Chapter – 3

ANALYSIS OF STOPPING STRATEGY
CHAPTER - 3

ANALYSIS OF STOPPING STRATEGY

3.1: Basic optimal stopping problem:

An interviewer inspects a known number of candidates which are presented before him one by one in random order for inspection. At any stage the interviewer is able to rank the candidates presented so far, in order of desirability. The total number of candidates is N and all N! permutations of N are equally likely. The candidate which is presented must be either accepted or rejected. If the candidate is accepted the process terminates & it results in the selection of that candidate. If the candidate is rejected then the next candidate is called for inspection. If the first (N-1) candidates are rejected then the last candidate is presented & it must be accepted. The interviewer’s objective is to have an optimal stopping rule and to select a candidate at a suitable stage. The general procedure that is usually followed is as follows:

The first r candidates are inspected without selecting any. Note the “best” among these first r candidates without selecting any. Inspect the candidates thereafter. Select the first candidate which is better than the “best” of first r candidates. If no candidate from the first (N-1) candidates is selected then select the last candidate (that is the Nth candidate).

As the choice of r characterizes the procedure, the main aim is to
choose an integer r. The guidelines for choosing r will be the optimality criterion. The above mentioned selection procedure is referred as “Original Secretary Problem”.

3.1.1: Basic definitions:

1) Units: Candidates or individuals or items under observation are known as Units. [Here onwards we call the candidates by ‘units’ only]

2) Real Ranks: Real ranks are the ranks associated with the units after observing all the N units.

3) Relative Ranks: Relative ranks of the i units are the ranks associated with the units after observing first i units (i = 1, 2, ..., N-2, N-1).

4) Best Unit: A unit corresponding to the maximum real rank is called as the best unit.

5) Optimal Policy: It is the optimum choice of r according to some desired optimality criterion. Deciding optimum r is the solution of secretary problem.

3.2: Various methods of solution of the secretary problems:

1] Lindley (1961) gave the solution of the problem from the point of view of utility function and using the technique of Dynamic programming.

2] The same solution has also been provided by the method of backward induction by Chow, Robbins, Siegmund (1971).
3] Govindarajulu (1975) slightly changed the mathematical formulation of the problem and got some sequence \( \{ X(i) \} \) which forms a Markov chain, and then arrived at the same solution.

4] Gilbert and Mosteller (1966) have also derived independently the same solution of the problem.

We present below the solution of the secretary problem along with some of the results stated by Gilbert and Mosteller (1966). We call the solution as Lindley’s solution because it was first neatly proposed by Lindley (1961).

3.2.1: Lindley’s solution:

The optimal policy is to inspect first \( r_0 \) units without selecting any of them \( (r_0 < N) \) and then to accept the first unit thereafter that is better than all the previous units, where \( r_0 \) is an integer satisfying the following inequality

\[
\frac{1}{r_0 + 1} + \frac{1}{r_0 + 2} + \ldots + \frac{1}{N-1} < 1 \leq \frac{1}{r_0} + \frac{1}{r_0 + 1} + \ldots + \frac{1}{N-1}
\]

If \( N \) is large, the value of \( r_0 \) approximately satisfies

\[
\int_{r_0}^{N} \frac{d \xi}{\xi} = 1
\]

Therefore,
\[
\log_e \frac{N}{r_0} = 1
\]

Hence,

\[
r_0 = Ne^{-1}
\]

Thus, for large \(N\) the rule is to inspect until \(Ne^{-1} \approx (0.368)N\) units without selecting any, and then to select a subsequent unit which is the best amongst those observed so far. The same result has been derived by Gilbert and Mosteller (1966).

Here optimality criterion is maximization of probability of selection of the best unit.

Further, Gilbert and Mosteller have also derived an explicit expression for \(P\), the probability of selecting the best unit when we pass the first \((r-1)\) units without selecting any and choose the first later unit, which is given by

\[
P = \frac{r-1}{N} \sum_{k=r}^{N} (k-1)^{-1}
\]

The same result has also been obtained by Chow et al. (1971).

