国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
一九八六年第二期　总第五十四期　No.2 1986　Sum. 54

允许拒绝的秘书问题

金治明

摘　要　本文考虑了一类特殊的秘书问题，每一名候选的姑娘都因某种理由以一定概率拒绝入选，而且假定这个拒绝概率只与候选人的绝对名次有关。本文给出了这一问题的最优停止规则，推广了文献[2]的结果。

一、引　言

秘书问题是最优停止理论中一个重要而又引人入胜的理论模型。它的意义当然不只限于选择秘书，实际中许多最优选择的问题都可以化为秘书问题来解决。

秘书问题的原型是这样的。一个经理要从N个姑娘中选用一名秘书。按照某种标准，我们可以把这N位姑娘排成第1，2，…，N名。经理每次会见一名姑娘，通过面试决定录用与否。如果录用当前这位姑娘，则停止下面的会见，否则进行下一次会见，而且假定每名姑娘一经拒绝就再也不能招回。在每次会见时，经理只能知道当前的姑娘与已会见的姑娘相比的相对名次，并不能知道她的绝对名次。我们假设姑娘到来的优劣次序是随机的，问经理应怎样决定他的录用策略，也就是采取什么方式来停止他的会见才是最优的。所谓最优，可以有不同的标准。比如，使得选中最好的姑娘（即选中第1名姑娘）的概率最大，或者使得选取姑娘的绝对名次的数学期望最小。在[4]中我们选用第二种标准，本文采用第一种标准。

令$\Omega=\{(a_1, a_2,\cdots,a_N)$：其中$(a_1,\cdots,a_N)$是1，2，…，$N$的一个排列$\}$，并规定每一个基本事件的概率是相等的，都等于$\frac{1}{N!}$. 令$\mathscr{F}$是$\Omega$的一切子集的全体，它是一个$\sigma$代数。我们用$b_1, b_2,\cdots,b_N$表示某一个排列，则$b_n$，$1\leqslant n\leqslant N$是随机变量。

设y_n等于$b_1, b_2,\cdots,b_n$中小于等于b_n的个数，即b_n的相对名次。

$$\mathscr{F}_n=\sigma(y_1,y_2,\cdots,y_n).$$
$$X_n=P(b_n=1\,|\mathscr{F}_n).$$

称取正整数值的随机变量t为停止规则，即指它满足$\{t=n\}\in\mathscr{F}_n$，且$t<\infty\ a.s.$ 容易算得

$$EX_t=P(b_t=1)$$

1985年9月收到

按照第一个标准，我们的目的是选一个停止规则t^*，使得

$$EX_{t^*}=\sup_{t\in T}EX_t \triangleq v^N$$

其中T是一切停止规则的全体，而称V^N为序列$\{X_n,\ 1\leqslant n\leqslant N\}$的值，即问题的最优值。

Chow, Y.S. 和 Robbins. H 在文献[1]中证明了，此时最优规则是

$$t^*=\inf\{n\geqslant r^*:\ y_n=1\}$$

其中$r^*\triangleq r^*(N)=\inf\left\{r:\ \sum_{k=r}^{N-1}\frac{1}{k}\leqslant 1\right\}$，而且证明了

$$\lim_{N\to\infty}\frac{r^*(N)}{N}=\lim_{N\to\infty}V^N=\frac{1}{e}\approx 0.368$$

经理该怎样运用这个最优规则呢？比如应试的姑娘有 100 名，那么经理对前 37 位只会见而不录用，而当接见第 38 位姑娘就看它是否比前 37 位都好。如果她比前 37 位好则录用第 38 位，否则会见下一位。如此下去，一直到遇见一位比前面都好的姑娘。最不幸的事情也可能发生，也就是从第 38 位之后再也找不到一位比前面都好的姑娘，于是只能录用最后一位应试者。但是容易算得这件事发生的概率是很小很小的。

