

Koliko stane zbiranje sličic?

↓↓↓
BOŠTJAN KUZMAN

→ Nogometna mrzlica svetovnega prvenstva v Rusiji 2018 se je zdaj že povsem polegla, nekateri navdušenci pa morda še vedno zbirajo in menjavajo zadnje manjkajoče sličice s podobami igralcev nogometnih reprezentanc. V albumu, ki je bil naprodaj v Sloveniji, je prostora za 600 različnih sličic. Te so se prodajale v paketkih s po 3 sličicami s ceno 0,70 EUR za paket. Zlahka torej izračunamo, da je treba za poln album kupiti vsaj 200 paketkov, kar nas stane najmanj 140 EUR (oz. skoraj 7 let naročnine na Presek). Seveda pa bo končni strošek višji, ker so paketki in sličice v njih bolj ali manj naključno premešane, vendar ga lahko po drugi strani zmanjšamo z menjavanjem sličic z drugimi zbiralci.

Preden nadaljujete z branjem, poskusite podati svojo grobo oceno, koliko naključnih sličic je treba kupiti, da bi brez menjavanja zbrali vseh 600 različnih. Je to morda 1000, 2000, 3000 ali celo več sličic? Koliko denarja bi torej tako porabili za poln album?

Ocena povprečnega števila kupljenih sličic, ki jih potrebujemo za napolnitev albuma brez menjavanja, predstavlja matematični problem s področja verjetnosti, ki je v angleščini znan tudi pod imenom *Coupon collector's problem* [1]. Preden se lotimo njegevega reševanja po teoretični poti, lahko oceno poskusimo dobiti eksperimentalno. Bralke in bralci, ki poznate osnove programiranja, boste zlahka napisali ustrezno programsko simulacijo poskusa, pri kateri program naključno izbira števila med 1 in 600, dokler ne zbere vseh različnih vrednosti.

Za šolsko rabo pa je zanimiva tudi poenostavljena verzija eksperimenta zbiranja 6 različnih sličic, ki ga lahko izvedemo s pomočjo običajne igralne kocke. Igralno kocko vržemo večkrat zapored in si zapisu-

Št.	Zaporedje padlih vrednosti	Št. metov	T_1	T_2	T_3	T_4	T_5	T_6
1.	66343143353153312	17	1	2	1	2	4	7
2.	5424113356	10	1	1	1	2	2	3
3.	4666633142664335	16	1	1	4	2	2	6
4.	6163452	7	1	1	2	1	1	1
5.	2152221535614	13	1	1	1	6	2	2
6.	12425444224663	15	1	1	1	2	8	2
7.	626662463652565332261	21	1	1	5	2	2	10
8.	662563163364	12	1	2	1	2	1	5
9.	133664442262616231133246265	27	1	1	2	2	3	18
10.	4211243624625	13	1	1	1	4	1	5
Povprečno število metov		15,1	1,0	1,2	1,9	2,5	2,6	5,9

TABELA 1.
Rezultati desetih ponovitev poskusa s kocko. Z modro je obarvana prva pojavitev posamezne vrednosti v zaporedju.

jemo zaporedje padlih vrednosti. Ko se v zaporedju pojavi vseh 6 različnih vrednosti, preštejemo število potrebnih metov in poskus ponovimo. Pri dovolj velikem številu ponovitev bo aritmetična sredina meritev predstavljala oceno iskanega povprečja za album s 6 sličicami. Pri 10 ponovitvah poskusa, kot so zapisane v tabeli, smo denimo dobili povprečno vrednost $151/10$ oz. 15,1.

