Application of R Programming to the Solution of Secretary Problem

Sadhana Rajendra Kolhekar * and Mayura Kiran Mathankar
Govt. Vidarbha Institute of Science & Humanities, Amravati, Maharashtra, India.
*Corresponding author email id: kolhekar_sr@yahoo.co.in

Date of publication (dd/mm/yyyy): 28/01/2020

Abstract **– Secretary problem is an example of optimal stopping problem. In such type of problems units are presented one by one before the observer, each time observer has to take decision whether the unit is accepted, and procedure is stopped. If the unit is rejected next unit is presented before the observer. In secretary problem decision is to stop procedure to either maximize profit or minimize cost and time required to continue it. Secretary problem is solved by many researchers with different methods. Kane S.P. (1988) solved secretary problem by defining two random variables X and Y and obtained probability distribution P(x, y/r, N). In the present paper R-program is developed corresponding to the solution of secretary problem and some interesting results are obtained.**

Keywords **– Optimal Stopping, Secretary Problem, Real Rank, Optimality Criteria, Best Unit.**

I. Introduction

Secretary problem is an optimal stopping problem. Basic secretary problem deals with stopping criteria when known number of units are to be presented one by one before the observer. The unit presented must be either selected or rejected. If the unit is accepted procedure stops by choosing that unit. If the unit is rejected next unit is called for observation. Here aim is to select the best unit among possible one. Secretary problem is also called as dowry problem, marriage problem or beauty contest problem.

1.1 *General Procedure*

Known, say N number of units are presented one by one. Inspect first ‘r’ units without selecting any. Select the unit there after which is better than the best among ‘r’ presented units. ‘r’ can be called as ‘experience period’. If none of (N-1) units are selected, then select the last i.e. N^{th} unit. Here ‘r’ is the parameter.

1.2 *Some Basic Definitions*

Unit: Candidate for interview or an item for observation which is presented before interviewer or observer is called as unit.

Real Rank: Rank associated with the unit after observing all N units.

Best Unit: A unit corresponding to maximum real rank is called as Best Unit.

1.3 *Review of Literature*

Secretary problem was solved by many researchers with different methods and concepts. Lindley (1961) (Citation: 241) produced solution of the problem using dynamic programming. Gilbert and Mosteller (1966) (Citation: 718) have derived independently solution of the problem which matches with the solution given by Lindley (1961). According to Dimitris et al (1999) (Citation: 6) the importance of this optimal stopping is that it provides an artificial idealized simulation of sequential decision process.

Detail review of the problem was given by Freeman (1983) (Citation: 447). Kane S.P. (1988) gave solution of the problem by defining two random variables X and Y and developed probability distribution P(x, y/ r, N). Ferguson (1989) (Citation: 709) discussed importance of Secretary Problem as a vast field in the area of mathematics, probability and optimization. Kane S.P. (2001) modified Secretary Problem by introducing concept of measurable character and found out solution of Secretary Problem with measurable character.

Gilbert & Mosteller (2006) (Citation: 720) discussed beauty contest problem, dowry problem as an example of Secretary Problem and gave solution of it. More recent papers on Secretary Problem include Moshe Babaioff et al (2008) (Citation: 171) in which generalization of Secretary Problem is discussed and applied to online auction problem. Metroid Secretary Problem is discussed by Michael Dinitz (2013) (Citation: 26). Here, Metroid Secretary Problem is viewed at a Weighted Secretary Problem.

II. Solution of the Secretary Problem by Kane S.P. (1988)

Kane S.P. (1988) defined two random variables for general procedure explained in section 1.1. X is defined real rank of selected units and Y is position at which selection is made. The values taken by X and Y are x and y respectively. P(x, y/ r, N) is the expression for probability that procedure of search stops at position y and x is rank of selected unit where N number of units are presented, and selection is not made up to ‘r’ units. Here it is assumed that N is known, and units are randomly presented before observer hence N! permutations are equally likely.