秘书问题近年来有许多深入的研究。比如，研究第二标准的秘书问题，无限个秘书的问题，多个选择的问题，姑娘到来的次序服从特定分布的问题，可招回的秘书问题等等。M.H.SMITH 在文献[2]中讨论了每一个候选的姑娘都有固定的概率 $1-p$ 拒绝入选的问题。[2]中证明了此时最优停止是

$$t^*=\inf\{n\geqslant r^*:\ y_n=1\}$$

其中

$$r^*=\inf\left\{r:\ \prod_{k=r}^{N-1}\left(1+\frac{1-p}{k}\right)\leqslant\frac{1}{p}\right\}$$

且

$$\lim_{N\to\infty}\frac{r^*}{N}=\lim_{N\to\infty}V^N=p^{\frac{1}{1-p}}$$

二、允许拒绝的秘书问题

本文考虑每位姑娘的拒绝概率与她的绝对名次有关的情形。当然更符合实际的应假定拒绝概率是绝对名次的递减函数。但是我们的研究表明不必作此假定。

设每位姑娘拒绝与否及其概率只与她的绝对名次有关，与她的相对名次以及其它姑娘是否拒绝是独立的。

令

$$Z_n=\begin{cases}0 & \text{第 } n \text{ 时刻应试的姑娘} b_n \text{ 拒绝入选}\\ 1 & \text{否则}\end{cases}\tag{1}$$

则$Y_1,\ Z_1,\ Y_2,\ Z_2,\cdots,Y_N,\ Z_N$是独立的随机变量，且假定

$$P(Z_n=1\mid b_n=k)=p_k$$

$$\mathscr{F}_n=\sigma(Y_1,\ Z_1,\cdots,Y_n,\ Z_n),\ 1\leqslant n\leqslant N\tag{2}$$

于是，当$i=1$，$k=1$时

$$
\begin{aligned}
&P(b_r=1,\ Z_r=1\,|Y_r=k,\ Z_r=i)\\
&=P(b_r=1,\ Z_r=1,\ Y_r=1)/p(Y_r=1,\ Z_r=1)\\
&=P(Z_r=1\,|b_r=1)p(b_r=1)/[p(Y_r=1)\cdot p(Z_r=1)]\\
&=rp_1\Big/\sum_{i=1}^{N}p_i
\end{aligned}
\tag{3}
$$

显然

$$
\begin{aligned}
X_r&\triangleq P(b_r=1,\ Z_r=1\,|\mathscr{F}_r)\\
&=P(b_r=1,\ Z_r=1\,|Y_r,\ Z_r)\\
&=\left(rp_1\Big/\sum_{i=1}^{N}p_i\right)I_{[Y_r=1,Z_r=1]}
\end{aligned}
\tag{4}
$$

对任何停时 t

$$
EX_t=\sum_{r=1}^{N}\int_{[t=r]}X_r=\sum_{r=1}^{N}\int_{[t=r]}I_{[b_r=1,Z_r=1]}=P(b_t=1,\ Z_t=1)\tag{5}
$$

我们的目的是求停时 t，使 EX_t 达到最大，亦即最好的姑娘被选中的概率最大。

令

$$
\gamma_N=X_N=\frac{Np_1}{\sum_{i=1}^{N}p_i}I_{[Y_N=1,Z_N=1]}\tag{6}
$$

$$
\gamma_n=X_n\vee E(\gamma_{n+1}\,|\mathscr{F}_n)\qquad n=N-1,\ N-2,\cdots,1\tag{7}
$$

容易由后退归纳法证明，r_{n+1} 与 $\mathscr{F}_n$ 独立，$n=N-1,\ N-2,\cdots,1$ 于是

$$
\gamma_n=X_n\vee E\gamma_{n+1}\triangleq X_n\vee v_n\tag{8}
$$

其中记 $v_n=E\gamma_{n+1}$，$n=0,1,2,\cdots,N-1$. 那么最优规则是

$$
t^*=\inf\{r\geqslant 1:\ X_r\geqslant v_r\}\tag{9}
$$

因

$$
v_{N-1}=E\gamma_N=EX_N=\frac{Np_1}{\sum_{i=1}^{N}p_i}p(Y_N=1,\ Z_N=1)=\frac{p_1}{N}
$$

$$
\begin{aligned}
v_{r-1}&=E\gamma_r=EX_rI_{[X_r\geqslant v_r]}+v_rP(X_r<v_r)\\
&=v_r+\left(\frac{rp_1}{\sum_{i=1}^{N}p_i}-v_r\right)^{+}\cdot\frac{\sum_{i=1}^{N}p_i}{Nr}
\end{aligned}
\tag{10}
$$