Pa si oglejmo, kako bi povprečno število potrebnih metov izračunali matematično. Pri posameznem metu poštene kocke je matematična verjetnost šestice enaka $p = 1/6$, saj je ugoden le eden od šestih enako verjetnih izidov. Statistično to pomeni, da bo pri velikem številu metov delež padlih šestic enak $1/6$, ali, povedano drugače, v povprečju bo šestica padla vsak šesti met. Torej bo od začetka poskusa do prve šestice potrebnih v povprečju 6 metov. **Podobno velja tudi pri neodvisnih zaporednih ponovitvah poskusa, pri katerem se ugoden dogodek zgodi z verjetnostjo $p > 0$. Povprečno število ponovitev poskusa do prvega uspeha je tedaj enako $1/p$.**

Od te ugotovitve ni več daleč do izračuna povprečnega števila potrebnih metov, s katerimi bi zbrali vse različne vrednosti. Razmišljamo namreč takole. Prvo od šestih vrednosti bomo zagotovo dobili v prvem metu. Ko eno vrednost že imamo, bo v naslednjem metu z verjetnostjo $5/6$ padla ena od vrednosti, ki je še nimamo, torej bomo za drugo vrednost v povprečju potrebovali še $6/5 = 1,2$ meta kocke. Ob tem smo seveda upoštevali, da so zaporedni meti med seboj neodvisni, torej predhodni meti na naslednjega ne vplivajo. Zdaj imamo že dve različni vrednosti, zato bo naslednja manjkajoča padla z verjetnostjo $4/6$ v posameznem metu, oz. v povprečju po $6/4 = 1,5$ meta, in tako dalje. Za zadnjo manjkajočo vrednost bomo potrebovali v povprečju kar 6 metov, saj je njena verjetnost v posameznem metu le $1/6$. Na ta način pridemo do števila

$$\blacksquare 1 + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = 6 \left(\frac{1}{6} + \frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} + \frac{1}{1} \right) = 14,7.$$

To število predstavlja povprečno število metov oz. kupljenih sličic pri velikem številu ponovitev zbiranja šestih različnih vrednosti. Izračunana vrednost se približno ujema tudi z našo eksperimentalno meritvijo. Prav tako se vrednosti členov v vsoti pribli-

žno ujemajo z izmerjenimi povprečnimi vrednostmi števil T_i iz tabele 1, ki pomenijo število metov od $(i - 1)$ do i -te nove vrednosti.



SLIKA 1.

Za zadnjo manjkajočo vrednost je v povprečju potrebnih kar 6 metov od skupnega povprečja 14,7 meta.

Morda bi kdo iz tega prenašlo sklepal, da je za album s 600 različnimi sličicami treba kupiti okoli 1470 sličic. To seveda ne velja, povprečno število kupljenih sličic je v tem primeru precej večje. S posnemanjem gornjega izračuna lahko ustrezni izraz zdaj zlahka zapišemo in s pomočjo računalnika tudi seštejemo:

$$\blacksquare 600 \left(\frac{1}{600} + \frac{1}{599} + \frac{1}{598} + \dots + \frac{1}{1} \right) \doteq 4184,987.$$

Če zaokrožimo navzgor, brez menjavanja sličic bi morali torej za napolnitev albuma s 600 sličicami v povprečju kupiti 4185 sličic. To pomeni 1395 paketkov, za katere bi odšteli skupaj kar 976,50 EUR. Precej denarja, kaj ne?

Zadnji izračun velja ob predpostavkah, da so paketi in sličice v njih enakomerno in neodvisno porazdeljeni, torej, da so pri nakupu nove sličice vse številke enako verjetne ne glede na predhodno kupljene sličice ali druge sličice v istem paketku. Če so denimo paketi treh sličic oblikovani tako, da v njih ni podvajanja sličic, potem zadnja domneva ne velja. Številni zbiralci tudi sumijo, da so nekatere sličice redkejšje od drugih, torej, da porazdelitev ni enakomerna, a v to se ne bomo spuščali.





Če torej zaupamo proizvajalcu, da natisne enako število vseh različnih sličic, ki so dobro premešane in neodvisno razporejene v paketke na prodajnih mestih, potem je splošna rešitev problema zbiralca sličic pri zbiranju n različnih sličic ob zgornjih predpostavkah enaka nH_n , kjer je

$$H_n = \frac{1}{1} + \frac{1}{2} + \dots + \frac{1}{n}$$

t. i. n -to harmonično število. Kot je razvidno iz nekaj vrednosti v tabeli, naraščanje števila nH_n v odvisnosti od n ni linearno. Bralcem z nekaj več matematičnega znanja omenimo, da je število nH_n velikostnega reda $O(n \log n)$, kar je posledica znanih lastnosti harmoničnih števil.

n	6	60	600	6000
nH_n	14,70	280,80	4185,99	55660,88

TABELA 2.