It is noted that random variable X and Y are discrete random variables taking value X = 1, 2…, N and Y = r+1, r+2…, N. Here joint probability distribution of X and Y i.e. P [X = x, Y = y] is given by,

$$P(x,y/r,N)=\begin{cases}\dfrac{r(N-y)!\,(x-1)!}{(y-1)N!\,(x-y)!}, & r+1\le y\le N-1\\ & y\le x\le N\\ \dfrac{r}{N(N-1)}, & y=N\\ & 1\le x\le N\\ 0, & otherwise\end{cases}$$

III. Application of R-programming to the Solution of Secretary Problem

Probability distribution mentioned in section 2 can be further analyzed using R-program. In this section we write R-program for above probability distribution P(x, y/ r, N). Here ‘r’ and N are parameters. Known number of units N are presented before the observer. ‘r’ is pre-decided number of units where selection is not made. We may call (1, r) interval as non-selection zone in this procedure. Following is the R-programme.

```
# Define variables N, x, y, r

N <- 10

x <- 1

y <- 1

r <- 3

P = P1 = P2 = P3 = 0

while (x < (N + 1)) {
```

```
 while (y < (N + 1)) {
  if (((r + 1) <= y & y <= (N - 1)) & (y <= x & x <= N))
  {
   P1 = (r * factorial (N - y) * factorial (x - 1)) / ((y - 1) * factorial (N) * factorial (x - y))
   print (paste (x, y, P1))
   P = P + P1
  } else
   if ((y == N) & (1 <= x & x <= N))
   {
    P2 = r / (N * (N - 1))
    print (paste (x, y, P2))
    P = P + P2
   } else
    print (paste (x, y, P3))
  P = P + P3
  y = y + 1
 }
 y = 1
 x = x + 1
}
 print (paste ("Total Probability = ", P))
```

Above program was particularly written for N = 10 and r = 3. We can run it for various values of N and r which gives us a common feature mentioned in the following table.

Table 3.1.

X \ Y	1	2	.	.	.	r	r+1	r+2	.	.	.	N-1	N
1	0	0	0	0	0	0	0	0	0	0	0	0	r/N(N-1)
2	0	0	0	0	0	0	0	0	0	0	0	0	r/N(N-1)
.	.	.	.	.	.	.	.	.	.	.	.	.	r/N(N-1)
.	.	.	.	.	.	.	.	.	.	.	.	.	r/N(N-1)
.	.	.	.	.	.	.	.	.	.	.	.	.	r/N(N-1)
r	0	0	0	0	0	0	0	0	0	0	0	0	r/N(N-1)

International Journal of Innovation in Science and Mathematics
Volume 8, Issue 1, ISSN (Online): 2347–9051

r+1	0	0	0	0	0	0	*	0	0	0	0	0	r/N(N-1)
r+2	0	0	0	0	0	0	*	*	0	0	0	0	r/N(N-1)
.	.	.	.	.	.	.	.	.	.	.	.	.	r/N(N-1)
.	.	.	.	.	.	.	.	.	.	.	.	.	r/N(N-1)
.	.	.	.	.	.	.	.	.	.	.	.	.	r/N(N-1)
N-1	0	0	0	0	0	0	*	*	.	.	.	*	r/N(N-1)
N	0	0	0	0	0	0	*	*	.	.	.	*	r/N(N-1)

Where '*' is equal to [r(N-y)! (x – 1)!] / [(y-1)N! (x – y)!].

For example, when number of candidates presented before the observer is 10 i.e. N = 10 and up to 'r' candidates' selection is not made, i.e. r = 3, and if R program mentioned above is run, the result obtained is given in Table 2. It gives probability of selection of unit with the real rank 1, 2 …N which are selected at the position Y = (r+1), (r+2) …N.

Table 3.2. Illustration

	X	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Y											
1		0	0	0	0	0	0	0	0	0	0.033
2		0	0	0	0	0	0	0	0	0	0.033
3		0	0	0	0	0	0	0	0	0	0.033
4		0	0	0	0.00119	0	0	0	0	0	0.033
5		0	0	0	0.004761	0.000595	0	0	0	0	0.033
6		0	0	0	0.0119	0.002976	0.000476	0	0	0	0.033
7		0	0	0	0.0238	0.008928	0.002857	0.000595	0	0	0.033
8		0	0	0	0.04166	0.02083	0.01	0.004166	0.00119	0	0.033
9		0	0	0	0.06666	0.04166	0.02666	0.01666	0.009523	0.004166	0.033
10		0	0	0	0.1	0.075	0.06	0.05	0.04285	0.0375	0.033
Total		0	0	0	0.249971	0.149989	0.099993	0.071421	0.053563	0.041666	0.33
Grand Total											0.996604

IV. Conclusion

1. It can be verified by running R-programme mentioned in section-3 that, $\sum_{x=1}^{N}\sum_{y=r+1}^{N} P(x,y/r,N) = 1$.
2. According to the procedure of Secretary problem, we can not stop observing till 'r' units. Hence, probability associated whenever (y < r) is always zero as seen from Table 3.1.
3. It can be observed that, selection zone starts from the position Y = (r + 1), (r + 2) … N.