这样 v_r 单调下降，且 $v_{N-1}=\frac{p_1}{N}>0$；$\frac{rp_1}{\sum_{i=1}^{N}p_i}$ 随 r 单调上升，且当 $r=N$ 时，$\frac{Np_1}{\sum_{i=1}^{N}p_i}>p_1>v_{N-1}$. 因此必有 r^*，使得当 $r\geqslant r^*$ 时，$v_r<\frac{rp_1}{\sum_{i=1}^{N}p_i}$.

由式(10)得

$$v_{r-1}-v_r=-v_r\frac{\sum_{i=1}^{N}p_i}{Nr}+\frac{p_1}{N},\quad r\geqslant r^* \tag{11}$$

我们来解上述差分方程。用试验法先求得它的一个特解。

令

$$f(r)=\frac{Cr}{N},\quad r\geqslant r^*$$

代回(11)式定出常数C，于是

$$f(r)=-\frac{rp_1}{N-\sum_{i=1}^{N}p_i}\qquad r\geqslant r^*$$

令$u_r=v_r-f(r)$，它满足齐次差分方程

$$u_{r-1}-u_r=-\frac{\sum_{i=1}^{N}p_i}{Nr}u_r\qquad r\geqslant r^*$$

边界条件

$$u_{N-1}=v_{N-1}-f(N-1)=\frac{p_1}{N}+\frac{(N-1)p_1}{N-\sum_{i=1}^{N}p_i}$$

从而解得

$$u_r=p_1\left(\frac{1}{N}+\frac{N-1}{N-\prod_{i=1}^{N}p_i}\right)\cdot\prod_{k=r+1}^{N-1}\left(1-\frac{\sum_{i=1}^{N}p_i}{Nk}\right)$$

经计算，

$$v_r=\frac{p_1r}{N-\sum_{i=1}^{N}p_i}\left[\frac{N^2-\sum_{i=1}^{N}p_i}{Nr}\prod_{k=r+1}^{N-1}\left(1-\frac{\sum_{i=1}^{N}p_i}{Nk}\right)-1\right]$$

$$=\frac{p_1r}{N-\sum_{i=1}^{N}p_i}\left[\prod_{k=r}^{N-1}\left(1+\frac{N-\sum_{i=1}^{N}p_i}{Nk}\right)-1\right],\qquad r\geqslant r^* \tag{12}$$

$$v_r=v_{r^*-1}\qquad r=0,1,\cdots,r^*-1$$

$$v_{N-1}=\frac{p_1}{N}$$

其中

$$r^*=\inf\left\{r\geqslant 1:\ \prod_{k=r}^{N-1}\left(1+\frac{N-\sum_{i=1}^{N}p_i}{Nk}\right)\leqslant\frac{N}{\sum_{i=1}^{N}p_i}\right\} \tag{13}$$

最优规则为

$$t^*=\inf\{n\geqslant r^*:\ y_n=1,\ Z_n=1\} \tag{14}$$

利用初等不等式

$$\left(\frac{k+1}{k}\right)^{1-\frac{\sum_{i=1}^{N}p_i}{N}}<1+\frac{1-\frac{1}{N}\sum_{i=1}^{N}p_i}{k}$$

$$<\left(\frac{(k+1)-\frac{1}{N}\sum_{i=1}^{N}p_i}{k-\frac{1}{N}\sum_{i=1}^{N}p_i}\right)^{1-\frac{\sum_{i=1}^{N}p_i}{N}}$$

以及 r^* 的定义，可得

$$\left(\frac{N-\frac{1}{N}\sum_{i=1}^{N}p_i}{r^*-1-\frac{1}{N}\sum_{i=1}^{N}p_i}\right)^{1-\frac{1}{N}\sum_{i=1}^{N}p_i}>\frac{N}{\sum_{i=1}^{N}p_i}>\left(\frac{N}{r^*}\right)^{1-\frac{1}{N}\sum_{i=1}^{N}p_i}$$

