_k^{(1)} - A_k) \ /\ (J_k-1)$$

$$> 0$$

(4) $(n-1)\, M_k^{(1)} - A_k > 0.$

We then check the non-emptiness of the interval in (3.15).

(1) $[\ W_{k+1} - V_k - (n\, M_k^{(1)} - A_k)\ ] \ /\ (J_k-1) - \delta_1$

$$- \{\ V_k - [\ W_{k+1} - V_k - (n\, M_k^{(2)} - B_k)\ ] \ /\ (J_k-1)\ \}$$

$$= [\ 2(W_{k+1} - V_k) - (n\, M_k - C_k) - (J_k-1)\, V_k\ ] \ /\ (J_k-1) - \delta_1$$

$$> 0$$

(2) $[\ W_{k+1} - V_k - (n\, M_k^{(1)} - A_k)\ ] \ /\ (J_k-1) - \delta_1 - M_k^{(1)}$

$$= [\ W_{k+1} - V_k - (n\, M_k^{(1)} - A_k) - (J_k-1)\, M_k^{(1)}\ ] \ /\ (J_k-1)\ ) - \delta_1$$

$$> 0$$

(3) $n\,M_k^{(1)} - A_k - \delta_2 - \{ V_k - [W_{k+1} - V_k - (n\,M_k^{(2)} - B_k)] / (J_k-1) \}$

> 0

(4) $n\,M_k^{(1)} - A_k - \delta_2 - M_k^{(1)} = (n-1)\,M_k^{(1)} - A_k - \delta_2 > 0$

(5) $V_k - M_k^{(2)} - \{ V_k - [W_{k+1} - V_k - (n\,M_k^{(2)} - B_k)] / (J_k-1) \}$

$= [W_{k+1} - V_k - (n\,M_k^{(2)} - B_k) - (J_k-1)M_k^{(2)}] / (J_k-1)$

$= [W_{k+1} - V_k - T_k + (n\,M_k^{(1)} - A_k) - (J_k-1)\,V_k + (J_k-1)(V_k - M_k^{(2)})]$
$/(J_k-1)$

$\geq [(n\,M_k^{(1)} - A_k) + (J_k-1)(V_k - M_k^{(2)})] / (J_k-1)$

> 0

(6) $V_k - M_k^{(2)} - M_k^{(1)} > 0$

Set $S_{1,k+1} = S_{1,k} \cup \{ (R_j, W_{k+1}^{*(1)}) : R_j \in S_{3,k} \}$, $S_{2,k+1} = S_{2,k} \cup \{(R_j, W_{k+1}^{*(2)}) : W_{k+1}^{*(2)} = V_k - W_{k+1}^{*(1)}, R_j \in S_{3,k} \}$ and $S_{3,k+1} = \{ R_{k+1} \}$. Thus, $C_{k+1}^{*} = C_k^{*} + n\,W_{k+1} - (n+J_k)\,V_k$. It can be proved that the conditions II1 - II7 hold at the end of stage k+1. That is,

$$(n-1)\,W_{k+1}^{*(1)} - (W_{k+1} - V_k) < A_{k+1} < n\,W_{k+1}^{*(1)},$$
$$(n-1)\,W_{k+1}^{*(2)} - (W_{k+1} - V_k) \leq B_{k+1} < n\,W_{k+1}^{*(2)},$$
$$A_{k+1} < (n-2)\,W_{k+1}^{*(1)},$$
$$M_k^{(1)} \leq W_{k+1}^{*(1)} \leq V_k - M_k^{(2)}.$$

where

$$A_{k+1} = (n-J_k)\,W_{k+1}^{*(1)} - (n\,M_k^{(1)} - A_k),$$
$$B_{k+1} = (n-J_k)\,W_{k+1}^{*(2)} - (n\,M_k^{(2)} - B_k).$$

It is because

$$A_{k+1} - (n-1)\,W_{k+1}^{*(1)} + (W_{k+1} - V_k)$$
$$= (n-J_k)\,W_{k+1}^{*(1)} - (n\,M_k^{(1)} - A_k) - (n-1)\,W_{k+1}^{*(1)} + (W_{k+1} - V_k)$$

$$= W_{k+1} - V_k - (n M_k^{(1)} - A_k) - (J_k-1) W_{k+1}^{*(1)}$$
$$\geq (J_k-1) \delta_1$$
$$> 0;$$

and

$$B_{k+1} - (n-1) W_{k+1}^{*(2)} + (W_{k+1} - V_k)$$
$$= W_{k+1} - V_k - (n M_k^{(2)} - B_k) - (J_k-1) V_k + (J_k-1) W_{k+1}^{*(1)}$$
$$\geq 0$$

CASE 15 : $0 < J_k \leq n-1$, $M_k < V_k < W_{k+1} < T_k + J_k V_k$.

From $W_{k+1} < T_k + J_k V_k$, it implies that $C_k^* + n W_{k+1} - (n+J_k) V_k < (n-1) W_{k+1}$. J_k elements in $S_{3,k}$ (= $W_{k+1}^{*(1)} + W_{k+1}^{*(2)}$) and W_{k+1} will be examined in the following three steps.

Step 1: find an $m(1 \leq m \leq J_k)$ such that

$$(n-2) W_{k+1} < C_k^* + n W_{k+1} - (n+m) V_k < (n-1) W_{k+1} \qquad (3.16)$$

is satisfied.

(i) $J_k = 1$.

Obviously, $(n-2) W_{k+1} < C_k^* + n W_{k+1} - (n+1) V_k < (n-1) W_{k+1}$. Define $C_k^{*1} = C_k^* + n (W_{k+1} - V_k)$. It is the CASE 13 and hence, choose $W_{k+1}^{*(1)}$ such that

$$\max (n M_k^{(1)} - A_k - 2 (W_{k+1} - V_k) + \delta_1 , M_k^{(1)}) \leq W_{k+1}^{*(1)} \leq$$
$$\min(V_k - (n M_k^{(2)} - B_k) + 2 (W_{k+1} - V_k) , n M_k^{(1)} - A_k - \delta_2 ,$$
$$V_k - M_k^{(2)}) \qquad (3.17)$$

where

$$0 < \delta_1 < \min (C_k^{*1} - V_k - (n-4) W_{k+1} - 2 V_k, \; V_k - M_k^{(2)} - n M_k^{(1)} + A_k + 2 (W_{k+1} - V_k), \; W_{k+1} - V_k)$$

and

$$0 < \delta_2 < \min ((n-1) M_k^{(1)} - A_k , W_{k+1} - V_k).$$

(ii) $1 < J_k \le n-1$.

We first check that whether m=1 makes the inequality (3.16) satisfy. If it does, choose $W_{k+1}^{*(1)}$ by the interval given in (3.17) and set m=1; otherwise, it is the CASE 14 and choose $W_{k+1}^{*(1)}$ such that

$$M_k^{(1)} \le W_{k+1}^{*(1)} \le \min (n M_k^{(1)} - A_k - \delta_1 , V_k - M_k^{(2)}) \qquad (3.18)$$

where

$$0 < \delta_1 < (n-1) M_k^{(1)} - A_k.$$

The remaining (J_k-1) elements in $S_{3,k}$ are examined one-by-one and it is followed by CASE 10 (or CASE 11). In either cases, $W_{k+1}^{*(1)}$ will be chosen with the identical weights which is chosen in (3.18). This step stops until CASE 11 is reached. That is, there is a m (> 1) such that (3.16) is satisfied. Denote $S_{3,k}^{'} \subseteq S_{3,k}$ be the set containing the records examined in this step.