The following inequalities for \(r_0\) has also been given by Gilbert and Mosteller (1966).

\[
\frac{N-(1/2)}{e} + \frac{1}{2} - \frac{3e-1}{2(2N+3e-1)} \leq r_0 + 1 \leq \frac{N-(1/2)}{e} + \frac{3}{2}
\]

It was further shown that the probability \(P\) asymptotically equals \(1/e\).
3.3: Analysis of secretary problem:

We note that the procedure described in section (3.1) leads to two random variables \( X \) and \( Y \), where \( Y \) is the number of unit at which the selection is made and \( X \) is real rank of the selected unit. Values taken by random variables \( X \) and \( Y \) are denoted by \( x \) and \( y \) respectively. We discuss explicitly the expression of the probability that we stop after examining \( y \) units and the selected unit has real rank \( x \), when there are in all \( N \) units available and selection is not done from first \( r \) units. This probability is denoted by \( P(x, y / r, N) \). This explicit expression of \( P(x, y / r, N) \) is of vital importance because from that we can derive the exact marginal distributions of \( X \) and \( Y \). Distribution of \( Y \) throws light on the expected cost.

We note that \( X \) and \( Y \) are two discrete random variates and

\[
X = 1, 2, 3, \ldots, N-1, N.
\]

\[
Y = r+1, r+2, \ldots, N-1, N.
\]

3.3.1: Joint distribution of \( X \) and \( Y \):

Theorem 3.3.1.1: The probability of \( (X = x, Y = y) \) is given by

\[
P(x, y / r, N) = \frac{r(N-y)!(x-1)!}{(y-1)N!(x-y)!}, \quad r+1 \leq y \leq N-1, \quad y \leq x \leq N
\]

\[
= \frac{r}{N(N-1)}, \quad y = N
\]

\[
= 0, \quad \text{otherwise}
\]

Proof: We divide the proof in two parts.
1) When \( Y \leq N - 1 \)

Let \( j \) be the maximum real rank of the first \( r \) units. Then real rank of the units in \((y-1)\) positions must be less than \( j \). The unit with real rank \( j \) can be at any of the \( r \) positions. Further, \( j \) can be at most \((y-1)\) and it cannot exceed \((x-1)\). Therefore, the number of permutations qualifying such condition is equal to

\[
r[(j-1)^{(y-2)}](N-y)!
\]

Where, \( a^{(b)} = a(a-1)(a-2)...(a-b+1) \), \( b \leq a \)

Therefore,

\[
P(x, y / r, N) = \frac{r(N-y)!}{N!} \sum_{j=1}^{x-1} (j-1)^{(x-2)}
\]

Using the result

\[
\sum_{n=a}^{b} a^{(x)} = \frac{1}{x+1} [(b+1)^{(x+1)} - a^{(x+1)}]
\]

\[
\sum_{j=y-1}^{x-1} (j-1)^{(y-2)} = \frac{1}{y-1} [(x-1)^{(y-1)} - (y-2)^{(y-1)}]
\]

Hence,
\[ P(x, y/r, N) = \frac{r(N-y)!}{(y-1)N!} [(x-1)^{(y-1)} - (y-2)^{(y-1)}] \]

Since,

\[ (y-2)^{(y-1)} = 0, \]

\[ P(x, y/r, N) = \frac{r(N-y)!}{(y-1)N!} (x-1)^{(y-1)} \]

However,

\[ (x-1)^{(y-1)} = (x-1)(x-2)\ldots(x-y+1) \]

\[ = \frac{(x-1)!}{(x-y)!} \]

Therefore,

\[ P(x, y/r, N) = \frac{r(N-y)!}{(y-1)N!(x-y)!} \frac{(x-1)!}{(x-y)!} \quad y \leq x \leq N \]

\[ \quad r+1 \leq y \leq N-1 \]

2) When \( Y = N \)
i) If \( Y = N \) and \( X = 1, 2, \ldots, N-1 \), then it is obvious that the best unit has already been appeared in the first \( r \) units. The last inspected unit can have rank \( x \). Hence excluding the best unit which is in first \( r \) units and the last inspected unit, the remaining \( (N-2) \) units can appear in \( (N-2)! \) ways. The best unit can be at any one of the first \( r \) positions. Therefore the number of permutations qualifying this condition is \( r(N-2)! \).

