The Marriage Problem / The Secretary Problem

Ryan Brill 7/29/2025

The Secretary Problem

n applicants are interviewed in uniformly random order. After each interview, you must immediately and permanently accept or reject. You may select only one applicant. Your goal is to maximize the probability of choosing one of the top K candidates (ranked 1 through K, where K is the best).

Finding a Wife via Dating

Interview = a series of date(s). Applicant = a girl.

Optimal Stopping Strategy

Observation phase: Skip the first r candidates. Then, select the first subsequent candidate whose observed rank is better than all previous (i.e, the best you've seen so far). You succeed if that candidate is among the top K.

Todo

Compute success probability as a function of n, k, r.
Compute the optimal threshold r as a function of n and K.
Choose reasonable values of n and K to produce high success prob.

Success probability

$$P(n,k,r) = \mathbb{P}(\text{the strategy successfully selects an applicant in the top } k)$$

$$= \sum_{j=1}^{k} \mathbb{P}(\text{the strategy selects applicant of rank } j)$$

where applicants are ranked from best to worst by $1, 2, \ldots, n$

$$= \sum_{j=1}^{k} \sum_{t=r+1}^{n} \mathbb{P}\left(\begin{array}{c}\text{the strategy} \\ \text{stops at } t\end{array} \text{ AND } \begin{array}{c}\text{the selected candidate} \\ \text{has rank } j\end{array}\right)$$

$$= \sum_{j=1}^{k} \sum_{t=r+1}^{n} \mathbb{P}(A \cap B \cap C)$$

$$\begin{cases} A = \text{candidate } t \text{ has true rank } j \\ B = \text{no one better than that candidate appears at time } 1, \ldots, t-1 \\ \quad (\text{candidate } t \text{ is local best among first } t) \\ C = \text{the local best in times } 1, \ldots, t-1 \text{ falls in} \\ \quad \text{the observation phase (times } 1, \ldots, r) \end{cases}$$

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(C \mid A \cap B)\, \mathbb{P}(B \mid A)\, \mathbb{P}(A)$$

$$\mathbb{P}(A) = 1/n$$

$$\mathbb{P}(C \mid A \cap B) = r/(t-1)$$

$$\mathbb{P}(B \mid A) = \binom{n-t}{j-1} \Big/ \binom{n-1}{j-1}$$

$$= \sum_{j=1}^{k} \sum_{t=r+1}^{n} \left(\frac{r}{t-1}\right)\cdot\left(\frac{1}{n}\right)\cdot \frac{\binom{n-t}{j-1}}{\binom{n-1}{j-1}}$$

$$= \frac{r}{n} \sum_{t=r+1}^{n} \frac{1}{t-1}\left(\sum_{j=0}^{k-1} \frac{\binom{n-t}{j}}{\binom{n-1}{j}}\right).$$

Continuous (large-n) approximation:

Letting $x = \frac{r}{n}$ and $u = \frac{t}{n}$,

$$\frac{\binom{n-t}{j}}{\binom{n-1}{j}} = \frac{(n-t)!}{\cancel{j!}\,(n-t-j)!}\cdot\frac{\cancel{j!}\,(n-1-j)!}{(n-1)!} = \prod_{m=0}^{j-1} \frac{n-t-m}{n-1-m}$$

$$= \prod_{m=0}^{j-1} \frac{1-\frac{t}{n}-\frac{m}{n}}{1-\frac{1}{n}-\frac{m}{n}} = \prod_{m=0}^{j-1} \underbrace{\frac{1-u-\frac{m}{n}}{1-\frac{1}{n}-\frac{m}{n}}}_{\text{as } n\to\infty,\ \to 1-u}$$

$$\longrightarrow (1-u)^{j} \quad \text{as } n \to \infty$$

and $$\sum_{j=0}^{k-1} (1-u)^{j} = \frac{1-(1-u)^{k}}{1-(1-u)} = \frac{1-(1-u)^{k}}{u}$$

so $$P(K,x) = \frac{r}{n}\sum_{t=r+1}^{n} \frac{1}{t-1}\left(\sum_{j=0}^{K-1}\frac{\binom{n-t}{j}}{\binom{n-1}{j}}\right)$$

$$\approx x\int_{u=x}^{u=1} \frac{1}{u}\cdot\frac{1-(1-u)^k}{u}\,du$$

$$= x\int_{x}^{1} \frac{1-(1-u)^k}{u^2}\,du \qquad \text{where } x=\frac{r}{n}.$$

