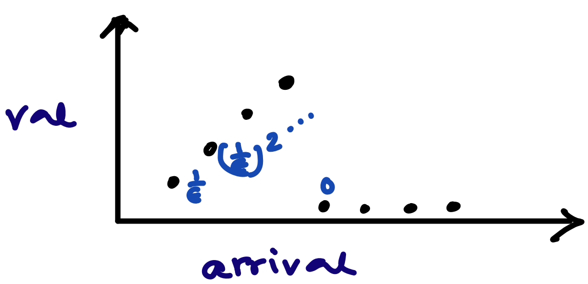
Secretary problem:

We saw in the class, online weighted matching & knapsack have unbounded C.R. under adversarial order arrival.

Even the following problem which is a special case of both the above problems is hopeless!

- A set of n candidates C. Candidate i has value v_i. → say, distinct
- Candidates arrive one by one;
  You either accept & stop;
  or reject & continue.
- Goal: Select the best candidate.
  (or even maximize the expected value of the accepted candidates).



if selected increases by $1/\epsilon$ factor, if rejected drops to 0.

So, we relax the order of arrival
- Secretary problem.

Candidates arrive according to a random permutation π, selected uniformly at random from the set of $n!$ possible permutations.

Idea:

Sample some initial candidates S to get an idea about the whole input I.

S too large — Best candidate can be in S (& we reject it)

S too small — may not get enough info.

- **Idea 1:**
  - Sample : Reject the first pn $(0 < p \leq 1)$ candidates [call it set S].
  - Exploit : Let s^* be the best in S. Then select the next appearing candidate with value $> s^*$.

$\mathbb{P}$[we select the best candidate]

$\geq \mathbb{P}$[the second best candidate is in S]

$\cdot\, \mathbb{P}$[the best candidate is in $I \setminus S$]

$$= \frac{pn}{n} \cdot \frac{(1-p)n}{n} = p(1-p).$$

By choosing $p = 1/2$, we get $\mathbb{P}[\text{success}] = 1/4$.

→ We improve further by more careful analysis.

Algorithm:

- Let S be the first $\lfloor n/e \rfloor$ randomly arriving candidates.
- Let $T := \max_{i \in S} val(i)$ be the best candi in S.
- For every arriving candi $c \in I \setminus S$, do:

  If $val(c) > T$ then accept c.

<u>Thm</u>: Let c^* be the best candidate in I.
Then $\mathbb{P}[c^* \text{ is selected}] \geq 1/e$.

<u>Proof</u>: $\mathbb{P}[c^* \text{ is selected}]$

$$= \sum_{i=\lceil n/e \rceil}^{n} \mathbb{P}[\underbrace{c^* \text{ appears on } i\text{th round}}_{A} \;\&\; \underbrace{\text{we accept in } i\text{th round}}_{B}]$$

Now, $\mathbb{P}[A \cap B]$

$= \mathbb{P}[B \mid A]\, \mathbb{P}[A]$ (Bayes' thm; $\mathbb{P}[A] \neq 0$)

Now, $\mathbb{P}[A] = 1/n$.

$\mathbb{P}[B \mid A]$

$= \mathbb{P}[\text{we accept in } i\text{'th round} \mid c^* \text{ appears on } i\text{th round}]$

$= \mathbb{P}[(i \geq \lceil n/e \rceil) \text{ and } (\text{no items in } \lceil n/e \rceil, \dots, i-1 \text{ had value} > T)]$

$= \mathbb{P}[\text{The best among prev. } (i-1) \text{ candidates appear in } S]$

$$= \frac{\lfloor n/e \rfloor}{i-1}.$$

Hence, $\mathbb{P}[C^* \text{ is selected}]$

$$= \sum_{i=\lceil n/e \rceil}^{n} \frac{1}{n} \cdot \frac{\lfloor n/e \rfloor}{i-1}$$

$$= \frac{\lfloor \frac{n}{e} \rfloor}{n} \sum_{i=\lfloor n/e \rfloor}^{n-1} \frac{1}{i} \geq \left(\frac{1}{e} - \frac{1}{n}\right) \cdot \ln\left(\frac{n}{\lfloor n/e \rfloor}\right)$$

$$\geq \left(\frac{1}{e} - \frac{1}{n}\right) \to 1/e \text{ as } n \to \infty.$$

Here, we used the fact

$$\sum_{i=\lfloor n/e \rfloor}^{n-1} 1/i \geq \int_{\lfloor n/e \rfloor}^{n} \frac{1}{x} dx = \ln\left(\frac{n}{\lfloor n/e \rfloor}\right). \qquad \blacksquare$$

$\rightarrow$ In a HW problem, we prove it to be tight.