Hence, the probability of this event is

\[
P(x, N/r, N) = \frac{r(N-2)!}{N!}
\]

Thus,

\[
P(x, N/r, N) = \frac{r}{N(N-1)}, \quad 1 \leq x \leq N-1
\]

ii) If \( Y = N \) and \( X = N \), then it is obvious that the second best unit has already been appeared in the first \( r \) units. Excluding the last unit and the second best unit, the remaining \( (N-2) \) units can appear in \( (N-2)! \) ways. Moreover the second best unit can beat any one of the \( r \) positions. Hence, the probability of this event becomes equal to

\[
\frac{r(N-2)!}{N!}
\]

Therefore,
\[ P(N, N/r, N) = \frac{r}{N(N-1)} \]

It can be easily seen that for other pairs of \( x \) and \( y \),

\[ P(x, y/r, N) \]’s are zero.

The table of probabilities \( P(x, y/r, N) \) is given below. The non zero values of \( P(x, y/r, N) \)’s are denoted by * in the table.
Table 3.1

Values of probabilities $P(x,y/r,N)$

<table>
<thead>
<tr>
<th>X</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>...</th>
<th>r</th>
<th>r+1</th>
<th>r+2</th>
<th>...</th>
<th>j</th>
<th>...</th>
<th>N-1</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>r+1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td>r+2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>*</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>...</td>
<td></td>
<td></td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N-1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
<td>*</td>
<td>*</td>
<td>...</td>
<td>*</td>
<td>...</td>
<td>*</td>
<td>r/N (N-1)</td>
</tr>
<tr>
<td>N</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>...</td>
<td>*</td>
<td>...</td>
<td>*</td>
<td>r/N (N-1)</td>
</tr>
</tbody>
</table>
Thus we have,

\[ P(x, y/r, N) = \frac{r(N-y)!(x-1)!}{(y-1)!N!(x-y)!}, \quad y \leq x \leq N \]

\[ \frac{r}{N(N-1)}, \quad y = N \]

\[ = 0, \quad \text{otherwise.} \]

It can be easily verified that

\[ \sum_{x=1}^{N} \sum_{y=r+1}^{N} P(x, y/r, N) = 1 \]

From above analysis i.e. from above joint probability function, we can obtain marginal function of Y which gives us the optimal stopping probability distribution.

3.4: ‘Optimal Stopping’ analysis of secretary problem:
(Marginal Probabilities of Y)

3.4.1: The distribution of Y:

The marginal distribution of random variable Y is given by

\[ P_y(y/r, N) = 0, \quad 1 \leq y \leq r \]

\[ = \frac{r}{y(y-1)}, \quad y = r+1, r+2, \ldots, N-1 \]

\[ = \frac{r}{N-1}, \quad y = N \]
Proof: When \( y \neq N \)

The marginal probabilities of \( Y \) denoted by \( P_y(y/r, N) \) are given by

\[
P_y(y/r, N) = \sum_{x=y}^{N} \frac{r(N-y)!}{(y-1)N!(x-y)!} \frac{x!}{(x-1)!}
\]

\[
= \sum_{x=y}^{N} \frac{r(N-y)!}{(y-1)N!(y-1)} \binom{x-1}{y-1}
\]

\[
= \frac{r(N-y)!(y-2)!}{N!} \sum_{x=y}^{N} \binom{x-1}{y-1}
\]

Using the result

\[
\sum_{n=a+1}^{b} \binom{n}{x} = \binom{b+1}{x+1} - \binom{a}{x+1}
\]