Nekaj vrednosti števila nH_n

Pozorni bralec je iz dosedanje razprave verjetno razbral, da je nadražje predvsem zbiranje zadnje peščice manjkajočih sličic. Samo za pridobitev zadnje manjkajoče sličice v albumu za nogometno prvenstvo 2018 bi jih morali v povprečju kupiti še 600, kar bi stalo 140 EUR. Zato je zanimivo omeniti, da založnik albuma omogoča, da zbiralec neposredno naroči 100 izbranih sličic po ceni 0,25 EUR za sličico. Za zadnjih 100 sličic bi sicer z naključnim kupovanjem potrebovali v povprečju

$$600H_{100} = 600 \left(\frac{1}{100} + \frac{1}{99} + \dots + \frac{1}{1} \right) \doteq 3112,43$$

naključnih sličic, oz. zaokroženo navzgor, 1038 paketkov, kar zneso 726,60 EUR. Z neposrednim naročilom zadnjih 100 manjkajočih sličic tako zbiralec v povprečju prihrani več kot 700 EUR in za napolnitev albuma porabi »le« 274,90 EUR (brez stroška pošiljanja sličic). Že samo z neposrednim naročilom zadnjih 20 manjkajočih sličic pa bi celotni strošek zmanjšali za približno polovico, kar lahko bralec premisli sam.

Pa si postavimo še nekoliko drugačno vprašanje. Denimo, da imate 210 EUR, ki jih boste porabili za nakup 300 paketkov s skupaj 900 sličicami. Koliko različnih sličic za vaš album lahko v tem primeru pričakujete?

Zdaj je razmislek nekoliko drugačen. Verjetnost, da naključno kupljena sličica ni sličica številka 1, je enaka $\frac{599}{600}$. Verjetnost, da nobena od 900 kupljenih sličic ni sličica številka 1, pa je enaka $\left(\frac{599}{600}\right)^{900} \doteq 0,22313$. Zato je verjetnost, da je med 900 kupljenimi tudi sličica številka 1, približno enaka 0,77687, kar lahko tolmačimo tudi kot povprečno število sličic številka 1 med 900 naključno kupljenimi. Da bi dobili povprečno število vseh različnih sličic med njimi, pa lahko seštejemo povprečno število sličic številka 1, številka 2, in tako dalje vse do 600. Zato izraz

$$600 \cdot 0,77687 \doteq 466,12$$

predstavlja povprečno število različnih sličic med 900 naključno kupljenimi izmed skupaj 600 različnih. Za menjavo nam v tem primeru torej ostane v povprečju 434 sličic, s katerimi bomo ob dovolj velikem številu drugih zbiralcev v sosesčini najbrž brez težav zbrali manjkajočih 134 za poln album.

V vsakem primeru pa je zbiranje sličic precej drag konjiček. Radovedna bralka in bralec lahko svoje razumevanje dosedanjih izračunov preizkusita z naslednjo nalogo.

Naloga. Izračunaj in primerjaj povprečni števili različnih sličic, ki jih pri prej navedenih cenah dobiš za 175 EUR, če:

- vseh 175 EUR porabiš za nakup paketkov,
- porabiš 140 EUR za nakup paketkov, od preostanka pa porabiš 25 EUR za neposredno naročilo 100 manjkajočih sličic in 10 EUR za poštnino.

Rešitev. V prvem primeru kupiš 750 sličic, med katerimi bo v povprečju približno 428 različnih. V drugem primeru pa kupiš 700 sličic, med katerimi bo v povprečju približno 479 različnih.

Literatura

- [1] Coupon collector's Problem, dostopno na en.wikipedia.org/wiki/Coupon_collector%27s_problem, ogled 14. 1. 2019
- [2] Harmonic number, dostopno na en.wikipedia.org/wiki/Harmonic_number, ogled 14. 1. 2019.

× × ×