4. The unit with rank X = (r + 1) will be selected at the most up to the position Y = (r + 1) in selection zone or at the last position Y = N with probability (r/N* (N - 1)). Similarly, unit with real rank X = (r + 2) will be selected at the most up to the position Y = (r + 2) in selection zone or at the last position Y = N and so on.

5. It is interesting to note that, if the best unit with rank X = N appears among first 'r' units then the observer is forced to stop at (Y = N) with probability [r/N (N - 1)].

6. Brevity of the R-programming language inspires to write program in R which gives brief programmes for probability distribution.

ACKNOWLEDGEMENT

Authors express their deep sense of gratitude and indebtedness to referees for their helpful comments which have resulted in the improvement of this paper.

REFERENCES

[1] Lindley, D.V. *"Dynamic programming and decision theory"*. Applied Statistics. Vol. 10, 1961, p. 39-52. Citation: 241.
[2] Gilbert, J and Mosteller, F. *"Recognizing the maximum of a sequence"*. J. Am. Statist. Assoc. Vol. 61, 1966, p. 35-73. Citation: 718.
[3] Dimitris, A Sardelis & Valahas, T.M. *"Decision Making: A Golden Rule"*. American Mathematical Monthly. 1999, P.215-226. Citation: 6.
[4] Freeman, P.R. *"The secretary problem and its extensions: A review"*. International Statistical Review. Vol. 51, 1983, p.189-206. Citation: 447.
[5] Kane, S.P. Ph.D. Thesis - *"A Study of Some Aspects of Secretary Problem and Stopping Rules"*. Submitted to Nagpur University, Nagpur. 1988.
[6] Ferguson, T.S. *"Who Solved Secretary Problem?"*. Statistical Science. Vol. 4, 1989, p.282-289. Citation: 709.
[7] Kane S.P. *"A Secretary Problem with Measurable Character"*. Journal of Indian Society for Probability & Statistics. Vol. 5, 2001, p.13-30.
[8] Gilbert J.P. & Frederick Mosteller *"Recognizing Maximum of Sequence"*. Part of Springer Series in Statistics, 2006, p. 355-398. Citation: 720.
[9] Moshe Babaioff, Nicol Immorlica, David Kempe, Robert Kleinberg *"Online auction & generalized secretary Problem"*. SIGecom Exch. June 2008, 7(2): 7: 1-7: 11. Citation: 171.
[10] Michael Dinitz, *"Recent Advances on the Metroid Secretary Problem"*. Vol. 44 No. 2, June 2013, p.126-142. Citation: 26.

AUTHOR'S PROFILE

First Author
Dr. Sadhana Rajendra Kolhekar, M.SC. (1985), M. PHIL. (1991), PH. D. (2005). Designation: Head of the Department of Statistics & Associate Professor, Department of Statistics, Govt. Vidarbha Institute of Science & Humanities, Amravati, Maharashtra, India. Posted as Lecturer through selection by Maharashtra Public Service Commission. 35 years of experience of teaching Statistics to Under-graduate and 32 years of teaching Statistics to Post-graduate and 4 years teaching Statistics to M.Phil. students. Guided 2 M.Phil. students. Member of board of studies at Sant Gadge Baba Amravati University, Amravati. Completed a Minor Research Project of UGC in the year of 2012.

Second Author
Dr. Mayura Kiran Mathankar, M.SC. (1985), M. PHIL. (1991), PH. D. (2007). Designation: Associate Professor, Department of Statistics, Govt. Vidarbha Institute of Science & Humanities, Amravati, Maharashtra, India. 34 years of experience of teaching Statistics to Under-graduate and 29 years of teaching Statistics to Post-graduate level. Member of board of studies at Sant Gadge Baba Amravati University, Amravati. Completed a Minor Research Project of University of Mumbai, India. email id: mayura.mathankar@gmail.com