We note that

\[
\sum_{x=y}^{N} \binom{x-1}{y-1} = \binom{N}{y}
\]

Therefore,

\[
P_y(y/r, N) = \frac{r(N-y)!(y-2)!N!}{N!y(N-y)!}
\]

\[
= \frac{r}{y(y-1)}
\]
(ii) When \( y = N \)

\[
P_y(N/r, N) = \frac{r}{N-1}
\]

Thus the probability that we stop after inspection of \( y \) units is

\[
P_y(y/r, N) = \frac{r}{y(y-1)} \quad , \quad y = r+1, r+2, \ldots, N-1
\]

\[
= \frac{r}{N-1} \quad , \quad y = N
\]

\[\ldots (3.4.1.1)\]

The table of probabilities \( P_y(y/r, N) \) is given below.

For calculation of the probabilities in the table 3.2 below, we have prepared a computer program in Visual Basic which is provided in Appendix I.
Table 3.2

Values of $P_y(y/r, N)$

<table>
<thead>
<tr>
<th>N</th>
<th>$r_0$</th>
<th>$y$</th>
<th>$P_y(y/r, N)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>7</td>
<td>8</td>
<td>0.1250</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>9</td>
<td>0.0972</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>10</td>
<td>0.0778</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>11</td>
<td>0.0636</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>12</td>
<td>0.0530</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>13</td>
<td>0.0449</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>14</td>
<td>0.0385</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>15</td>
<td>0.0333</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>16</td>
<td>0.0292</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>17</td>
<td>0.0257</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>18</td>
<td>0.0229</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>19</td>
<td>0.0205</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>20</td>
<td>0.3684</td>
</tr>
</tbody>
</table>
Since
\[ \frac{r}{y(y-1)} = \frac{r}{y-1} - \frac{r}{y} , \]

We easily note that,
\[ \sum_{y} P_{y}(y/r, N) = 1 \]

3.4.2: Expected value of $y$:

Now we find the expression for the expected value of $y$ with the help of (3.4.1.1).

Denoting the expected value of $y$ by $E(y/r, N)$

\[
E(y/r, N) = \sum_{y=r+1}^{N} y P_{y}(y/r, N) \\
= (r+1) \frac{r}{r(r+1)} + (r+2) \frac{r}{(r+1)(r+2)} + \ldots + (N-1) \frac{r}{(N-2)(N-1)} + \frac{Nr}{(N-1)} \\
= r \left[ \frac{1}{r} + \frac{1}{r+1} + \ldots + \frac{1}{N-2} + \frac{N}{N-1} \right] \\
\]

Since,
\[ \frac{N}{N-1} = \frac{1}{N-1} + 1 , \]
\[ E(y/r, N) = r + 1 + r \left[ \frac{1}{r+1} + \frac{1}{r+2} + \ldots + \frac{1}{N-1} \right] \]

3.4.3: Properties of expected value of Y:

1) \( E(y/r, N) \) is an increasing function of \( r \).

Proof: We have

\[
E(y/r+1, N) - E(y/r, N) \\
= (r+1) + (r+1) \left[ \frac{1}{r+2} + \frac{1}{r+3} + \ldots + \frac{1}{N-1} \right] - \left[ r + 1 + r \left( \frac{1}{r+1} + \frac{1}{r+2} + \ldots + \frac{1}{N-1} \right) \right] \\
= r + 2 + r \left[ \frac{1}{r+2} + \frac{1}{r+3} + \ldots + \frac{1}{N-1} \right] + \left[ \frac{1}{r+2} + \frac{1}{r+3} + \ldots + \frac{1}{N-1} \right] - r - 1 \\
- r \left[ \frac{1}{r+2} + \frac{1}{r+3} + \ldots + \frac{1}{N-1} \right] - \frac{r}{r+1} \\
= 1 + \left[ \frac{1}{r+2} + \frac{1}{r+3} + \ldots + \frac{1}{N-1} \right] - \frac{r}{r+1} > 0
\]

Hence the result.