Step 2: temporary omit the (J_k-m) elements in $S_{3,k}$ and examine W_{k+1} $(= W_{k+1}^{(1)} + W_{k+1}^{(2)})$.

After step 1, (3.16) holds. It resembles the consideration of the maximum weight in CASE 13 (or CASE 11). Since $(n-2)\ W_{k+1} < C_k^{*} + n\ W_{k+1} - (n+m)\ V_k < (n-1)\ W_{k+1}$ (or equivalently, $(n-2)\ W_{k+1} < C_{k+1}^{'*} < (n-1)\ W_{k+1}$), we re-label the index as $(n-2)\ V_k' < C_k^{'*} < (n-1)\ V_k'$ for convenience. From the previous intermediate step, we have $(n-2)\ V_k' < C_k^{'*} < (n-1)\ V_k'$, $(n-2)\ M_k^{'(1)} - 2\ (V_k' - M_k') < A_k' < (n-1)\ M_k^{'(1)}$ and $(n-2)\ M_k^{'(2)} - 2\ (V_k' - M_k') \le B_k' < (n-1)\ M_k^{'(2)}$. Also, $A_k' < (n-2)\ M_k^{'(1)}$. Note that A_k', B_k', C_k', $M_k^{'(1)}$, $M_k^{'(2)}$, M_k' and V_k' are updated after step 1. Choose $W_{k+1}^{(1)}$ such that

$$\max\ (\ (n\ M_k^{'(1)} - A_k')/2 + \delta_1\ , W_{k+1} - (n\ M_k^{'(2)} - B_k') + \delta_2\ ,\ M_k^{'(1)}\ ) \le W_{k+1}^{(1)}$$
$$\le \min(\ n\ M_k^{'(1)} - A_k' - \delta_3\ ,\ W_{k+1} - (n\ M_k^{'(2)} - B_k')/2\ ,\ W_{k+1} - M_k^{'(2)}\ ) \qquad (3.19)$$

where

$$0 < \delta_1 < \min\ (\ (n\ M_k^{'(1)} - A_k')/4\ ,\ W_{k+1} - (n\ M_k' - C_k')/2\ ,\ (\ 2(W_{k+1} - M_k') - (n-2)\ M_k^{'(1)} + A_k'\ )/2\ )\ ,$$
$$0 < \delta_2 < \min\ (\ (n\ M_k' - C_k' - W_{k+1})/2,\ (n\ M_k^{'(2)} - B_k')/2,\ (n-1)\ M_k^{'(2)} - B_k')$$

and

$$0 < \delta_3 < \min\ (\ (n\ M_k^{'(1)} - A_k')/4\ ,\ (n\ M_k' - C_k' - W_{k+1}\ )/2\ ,\ (n-1)\ M_k^{'(1)} - A_k\ )\ .$$

If we temporary omit the $(J_k - m)$ elements in $S_{3,k}$ after step 2, condition I1 - I5 are satisfied.

Step 3: re-examine the remaining $(J_k - m)$ elements in $S_{3,k}$.

In this case, since the maximum weight has been examined in step 2 and conditions I1 - I5 are satisfied, the remaining (J_k-m) elements in $S_{3,k}$ are simply the CASE 1, CASE 2 or CASE 3. We can examine them one-by-one such that conditions I1 - I5 hold at the end of this step. Denote $S_{3,k} \setminus S'_{3,k}$ contains the records examined in this step.

$S_{1,k+1} = S_{1,k} \cup \{ (R_j, W^{*(1)}_{k+1}) : R_j \in S'_{3,k} \} \cup \{ (R_{k+1}, W^{(1)}_{k+1}) \} \cup \{ (R_j, W^{(1)}_j) : R_j \in S_{3,k} \setminus S'_{3,k} \}$, $S_{2,k+1} = S_{2,k} \cup \{ (R_j, W^{*(2)}_{k+1}) : , W^{*(2)}_{k+1} = V_k - W^{*(1)}_{k+1}, R_j \in S'_{3,k} \} \cup \{ (R_{k+1}, W^{(2)}_{k+1}), W^{(2)}_{k+1} = W_{k+1} - W^{(1)}_{k+1} \} \cup \{ (R_j, W^{(1)}_j) : R_j \in S_{3,k} \setminus S'_{3,k}, W^{(2)}_j = W_j - W^{(1)}_j \}$ and $S_{3,k+1} = \emptyset$.

§3.3 Initialization

After examining each case, we consider the initialization step - to guarantee that the conditions (CASE I or CASE II in section 3.1) are satisfied at the end of stage k, where $k \le 2n$. We rearrange the first 2n-1 weights such that $W_1 \le W_2 \le \ldots \le W_{2n-1}$ and examine these weights one-by-one in non-descending order. The order of the remaining weights are preserved and are denoted as $W_{2n}, W_{2n+1} \ldots$. Choose U such that $0 < U < W_1$ and set $W^{(1)}_1 = W^{(1)}_2 = U$. Hence,

$$S_{1,2} = \{ (R_1, U), (R_2, U) \},$$

$$S_{2,2} = \{ (R_1, W_1 - U), (R_2, W_2 - U) \}$$

and

$$S_{3,2} = \emptyset.$$

Besides, we have

$$C_2^* = C_2 = n\ W_2 - (\ W_1 + W_2\ ) = (n-1)\ W_2 - W_1 \geq 0,$$

$$A_2 = n\ M_2^{(1)} - 2\ M_2^{(1)} = (n-2)\ M_2^{(1)},$$

$$B_2 = n\ (W_2 - U) - (\ (W_1 - U) + (W_2 - U)\ )$$

$$= (n-1)\ (W_2 - U) - (W_1 - U).$$

Thus,

$$(n-2)\ W_2 \leq C_2 < (n-1)\ W_2,$$

$$(n-2)\ M_2^{(1)} \leq A_2 < (n-1)\ M_2^{(1)},$$

$$(n-2)\ (W_2 - U) \leq B_2 < (n-1)\ (W_2 - U).$$

Clearly, $A_2 = (n-2)\ M_2^{(1)}$ contradicting condition I3. Moreover, condition I5 may also fail. In stage three, the third record is read in. It will be followed by one of the following cases: CASE 1, CASE 3, CASE 4, CASE 6 or CASE 7. If it is followed by CASE 1, CASE 4 or CASE 6, W_3 can be divided into two parts as discussed in the previous section such that

$$S_{1,3} = S_{1,2} \cup \{\ (R_3,\ W_3^{(1)})\ \},$$

$$S_{2,3} = S_{2,2} \cup \{\ (R_3,\ W_3 - W_3^{(1)})\ \}$$

and

$$S_{3,3} = \emptyset, \text{ where } 0 < W_3^{(1)} \leq W_3.$$

All conditions I1 - I5 are satisfied after stage three because from stage two to stage three, there is no need for the conditions I3 or I5 to hold at stage two. If it is followed by CASE 3, conditions I1 - I4 will hold after stage three because, again, from stage two to stage three, there is no need for the condition I3 to hold at stage two.

However, condition I5 still fails and will hold if (i) the next incoming weight is strictly greater than the current maximum weight or (ii) $W_3 = W_4 = \dots = W_{n+1}$. Then, $C_k < 0$ and there is no need to satisfy condition I5. Either one of the above two cases must occur.

On the other hand, if it is followed by CASE 7, W_3 cannot be divided into two parts in stage three and we have $S_{1,3} = S_{1,2}$, $S_{2,3} = S_{2,2}$ and $S_{3,3} = \{ R_3 \}$. The following records are read in one-by-one. There will be two main cases.

