Decipher Multi-Objectives gathering with liberate Dates Problem

Tanveer A. Tarray\textsuperscript{a}, Zahoor A. Ganie\textsuperscript{b} and Mohd. Khalid\textsuperscript{c}

\textsuperscript{a}Department of Mathematical Sciences, Islamic University of Science and Technology, Awantipora, Jammu and Kashmir, India. (E-mail: tanveer.tarray@islamicuniversity.edu.in)
\textsuperscript{b}Department of Electrical Engineering, Islamic University of Science and Technology, Awantipora, Jammu and Kashmir, India. (E-mail: zahoor.ganie@islamicuniversity.edu.in)
\textsuperscript{c}Department of Statistics & Operations Research (Women’s College), Aligarh Muslim University, Aligarh, India. (E-mail: mohdkhalidkhan4870@gmail.com)

Abstract: This crux of this paper is to suggest a mathematical problem using randomized response technique. We present some special cases applying branch and bound algorithm in order to find the exact (optimal) solution for it. By some heuristics methods we find the upper bound using different algorithms.

Keywords: Stratified Sampling; Randomized Response Technique, Multi-Objective Problem, Branch and Bound method, LINGO.

1. Introduction

In order to reduce non-response and response bias, a survey technique different from open or direct surveys was needed that made people comfortable and encouraged truthful answers. Warner (1965) developed such an alternative survey technique that is called randomized response technique (RRT). This pioneering work of Warner’s (1965) led to modifications and developments in various directions. Feeling that the cooperation of the respondent might be further enhanced if one of the two questions referred to a non-sensitive, innocuous attribute, unrelated to sensitive attribute, Horvitz et al. (1967) proposed an unrelated question randomized response model (U-model). Theoretical details for this model were given by Greenberg et al. (1969). This technique has generated much interest in the statistical literature since the publication of Warner’s randomized response model. Subsequently, several other workers have proposed different strategies for instance, see Land et al. (2012), Singh and Tarray (2014a, 2014b, 2015), Tarray and Singh (2015,2016a, 2016b, 2017, 2018a,2018b) and Tarray et al. (2018). An RRT using a stratified random sampling gives the group characteristics related to each stratum estimator. In recent years, Lawler (1973) suggested a method for reducing the maximum cost ($f_{\text{max}}$). After that, scheduling difficulties sparked a lot of attention, resulting in a significant number of studies introducing excellent approaches for determining optimality see Blazewicz (2007), Chen et al. (2007), Graham (1979), Johnson (1954) etc.

2. Problem Formulation:

The issue addressed in this work is that of scheduling the set $N$
of n jobs, \( N = \sum_{i=1}^{j} N_i \) on a one-machine”. Each job \( i \in N \) has a processed time that is an integer \( z_i \), a release date \( r_i \), and due date \( d_i \). Given a schedule \( \sigma=(1,\ldots,n) \), the flowing time of the job \( i \), \( F_i \), can be defined as \( F_i = \sigma - r_i \), where \( \sigma \) is completion time for job \( i \), given by relationship:

\[
\sigma_i = \tau_i + \zeta_i, \quad \zeta = \max\{r_i, \zeta_1\} + z_i, \quad \text{for} \ i=2,\ldots,n.
\]

Job \( i \)'s tardiness is defined by \( T_i = \max\{\sigma_i - d_i, 0\} \), and earliness by \( E_i = \max\{d_i - \sigma_i, 0\} \). The late work of job \( i \) given by \( V_i = \min\{T_i, Z_i\} \). Our issue \( (Z) \) has the following mathematical form:

\[
M = \min \{ F(\sigma) = \min_{\sigma \in \delta} \left\{ \sum_{i=1}^{n} (F_\sigma(i) + T_\sigma(i) + E_\sigma(i) + V_\sigma(i)) \right\} \}
\]

Subject to:

\[
C_\sigma(1) = r_\sigma(1) + z_\sigma(1)
\]

\[
C_\sigma(i) = \max\{r_\sigma(i), C_\sigma(i-1) + z_\sigma(i)\}, \quad i = 2,\ldots,n
\]

\[
T_\sigma(i) = \max\{C_\sigma(i) - d_\sigma(i), 0\}, \quad i = 1,\ldots,n
\]

\[
E_\sigma(i) = \max\{d_\sigma(i) - C_\sigma(i), 0\}, \quad i = 1,\ldots,n
\]

\[
V_\sigma(i) = \max\{T_\sigma(i), z_\sigma(i)\}, \quad i = 2,\ldots,n
\]

The purpose is to find a processing sequence \( \sigma = (\sigma(1), \ldots, (n)) \) (The sum of overall flow times, total tardiness, total earliness, and total late work for the problem \( z_i \)) to reduce.