2) \( E(y/r, N) \) is an increasing function of \( N \).

Proof: For \( E(y/r, N) \) to be an increasing function of \( N \), we must have
$E(y/r, N+1)$ to be greater than $E(y/r, N)$.

We note that,

$$E(y/r, N+1) = E(y/r, N) + \frac{r}{N}$$

Therefore,

$$E(y/r, N+1) - E(y/r, N) = \frac{r}{N} > 0$$

Hence the result.

3.4.4: Remark:

We have,

$$E(y/r, N) = r+1 + \left[ \frac{1}{r+1} + \frac{1}{r+2} + \ldots + \frac{1}{N-1} \right]$$

If $N$ is large then,

$$\left[ \frac{1}{r+1} + \frac{1}{r+2} + \ldots + \frac{1}{N-1} \right]$$

is equal to

$$\int^N_x \frac{dx}{x} = \log_e \left( \frac{N}{r} \right)$$

Therefore for large $N$

$$E(y/r, N) \approx r+1 + r \log_e \left( \frac{N}{r} \right)$$
Further, when
\[ r = r_0 = Ne^{-1} \] then
\[ E(y/r_0, N) \approx r_0 + 1 + r_0 \]
\[ \approx 1 + 2r_0 \]
\[ \approx 1 + 2Ne^{-1} \]

Therefore,
\[
\lim_{N \to \infty} \left( \frac{E(y/r_0, N)}{N} \right) \approx 2e^{-1} = 0.736 \approx \frac{3}{4}
\]

This shows that when \( N \) is large and \( r \) is chosen equal to \( r_0 \) then nearly \((3/4)\) of the units will be observed. This fact is clearly visible in the following table.

For calculation of the expected values \( c_n^A \) for given \( N \) and \( r_0 \) in the table 3.3 below, we have prepared a computer program in Visual Basic which is provided in Appendix VII.
Table 3.3

Values of $[E(y/r_0, N)]$ for given $N$ and $r_0$

<table>
<thead>
<tr>
<th>$N$</th>
<th>$r_0$</th>
<th>$[E(y/r_0, N)]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4</td>
<td>7.9825</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
<td>14.68418</td>
</tr>
<tr>
<td>30</td>
<td>11</td>
<td>22.35954</td>
</tr>
<tr>
<td>40</td>
<td>15</td>
<td>30.029708</td>
</tr>
<tr>
<td>50</td>
<td>18</td>
<td>36.71375</td>
</tr>
<tr>
<td>60</td>
<td>22</td>
<td>44.392596</td>
</tr>
<tr>
<td>70</td>
<td>26</td>
<td>52.067414</td>
</tr>
<tr>
<td>80</td>
<td>29</td>
<td>58.748448</td>
</tr>
<tr>
<td>90</td>
<td>33</td>
<td>66.427803</td>
</tr>
<tr>
<td>100</td>
<td>37</td>
<td>74.10428</td>
</tr>
</tbody>
</table>
3.5: A different method of analysis of the optimal stopping problem:

In the literature of the study of secretary problem up till now, the entire emphasis is given to mainly the 'stopping position' in the selection procedure. In other words one can find the probability of stopping at the $Y^\text{th}$ position. It should be noted that no one has thought of analysis of the problem which estimates the time required to stop after starting the inspection procedure. That is the time required to terminate the process. We have tried to analyze the secretary problem from this point of view in this section.

Hence, now we may say that the secretary problem involves three random variables instead of two well defined random variables $X$ & $Y$. The third random variable is time required to terminate the process which is denoted by $Z$ (say).

So we start developing the analysis from 'stopping time' point of view instead of simply studying 'stopping position' random variable i.e. $Y$.

Further it should be noted that the random variables $X$ and $Y$ are jointly distributed as discussed in section 3.1, but $Z$ has its own new distribution which is thought of independent of $X$ & $Y$. Hence the distribution of $Z$ is separately derived.