(i) $W_3 = W_4 = \dots = W_{n+2}$. It is the CASE 9 and the weights W_3, W_4, ..., W_{n+2} are divided into two parts according CASE 9. After the stage, condition I1 - I5 will hold even conditions I3 and I5 are not satisfied at stage two.

(ii) $W_3 = W_4 = \dots = W_{m+3}$, for $1 \le m < n-1$ and V_{m+4} ($> W_3$). Then, it will come to one of the following cases: CASE 12 - CASE 15. In CASE 12, V_{m+4} is divided into two parts accordingly. After this stage, condition II1 - II7 will hold even conditions I3 or I5 are not satisfied at stage two. In CASE 13, V_{m+4} is divided into two parts and followed by the weights W_j, where $R_j \in S_{3,m+4}$. All conditions I1 - I5 are satisfied after this stage because from stage two to this stage, there is no need for the conditions I3 or I5 to hold at stage two. In CASE 14, the weights W_j, where $R_j \in S_{3,m+4}$, are divided into two parts and after this stage, conditions II1 - II7 will hold though conditions I3 or I5 are not satisfied at stage two. Lastly, in CASE 15, the weights W_j, where $R_j \in S'_{3,m+4}$, are divided into two parts, followed by V_{m+4} and the weights W_j, where $R_j \in S_{3,m+4} \setminus S'_{3,m+4}$. Conditions I1 - I5 will satisfy after this stage even conditions I3 or I5 are not satisfied at stage two.

The reason why the first 2n-1 weights are sorted is as follows. To guarantee that at stage k', there are at most (n-2) self-selective units in $S_{1,k'}$, we have to ensure that the $(n-1)^{th}$ largest weight in $S_{1,k'}$ is smaller than the sum of the weights which are included in $S_{1,k'}$ and smaller than that weight. The proof will be given in Theorem 3.4. Since for the first two units in $S_{1,2}$, we include two identical weights U. As the incoming weights are sorted in non-descending order, a new maximum weight, which is smaller than the sum of the weights already in $S_{1,k'}$ will be included in $S_{1,k}$ at stage k (>2). For the worst case, after 2n-1 units are examined, there are n units included in $S_{1,2n-1}$ and n-1 units in $S_{3,2n-1}$ with identical weights. For the n units included in the $S_{1,2n-1}$, the $(n-1)^{th}$ unit is of weight at least U. When the $(2n)^{th}$ unit is examined, at least an element will divide into two parts out of the CASE 9 - CASE 15 and is included in $S_{1,2n}$ and $S_{2,2n}$ respectively, no matter what the $(2n)^{th}$ weight is. After examining at most 2n units, the $(n-1)^{th}$ largest weight in $S_{1,2n}$ is smaller than the sum of the weights which are included in $S_{1,2n}$ and smaller than that weight. Besides, $S_{1,2n}$ contains at least (n+1) elements when the initialization step finishes.

There is a property when the initialization steps are finished. Define $W^{(1)}_{k,(i)}$ be the i^{th} largest weight included in $S_{1,k}$ and therefore, we have $W^{(1)}_{k,(1)} \geq W^{(1)}_{k,(2)} \geq \dots \geq W^{(1)}_{k,(n-1)}$. Also, define the following function,

$$I_{(W^{(1)}_{k,(i)} < W^{(1)}_{k,(j)})} = 1 \quad \text{if} \quad W^{(1)}_{k,(i)} < W^{(1)}_{k,(j)};$$

$$= 0 \quad \text{otherwise.}$$

Recall the initialization step, records are read in with non-descending order and the procedure continues. Note that at stage j ($3 \le j$), the weight W_u ($3 \le u \le j$) is divided into two parts, $W^{*(1)}_u$ and $W^{*(2)}_u$ with $M^{(1)}_j = W^{*(1)}_u$ and $M^{(2)}_j = W^{*(2)}_u$. Since $A_j < (n-2)\ M^{(1)}_j$ is satisfied, we have $n\ M^{(1)}_j - \sum_{i \in L_j} W^{(1)}_i < (n-2)\ M^{(1)}_j$ or equivalently, $M^{(1)}_j < \sum_{i \in L_j \setminus \{u\}} W^{(1)}_i$. For each of the weight divided, it satisfies I3 (or II3). Therefore, after stage k' ($\le 2n$), where the size of $L_{k'}$ is at least n+1, we have

$$W^{(1)}_{k',(j)} < \sum_{i \in L_{k'}} W^{(1)}_{k',(i)}\ I_{(W^{(1)}_{k',(i)} < W^{(1)}_{k',(j)})}, \text{ for } j = 1, \dots, n-1$$

where

$$L_{k'} \subseteq \{ 1, \dots, k' \},\ 1 \le k' \le N.$$

§ 3.4 Final step

As described previously, Chao's algorithm is applied on $S_{1,k}$ and $S_{2,k}$ respectively to select two samples of size n. Thus, after the file is read through one time, two samples of size n, S_1 and S_2, say, are obtained from $S_{1,N}$ and $S_{2,N}$. We have the following five possible cases:

(i) $J_N = 0,\ C_N > 0,$

(ii) $J_N = 0$, $C_N \le 0$,

(iii) $0 < J_N < n-1$,

(iv) $J_N = n-1$, $(n-1)\ V_N < C_N^*$, and

(v) $J_N = n-1$, $(n-1)\ V_N = C_N^*$.

For case (i), $n\ M_N\ /\ T_N = (C_N/T_N) + 1 > 1$. By Theorem 3.1, no sample exists. For case (iii), there are J_N elements in $S_{3,N}$ and $(n-1)\ V_N \le C_N^* < n\ V_N$. Note that C_N^* is the least weights required for the sample to exist. $\dfrac{n\ V_N}{\sum_{j=1}^{N} W_j} = \dfrac{n\ V_N}{(n+J_N)\ V_N - C_N^*} \ge$

$\dfrac{n\ V_N}{(n+(n-2))\ V_N - (n-1)\ V_N} > 1$. From Theorem 3.1, no sample exists.

Similarly, for case (iv), $\dfrac{n\ V_N}{\sum_{j=1}^{N} W_j} = \dfrac{n\ V_N}{(n+n-1)\ V_N - C_N^*} >$

$\dfrac{n\ V_N}{(2n-1)\ V_N - (n-1)\ V_N} = 1$. By Theorem 3.1, it implies that the sample does not exist.

For case (ii), define $0 < p \equiv \dfrac{\sum_{u=1}^{N} W_u^{(1)}}{\sum_{j=1}^{N} W_j} < 1$. Then, with probability p, we select S_1 as our sample and with probability 1-p, select S_2 as our sample. As will be proved in Theorem 3.7, a weighted sample with positive second order inclusion probabilities is obtained.

