

# 最优停止理论与导师选取

满洋

哈尔滨工业大学（深圳） 计算机科学与技术学院 200111525 \*

2021 年 11 月

## 1 问题背景

最近我阅读到了一篇文章 [1]，描述最优停止理论在研究生导师选取上的应用。文章使用了经典麦穗问题的结论，我认为对于这个场景，有更合适的最优停止策略。

想象有以下情景：某同学在参加保研面试。面试之前他与这些老师互不了解，聊完之后他对老师们有了自己的评价；但出乎他意料的是，每个老师在面试结束时都表达出对某同学的满意，当即给出 offer，并要求他现场决定是否接受。某同学只有一次选择机会，offer 接受后不能再反悔拒绝，拒绝后也不能再反悔接受。在这样的情境下，这位同学该如何在他们中选出评价最高（即最喜欢）的一位呢？

此类问题其实是非常古老的问题，被人们研究了很多年了，策略通常为“观望-选择”，因主要研究最佳的停止观望时间，此类问题称为最优停止问题。最优停止问题又称麦穗问题、秘书问题、博士结婚问题等，大家应该对以下的故事有所耳闻：

在很久之前，苏格拉底的三个弟子向他请教，问他怎样才能找到自己理想的伴侣。他什么话都没说，直接把三个弟子带到一片麦田里，让弟子们每人挑一根最大的麦穗。做这一切的前提是，不能走回头路，也就意味着他们只有一次选择的机会。

这三个弟子呢，颇有意思，因为他们的选择大不相同。

第一个弟子刚走进去没多久，就摘了一根自以为很大的麦穗，心里不禁洋洋得意。但他越走越走发现自己做错了选择，因为后面的麦穗远大于他手里的麦穗。

第二个弟子不同于第一个弟子，他恰好相反，他一直往前走，总觉得最大的麦穗都在前面，所以他看到比较大的麦穗早就错过了。

第三个弟子，也是三个人里比较聪明的一个了，他不慌不忙地边走边思考，然后把麦田分成了三份，走第一段路时，他只看不摘；继续往前走时，他会将现在的麦穗与之前看到过的麦穗进行对比。最后他挑选了一根他认为最大的麦穗。当然，他也是这三人中拿到最大麦穗的人。[2]

对于经典的麦穗问题，最受公认的最优停止时间是麦穗总数  $N$  的  $\frac{1}{e}$ ，也即  $\frac{N}{e}$  [3]。以下我们将会讨论这个结论的来源和我们“导师选取”问题的结论。

## 2 经典结论

我们将这个问题用数学语言描述：考虑一组  $N$  个独立服从  $[0, 1]$  上均匀分布的随机变量  $X_1, X_2, \dots, X_N$ 。然后我们按编号从小到大遍历这些变量，对于每个变量，我们只能做出“保留”或者“丢弃”两种

---

\*Email: me@eastonman.com

选择，并且我们只有一次保留机会。那么应该如何制定策略，使得我们选取到最大变量的概率最大？

我们主要关注于一种简单的策略，就是通过某种方法设定一个阈值  $c$ ，在选取过程中一旦出现值超过阈值  $c$  的变量，那么立即选取。很自然地有这样一个策略：对于前  $m$  个变量，我们一律不选，对于从  $m+1$  号开始的变量，一旦有比前面更大的变量，我们就立即选取。这种先观望，后选取的策略，以下称为  $m$  策略。

现在讨论如何选取  $m$  的值。

对于某个特定的  $m$ ，如果最大值出现在第  $i$  个位置， $m$  策略选取到最大值的概率为

$$P(m, i) = \begin{cases} 0 & i \leq m \\ \frac{m}{i-1} & i > m \end{cases}$$

所以如果我们使用  $m$  策略，成功选取到最大值的概率为

$$P(m) = \sum_{i=m+1}^n \frac{1}{n} P(m, i) = \frac{m}{n} \sum_{i=m+1}^n \frac{1}{i-1}$$

当  $n$  和  $m$  都较大的时候，令  $r = \frac{m}{n}$ ，可将上式右侧的调和级数近似为  $\ln(\frac{n}{m})$ ，可得：

$$\begin{aligned} P(m) &\approx \frac{m}{n} \ln \frac{n}{m} \\ \Rightarrow P(r) &= -r \ln r \end{aligned}$$