In the secretary problem selection procedure, the interviewer can stop at any time after inspecting first $r$ units. It is further to be noted that the time taken to inspect a unit may vary from unit to unit (it may not be same for each unit).
We assume that the time of stopping the procedure starts after inspecting first ‘r’ units. So we start our time scale after r units are observed.

Hence after inspecting r units our time starts (i.e. at r\textsuperscript{th} position, z = 0) Although the time is continuous random variable, for the analytical simplicity we make it discrete with equal small intervals of time segments. If the process terminates in any time interval we assume that it is terminated at the end of that time interval.

Let the time required to stop the selection process be called as 'Optimal Stopping Time' (OST). Thus OST is a random variable Z taking values from zero onwards.

We can find the OST distribution as a probability distribution.

Let p be the probability of not terminating the process in a unit time. Hence (1 - p) = q is the probability of stopping the process in a unit time.

Therefore the probability that an observer utilizes ‘z’ units of time before stopping the process i.e.

\[ P(Z = z) = pq^z, \quad z = 0, 1, \ldots \]

\[ 0 \leq p \leq 1 \]

\[ q = 1 - p \]

\[ = 0, \quad \text{otherwise} \]
In case of secretary problem we have finite number of units presented before an observer; whereas in geometric distribution, \( z \) takes values from 0 to \( \infty \). So in the secretary problem framework we have to consider a truncated geometric distribution, truncated to the right at \( z = A \) (say). Moreover \( p \) varies from unit to unit as well as from time to time in secretary problem since all units are different from each other with respect to their ranks which is the only criterion of selection (i.e. stopping). Hence, it is necessary to consider \( p \) as a random variable instead of a parameter in this case. The suitable distribution of \( p \) can be assumed as a beta distribution. Hence we arrive at a situation where we must consider a compound right truncated geometric beta distribution. The geometric distribution is truncated to the right at \( z = A \).

### 3.5.1: About Compound Geometric Beta Distribution:

Let \( Z \) follows Geometric distribution with parameter \( p \).

\[
\text{i.e.} \quad p[Z = z] = p^{z-1}(1-p), \quad z = 1, 2, \ldots
\]

Here \( p \) is the probability of not stopping in a unit time & hence \( p \) is random variable which ranges from 0 to 1. A beta distribution can be assumed as a probability distribution of \( p \) with parameters \( a \) & \( b \), having p.d.f. as:

\[
f(p) = \frac{1}{B(a,b)} p^{a-1}(1-p)^{b-1}, \quad a, b > 0 \quad \text{and} \quad 0 < p < 1
\]
Therefore Compound Geometric Beta distribution comes out as:

\[ p(Z = z) = \int_0^1 f(p) p^{z-1} (1-p) \, d_z \]

\[ = \left[ \frac{1}{B(a,b)} \right] \int_0^1 p^{a-1} (1-p)^{b-1} p^{z-1} (1-p) \, d_z \]

\[ = \left[ \frac{1}{B(a,b)} \right] \int_0^1 p^{a+z-2} (1-p)^b \, d_z \]

\[ = \left[ \frac{1}{B(a,b)} \right] B(a + z - 1, b + 1) \]

Thus the probability distribution of \( z \) is given by

\[ p(Z = z) = \frac{B(a + z - 1, b + 1)}{B(a, b)} \quad , \quad z = 1, 2, \ldots \]

\[ a, b > 0 \]

Now we know that

\[ B(a,b) = \int_0^1 x^{a-1} (1-x)^{b-1} \, dx \]
\[
\begin{align*}
\Gamma(a)\Gamma(b) &= \Gamma(a + b) \\
&= \frac{(a - 1)! (b - 1)!}{(a + b - 1)!}
\end{align*}
\]