For case (v), $n\ W_i\ /\ T_N = 1$ for all i such that $R_i \in S_{3,N}$. We have (n-1) self-selective units. By Theorem 3.2, no sample with all second order inclusion probability positive may exist. Recall $T_N = \sum_{i=1}^{N} W_i$, $T_N^{(1)} = \sum_{i=1}^{N} W_i^{(1)}$, $T_N^{(2)} = \sum_{i=1}^{N} W_i^{(2)}$. We need to perform one additional step. Because $(n-1)\ V_N = C_N^*$, we have $V_N = T_N$. Consider the updated C_N, C_N' say,

$$C_N' = n\ V_N - (\ T_N + (n-1)\ V_N\ ) = 0.$$

If, further, V_N $(= W^{*(1)} + W^{*(2)})$ is divided into two parts, $W^{*(1)} = T_N^{(1)}$ and $W^{*(2)} = T_N^{(2)}$, then

$$A_N' = n\ T_N^{(1)} - (\ T_N^{(1)} + (n-1)\ T_N^{(1)}\ ) = 0,$$

$$B_N' = n\ T_N^{(2)} - (\ T_N^{(2)} + (n-1)\ T_N^{(2)}\ ) = 0.$$

Update the new $S_{1,N}$ by $S_{1,N} \cup \{\ (R_j, W^{*(1)}) : R_j \in S_{3,N}\ \}$ ($= S_{1,N}'$ say), the new $S_{2,N}$ by $S_{2,N} \cup \{\ (R_j, W^{*(2)}) : R_j \in S_{3,N}, W^{*(2)} = V_N - W^{*(1)}\ \}$ ($= S_{2,N}'$ say), and $S_{3,N}' = \emptyset$. Thus, there are no more records in $S_{3,N}'$. Define $0 < p = \dfrac{\sum_{u=1}^{N} W_u^{(1)}}{\sum_{j=1}^{N} W_j} < 1$. Then, with probability p, we select S_1 as our sample and with probability 1-p , select S_2 as our sample. As will be proved in Theorem 3.6, a weighted sample is obtained.

§3.5 Theorems

The following Theorem 3.1 is a well-known theorem for the necessary and sufficient conditions of the existence of a weighted sample.

Theorem 3.1

Given that $W_i > 0$ for $i = 1, \ldots, N$, a weighted sample of size n ($1 \le n \le N$) exists if, and only if $\frac{n W_i}{\sum_{j=1}^{N} W_j} \le 1$ for $i = 1, \ldots, N$.

Theorem 3.2

Given that a weighted sample of size n ($2 \le n < N$) exists from N units, with $W_i > 0$ for $i = 1, \ldots, N$. Then, there is a weighted sample of size n with all the second order inclusion probability strictly positive if and only if there are at most n-2 self-selective units in the sample.

Proof of Theorem 3.2:

Recall that if $\frac{n W_i}{\sum_{j=1}^{N} W_j} = 1$, W_i is called a self-selective unit.

Without loss of generality, suppose that $W_1 \le W_2 \le \ldots \le W_N$ and there are k self-selective units in the population. Assume $0 \le k \le n-2$. Denote

$T_j = \sum_{i=1}^{j} W_i$. We have $T_N = T_{N-k} + k\, W_N$ and $W_i = \frac{T_N}{n}$ for $i = N-k+1, \ldots, N$. As the k self-selective units must be included in the sample, we need only to select n-k units from the remaining N-k units. Let

S_1: a simple random sample of size n-k without replacement from the remaining N-k units

S_2: a weighted sample of size n-k without replacement with weight W_i^*, for $i = 1, \ldots, N-k$,

where

$$W_i^* = W_i - \frac{p\, T_{N-k}}{N-k}, \qquad i = 1, \ldots, N-k$$

and

$$0 < p < \min\left(\frac{(n-k)\, W_1}{T_{N-k}}, \frac{N-k}{N-n}\left(1 - \frac{(n-k)\, W_{N-k}}{T_{N-k}} \right) \right) < 1.$$

Let S_3 be the a set containing those k self-selective units. We accept $S_1^* = S_1 \cup S_3$ as our sample S with probability p and accept $S_2^* = S_2 \cup S_3$ with probability 1-p.

S_1^* exists as the remaining n-k units are drawn using simple random sampling without replacement. Besides, as $p < \frac{(n-k)\, W_1}{T_{N-k}}$, that is, $(n-k)\, W_1 - p\, T_{N-k} > 0$ and thus, $W_1^* > 0$. Because $W_1^* \le W_2^* \le \ldots \le W_{N-k}^*$, $W_i^* > 0$ for $i = 1, \ldots, N-k$. Consider

$$(n-k)\, W_i^* - \sum_{j=1}^{N-k} W_j^*$$

$$= (n-k)\left(W_i - \frac{p\, T_{N-k}}{N-k} \right) - (1-p)\, T_{N-k}$$

$$\le \left((n-k)\, W_{N-k} - T_{N-k} \right) + \frac{p\,(N-n)\, T_{N-k}}{N-k}$$

$$< \left((n-k)\, W_{N-k} - T_{N-k} \right) + T_{N-k} \left(1 - \frac{(n-k)\, W_{N-k}}{T_{N-k}} \right)$$

$$= 0.$$

Thus, $(n-k)\, W_i^* - \sum_{i=1}^{N-k} W_i^* < 0$ for $i = 1, \ldots, N-k$. That is,

$$\frac{(n-k)\, W_i^*}{\sum_{i=1}^{N-k} W_i^*} < 1 \text{ , for } i = 1, \ldots, N-k.$$

From Theorem 3.1, it implies that the sample from S_2^* exists. For $N-k+1 \le i \le N$,

$$\Pr(\, i \in S \,) = 1 = \frac{n\, W_i}{T_N} .$$

For $1 \le i \le N-k$,

$$\Pr(\, i \in S \,) = p \Pr(\, i \in S_1 \,) + (1-p) \Pr(\, i \in S_2 \,)$$

$$= p \frac{n-k}{N-k} + (1-p) \frac{(n-k)\, W_i^*}{\sum_{i=1}^{N-k} W_i^*}$$

$$= p \frac{n-k}{N-k} + (1-p) \frac{(n-k)\left((N-k)\, W_i - p\, T_{N-k} \right)}{(N-k)\,(1-p)\, T_{N-k}}$$

$$= \frac{(n-k)\, W_i}{T_{N-k}}$$

$$= \frac{n\, W_i}{T_N} .$$

A sample with probabilities proportional to weight is obtained. Next we consider the second order inclusion probability.

For $N-k+1 \le i,j \le N$, $Pr(i,j \in S) = 1$.

For $N-k+1 \le i \le N$ and $1 \le j \le N-k$,

$$Pr(i,j \in S) = Pr(j \in S) = \frac{n\, W_j}{T_N} > 0.$$

For $1 \le i,j \le N-k$,

$$\begin{aligned} & Pr(i,j \in S) \\ &= p\, Pr(i,j \in S_1) + (1-p)\, Pr(i,j \in S_2) \\ &\ge p\, Pr(i,j \in S_1) \\ &> 0. \end{aligned}$$

The last statement holds as $n-k \ge 2$. Thus, if there are at most n-2 self-selective units in the sample, there is a weighted sample such that the second order inclusion probabilities are strictly positive.

The 'only if' part can be shown by contradiction. Assume that there are more than n-2 self-selective units, that is, there are n-1 self-selective units. Because only one unit is selected from N-n+1 (≥ 2) non-self-selective units, the second order inclusion probability of any two non-self-selective units is zero. Q.E.D.

Theorem 3.3

Suppose at stage k ($n < k \leq N$) of the Chao's algorithm, the k weights are sorted in descending order and are denoted as $W_{(1)} \geq W_{(2)} \geq \ldots \geq W_{(k)}$. If

$$Q_j \equiv \frac{(n-j+1)\, W_{(j)}}{\sum_{t=j}^{k} W_{(t)}} < 1 , \qquad j = 1, \ldots, n-1,$$

then $W_{(j)}$ is a non-self-selective unit.

Proof of Theorem 3.3:

First, if $W_{(j)}$ is a non-self-selective unit, $W_{(j+1)}$ is also a non-self-selective unit.