3. Solution Procedure:

To get the full benefit from stratification, it is assumed that the number of units in each stratum is known. In the stratified Warner’s randomized response model, an individual respondent in the sample from stratum ‘\( i \)’ is instructed to use the randomization device which consists of a sensitive question (S) card with probability \( P_i \) and its negative question (\( S \)) card

\[
\text{Institute of Science, BHU Varanasi, India}
\]

with probability \((1-P_i)\). The respondent answers the question with “Yes” or “No” without reporting which question card he or she has. A respondent belonging to the sample in different strata will perform different randomization device, each having different pre-assigned probabilities. Under the assumption that these “Yes” or “No” reports are made truthfully and \( P_i \) is set by the researcher, the probability of “Yes” answers in stratum ‘\( i \)’ for the stratified Warner’s RR model is:

\[
Z_i = P_i \pi_i + \left(1 - P_i \right) \left(1 - \pi_i \right), \quad \text{for} \ i = 1, 2, \ldots, k \), the variance / MSE of \( \hat{\pi}_S \) is given by

\[
V(\hat{\pi}_S) = \frac{1}{n} \sum_{i=1}^{k} \left( \pi_i \left(1 - \pi_i \right) + \frac{P_i \left(1 - P_i \right)}{2P_i - 1} \right)^2
\]

The sampling cost function is of the form \( \sum_{h=1}^{k} c_h n_h \), the cost is proportional to the size of the sample within any stratum.

We define \( C^0 = C - C^0 \). The linear cost function is

\[
C = C^0 + \sum_{h=1}^{k} c_h n_h, \quad \text{where} \ C^0 \ \text{is the over head cost,} \ c_h \ \text{is the per unit cost of measurement in} \ h^\text{th} \ \text{stratum,} \ C \ \text{is the available fixed budget for the survey.}
\]

The problem of optimum allocation can be formulated as a non linear programming problem (NLPP) for fixed cost as

\[
\begin{aligned}
\text{Minimize} & \quad V(\hat{\pi}_S) = \sum_{h=1}^{k} \frac{w_h}{n_h} V_h \\
\text{subject to} & \quad \sum_{h=1}^{k} c_h n_h \leq c^0 \\
& \quad 2 \leq n_h \leq N_h \quad \text{and} \ n_h \ \text{integers,} \ h = 1, 2, \ldots, k
\end{aligned}
\]

The above NLPP can be solved using non linear integer programming technique. We can now apply Branch and Bound method to solve the challenge \( z_i \). The branch and bound approach is the major way for solving the problem, whereas the bat algorithm and the Gray Wolf algorithm (GW) are used. The
upper bound will be found by applying the bat algorithm in which the parameter \( x \) and \( v \) refers to position and velocity of solution respectively, the following equation describe the updating the value of both \( x \) and \( v \),

\[
(t) = \nu(t) + \left( x'(t) - Bval \times x(t) - 1 \right) \times f(t) .
\]

\[
(t) = x'(t) + \nu(t) .
\]

and

\[
x'(t) = x(t) \times val .
\]

\[
f(t) = fmin + \left( max - fmin \right) \times B(t) .
\]

\[
(t) = R(t)(1 - e^{-n}) .
\]

\[
A(t) = A(t) \times aBval/vol
\]

Where: \( val \) is the current local solution. And \( (t) \) is a random vector, \( and \), \( \epsilon \) and \( \gamma \subseteq (0,1) \). The algorithm terminates after finish all iterations \( CNan and N \), or if the procedure exceeds fixed period of time. Now for the lower bound The problem \( (z) \) can be broken into two sub problems in order to have a less complex structure \((z_i)\) and \((z_2)\).

\[
M_1 = \text{Min} \sum_{i=1}^{n} \left( F_{\sigma(i)} + T_{\sigma(i)} + E_{\sigma(i)} \right)
\]

Subject to:

\[
C_{\sigma(1)} = r_{\sigma(1)} + z_{\sigma(1)}
\]

\[
C_{\sigma(i)} = \text{Max} \{ r_{\sigma(i)}, C_{\sigma(i-1)} + z_{\sigma(i)} \}, \quad i = 2...n
\]

\[
T_{\sigma(i)} = \text{Max} \{ C_{\sigma(i)} - d_{\sigma(i)}, 0 \}, \quad i = 1...n
\]

\[
E_{\sigma(i)} = \text{Max} \{ d_{\sigma(i)} - C_{\sigma(i)}, 0 \}, \quad i = 1...n
\]

\[
M_2 = \text{Min} \sum_{i=1}^{n} \left( V_{\sigma(i)} \right)
\]

Subject to:

\[
C_{\sigma(1)} = r_{\sigma(1)} + z_{\sigma(1)}
\]

\[
C_{\sigma(i)} = \text{Max} \{ r_{\sigma(i)}, C_{\sigma(i-1)} + z_{\sigma(i)} \}, \quad i = 2...n
\]

\[
T_{\sigma(i)} \geq 0, \quad i = 1...n
\]

\[
V_{\sigma(i)} = \text{Max} \{ T_{\sigma(i)}, z_{\sigma(i)} \}, \quad i = 2...n
\]