从而，

$$\lim_{N\to\infty}\frac{r^*}{N}=\lim_{N\to\infty}\left(\frac{\sum_{i=1}^{N}p_i}{N}\right)^{\frac{1}{1-\frac{1}{N}\sum_{i=1}^{N}p_i}} \tag{15}$$

（只要右方极限存在）。

如果再假定

$$\lim_{n\to\infty}p_n=p$$

则

$$\lim_{N\to\infty}\frac{r^*}{N}=p^{\frac{1}{1-p}} \tag{16}$$

而

$$\lim_{N\to\infty}V^N=\lim_{N\to\infty}v_0=\lim_{N\to\infty}v_{r^*}=\lim_{N\to\infty}\frac{r^*}{N}\cdot\frac{N}{\sum_{i=1}^{N}p_i}p_1=p_1p^{\frac{p}{1-p}}$$

最后我们给出最优值的表达式

$$V^N=v_0$$

$$=\frac{p_1(r^*-1)}{N-\sum_{i=1}^{N}p_i}\left[\prod_{k=r^*-1}^{N-1}\left(1+\frac{N-\sum_{i=1}^{N}p_i}{Nk}\right)-1\right]$$

$$=\frac{p_1}{N-\sum_{i=1}^{N}p_i}\left[(r^*-1)\left(1+\frac{N-\sum_{i=1}^{N}p_i}{N(r^*-1)}\right)\prod_{k=r^*}^{N-1}\left(1+\frac{N-\sum_{i=1}^{N}p_i}{Nk}\right)-r^*+1\right]$$

$$=\frac{p_1}{N-\sum_{i=1}^{N}p_i}\left[\left(r^*-\frac{1}{N}\sum_{i=1}^{N}p_i\right)\prod_{k=r^*}^{N}\left(1+\frac{N-\sum_{i=1}^{N}p_i}{Nk}\right)-r^*+1\right]$$

由(13)，(14)，(17)式可见，最优规则只与平均值$\frac{1}{N}\sum_{i=1}^{N}p_i$有关，而最优值$V^N$却与$p_1$的值有关。这个事实是有趣的。当$p_i=p$，$1\leqslant i\leqslant N$，就化为[2]中所讨论的问题。

下面的表1给出了在

(1) $p_1=p_2=\cdots=p_N=0.5$

(2) $p_1=0.8$，$\frac{1}{N}\sum_{i=1}^{N}p_i=0.5$

两种情形下v_0的值与r^*值的比较。

表 1

N	$p_1=p_2=\cdots=p_N=0.5$		$p_1=0.8$，$\sum_{i=1}^{N}p_i/N=0.5$	
	r^*	V^N	r^*	V^N
2	1	0.37500	1	0.60000
3	1	0.31250	1	0.50000
4	2	0.29687	2	0.47500
5	2	0.29218	2	0.46750
10	3	0.26985	3	0.43177
25	7	0.25770	7	0.41232
50	13	0.25379	13	0.40606
100	26	0.25187	26	0.40301
1000	251	0.25018	251	0.40030

参 考 文 献

[1] [美]Chow.Y.S, H. Robbins, D. Sigmund 著，何声武、汪振鹏译，最优停止理论，上海科技出版社 83.4

[2] M.H. SMITH, A secretary problem with uncertain employment, J. Appl. prob. 12(1975),620~624.

[3] CHOW, Y.S. MORIGUTI, S. Robbins, H. AND Samuels, S. M. (1964), Optimum seclection based on relative rank(THE'Sec retary prob lem'), Israel. J. Math, 2, 81—90.

[4] 金治明，平均名次最小的可拒绝秘书问题，应用概率与统计，1986.4.

A Secretay Problem With Refusing

Jing Zhiming

Abstract

In this paper, we comsider a class of special secretary problem. Every applicant can refuse an offier of employment with a probability which depend only on her absolute rank. We derive the optimal stopping rule which maximizes the probabitily of employing the best one.