To determine whether a weight, $W_{(j)}$, is self-selective or non-self-selective at stage k ($n < k \leq N$), we first check the value

$$Q_1 = \frac{n\, W_{(1)}}{\sum_{t=1}^{k} W_{(t)}} .$$

If this value is less than one, $W_{(1)}$ is a non-self-selective unit. With the above argument, all weights are also non-self-selective units. If this value is greater than or equal to one, $W_{(1)}$ is a self-selective unit and we further consider the value (by reducing the sample size by one)

$$Q_2 = \frac{(n-1)\, W_{(2)}}{\sum_{t=2}^{k} W_{(t)}} = \frac{(n-2+1)\, W_{(2)}}{\sum_{t=2}^{k} W_{(t)}} .$$

Again, if this value is less than one, $W_{(2)}$ (and those less than or equal to $W_{(2)}$) is a non-self-selective unit. Otherwise, it is a self-selective unit and continues the procedure. Suppose $Q_j < 1$. Let i be the smallest integer such that $1 \le i \le j$ and $Q_i < 1$. From above argument, $W_{(i)}$ is a non-self-selective unit. Hence, $W_{(j)}$ is a non-self-selective unit. Q.E.D.

Theorem 3.4

Suppose at stage k ($n < k \le N$) of the Chao's algorithm, the k weights are sorted in descending order and are denoted as $W_{(1)} \ge W_{(2)} \ge \ldots \ge W_{(k)}$. If $W_{(n-1)} < \sum_{t=n}^{k} W_{(t)}$, there are at most n-2 self-selective units in the sample at stage k.

Proof of Theorem 3.4:

At stage k, if $W_{(n-1)} < \sum_{t=n}^{k} W_{(t)}$, then $2\, W_{(n-1)} < \sum_{t=n-1}^{k} W_{(t)}$, and thus $\frac{2\, W_{(n-1)}}{\sum_{t=n-1}^{k} W_{(t)}} < 1$. By Theorem 3.3, $W_{(n-1)}$ is a non-self-selective unit. Therefore, there are at most n-2 self-selective units in the sample at stage k. Q.E.D.

Theorem 3.5

For the proposed algorithm, when $n < k \le N$ and the size of L_k is at least n+1, we have

$$W^{(1)}_{k,(j)} < \sum_{i \in L_k} W^{(1)}_{k,(i)} \, I_{(W^{(1)}_{k,(i)} < W^{(1)}_{k,(j)})}, \quad j = 1, \ldots, n-1$$

where

$W^{(1)}_{k,(i)}$: the i^{th} largest weight included in $S_{1,k}$,

$W^{(1)}_{k,(1)} \ge W^{(1)}_{k,(2)} \ge \ldots \ge W^{(1)}_{k,(n-1)}$,

$L_k \subseteq \{1, \ldots, k\}$, $1 \le k \le N$, which is defined in section 3.1

and

$$I_{(W^{(1)}_{k,(i)} < W^{(1)}_{k,(j)})} = 1 \quad \text{if } W^{(1)}_{k,(i)} < W^{(1)}_{k,(j)};$$
$$= 0 \quad \text{otherwise.}$$

Proof of Theorem 3.5:

It can be proved by mathematical induction on k. Let k' be the smallest stage number where there are exactly n+1 elements in $S_{1,k'}$. From section 3.3, we have

$$W^{(1)}_{k',(j)} < \sum_{i \in L_{k'}} W^{(1)}_{k',(i)} \, I_{(W^{(1)}_{k',(i)} < W^{(1)}_{k',(j)})} \quad \text{for } j = 1, \ldots, n-1.$$

Assume that the theorem is true for k=m. At stage m ($n < k' \le m < N$),

$$W^{(1)}_{m,(j)} < \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I_{(W^{(1)}_{m,(i)} < W^{(1)}_{m,(j)})} \quad j = 1, \ldots, n-1.$$

It can be shown that at each step in a particular stage, when the condition I3 or II3 holds, the largest weight included in $S_{1,k}$ is smaller than the sum of the weights smaller than the largest weight included. Consider k = m+1. If $S_{1,m+1}$ and $S_{2,m+1}$ are unchanged, clearly the inequalities hold. We will consider the stages where only one element enters $S_{1,m+1}$. The proof when more than one elements enter $S_{1,m+1}$ is similar and therefore is omitted. At stage m+1, let $W_{m+1} = W^{*(1)}_{m+1} + W^{*(2)}_{m+1}$ and $(R_{m+1}, W^{*(1)}_{m+1})$, $(R_{m+1}, W^{*(2)}_{m+1})$ be the elements included in $S_{1,m}$, $S_{2,m}$ to form $S_{1,m+1}$, $S_{2,m+1}$ respectively. There are three possible cases.

(i) $W^{*(1)}_{m+1} > M^{(1)}_{m} = W^{(1)}_{m,(1)}$.

Since $A_{m+1} < (n-2)\, M^{(1)}_{m+1} = (n-2)\, W^{*(1)}_{m+1}$, we have $W^{*(1)}_{m+1} < \sum_{i\in L_m} W^{(1)}_{m,(i)}$.

Consider

$$\begin{aligned} W^{(1)}_{m+1,(n-1)} &= W^{(1)}_{m,(n-2)} \\ &< \sum_{i\in L_m} W^{(1)}_{m,(i)}\, I\,(W^{(1)}_{m,(i)} < W^{(1)}_{m,(n-2)}) \\ &= \sum_{i\in L_{m+1}} W^{(1)}_{m+1,(i)}\, I\,(W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(n-1)}) . \end{aligned}$$

Also, $W^{(1)}_{m+1,(j)} = W^{(1)}_{m,(j-1)}$ for j = 2,... , n-2. Therefore, we have

$$W^{(1)}_{m+1,(j)} < \sum_{i\in L_{m+1}} W^{(1)}_{m+1,(i)}\, I\,(W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(j)}) \text{ for } j = 1 ,\dots,n-1.$$

(ii) $W^{(1)}_{m,(n-1)} < W^{*(1)}_{m+1} \le M^{(1)}_{m}$ (or $W^{(1)}_{m,(n-1)} \le W^{(1)}_{m,(t+1)} < W^{*(1)}_{m+1} \le W^{(1)}_{m,(t)} \le M^{(1)}_{m}$ for some positive integer t). We first consider $W^{(1)}_{m+1,(t)}$,

$W^{(1)}_{m+1,(t+1)}$ and $W^{(1)}_{m+1,(j)}$ for j = t+2, ... , n-1.

Consider

$$\begin{aligned} W^{(1)}_{m+1,(t)} &= W^{(1)}_{m,(t)} \\ &< \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I\left(W^{(1)}_{m,(i)} < W^{(1)}_{m,(t)} \right) \\ &< \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I\left(W^{(1)}_{m,(i)} < W^{(1)}_{m,(t)} \right) + W^{*(1)}_{m+1} \\ &= \sum_{i \in L_{m+1}} W^{(1)}_{m+1,(i)} \, I\left(W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(t)} \right) . \end{aligned}$$

Consider

$$\begin{aligned} W^{(1)}_{m+1,(t+1)} &= W^{*(1)}_{m+1} \\ &\le W^{(1)}_{m,(t)} \\ &< \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I\left(W^{(1)}_{m,(i)} < W^{(1)}_{m,(t)} \right) \\ &= \sum_{i \in L_{m+1}} W^{(1)}_{m+1,(i)} \, I\left(W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(t+1)} \right) . \end{aligned}$$

For j = t+2, ... , n-1, consider

$$\begin{aligned} W^{(1)}_{m+1,(j)} &= W^{(1)}_{m,(j-1)} \\ &< \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I\left(W^{(1)}_{m,(i)} < W^{(1)}_{m,(j-1)} \right) \\ &= \sum_{i \in L_{m+1}} W^{(1)}_{m,(i)} \, I\left(W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(j)} \right) . \end{aligned}$$

Also, $W^{(1)}_{m+1,(j)} = W^{(1)}_{m,(j)}$ for j = 1, ..., t. Therefore, we have

$$W^{(1)}_{m+1,(j)} < \sum_{i \in L_{m+1}} W^{(1)}_{m+1,(i)} \, I \, (W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(j)}) \text{ for } j = 1 ,\ldots,n-1.$$

(iii) $W^{*(1)}_{m+1} \le W^{(1)}_{m,(n-1)}$.