The lower bound for \( M_1 \) and \( M_2 \) into two sub problems can be written as,

\[
\text{Min} Z(\delta) = \sum_{i=1}^{n} \left( E_i + T_{\delta_i} + F_{\delta_i} \right)
\]

\[
= \sum_{i=1}^{n} \left( \text{Max} \{ d_{\delta_i} - r_{\delta_i}, 2c_{\delta_i} - d_{\delta_i} - r_{\delta_i}, c_{\delta_i} - r_{\delta_i} \} \right)
\]

Since the third term \( c_{\delta_i} - r_{\delta_i} \) is always between \( d_{\delta_i} - r_{\delta_i} \) and \( 2c_{\delta_i} - d_{\delta_i} - r_{\delta_i} \), then we can write the objective function \( Z(\delta) \) as:

\[
= \sum_{i=1}^{n} \left( \text{Max} \{ d_{\delta_i} - r_{\delta_i}, 2c_{\delta_i} - d_{\delta_i} - r_{\delta_i} \} \right)
\]

This means that the cost of scheduling job \( \delta_j \) is \( (\delta_j) \), given by:

\[
Z(\delta) = \begin{cases} 
\delta_{\delta_i} - r_{\delta_i} & \text{if } c_{\delta_i} - r_{\delta_i} \leq d_{\delta_i} - r_{\delta_i} \\
2c_{\delta_i} - d_{\delta_i} - r_{\delta_i} & \text{otherwise}
\end{cases}
\]

i.e., \( z(\delta_i) \) is equal to \( d_{\delta_i} \) if job is early and \( z(\delta_i) \) is equal to \( 2c_{\delta_i} - d_{\delta_i} - r_{\delta_i} \) if job \( i \) is tardy.

Also, we can write the objective of the \( M_1 \) in other form as the following:

\[
\sum_{i \in \mathcal{E}} \left( E_i + T_i + F_i \right) = \sum_{i \in \mathcal{E}} \left( E_i + T_i + C_i - r_i \right)
\]

\[
= \sum_{i \in \mathcal{E}} \left( d_i - r_i + 2 \right) \sum_{i \in \mathcal{L}} \sum_{i \in \mathcal{L}} \left( d_i - r_i \right)
\]

For problem \( P_1 \) since:
\[ \geq \min_{\delta \in S} \left\{ \max \left\{ \sum_{i=1}^{n} d_{\delta i} \right\} - r_{\delta i}, \sum_{i=1}^{n} \max \left\{ 2c_{\delta i} - d_{\delta i} - r_{\delta i}, c_{\delta i} \right\} - r_{\delta i} \right\} \]

Put \( y_{\delta i} = \max \left\{ 2c_{\delta i} - d_{\delta i} - r_{\delta i}, c_{\delta i} - r_{\delta i} \right\} \)

To show that

\[ : \quad LB \]

\[ = \min_{\delta \in S} \left\{ \max \left\{ \sum_{i=1}^{n} d_{\delta i} \right\} - r_{\delta i}, \sum_{i=1}^{n} \max \left\{ 2c_{\delta i} - d_{\delta i} - r_{\delta i}, c_{\delta i} \right\} - r_{\delta i} \right\} \]

The mathematical form of problem \( P_2 \) is as follows:

\[ \min w(\delta) = \min_{\delta \in S} \left( \sum_{i=1}^{n} v_{\delta(i)} \right) \]

Clear that:

\[ V_{\max}(\delta) \leq \sum_{i=1}^{n} v_{\delta(i)} \quad \forall \delta \in S \]

\[ : \quad V_{\max}(lw) \leq V_{\max}(\delta), \quad \forall \delta \in S \text{ when } lw \text{ is the Lower sequence for } V_{\max}. \]

4. Experimental Results:

<table>
<thead>
<tr>
<th>n</th>
<th>UB</th>
<th>time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>200</td>
<td>0.00363</td>
</tr>
<tr>
<td>2</td>
<td>212</td>
<td>0.001456</td>
</tr>
<tr>
<td>3</td>
<td>160</td>
<td>0.002207</td>
</tr>
</tbody>
</table>

Table (1) shows that the upper values of the mathematical problem when \( n=10 \), one can find different values with time if \( n=20, 30 \), etc. So, it clearly shows that the proposed mathematical model is dexterous than the existing one.

5. Conclusion:

A stratified randomized response method assists to solve the limitations of randomized response that is the loss of individual characteristics of the respondents. Formulating nonlinear programming problem (NLPP) of optimum allocation in stratified sampling with linear cost function in presence of non responses using Branch and Bound / Gray Wolf algorithm provides the optimum integer solution.

Acknowledgement

We sincerely thank the anonymous referees for their valuable comments and suggestions. The research is supported by JKST&IC under the research project number JKST&IC/SRE/397-400.

References


***