PROBLEM IZBIRE NAJBOLJŠE TAJNICE

MATIJA VIDMAR

Fakulteta za matematiko in fiziko, Univerza v Ljubljani
Inštitut za matematiko, fiziko in mehaniko

Math. Subj. Class. (2010): 60G40, 62L15

V problemu izbire najboljše tajnice zaporedoma intervjuvamo za eno samo odprto delovno mesto $n \in \mathbb{N}_{\geq 2}$ kandidatov s ciljem, da izberemo najboljšega med njimi. O zavrnitvi ali sprejemu kandidata se moramo odločiti takoj po njegovem intervjuju. Za velike n je, v približku, optimalno zavrniti prvih n/e kandidatov in nato vzeti prvega, ki je boljši od vseh prejšnjih; najboljšega tako izberemo z verjetnostjo $1/e$. Problem ima natančno in eksplicitno rešitev za vse n.

SECRETARY PROBLEM

In the secretary problem we consecutively interview, for a single open position, $n \in \mathbb{N}_{\geq 2}$ applicants, with the goal of choosing the best. The decision of whether to reject or accept a given applicant must be made immediately following his interview. For large n, approximately, it is optimal to decline the first n/e candidates and then to hire the first one that is better than all his predecessors; the best being thus chosen with probability $1/e$. The problem admits a precise and explicit solution for arbitrary n.

Uvod in zamejitev problema

Pričnimo z natančnim opisom klasičnega problema izbire tajnice.

(i) Na voljo je eno prosto delovno mesto (tajnice).

(ii) Imamo n kandidatov, $n \in \mathbb{N}_{\geq 2}$ je znan *a priori* (je neslučajen). Nobena dva kandidata nista enako sposobna. Stopnja sposobnosti kandidatov nam ni znana *a priori*.

(iii) Kandidate intervjuvamo zaporedoma, pri čemer so vsi vrstni redi enako verjetni.

(iv) Po vsakem intervjuju lahko že intervjuvane kandidate linearno uredimo od najboljšega do najslabšega. Pravkar intervjuvanega kandidata takoj bodisi zavrnemo ali sprejmemo na delovno mesto in odločitve ne moremo več spremeniti. Na koncu ni nujno, da mesto zapolnimo.

(v) Odločitev o tem, ali kandidata sprejeti ali zavrniti, je odvisna samo od relativne razvrstitve kandidatov, ki so že bili intervjuvani.

(vi) Maksimizirati želimo verjetnost, da izberemo najboljšega kandidata.

Opisano je dobro znan problem iz teorije optimalnega ustavljanja, ležeč na preseku verjetnosti in optimizacije. Klasična monografija s tega področja je [1]. Natančneje ga uvrščamo na področje optimalnega ustavljanja verig Markova v diskretnem času, glej npr. [5, str. 135, primer 8.16], [7, str. 485]. Problem ima relativno kratko, a bogato zgodovino, katere objavljeni začetki segajo v drugo polovico 20. stoletja. Originalno avtorstvo ni jasno. Bralca, ki ga zanima natančnejši zgodovinski opis, napotimo npr. na pregledna članka [8, razdelek 3] in [9, razdelek 1.2]. V tem pogledu navedimo še, da je problem poznan tudi pod drugimi imeni: problem poroke oz. dote, problem lepotnega tekmovanja, problem najboljše izbire, če jih navedemo le nekaj. Njegova privlačnost je v tem, da ima preprosto, elementarno in eksplicitno rešitev. Zares imamo sledečo trditev, ki jo bomo dokazali v naslednjem razdelku.

Trditev 1. *Definirajmo*

$$r_n := \min\left\{r \in \{1,\dots,n-1\} : \sum_{i=r}^{n-1} \frac{1}{i} \leq 1\right\} \quad \textit{in}$$

$$P_n := \begin{cases} 1/2, & \textit{za } n = 2 \\ \dfrac{r_n - 1}{n} \displaystyle\sum_{i=r_n-1}^{n-1} \frac{1}{i}, & \textit{za } n \geq 3 \end{cases}.$$

Potem je

> *»Zavrni prvih $r_n - 1$ kandidatov in nato izberi prvega, ki je boljši od vseh svojih predhodnikov, če ta obstaja.«*

optimalna strategija za obravnavani problem; pri tem je verjetnost, da izberemo najboljšega kandidata, enaka P_n. Zaporedje $(P_n)_{n\in\mathbb{N}_{\geq 2}}$ pada proti in ima člene strogo večje od e^{-1}. Zaporedje $(\frac{r_n}{n})_{n\in\mathbb{N}_{\geq 2}}$ pa ima člene prav tako strogo večje od e^{-1} in konvergira proti e^{-1}. Na kratko: $e^{-1} < P_n \downarrow e^{-1}$ in $e^{-1} < \frac{r_n}{n} \to e^{-1}$.