Hence we can write

\[
p(Z = z) = f_z
\]

\[
= \frac{B(a + z - 1, b + 1)}{B(a, b)} , \quad z = 1, 2, \ldots
\]

\[a, b > 0\]

\[
= \frac{(a + z - 2)! (b)! (a + b - 1)!}{(a + b + z - 1)! (a - 1)! (b - 1)!}
\]

\[
= \frac{b [(a + z - 2)(a + z - 3) \ldots 1][a + b - 1](a + b - 2) \ldots 1]}{[(a + b + z - 1)(a + b + z - 2) \ldots (a + b + z - (z + 1))(a + b + z - (z + 1)) \ldots (a - 1)(a - 2) \ldots ]}
\]

\[
= \frac{b [(a + z - 2)(a + z - 3) \ldots 1]}{[(a + b + z - 1)(a + b + z - 2) \ldots (a + b) (a - 1) \ldots 1]}
\]
\[
\frac{b[(a+z-2)(a+z-3)(a+z-4)...(a+z-(z+1))]}{[(a+b+z-1)(a+b+z-2)(a+b)(a-1)...1]}
\]

\[
= \frac{b[(a+z-2)(a+z-3)...(a)(a-1)...1]}{[(a+b+z-1)(a+b+z-2)...(a+b)](a-1)...1}
\]

\[
= \frac{b[(a+z-2)(a+z-3)...a]}{[(a+b+z-1)(a+b+z-2)...(a+b)]}
\]

Now using the relation

\[X^{(j)} = X(X-1)(X-2)...(X-Y+1)\]

in above we get

\[
P[Z = z] = \frac{b(a+z-2)^{(z-1)}}{(a+b+z-1)^{(z)}}
\]

This is the probability mass function of compound geometric beta distribution. But since we are concerned with the compound truncated geometric beta distribution
where geometric distribution is truncated to the right at \( z = A \), we proceed to derive its probability mass function as follows:

3.5.2: Compound right truncated Geometric Beta distribution:

Let the Geometric distribution be truncated at \( z = A \).
We have p.m.f. of compound geometric beta distribution as follows

\[
f_z = \frac{B(a + z - 1, b + 1)}{B(a, b)} , \quad z = 1, 2, \ldots, \quad a, b > 0
\]

\( : \) We write p.m.f. of right truncated compound Geometric Beta distribution as

\[
f_z = \frac{KB(a + z - 1, b + 1)}{B(a, b)} , \quad z = 1, 2, \ldots, A-1, A
\]

\[= 0 \quad , \quad z > A\]

Now \( \sum_{z=1}^{n} f_z = 1 \)

\( \Rightarrow \sum_{z=1}^{A} f_z + \sum_{z=A+1}^{\infty} f_z = 1 \)
\[ i.e. \sum_{z=1}^{A} f_z = \sum_{z=1}^{A} K \cdot \frac{B(a + z - 1, b + 1)}{B(a, b)} = 1 \]

\[ \Rightarrow K = \frac{B(a, b)}{\sum_{z=1}^{A} B(a + z - 1, b + 1)} \]

\[ \therefore f_z = \frac{B(a, b)}{\sum_{z=1}^{A} B(a + z - 1, b + 1)} \cdot \frac{B(a + z - 1, b + 1)}{B(a, b)} \]

\[ = \frac{B(a + z - 1, b + 1)}{\sum_{z=1}^{A} B(a + z - 1, b + 1)} \quad \ldots (3.5.2.1) \]

Consider from equation (3.5.2.1) \( \sum_{z=1}^{A} B(a + z - 1, b + 1) \) only

\[ \therefore \sum_{z=1}^{A} B(a + z - 1, b + 1) = B(a, b + 1) + B(a + 1, b + 1) + B(a + 2, b + 1) + \ldots + B(a + A - 1, b + 1) \]

\[ = \int_{0}^{1} y^{a-1}(1 - y)^{b} \, dy + \int_{0}^{1} y^{b}(1 - y)^{b} \, dy + \int_{0}^{1} y^{a+1}(1 - y)^{b} \, dy + \ldots + \int_{0}^{1} y^{a+A-2}(1 - y)^{b} \, dy \]

\[ = \int_{0}^{1} y^{a+1}(1 + y + y^2 + \ldots + y^{A-1})(1 - y)^{b} \, dy \]
\[
\begin{align*}
&= \int_0^1 y^{a-1} \frac{1-y^b}{1-y} (1-y)^b \, dy, \quad y \neq 1 \\
&= \int_0^1 y^{a-1}(1-y)^{b-1} \, dy - \int_0^1 y^{a+1}(1-y)^{b-1} \, dy \\
&= B(a,b) - B(A+a,b) \\
&= B(a,b) - B(b,A+a) \quad \ldots (3.5.2.2)
\end{align*}
\]

