

9 Secretary Problem

This refers to the problem of choosing a candidate out of n rankable candidates, arriving in random order. The selection has to be made as soon as the candidate is seen and the decision can not be reverted. We analyze the following two questions:

1. Finding the best secretary
2. Finding an expected good secretary

9.1 Finding the best

Assumption: The candidates arrive in a uniformly random distribution.

9.1.1 Preliminary approaches:

1. Take first strategy:
 $\mathbb{P}(\text{finding best person}) = \frac{1}{n}$
2. Explore the first $n/2$ persons, then take the first person that is better than the first $n/2$ (or if no one is found, take the last person)

$$\begin{aligned}\mathbb{P}(\text{finding best person}) &\geq \mathbb{P}(\text{second best person in first half}) \cdot \mathbb{P}(\text{best person in second half}) \\ &= (1/2)(1/2)\end{aligned}$$

$$\mathbb{P}(\text{finding best person}) \geq 1/4$$

9.1.2 Optimal strategy

We start accepting at position r .

$$\begin{aligned}
 \mathbb{P}(\text{picking the best}) &= P(r) \\
 &= \sum_{i=1}^n \mathbb{P}(\text{picking } i \cap i \text{ is the best}) \\
 &= \sum_{i=1}^n \mathbb{P}(\text{picking } i \mid i \text{ is the best}) \mathbb{P}(i \text{ is the best}) \\
 &= \sum_{i=1}^{r-1} 0 + \frac{1}{n} \sum_{i=r}^n \mathbb{P}(\text{second best of the first } i \text{ applicants is in the first } r-1 \mid i \text{ is the best}) \\
 &= \frac{1}{n} \sum_{i=r}^n \frac{r-1}{i-1} \\
 &= \frac{r-1}{n} \sum_{i=r}^n \frac{1}{i-1}
 \end{aligned}$$

We need to optimize $\mathbb{P}(r)$, i.e., we need to find the best $\frac{r}{n}$ of applicants to explore. Consider $x = \frac{r}{n}$, $t = \frac{i}{n}$, $dt = \frac{1}{n}$, $0 < x < 1$

As, $n \rightarrow \infty$, the summation can be expressed as

$$x \int_x^1 \frac{1}{t} dt = -x \log(x)$$

The maximum value is obtained when $x^* = \frac{1}{e}$ and the associated $\mathbb{P}(x^*) = \frac{1}{e}$. Thus, the optimal strategy is to start accepting from $\frac{n}{e}$ applicants.

9.2 Finding an expected good secretary

In this analysis, every person is assumed to be a random variable coming from uniform distribution. We have a sequence x_1, x_2, \dots, x_n , where $\forall i, x_i \sim \text{Uni}(0, 1)$ and we want to maximize the expectation of the person hired. The strategy is to explore first $c-1$ applicants and take the best after $c-1$ applicants.

Aside: For x_1, x_2, \dots, x_t , where $\forall i, x_i \sim \text{Uni}(0, 1)$, let $X_t = \max(x_1, x_2, \dots, x_t)$

$$\begin{aligned}
 \mathbb{E}(X_t) &= \frac{t}{t+1} \\
 \mathbb{E}(x_i) &= \frac{1}{2}
 \end{aligned}$$

This can be proved as follows

$$F(s) = \mathbb{P}(X_t \leq s) = \prod_i \mathbb{P}(X_i \leq s) = s^t$$

$$\mathbb{P}(X_t = s) = \frac{dF}{ds} = p(s)$$

$$\mathbb{E}(X_t = s) = \int_0^1 s.p(s) ds = \frac{t}{t+1}$$

Expected value of the person chosen = $V_c(n)$

$$V_c(n) = \sum_{t=c}^{n-1} (\mathbb{P}(\text{not taking persons } c \text{ to } t-1))(\mathbb{P}(\text{taking person } t))(\mathbb{E}(\text{person } t)) \\ + \mathbb{P}(\text{taking last person})\mathbb{E}(\text{last person})$$

$$\mathbb{P}(\text{not taking persons } c \text{ to } t-1) = \prod_{s=c}^{t-1} \left(\frac{s-1}{s}\right)$$

$$\mathbb{P}(\text{taking person } t) = \frac{1}{t}$$

$$\mathbb{P}(\text{taking last person}) = \prod_{s=c}^{n-1} \left(\frac{s-1}{s}\right)$$

$$\mathbb{E}(\text{person } t) = \frac{t}{t+1}$$

$$\mathbb{E}(\text{last person}) = \frac{1}{2}$$

Using these, we get $V_c(n)$.

$$V_c(n) = \frac{2cn - c^2 + c - n}{2cn}$$

$$\frac{dV_c}{dc} = \frac{-c^2 + n}{2c^2n}$$

$$\frac{d^2V_c}{dc^2} \geq 0; 1 \leq c \leq n$$

To get the maximum value, we equate the derivative to zero. The optimal value of $c = \sqrt{n}$

Variants of this method include cases when n is unknown and when the candidates are modeled according to other distributions.