Intuitivno je dokaj jasno, da P_n pada z naraščajočim n. Morda presenetljivo pa je, da verjetnost P_n ne pada proti nič (marveč proti e^{-1}).

Problem je bil študiran v številnih razširitvah in variantah: število kandidatov je lahko slučajno namesto deterministično; na voljo je lahko več kot eno mesto in posledično želimo izbrati nekaj najboljših kandidatov; kvalitete kandidatov so lahko merljive in njihove vrednosti (vendar ne pripadnost posameznemu kandidatu) poznane *a priori*; odločitve lahko z določenim stroškom spreminjamo za nazaj; maksimiziramo lahko kvaliteto izbranega kandidata, kadar je ta merljiva; intervjuvamo v zveznem času itd.: glej članka [8, razdelki 4–7] in [9, razdelki 2–9] za nadrobnejši opis (v ruščini je dostopna tudi knjiga [4], ki je v celoti posvečena problemu tajnice in njegovim variacijam). Nekaj zanimivih verzij bo bralec našel v člankih [10, 11, 13, 6, 2, 14].

Posebej zabavna je primerjava s problemom podoktoranda, v katerem želimo, *ceteris paribus*, izbrati drugega najboljšega kandidata (»najboljši bo šel na univerzo Harvard«). Naj bo

$$k_n := [n/2] = \begin{cases} n/2, & \text{če je } n \text{ sod} \\ (n-1)/2, & \text{če je } n \text{ lih} \end{cases}.$$

Optimalna strategija v tem primeru je:

> »Zavrni prvih k_n kandidatov in nato izberi prvega, ki je drugi najboljši od vseh svojih predhodnikov, če ta obstaja.«

Pri tem je verjetnost, da izberemo drugega najboljšega kandidata, enaka $\frac{k_n(n-k_n)}{n(n-1)} \approx 1/4$. Lažje je izbrati najboljšega kot drugega najboljšega kandidata! Za dokaz glej [15] ali [13].

Rešitev problema izbire najboljše tajnice in njena analiza

Začnimo z nekaj uvodnimi opažanji in dogovori.

• Ker je odločitev, ali izbrati kandidata (in ustaviti intervjuje) ali nadaljevati intervjuje, odvisna samo od relativnega ranga kandidatov, ki so že bili intervjuvani, je različnih možnih pravil ustavljanja (tj. pravil izbire kandidata za odprto delovno mesto) končno mnogo. V posebnem optimalno pravilo obstaja.

• Za intervjuvanega kandidata bomo rekli, da je relativno najboljši, če je boljši od vseh kandidatov, ki so bili intervjuvani pred njim. Kadar pa bomo rekli, da je kandidat najboljši, bomo imeli v mislih, da je on najboljši izmed sploh vseh kandidatov. Razvidno je, da lahko v optimalni strategiji izbiramo vedno samo relativno najboljše kandidate.

• Denimo, da je j-ti kandidat relativno najboljši. Potem bomo v optimalni strategiji le-tega izbrali, če bo, pogojno glede na informacijo, ki smo jo prejeli do vključno njegovega intervjuja, verjetnost, da je on najboljši – označimo jo s P_j – vsaj tako velika kot verjetnost – označimo jo s Q_j –, da izberemo najboljšega kandidata, če z intervjuji nadaljujemo (in izbiramo optimalno). Sledi, da obstaja optimalna strategija oblike:

> »Izberi kandidata j, če je ta relativno najboljši in je $P_j \geq Q_j$, sicer nadaljuj.«

- Končno trdimo, da obstaja optimalno pravilo oblike:

> S(r): »Zavrni prvih $r-1$ kandidatov in nato izberi prvega relativno najboljšega, če ta obstaja.«

za neki $r \in \{1, \ldots, n\}$. (Vsa možna pravila ustavljanja seveda niso take oblike. Za $i \in \{1, \ldots, n\}$ lahko npr. vedno izberemo po vrsti i-tega ali pa i-tega relativno najboljšega kandidata.)