Using the standard result that

\[B(a+1,b) = \frac{a}{a+b} B(a,b)\]

We can write

\[B(b+1,a+A) = \frac{b}{(a+A+b)} B(b,a+A)\]

\[\therefore B(b,a+A) = \frac{(a+A+b)}{b} B(b+1,a+A)\]

Putting above value of \(B(b,a+A)\) in equation (3.5.2.2) we get
\[
\sum_{z=1}^{1} B(a+z-1, b+1) = B(a, b) - \frac{(a+A+b)}{b} B(b+1, A+a)
\]

\[
\therefore \sum_{z=1}^{1} B(a+z-1, b+1) = \frac{bB(a,b) - (a+A+b)B(b+1, A+a)}{b}
\]

Therefore we can write equation (3.5.2.1) as

\[
f_z = \frac{bB(a+z-1, b+1)}{bB(a,b) - (a+b+A)B(b+1, A+a)} \quad z = 1, 2, \ldots, A-1, A
\]

\[a, b > 0\]

\[
\quad \ldots \quad (3.5.2.3)
\]

Which is a p.m.f. of right truncated compound Geometric Beta distribution.

Below in table 3.4 we have calculated some probability values for various values of \(z, a, b\) and \(A\).

For calculation of the probability values for various values of \(z, a, b\) and \(A\) in the table 3.4 below, we have prepared a computer program in Visual Basic which is provided in Appendix II.
### Table 3.4

Values of $f_z$ with finite range of $z$

<table>
<thead>
<tr>
<th>$z$</th>
<th>$f_z$</th>
<th>$f_z$</th>
<th>$f_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$b = 10$</td>
<td>$a = 5$</td>
<td>$a = 15$</td>
</tr>
<tr>
<td>2</td>
<td>0.208333</td>
<td>0.230770</td>
<td>0.198422</td>
</tr>
<tr>
<td>4</td>
<td>0.028595</td>
<td>0.083028</td>
<td>0.099070</td>
</tr>
<tr>
<td>6</td>
<td>0.005418</td>
<td>0.032639</td>
<td>0.051567</td>
</tr>
<tr>
<td>8</td>
<td>0.001290</td>
<td>0.013819</td>
<td>0.027850</td>
</tr>
<tr>
<td>10</td>
<td>0.000365</td>
<td>0.006232</td>
<td>0.015544</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$z$</th>
<th>$f_z$</th>
<th>$f_z$</th>
<th>$f_z$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a = 10$</td>
<td>$b = 5$</td>
<td>$b = 15$</td>
</tr>
<tr>
<td>2</td>
<td>0.208388</td>
<td>0.230769</td>
<td>0.198413</td>
</tr>
<tr>
<td>4</td>
<td>0.089893</td>
<td>0.040293</td>
<td>0.018628</td>
</tr>
<tr>
<td>6</td>
<td>0.043054</td>
<td>0.008429</td>
<td>0.002173</td>
</tr>
<tr>
<td>8</td>
<td>0.022366</td>
<td>0.002039</td>
<td>0.000303</td>
</tr>
<tr>
<td>10</td>
<td>0.012398</td>
<td>0.000556</td>
<td>0.000049</td>
</tr>
</tbody>
</table>
These types of calculations are useful in estimating the values of a & b precisely and make valid conclusions on the 'time taken' to stop along with the 'position' at which we stop.

We have studied this aspect of the problem in the next chapter using some empirical results.