2003年日本オペレーションズ・リサーチ学会
春季研究発表会

応募者数が一般化された一様分布の下でのGusein-Zade問題

02102643　愛知大学　*川合 益代　KAWAI Masuyo
01303783　　　　　　玉置 光司　TAMAKI Mitsushi

1. はじめに

最も単純な秘書問題は，以下のように記述できる．ある雇用者が1人の秘書を採用するために面接を行い，これに応募してきた n 人に毎時1人ずつ面接を行う．雇用者は，全応募者の中で最良の応募者を採用したいと考えており，実際に採用した応募者が最良 (best) のとき成功とみなす．このとき，成功確率を最大にする政策とその下での成功確率を求める問題が古典的秘書問題である．ただし，雇用者は，最良の応募者から順に $1, 2, \cdots, n$ と順位付けすることができ，面接を行ったら直ちにその応募者を採用するかまたは採用せずに次の応募者を面接するかを決定しなければならない．採用するか否かの意思決定は，応募者の相対順位 (今まで面接した中での順位) にのみ依存して行い，一度不採用にした応募者は，後で採用することができない．もし $(n-1)$ 番目の応募者まで採用しないならば，必ず最後の応募者を採用しなければならないものとする．ここでは，相対順位1の応募者を候補者と呼ぶことにする．この問題において n が十分に大きい場合の最適政策は，最初の $e^{-1}n$ 人の応募者を採用せず，それ以降に出現する最初の候補者を採用することである．この政策下での成功確率もまた e^{-1} に近づく．

この論文では，1人だけ採用することができ，採用した応募者が最良 (best) または2番目に良い (2nd best) とき，成功とみなす．ここでは，相対順位1（2）の応募者を候補者1（2）と呼び，区別する必要がない場合は単に候補者と呼ぶ．また，応募者数 N が確率変数で区間 $[m, n]$ 上の一様分布に従うものとして問題を考察する．ただし，m, n は，与えられた正整数で $m \leq n$ を満足する．$m = 1$ の場合は，既に Presman and Sonin によって研究されている．$m = n$ の場合は，古典的秘書問題に他ならない．このように我々の問題は，古典的秘書問題と従来考えられていた一様分布の問題を特殊な場合として包含するものであり，この意味でこの一様分布を一般化された一様分布と呼ぶ．一般化された一様分布においても，最適政策は，閾値ルール (threshold rule) となることが示された．

2. 定式化

応募者数 N が有限な確率変数で，分布 $p_r = P\{N = r\}, 1 \leq r \leq n$ をもつの場合の最適方程式は以下のように与えられる．また，$\pi_r = \sum_{k=r}^{n} p_k$ とおく．

$$f_i(r) = \max\{f_i^A(r), f^R(r)\}, \qquad i = 1, 2, \qquad 1 \leq r \leq n$$

ただし

$$\begin{cases} f_1^A(r) = \dfrac{r}{\pi_r} \displaystyle\sum_{k=r}^{n} \dfrac{p_k(2k - r - 1)}{k(k-1)}, \\ f_2^A(r) = \dfrac{r(r-1)}{\pi_r} \displaystyle\sum_{k=r}^{n} \dfrac{p_k}{k(k-1)}, \\ f^R(r) = \dfrac{r(r-1)}{\pi_r} \displaystyle\sum_{k=r+1}^{n} \dfrac{\pi_k[f_1(k) + f_2(k)]}{k(k-1)(k-2)}. \end{cases}$$

ここで

$f_i(r)$: r 番目の応募者が候補者 $i, i = 1, 2$ のとき，以後最適にふるまって成功する確率．
$f_i^A(r)$: r 番目の応募者が候補者 $i, i = 1, 2$ のとき，その人を採用して成功する確率．
$f^R(r)$: r 番目の応募者が候補者のとき，その人を採用せず以後最適にふるまって成功する確率．

3. N が $[m,n]$ 上の一様分布に従うケース

これ以降は，応募者数 N が $[m,n]$ 上の一様分布に従うものとして問題を考察する．このとき，p_k は以下のようになる．

$$p_k = \begin{cases} \dfrac{1}{n-m+1}, & m \le k \le n \\ 0, & 1 \le k < m \end{cases}$$

補題1　$f_i^A(r)$, $i=1,2$ は，r について非減少関数である．

補題2　$f_1^A(r) \ge f_2^A(r)$, $1 \le r \le n$.