Res. Po eni strani lahko izrazimo

$$P_j = \mathbb{P}(j\text{-ti kandidat je najboljši}|j\text{-ti kandidat je relativno najboljši})$$

$$= \frac{\mathbb{P}(j\text{-ti kandidat je najboljši})}{\mathbb{P}(j\text{-ti kandidat je relativno najboljši})} = \frac{1/n}{1/j} = j/n.$$

Po drugi strani je Q_j preprosto enak optimalni (tj. maksimalni, uporabljajoč najboljšo strategijo) verjetnosti, da izberemo najboljšega kandidata, če se že na začetku omejimo na strategije, v katerih izbiramo samo med kandidati z zaporednimi številkami $j+1, \ldots, n$: namreč če nam za $i \in \{1, \ldots, n\}$ slučajna spremenljivka X_i pove relativni rang i-tega kandidata med prvimi i kandidati, potem je (ker so vsi vrstni redi enako verjetni) porazdelitev $X_{j+1}, \ldots, X_n$, pogojno na $X_1, \ldots, X_j$, enaka brezpogojni porazdelitvi $X_{j+1}, \ldots, X_n$.

V posebnem vidimo, da sta P_j in Q_j odvisna samo od j in n in gotovo je $Q_{j+1} \leq Q_j$ za $j < n$. Naj bo še $S := \{j \in \{1, \ldots, n\} : P_j \geq Q_j\}$. Če je $j \in S$ in je $j < n$, potem je $Q_j \leq P_j = j/n$ in zato toliko bolj $Q_{j+1} \leq Q_j \leq j/n < (j+1)/n = P_{j+1}$, torej $j+1 \in S$. Sledi, da je $S = \{r, \ldots, n\}$ za neki $r \in \{1, \ldots, n+1\}$. Primer, ko je $r = n+1$, torej strategija ne-izbire sploh kateregakoli kandidata, je očitno strogo suboptimalna (saj npr. že strategija »izberi prvega kandidata« da strogo pozitivno verjetnost izbire najboljšega kandidata).

Do sedaj smo torej identificirali relativno preprost enoparametričen razred $\{\mathrm{S}(r) : r \in \{1, \ldots, n\}\}$ pravil ustavljanja, v katerem leži tudi (neko, morda jih je več) optimalno pravilo. Pot naprej je jasna: določiti verjetnost, označili jo bomo s P_r^n, da zmagamo (tj. da izberemo najboljšega kandidata) uporabljajoč pravilo S(r), in poiskati maksimum P_r^n za $r \in \{1, \ldots, n\}$.

V ta namen naj bo $r \in \{1, \dots, n\}$ in denimo, da zasledujemo strategijo $\mathrm{S}(r)$. Očitno je $P_1^n = \mathbb{P}(\text{prvi kandidat je najboljši}) = 1/n$. Naj bo $r > 1$. Dogodek, da zmagamo, razpade na disjunktno unijo dogodkov, da izberemo i-tega kandidata po vrsti in da je ta najboljši; seveda izbiramo samo kandidate z indeksi $i \in \{r, \dots, n\}$. Iz končne aditivnosti verjetnosti sledi

$$P_r^n = \sum_{i=r}^{n} \mathbb{P}(\text{izberemo } i\text{-tega kandidata in ta je najboljši}).$$

Denimo, da je i-ti kandidat najboljši za neki $i \in \{r, \dots, n\}$. Potem je, na dogodku, da je i-ti kandidat najboljši, to, da ga izberemo, istovetno s tem, da noben izmed kandidatov z indeksi $r, \dots, i-1$ ni bil izbran, tj. da je najboljši izmed prvih $i-1$ kandidatov med prvimi $r-1$ kandidati. Dobimo, da je

$$\begin{aligned}
P_r^n &= \sum_{i=r}^{n} \mathbb{P}(i\text{-ti kandidat je najboljši in najboljši izmed prvih} \\
&\qquad\qquad i-1 \text{ kandidatov je med prvimi } r-1 \text{ kandidati}) \\
&= \sum_{i=r}^{n} \mathbb{P}(\text{najboljši izmed prvih } i-1 \text{ kandidatov je med prvimi} \\
&\qquad\qquad r-1 \text{ kandidati} | i\text{-ti kandidat je najboljši}) \cdot \\
&\qquad\qquad \cdot \mathbb{P}(i\text{-ti kandidat je najboljši}) \\
&= \sum_{i=r}^{n} \frac{r-1}{i-1} \cdot \frac{1}{n} = \frac{r-1}{n} \sum_{i=r}^{n} \frac{1}{i-1} = \frac{r-1}{n} \sum_{i=r-1}^{n-1} \frac{1}{i}.
\end{aligned}$$