Consider

$$\begin{aligned} W^{(1)}_{m+1,(n-1)} &= W^{(1)}_{m,(n-1)} \\ &< \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I \, (W^{(1)}_{m,(i)} < W^{(1)}_{m,(n-1)}) \\ &< \sum_{i \in L_m} W^{(1)}_{m,(i)} \, I \, (W^{(1)}_{m,(i)} < W^{(1)}_{m,(n-1)}) + W^{*(1)}_{m+1} \\ &= \sum_{i \in L_{m+1}} W^{(1)}_{m,(i)} \, I \, (W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(n-1)}) . \end{aligned}$$

Also, $W^{(1)}_{m+1,(j)} = W^{(1)}_{m,(j)}$ for $j = 1, \ldots, n-2$. Therefore, we have

$$W^{(1)}_{m+1,(j)} < \sum_{i \in L_{m+1}} W^{(1)}_{m+1,(i)} \, I \, (W^{(1)}_{m+1,(i)} < W^{(1)}_{m+1,(j)}) \text{ for } j = 1,\ldots,n-1.$$

By mathematical induction,

$W^{(1)}_{k,(j)} < \sum_{i \in L_k} W^{(1)}_{k,(i)} \, I \, (W^{(1)}_{k,(i)} < W^{(1)}_{k,(j)})$ for $j = 1 , \ldots , n-1$, is true for all positive integer k, where $n < k \le N$ and the size of L_k is at least n+1. Q.E.D.

<u>Theorem 3.6</u>

A weighted sample of size n (≥ 2) is drawn from N units, where $W_i > 0$ for $i = 1, \ldots , N$ and $n < N$. The proposed algorithm can find a weighted sample if the following conditions are satisfied:

(i) $J_N = 0$ and $C_N \le 0$; or

(ii) $J_N = n-1$ and $(n-1)\, V_N = C^*_N$.

<u>Proof of Theorem 3.6:</u>

Without loss of generality, assume that $W_1 \le W_2 \le \ldots \le W_N$. Consider (i) $J_N = 0$ and $C_N \le 0$. Since $C_N \le 0$, we have $A_N \le 0$ and $B_N \le 0$. It implies that $\dfrac{n\, W_N^{(1)}}{\sum_{j=1}^{N} W_j^{(1)}} \le 1$ and $\dfrac{n\, W_N^{(2)}}{\sum_{j=1}^{N} W_j^{(2)}} \le 1$. By Theorem 3.1, both samples, S_1 and S_2, which are the samples drawn from $S_{1,N}$ and $S_{2,N}$ respectively, exist. Define $0 < p = \dfrac{\sum_{u=1}^{N} W_u^{(1)}}{\sum_{j=1}^{N} W_j} < 1$. We accept S_1 as our sample S with probability p and accept S_2 with probability 1-p. That is,

$$\Pr(\, i \in S \,) = p \Pr(\, i \in S_1 \,) + (1-p) \Pr(\, i \in S_2 \,)$$

$$= \frac{\sum_{u=1}^{N} W_u^{(1)}}{\sum_{j=1}^{N} W_j} \frac{n\, W_i^{(1)}}{\sum_{j=1}^{N} W_j^{(1)}} + \frac{\sum_{u=1}^{N} W_u^{(2)}}{\sum_{j=1}^{N} W_j} \frac{n\, W_i^{(2)}}{\sum_{j=1}^{N} W_j^{(2)}}$$

$$= \frac{n\, W_i}{\sum_{j=1}^{N} W_j} .$$

Therefore, a sample with probabilities proportional to weight, is obtained.

For (ii) $J_N = n-1$ and $(n-1)\, V_N = C^*_N$. In section 3.4, we have performed one additional step such that $\frac{n W_N}{\sum_{j=1}^{N} W_j^{(1)}} = 1$, $\frac{n W_N^{(1)}}{\sum_{j=1}^{N} W_j^{(1)}} = 1$

and $\frac{n W_N^{(2)}}{\sum_{j=1}^{N} W_j^{(2)}} = 1$. By Theorem 3.1, both samples, S_1 and S_2 exist. With the same argument in (i), a sample with probabilities proportional to weight exists.

Q.E.D.

Theorem 3.7

Suppose $2 \le n < N$ and $W_i > 0$ for $i = 1, \dots, N$. The proposed algorithm can find a sample of size n with positive second order inclusion probability if $J_N = 0$ and $C_N \le 0$.

Proof of Theorem 3.7:

Since $C_N \le 0$ and $J_N = 0$, we have $A_N \le 0$ and $B_N \le 0$. By Theorem 3.1, samples drawn from $S_{1,N}$ and $S_{2,N}$ exist and denote the samples as S_1 and S_2. By Theorem 3.6, a weighted sample is obtained. Besides, at each stage k (where the size of L_k is at least n+1), $A_k < (n-2)\, M_k^{(1)}$. By Theorem 3.4, there are at most n-2 self-selective units in the sample drawn from $S_{1,k}$ at stage k, and by Theorem 2.1, the second order inclusion probability are strictly positive in the sample drawn from $S_{1,N}$. We accept the sample drawn from $S_{1,N}$ with probability p which is defined in Theorem 3.6, and from $S_{2,N}$ with probability 1-p. Thus,

$\Pr(i,j \in S)$

$= p \Pr(i,j \in S_1) + (1-p) \Pr(i,j \in S_2)$

$\geq p \Pr(i,j \in S_1)$

$> 0.$

Therefore, the proposed algorithm can find a sample of size n with positive second order inclusion probabilities. Q.E.D.

§3.6 Worked example

The data file contains ten records R_1, ... , R_{10} (N = 10) with weights 60, 20, 60, 10, 20, 30, 20, 230, 240 and 50, denoted as W_1, ... , W_{10}. We want to draw a weighted sample of size three (n = 3).

We first sort the leading five (that is 2n-1) units in ascending order. The weights are 10, 20, 20, 60 and 60 and are denoted as $W_{(1)}$, $W_{(2)}$, $W_{(3)}$, $W_{(4)}$ and $W_{(5)}$. Consider the first two units and the algorithm starts from stage two.

Stage Two:

Choose U such that $0 < U < W_{(1)}$, say U = 5. Then, set $W^{(1)}_{(1)} = W^{(1)}_{(2)} = 5$.

$S_{1,2} = \{ (R_4, 5), (R_2, 5) \}$,

$S_{2,2} = \{ (R_4, 5), (R_2, 15) \}$

and

$S_{3,2} = \emptyset$.