補題3 a　s_2 を以下のように定義する．

$$s_2 = \min \begin{cases} 1 \le r \le m : \dfrac{3(n-r+1)}{2n} - \displaystyle\sum_{k=r}^{n} \frac{1}{k} \ge 0, \\ m \le r \le n : \dfrac{3(r-1)(n-m+1)}{2n(m-1)} - \displaystyle\sum_{k=m}^{n} \frac{1}{k} \ge 0 \end{cases}$$

このとき，$s_2 \le r$ に対して次式が成立する．(i)，$f_1(r) = f_1^A(r)$,　(ii)，$f_2(r) = f_2^A(r)$.

補題3 b　s_1 を以下のように定義する．

$$s_1 = \min \begin{cases} s_2 \le m \le r \le n : -2\displaystyle\sum_{j=r}^{s_2-1} \frac{1}{j}\left(\sum_{k=j+1}^{n} \frac{1}{k} - \frac{3}{2}\right) + \frac{2r-3s_2+1}{n} + 1 \ge 0, \\ 1 \le r < m \le s_2 : 2\displaystyle\sum_{k=m}^{n} \frac{1}{k}\sum_{j=r}^{m-1}\frac{1}{j} + 2\sum_{j=m}^{s_2-1}\frac{1}{j}\left(\sum_{k=j+1}^{n}\frac{1}{k} - \frac{3}{2}\right) + \frac{3s_2-2r-1}{n} + \frac{2(r-1)}{m-1} - 1 \le 0, \\ 1 \le r < m, s_2 < m : \displaystyle\sum_{k=m}^{n}\frac{1}{k}\left(\sum_{j=r}^{s_2-1}\frac{1}{j} - \frac{(1-s_2)(n-m)}{m-1}\right) - (n-m+1)\left(\frac{2r-s_2-1+2(1-s_2)(n-m+1)}{2n(m-1)}\right) \le 0 \end{cases}$$

このとき，$s_1 < s_2$ であるならば，$s_1 \le r \le s_2 - 1$ に対して次式が成立する．(i)，$f_1(r) = f_1^A(r)$, (ii)，$f_2(r) = f^R(r)$.

補題3 c　$1 \le r < s_1$ に対して次式が成立する．(i)，$f_1(r) = f^R(r)$,　(ii)，$f_2(r) = f^R(r)$.

補題4　$s_1 \le s_2$,　$s_1 \ne 1$

定理　最適政策は，閾値 (s_1, s_2) の閾値ルールとなる．すなわち，正整数 s_1, s_2 が存在して，最初の (s_1-1) 人の応募者を採用せず，s_1 以降に出現する最初の候補者1を採用する．ただし，もし候補者1が出現しないならば，s_2 以降に出現する最初の候補者を採用することである．

補題5　最適政策下での成功確率 V は以下のように与えられる．

$$V = \begin{cases} \dfrac{s_1}{n-m+1}\left[2\displaystyle\sum_{j=s_1}^{s_2-1}\frac{1}{j}\sum_{k=j+1}^{n}\frac{1}{k} - \sum_{k=s_1}^{s_2-1}\frac{1}{k} + 2\sum_{k=s_2}^{n-1}\frac{1}{k} + \frac{3s_2-s_1}{n} - 2\right], & 1 < m < s_1 < s_2 < n \\ s_1\left[\dfrac{2}{n-m+1}\displaystyle\sum_{k=m}^{n}\frac{1}{k}\left(\sum_{j=s_1}^{s_2-1}\frac{1}{j} + \frac{s_2(n-m)-n+2m-1}{m-1}\right) + \frac{s_1-s_2+2(1-s_2)(n-m+1)}{n(m-1)}\right], & 1 < s_1 < s_2 < m < n \\ \dfrac{s_1}{n-m+1}\left[2\displaystyle\sum_{k=m}^{n}\frac{1}{k}\sum_{j=s_1}^{m-1}\frac{1}{j} + 2\sum_{j=m}^{s_2-1}\frac{1}{j}\sum_{k=j+1}^{n}\frac{1}{k} - \sum_{k=m}^{s_2-1}\frac{1}{k} + 2\sum_{k=s_2}^{n-1}\frac{1}{k} + \frac{3s_2-s_1}{n} + \frac{s_1-1}{m-1} - 1\right], & 1 < s_1 < m < s_2 < n \end{cases}$$

参考文献

[1] Gusein-Zade, S. M. The problem of choice and the optimal stopping rule for a sequence of indipendent trials, *Theo.Prob.Appl.* **11**, 472-476. (1966)

[2] 川合 益代，玉置 光司，応募者数が一般化された一様分布に従う秘書問題，京都大学数理解析研究所講究録 1194, p129-134. (2000)