Torej je

$$P_r^n = \begin{cases} \frac{r-1}{n} \sum_{i=r-1}^{n-1} \frac{1}{i}, & \text{za } r \in \{2, \dots, n\} \\ 1/n, & \text{za } r = 1 \end{cases}.$$

Raziščimo odvisnost P_r^n od r. Naj bo $\diamond \in \{>, =\}$. Potem je za vse $r \in \{2, \dots, n-1\}$:

$$\begin{aligned}
P_{r+1}^n \diamond P_r^n &\Leftrightarrow \frac{r}{n} \sum_{i=r}^{n-1} \frac{1}{i} \diamond \frac{r-1}{n} \sum_{i=r-1}^{n-1} \frac{1}{i} \\
&\Leftrightarrow (r-1+1) \sum_{i=r}^{n-1} \frac{1}{i} \diamond \left(1 + (r-1) \sum_{i=r}^{n-1} \frac{1}{i}\right) \\
&\Leftrightarrow \sum_{i=r}^{n-1} \frac{1}{i} \diamond 1.
\end{aligned}$$

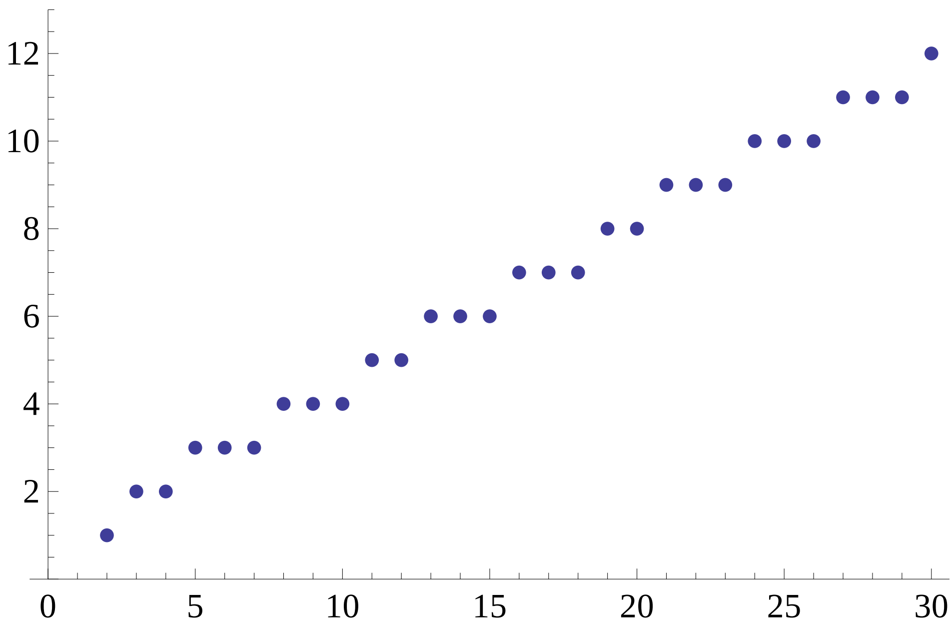
Ekvivalenca $P^n_{r+1} \diamond P^n_r \Leftrightarrow \sum_{i=r}^{n-1} \frac{1}{i} \diamond 1$ velja tudi pri $r = 1$, saj je $\frac{1}{n} \sum_{i=1}^{n-1} \frac{1}{i} \diamond \frac{1}{n} \Leftrightarrow \sum_{i=1}^{n-1} \frac{1}{i} \diamond 1$. Če povzamemo: za vse $r \in \{1, \dots, n-1\}$ velja $P^n_{r+1} > P^n_r$, če in samo če je $\sum_{i=r}^{n-1} \frac{1}{i} > 1$, in velja $P^n_{r+1} = P^n_r$, če in samo če je $\sum_{i=r}^{n-1} \frac{1}{i} = 1$.