Exact solution:

$$\boxed{P(n,k,r) = \frac{r}{n}\sum_{t=r+1}^{n} \frac{1}{t-1}\left(\sum_{j=0}^{K-1}\frac{\binom{n-t}{j}}{\binom{n-1}{j}}\right)}$$

Approximate (large n) solution:

$$\boxed{P(K,X) = x\int_{x}^{1} \frac{1-(1-w)^k}{w^2}\,dw \qquad \text{where } x=\frac{r}{n}}$$

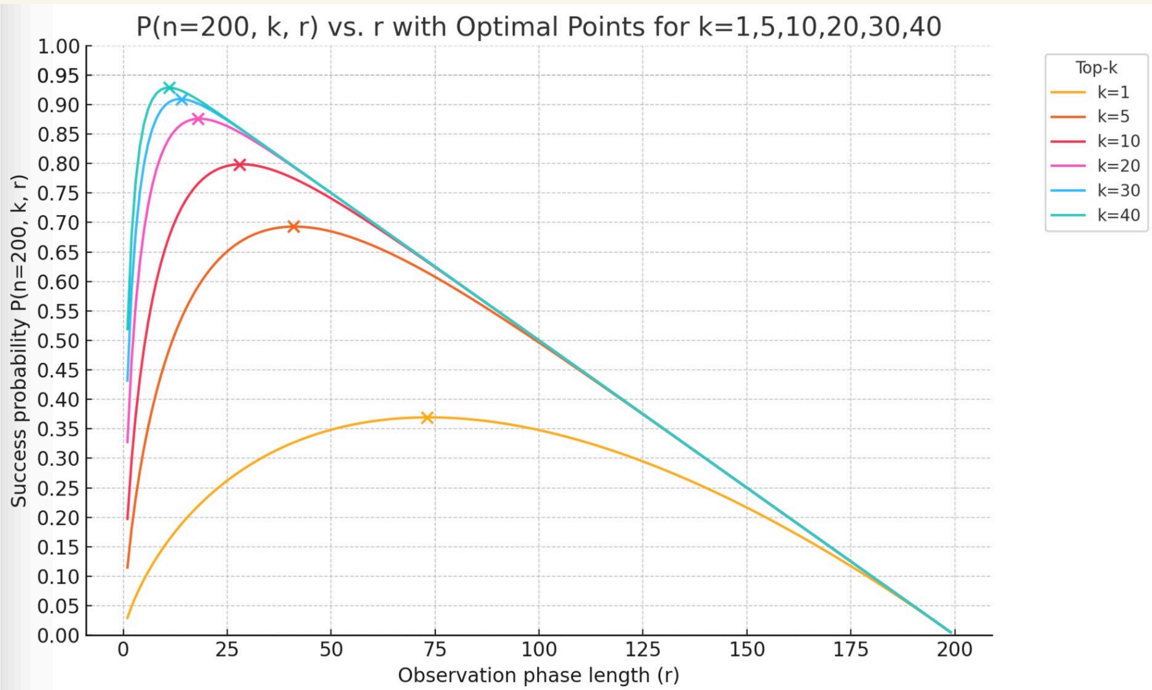
Results

I am nearly 28 years old and it would be nice to have a wife by 34.

Suppose I have the bandwidth to date about 3 new girls per month, on average.

That's about $n=216$ girls, say $n=200$ to make it Round.

Lets visualize $P(n=200, k, r)$.



Note the agreement with the solution to the classical secretary problem ($k=1$, ending up with the best girl): use threshold $r_* = \frac{n}{e} = \frac{200}{e} \approx 73$ which has success probability $\frac{1}{e} \approx 37\%$.

An observation period of $r \approx 30$ girls produces an $\approx 80\%$ chance of successfully attaining a top 5% (10/200) girl.

An observation period of $r \approx 15$ girls produces a $\approx 90\%$ chance of successfully attaining a top 15% (30/200) girl.

Not bad!

In Practice

You could fall in love during the observation period! That would be a success! So these are underestimates. $r = 22$ seems like a good choice.