Also, we have $A_2 = 5$, $B_2 = 25$, $C_2 = 30$; $M^{(1)}_2 = 5$, $M^{(2)}_2 = 15$, $M_2 = 20$; $T^{(1)}_2 = 10$, $T^{(2)}_2 = 20$, and $T_2 = 30$.

Stage Three:

Handle $W_{(3)} = 20$.

Since $W_{(3)} = M_2$ and $M_2 < C_2 < M_2 + W_{(3)}$, it is the CASE 3. Choose $W^{(1)}_{(3)}$ such that

$$\max(\delta_1, \delta_2, 5) \le W^{(1)}_{(3)} \le \min(10 - \delta_3, 5)$$

where

$$0 < \delta_1 < \min(5, 5),$$

$$0 < \delta_2 < \min(5, 5)$$

and

$$0 < \delta_3 < \min(5, 5, 5).$$

Take $W_5^{(1)} = 5$ and $W_5^{(2)} = 15$. Thus, after this stage,

$$S_{1,3} = S_{1,2} \cup \{ (R_5, 5) \},$$

$$S_{2,3} = S_{2,2} \cup \{ (R_5, 15) \}$$

and

$$S_{3,3} = \emptyset.$$

Also, we have $A_3 = 0$, $B_3 = 10$, $C_3 = 10$; $M_3^{(1)} = 5$, $M_3^{(2)} = 15$, $M_3 = 20$; $T_3^{(1)} = 15$, $T_3^{(2)} = 35$, and $T_3 = 50$.

Stage Four:

Handle $W_{(4)} = 60$.

Since $W_{(4)} > M_3$ and $W_{(4)} > T_3$, it is the CASE 7. Set $S_{1,4} = S_{1,3}$, $S_{2,4} = S_{2,3}$ and $S_{3,4} = \{ R_1 \}$. We have $A_4 = 0$, $B_4 = 10$, $C_4 = 10$; $M_4^{(1)} = 5$, $M_4^{(2)} = 15$, $M_4 = 20$; $T_4^{(1)} = 15$, $T_4^{(2)} = 35$, $T_4 = 50$; $C_4^* = 130$, $V_4 = 60$ and $J_4 = 1$.

Stage Five:

Handle $W_{(5)} = 60$.

Since $W_{(5)} = V_4$ and $J_4 = 1$, it is the CASE 8. Set $S_{1,5} = S_{1,4}$, $S_{2,5} = S_{2,4}$ and $S_{3,5} = \{ R_1, R_3 \}$. We have $A_5 = 0$, $B_5 = 10$, $C_5 = 10$; $M_5^{(1)} = 5$, $M_5^{(2)} = 15$, $M_5 = 20$; $T_5^{(1)} = 15$, $T_5^{(2)} = 35$, $T_5 = 50$; $C_5^* = 130$, $V_5 = 60$ and $J_5 = 2$.

Stage Six:

Handle $W_6 = 30$.

Since $J_5 = 2$, $M_5 < W_6$ and $V_5 - T_5 < W_6 < V_5$, it is the CASE 13. Choose $W_6^{(1)}$ such that

$$\max (-45 + \delta_1, 5) \le W_6^{(1)} \le \min (55, 15 - \delta_2, 15)$$

where

$$0 < \delta_1 < \min (100, 60, 30) \text{ and } 0 < \delta_2 < \min (10, 30).$$

Take $W_6^{(1)} = 14$ and $W_6^{(2)} = 16$. Denote the updated values of A, B, C, $T^{(1)}$, $T^{(2)}$, T, $M^{(1)}$, $M^{(2)}$, M, C^*, V and J as A_6', B_6', C_6', $T_6'^{(1)}$, $T_6'^{(2)}$, T_6', $M_6'^{(1)}$, $M_6'^{(2)}$, M_6' , $C_6'^*$, V_6' and J_6'. After this intermediate step, we have $A_6' = 13$, $B_6' = -3$, $C_6' = 10$, $T_6'^{(1)} = 29$, $T_6'^{(2)} = 51$, $T_6' = 80$, $M_6'^{(1)} = 14$, $M_6'^{(2)} = 16$, $M_6' = 30$, $C_6'^* = 100$, $V_6' = 60$ and $J_6' = 1$. Consider the two units in $S_{3,6}$, that is V_6'. Let $V_6' = W_6'^{(1)} + W_6'^{(2)}$. Choose $W_6'^{(1)}$ such that

$$\max (14.5 + \delta_1, 9 + \delta_2, 14) \le W_6'^{(1)} \le \min (29 - \delta_3, 34.5, 44)$$

where

$$0 < \delta_1 < \min (7.25, 20, 29.5)$$

$$0 < \delta_2 < \min (10, 25.5, 35)$$

and

$$0 < \delta_3 < \min (7.25, 10, 15).$$

Take $W_6^{'(1)} = 28$ and $W_6^{'(2)} = 32$. Thus, after this stage,

$$S_{1,6} = S_{1,5} \cup \{ (R_6, 14) \} \cup \{ (R_1, 28), (R_3, 28) \},$$

$$S_{2,6} = S_{2,5} \cup \{ (R_6, 16) \} \cup \{ (R_1, 32), (R_3, 32) \}$$

and

$$S_{3,6} = \emptyset.$$

Also, we have $A_6 = -1$, $B_6 = -19$, $C_6 = -20$; $M_6^{(1)} = 28$, $M_6^{(2)} = 32$, $M_6 = 60$; $T_6^{(1)} = 85$, $T_6^{(2)} = 115$, $T_6 = 200$ and $J_6 = 0$.

Stage Seven:

Handle $W_7 = 20$.

Since $C_6 < W_7 < M_6$, it is the CASE 1. Choose $W_7^{(1)}$ such that

$$\max (-1 + \delta_1, \delta_2, -12) \le W_7^{(1)} \le \min (39 - \delta_3, 28, 20)$$

where

$$0 < \delta_1 < \min (20, 29, 21),$$

$$0 < \delta_2 < \min (28, 20, 19.5)$$

and

$$0 < \delta_3 < \min (20, 51, 19.5).$$

Take $W_7^{(1)} = 20$ and $W_7^{(2)} = 0$. Thus,

$$S_{1,7} = S_{1,6} \cup \{ (R_7, 20) \},$$

$$S_{2,7} = S_{2,6} \cup \{ (R_7, 0) \}$$

and

$$S_{3,7} = \emptyset.$$

Also, we have $A_7 = -21$, $B_7 = -19$, $C_7 = -40$; $M_7^{(1)} = 28$, $M_7^{(2)} = 32$, $M_7 = 60$; $T_7^{(1)} = 105$, $T_7^{(2)} = 115$, and $T_7 = 220$.

Stage Eight:

Handle $W_8 = 230$.

Since $W_8 > M_7$ and $W_8 > T_7$, it is the CASE 7. Set $S_{1,8} = S_{1,7}$, $S_{2,8} = S_{2,7}$ and $S_{3,8} = \{ R_8 \}$. We have $A_8 = -21$, $B_8 = -19$, $C_8 = -40$; $M_8^{(1)} = 28$, $M_8^{(2)} = 32$, $M_8 = 60$; $T_8^{(1)} = 105$, $T_8^{(2)} = 115$, and $T_8 = 220$. Also, $C_8^* = 470$, $V_8 = 230$ and $J_8 = 1$.

Stage Nine:

Handle $W_9 = 240$.