求导得：

$$P'(r) = -1 - \ln r$$

当  $r = \frac{1}{e}$  时， $P'(r) = 1$ ， $P(r)$  达到最大值  $\frac{1}{e}$ 。

所以对于麦穗问题的  $m$  策略， $m$  的最优值就是  $\frac{N}{e}$ （取整数）。

然而，选取研究生导师和麦穗问题还有不同之处：1) 我们并不需要选到最好的导师，只是希望选取到好的导师。2) 我们不能允许我们糟糕的选择策略使得我们没有导师！

### 3 导师选取问题

首先用数学语言描述“导师选取”问题：对于  $N$  个导师，我们有主观评价分  $X_1, X_2, \dots, X_N$ ，然后我们希望通过某种策略选取导师。这种策略应当使选取的导师的评价分的期望尽可能高，同时也能保证我们能够选到老师（有学上）。除此之外，在现实中，评价导师并不能使用量化的分数，而只能通过比较，认为某导师较好。因此，我们设定虽然每位导师都有评价分  $X_i$ ，但我们选择的时候并不能得知具体的分  $X_i$ ，而是只能得到导师评价分之间的关系。

因此我们无法使用任何动态阈值或使用不同于任何导师评价分的阈值。我们只能使用类似  $m$  策略的方式，为了区分，下称  $t$  策略。 $t$  策略是指通过遍历前  $m$  个导师确定一个阈值  $t = \max(X_1, X_2, \dots, X_m)$ ，然后从第  $m+1$  开始，只要某导师的评价高于  $t$ ，那么立即选取该导师，如果直到第  $N-1$  位导师都没有找到评价大于  $t$  的，那么选取最后一位导师。

我们希望选取合适的  $m$  使得最终我们选取的导师的评价分的期望尽量的大。

首先我们计算阈值  $t$  的分布函数

$$\begin{aligned} F(t) &= F_{\max}(t) \\ &= P(\max(X_i) \leq t) \\ &= \prod P_i(X \leq t) \\ &= t^m \end{aligned}$$

由此可以得到阈值  $t$  的概率密度函数

$$\begin{aligned} f(t) &= F'(t) \\ &= mt^{m-1} \end{aligned}$$

对于确定的阈值  $t$ ，有两种情况：1) 我们成功选到了心仪的导师。2) 我们只好选了最后一位导师。

分别有概率

$$\begin{aligned} P_2 &= t^{N-m-1} \\ P_1 &= 1 - P_2 = 1 - t^{N-m-1} \end{aligned}$$

而这两种情况对应的评价分的期望为  $E_1 = \frac{1+t}{2}$ ， $E_2 = \frac{1}{2}$

于是我们可以计算出对于确定的  $m$ ，选取导师的评价分的期望为

$$\begin{aligned} E(x) &= \int_0^1 f(t)(P_1 E_1 + P_2 E_2) dt \\ &= \int_0^1 mt^{m-1} \left[ \frac{1}{2} t^{N-m-1} + (1 - t^{N-m-1}) \left( \frac{1+t}{2} \right) \right] dt \\ &= \frac{2m+1}{2(m+1)} - \frac{m}{2N} \end{aligned}$$

当期望  $E(x)$  最大时，有

$$\frac{\partial E}{\partial m} = \frac{1}{2(m+1)^2} - \frac{1}{2N} = 0$$

解得

$$m = \sqrt{N} - 1$$

可见如果我们希望以期望为指标，那么我们的最优停止时间将是  $\sqrt{N} - 1$ 。也就是从第  $\sqrt{N}$  位导师开始，只要出现比前面评价分高的导师，我们就立即选取。

## 4 讨论

我们发现，麦穗问题和“导师选取”问题的结论其实有很大不同，分别为  $m = \frac{N}{e}$  和  $m = \sqrt{N} - 1$ 。这里显然有一个增长速率上的差异。当  $N$  足够大时，按照  $\sqrt{N} - 1$  进行的策略需要观察的元素数量远小于按照  $\frac{N}{e}$  的策略。

这是因为选择到最大元素  $X_k$  要求在  $[1, k-1]$  的元素中，我们选定的阈值  $c$  既是  $[1, m]$  中最大的，也是  $[1, k-1]$  中最大的，这样的情况下要求我们选定的阈值较高，才更有可能选取到最大元素。而在按照期望为指标的策略中，我们并不需要选取最大的元素，而是期望较大就可以了，因此需要的观望次数也较少。

现实生活中有许多可以应用最优停止理论的地方，其中有相当一部分并不完全符合经典麦穗问题类的前提，而是与本文提出的修正情况相符合，因此本文得出的结论具有一定的现实意义。

## 5 结语

当然以上提出的情景和我没有任何关系，因为我在写概率论论文，没有导师给我发 offer。希望我能够有应用这个结论的场景，因为这意味着我至少不会失学了。同时也祝愿所有同学都不失学，并且能找到对象。

## 参考文献

- [1] P. Rain, “如何选择心仪对象,” Jun 2021. [Online]. Available: <https://blog.lhp-pku.top> 1
- [2] 庭前云落, “麦穗理论 | 不求最好, 但求合适,” Nov 2021. [Online]. Available: <https://zhuanlan.zhihu.com/p/113593894> 1
- [3] S. Demers, “The duration of optimal stopping problems,” 2019. 1