Sedaj imamo, kot sledi. Za $n = 2$ sta optimalna r dva, $1 = r_2$ in 2, verjetnost zmage je $1/2$. Za $n > 2$ vsota

$$S_{n,r} := \sum_{i=r}^{n-1} \frac{1}{i}$$

ni enaka 1 za noben $r \in \{1, \dots, n-1\}$ (segmenti harmonične vrste se ne seštejejo v ena, razen prvega člena: glej [12, str. 157], [3]), strogo pada v r, je strogo večja od 1 za $r = 1$ in je enaka $\frac{1}{n-1}$ (torej strogo manjša od 1) za $r = n - 1$. Sledi, da je edini optimalen r enak

$$\min \left\{ r \in \{1, \dots, n-1\} : \sum_{i=r}^{n-1} \frac{1}{i} < 1 \right\} =$$
$$\min \left\{ r \in \{1, \dots, n-1\} : \sum_{i=r}^{n-1} \frac{1}{i} \leq 1 \right\} = r_n.$$



Slika 1. Optimalen r_n kot funkcija n: raste, vsakič največ za 1.

Prvih nekaj vrednosti r_n s pripadajočimi optimalnimi verjetnostmi $P^n_{r_n} = P_n$ je zbranih v spodnji tabeli.

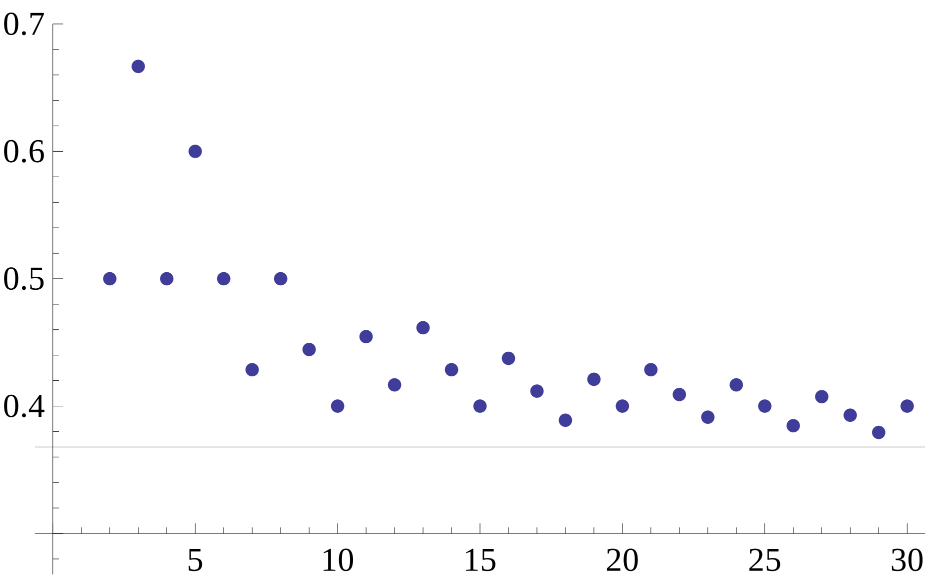
n	2	3	4	5	6	7	8	9	10
r_n	1	2	2	3	3	3	4	4	4
P_n	0,500	0,500	0,458	0,433	0,428	0,414	0,410	0,406	0,399

Končno raziščimo nekaj lastnosti optimalne rešitve in tako zaključimo dokaz trditve iz uvoda.

(A) Funkcija $\mathbb{N}_{\geq 2} \ni n \mapsto r_n$ narašča z n, vsakič največ za 1 (tj. $0 \leq r_{n+1} - r_n \leq 1$ za vse $n \in \mathbb{N}$). To sledi neposredno iz definicije r_n: vsota $S_{n,r}$ narašča v n in $S_{n+1,r+1} \leq S_{n,r}$.