Since $M_8 < V_8 < W_9 < T_8 + V_8$, it is the CASE 15. Consider V_8 and let $V_8 = W_9^{*(1)} + W_9^{*(2)}$. Choose $W_9^{*(1)}$ such that

$$\max (85 + \delta_1, 28) \le W_9^{*(1)} \le \min (135, 105 - \delta_2, 198)$$

where

$$0 < \delta_1 < \min (50, 113, 10) \text{ and } 0 < \delta_2 < \min (77, 10).$$

Take $W_9^{*(1)} = 104$ and $W_9^{*(2)} = 126$. Denote the updated values of A, B, C, $T^{(1)}$, $T^{(2)}$, T, $M^{(1)}$, $M^{(2)}$, M, C^*, V and J as A_9', B_9', C_9', $T_9'^{(1)}$, $T_9'^{(2)}$, T_9', $M_9'^{(1)}$, $M_9'^{(2)}$, M_9' , $C_9'^*$, V_9' and J_9'. After this intermediate step, we have $A_9' = 103$, $B_9' = 137$, $C_9' = 240$, $T_9'^{(1)} = 209$, $T_9'^{(2)} = 241$, $T_9' = 450$, $M_9'^{(1)} = 104$, $M_9'^{(2)} = 126$, $M_9' = 230$, $C_9'^* = 270$. Then consider W_9 and choose $W_9^{(1)}$ such that

$$\max (104.5 + \delta_1, -1 + \delta_2, 104) \le W_9^{(1)} \le \min (209 - \delta_3, 119.5, 114)$$

where

$$0 < \delta_1 < \min (52.25, 15, 9.5),$$

$$0 < \delta_2 < \min (105, 120.5, 115)$$

and

$$0 < \delta_3 < \min (52.25, 105, 105).$$

Take $W_9^{(1)} = 114$ and $W_9^{(2)} = 126$. Thus, after this stage,

$$S_{1,9} = S_{1,8} \cup \{ (R_8, 104), (R_9, 114) \},$$

$$S_{2,9} = S_{2,8} \cup \{ (R_8, 126), (R_9, 126) \}$$

and

$$S_{3,9} = \emptyset.$$

Also, we have $A_9 = 19$, $B_9 = 11$, $C_9 = 30$; $M_9^{(1)} = 114$, $M_9^{(2)} = 126$, $M_9 = 240$; $T_9^{(1)} = 323$, $T_9^{(2)} = 367$, $T_9 = 690$ and $J_9 = 0$.

Stage Ten:

Handle $W_{10} = 50$.

Since $J_9 = 0$ and $C_9 < W_{10} < M_9$, it is the CASE 1. Choose $W_{10}^{(1)}$ such that

$$\max (19 + \delta_1, \delta_2, -76) \le W_{10}^{(1)} \le \min (39 - \delta_3, 114, 50)$$

where

$$0 < \delta_1 < \min (10, 95, 31),$$

$$0 < \delta_2 < \min (114, 50, 19.5)$$

and

$$0 < \delta_3 < \min (10 , 115, 19.5).$$

Take $W_{10}^{(1)} = 38$ and $W_{10}^{(2)} = 12$. Thus,

$$S_{1,10} = S_{1,9} \cup \{ (R_{10}, 38) \},$$

$$S_{2,10} = S_{2,9} \cup \{ (R_{10}, 12) \}$$

and

$$S_{3,10} = \emptyset.$$

We have $A_{10} = -19$, $B_{10} = -1$, $C_{10} = -20$; $M_{10}^{(1)} = 114$, $M_{10}^{(2)} = 126$, $M_{10} = 240$; $T_{10}^{(1)} = 361$, $T_{10}^{(2)} = 379$, and $T_{10} = 740$.

When four units have been included in the set $S_{1,6}$ and $S_{2,6}$, we begin the selection of two samples of size three from $S_{1,6}$ and $S_{2,6}$ respectively by Chao's method. Whenever there is a new unit included in $S_{1,k}$ and $S_{2,k}$ at stage k, we select samples of size three from $S_{1,k}$ and $S_{2,k}$ respectively, by Chao's method. At the end of the file, we select the sample drawn from $S_{1,10}$ as our sample with probability 361/740 and select the sample drawn from $S_{2,10}$ as our sample with probability 379/740. At the end of the file, $C_{10} < 0$ and $S_{3,10}$ is empty and hence a weighted sample exists with positive second order inclusion probabilities.

CHAPTER 4. CONCLUSION

We have presented a new one-pass algorithm that selects an unequal probability sample proportional to weight without replacement. Since the algorithm incorporates with the Chao's algorithm, it still retain the advantages of the Chao's algorithm. More importantly, the proposed algorithm guarantees the positive second order inclusion probabilities and there is an unbiased variance estimate whenever it is possible.

At each stage, one of the fifteen cases will occur and examine. Comparing with the Chao's algorithm, the proposed algorithm makes much considerations and seems much complicated. For each weight, it is either divided into two parts and included in $S_{1,k}$ and $S_{2,k}$ or its unit is added to $S_{3,k}$. In each of $S_{1,k}$ and $S_{2,k}$, a sample is selected based on the Chao's algorithm. Obviously, the storage requirement, including the working storage for the three sets, is at least two times as much as that for the Chao's algorithm. Since the Chao's algorithm is performed two times at each stage, the computer time required is nearly double the Chao's algorithm. However, the overall computer time can be reduced. At each stage, there is an interval in choosing the weight divided into $S_{1,k}$. If we choose the upper bound of the interval, the weight divided into $S_{2,k}$ may be close to zero in some cases. When the new weight in $S_{2,k}$ is zero, the sample at previous stage retains and there is no need to perform the Chao's algorithm at this stage.

Overall to say, the proposed algorithm is most suitably implemented on the computer to select a sample with probabilities proportional to weight without replacement and it can guarantee the positive second order inclusion probabilities.

REFERENCES:

Bethlehem, J.G. and Schuerhoff, M.H. (1984). Second-order inclusion probabilities in sequential sampling without replacement with unequal probabilities. *Biometrika* **71**, 642-644.

Brewer, K.R.W. and Hanif, M. (1983). *Sampling with unequal probabilities*, Springer-Verlag, New York.

Chao, M.T. (1982) A general purpose unequal probability sampling plan. *Biometrika* **69**, 653-656.

Chauduri, A. and Vos, J.W.E. (1988). *Unified Theory and Strategies of Survey Sampling*, North-Holland, Amsterdam.

Cochran, W.G. (1977). *Sampling Techniques*, 3rd edition, Wiley, New York.

Foreman, E.K. (1991). *Survey sampling principles*, Marcel Dekker, New York.

Hansen, M.H. and Hurwitz, W.N. (1943). On the theory of sampling from a finite population. *Ann. Math. Statist.* **14**, 333-362.

Horvitz, D.G. and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* **47**, 663-685.

Li, K.H. (1992) A computer implementation of the Yates-Grundy draw by draw procedure. (Submitted for publication)

Madow, W.G. (1949). On the theory of systematic sampling, II. *Ann. Math. Statist.* **20**, 333-354.

Raj. D. (1956). Some estimators in sampling with varying probabilities without replacement. *J. Amer. Statist. Assoc.* **51**, 269-284.

Richardson, S.C. (1989). One-pass selection of a sample with probability proportional to size. *Appl. Stat.* **38**, 517-520.

Schultz, D.M. (1990). Ordered random selection with probabilities proportional to size. *South African Statist. J.* **24**, 35-44.

Sen, A.R. (1953). On the estimate of the variance in sampling with varying probabilities. *J. Indian Soc. Agric. Statist.* **5**, 119-127.

Yates, F. and Grundy, P.M. (1953). Selections without replacement from within strata with probability proportional to size, *JRSS B* **15**, 253-261.