(B) Funkcija $\mathbb{N}_{\geq 2} \ni n \mapsto P_n = P^n_{r_n}$ pada, strogo na $\mathbb{N}_{\geq 3}$. Naj bo $n \in \mathbb{N}$ in označimo $r := r_n$. Velja $P_2 = 1/2 = P_3$; preveriti moramo še, da je $P_n > P_{n+1}$ za $n \geq 3$. V slednjem primeru je $r > 1$. Iz (A): $r_{n+1} = r$ ali pa $r_{n+1} = r + 1$. Naj bo najprej $r_{n+1} = r$. Pokazati moramo, da je $\frac{r-1}{n} \sum_{i=r-1}^{n-1} \frac{1}{i} > \frac{r-1}{n+1} \sum_{i=r-1}^{n} \frac{1}{i}$, kar sledi iz

$$\frac{1}{n} \sum_{i=r-1}^{n-1} \frac{1}{i} > \frac{1}{n+1} \left(\sum_{i=r-1}^{n-1} \frac{1}{i} + \frac{1}{n} \right) \Leftarrow \sum_{i=r-1}^{n-1} \frac{1}{i} > 1,$$



Slika 2. Optimalno razmerje r_n/n kot funkcija n: konvergira k, in je strogo nad, e^{-1} (horizontalna linija).

in to je res po definiciji r_n. Naj bo sedaj $r_{n+1} = r + 1$. Pokazati moramo, da je $\frac{r-1}{n} \sum_{i=r-1}^{n-1} \frac{1}{i} > \frac{r}{n+1} \sum_{i=r}^{n} \frac{1}{i}$, kar sledi iz

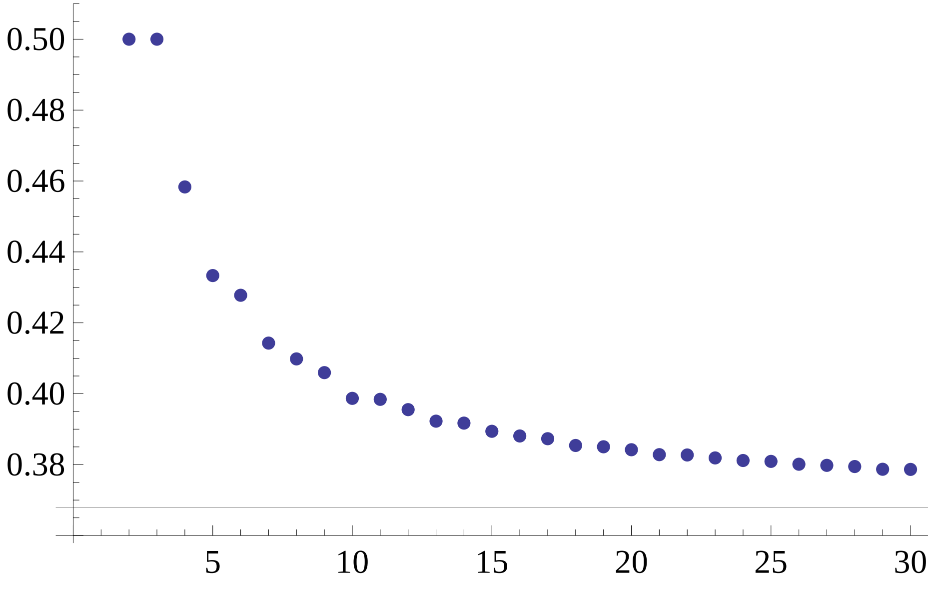
$$\frac{r-1}{n}\left(\frac{1}{r-1} + S_{n,r}\right) > \frac{r}{n+1}\left(S_{n,r} + \frac{1}{n}\right)$$
$$\Leftarrow \left(\frac{1}{n} - \frac{r}{n\,(n+1)}\right) > S_{n,r}\left(\frac{r}{n+1} - \frac{r-1}{n}\right) \Leftarrow 1 > S_{n,r},$$

in to je res po definiciji r_n.

(C) Imamo: $e^{-1} < r_n/n \overset{n\to\infty}{\longrightarrow} e^{-1}$. Res, vsoto $S_{n,r} = \sum_{i=r}^{n-1} \frac{1}{i}$ moremo za $r \in \{1, \ldots, n-1\}$ takole oceniti:

$$\ln\left(\frac{n}{r}\right) = \int_r^n \frac{dx}{x} \le S_{n,r} \le \int_r^n \frac{dx}{x-1} = \ln\left(\frac{n-1}{r-1}\right),$$

kjer razumemo $\ln(a/0) = \infty$ za $a > 0$. Ker po definiciji r_n velja $S_{n,r_n} \le 1$, sledi iz zgornje ocene, da je $\ln\left(\frac{n}{r_n}\right) \le 1$, torej $n/e \le r_n$. Po drugi strani za $r := \lceil \frac{n-1}{e} + 1 \rceil$ (tu je za $x \in \mathbb{R}$, $\lceil x \rceil := \min\{k \in \mathbb{Z} : k \ge x\}$) velja $\ln\left(\frac{n-1}{r-1}\right) \le 1$ in zato iz zgornje ocene (če je le $r \le n-1$) $S_{n,r} \le 1$, od koder, spet po definiciji r_n sledi r_n, $r_n \le \lceil \frac{n-1}{e} + 1 \rceil$. Skratka, $n/e \le r_n \le \lceil \frac{n-1}{e} + 1 \rceil$



Slika 3. Optimalna verjetnost $P^n_{r_n}$ kot funkcija n: strogo pada na $\mathbb{N}_{\ge 3}$ proti e^{-1} (horizontalna linija).

in želena konvergenca optimalnega razmerja r_n/n sledi. Stroga neenakost $e^{-1} < r_n/n$ velja preprosto zaradi iracionalnosti e.

(D) $e^{-1} < P_n \overset{n\to\infty}{\longrightarrow} e^{-1}$. Konvergenca sledi iz (C), ker za $n > 2$ velja $P^n_{r_n} = \frac{r_n - 1}{n} \sum_{i=r_n-1}^{n-1} \frac{1}{i}$ in ker po definiciji r_n velja $\sum_{i=r_n-1}^{n-1} \frac{1}{i} \to 1$ (saj $r_n \to \infty$ po (C)), ko gre $n \to \infty$. Da je $e^{-1} < P_n$, lahko potem ugotovimo iz (B).

Rezultate si naposled predstavimo še grafično: glej slike 1, 2 in 3.

LITERATURA

[1] A. B. Aries in A. N. Shiryaev, *Optimal stopping rules*, Stochastic Modelling and Applied Probability, Springer Berlin Heidelberg, 2007.

[2] J. N. Bearden, *A new secretary problem with rank-based selection and cardinal payoffs*, Journal of Mathematical Psychology **50** (2006), 1, 58–59.

[3] H. Belbachir in A. Khelladi, *On a sum involving powers of reciprocals of an arithmetical progression*, Annales Mathematicae et Informaticae **34** (2007), 29–31.

[4] B. A. Berezovskiy in A. V. Gnedin, *Problemi najboljše izbire (v ruščini)*, Akademia Nauk USSR, 1984.

[5] P. Billingsley, *Probability and measure*, Wiley Series in Probability and Statistics, Wiley, 2012.

[6] F. R. Bruss, *Sum the odds to one and stop*, The Annals of Probability **28** (2000), 3, 1384–1391.

[7] E. B. Dynkin, *The optimum choice of the instant for stopping a Markov process*, Selected Papers of E. B. Dynkin with Commentary (A. A. Yushkevich, G. M. Seitz, in A. L. Onishchik, ur.), CWorks / American Mathematical Society, American Mathematical Society, 2000.

[8] T. S. Ferguson, *Who solved the secretary problem?*, Statistical Science **4** (1989), 3, 282–289.

[9] P. R. Freeman, *The secretary problem and its extensions: A review*, International Statistical review **51** (1983), 2, 189–206.

[10] J. P. Gilbert in F. Mosteller, *Recognizing the maximum of a sequence*, Journal of the American Statistical Association **61** (1966), 313, 35–73.

[11] M. Hlynka in J. N. Sheahan, *The secretary problem for a random walk*, Stochastic Processes and their Applications **28** (1988), 2, 317–325.

[12] P. Hoffman, *The man who loved only numbers: The story of Paul Erdos and the search for mathematical truth*, Hyperion Books, 1998.

[13] J. S. Rose, *A problem of optimal choice and assignment*, Operations Research **30** (1982), 1, 172–181.

[14] D. A. Sardelis in T. M. Valahas, *Decision making: A golden rule*, The American Mathematical Monthly **106** (1999), 3, 215–226.

[15] R. J. Vanderbei, *The postdoc variant of the secretary problem*, 2012, dostopno na: `www.princeton.edu/~rvdb/tex/PostdocProblem/PostdocProb.pdf`, ogled: 5. 2. 2017.