

ZADACI I PROBLEMI

Pet zadataka iz vjerojatnosti

ZVONIMIR ŠIKIĆ, Zagreb

1. Koliko puta, u prosjeku, morate baciti kocku da biste dobili 6 ?
2. U svakoj kutiji deterdženta *DEX* nalazi se jedan od kupona označenih brojevima 1, 2, 3, 4 ili 5. Kada sakupite svih pet različitih kupona, dobivate nagradu. Koliko *DEX*-a, u prosjeku, trebate kupiti da biste dobili nagradu?
3. Uzastopnim bacanjem kovanica dobivate blokove glava i blokove pisama različitih duljina. Kolika je prosječna duljina bloka?
4. Samuel Pepys zamolio je Isaaca Newtona da mu izračuna koji je od sljedeća tri događaja vjerojatniji: barem 1 šestica u 6 bacanja kocke, barem 2 šestice u 12 bacanja kocke, barem 3 šestice u 18 bacanja kocke? Što je vaš odgovor?
5. Koja je najmanja skupina ljudi u kojoj je vjerojatnost da 2 osobe rođendan imaju istog datuma (godine se ne moraju poklapati) veća od 50 % ?

ZADACI I PROBLEMI

Rješenja pet zadataka iz vjerojatnosti

ZVONIMIR ŠIKIĆ

U Poučku br. 21 objavili smo 5 zadataka iz vjerojatnosti koje ste, nadamo se, uspješno riješili. Ovdje nudimo naša rješenja u koja smo upleli dosta zanimljive matematike (geometrijski red, Eulerovu konstantu γ , beskonačno protegnutu površinu Nikole Oresma, razvoj od e^{-x} u red potencija itd.). Ukoliko su vaša rješenja zanimljiva i bitno različita, javite nam.

1. Većina ljudi nekako osjeća da je odgovor 6. To je jedan od rijetkih slučajeva u kojima nas vjerojatnosni osjećaj ne vara.

Mi ćemo dokazati nešto općenitiji rezultat. Pretpostavimo da je vjerojatnost željenog ishoda nekog eksperimenta p . Tada neželjeni ishod ima vjerojatnost $q = 1 - p$. Vjerojatnost da se željeni ishod poluči tek u n -tom ponavljanju međusobno nezavisnih eksperimenata je $P_n = pq^{n-1}$ ($n - 1$ puta mora se zbiti neželjeni ishod čija je vjerojatnost q , a zatim se jedanput mora zbiti željeni ishod čija je vjerojatnost p). Očekivana vrijednost potrebnog broja pokušaja zato je

$$E = 1P_1 + 2P_2 + \dots = \sum_{n=1}^{\infty} nP_n = p + 2pq + 3pq^2 + \dots = \sum_{n=1}^{\infty} npq^{n-1}.$$

Očekivanu vrijednost E možemo izračunati metodom kojom se obično računa suma geometrijskog reda:

$$\begin{aligned} E &= p + 2pq + 3pq^2 + 4pq^3 + 5pq^4 + \dots \\ qE &= pq + 2pq^2 + 3pq^3 + 4pq^4 + \dots \\ E - qE &= p + pq + pq^2 + pq^3 + pq^4 + \dots \end{aligned}$$

Primjenom formule za sumu geometrijskog reda sada lako nalazimo:

$$E(1 - q) = p(1 + q + q^2 + q^3 + q^4 + \dots) = \frac{p}{1 - q} = \frac{p}{p} = 1.$$

Dakle, $E = 1/(1 - q) = 1/p$. To znači da, u prosjeku, eksperiment moramo ponoviti $1/p$ puta da bismo došli do željenog ishoda. Ako je $p = 1/6$, kao u našem zadatku, eksperiment trebamo provjeriti prosječno 6 puta.

Isti problem možemo riješiti i bez računanja beskonačnih suma. Uočimo da prvi u nizu eksperimenata može dati neželjeni ishod (s vjerojatnošću q) ili željeni ishod (s vjerojatnošću p). U prvom slučaju krećemo ispočetka, tj. očekivani broj eksperimenata do željenog ishoda sada je $1 + E$. U drugom slučaju došli smo do željenog ishoda, tj. očekivani broj eksperimenata do željenog ishoda je 1. Sada je

$$E = (1 + E)q + 1p,$$

odakle odmah slijedi

$$E(1 - q) = q + p = 1, \quad E = \frac{1}{1 - q} = \frac{1}{p}.$$

2. U prvoj kutiji sigurno ćete naći jedan broj. Vjerojatnost da u sljedećoj kutiji nađete novi broj iznosi $4/5$. Prema rješenju prethodnog zadatka, prosječno trebate kupiti $1/(4/5) = 5/4$ kutije da bi se to desilo. Vjerojatnost da u sljedećoj kutiji nađete novi broj je $3/5$ i u prosjeku trebate kupiti $1/(3/5) = 5/3$ kutija da bi se to desilo. Zadnja dva broja, u prosjeku, zahtijevaju kupnju $1/(2/5) = 5/2$ i $1/(1/5) = 5$ kutija. Ukupan (prosječni) broj kutija koje trebate kupiti je

$$1 + \frac{5}{4} + \frac{5}{3} + \frac{5}{2} + 5 = 5 \left(\frac{1}{5} + \frac{1}{4} + \frac{1}{3} + \frac{1}{2} + \frac{1}{1} \right) \approx 11.4.$$

Kada bi DEX u svojoj nagradnoj igri koristio n brojeva, prosječan broj kutija koji biste trebali kupiti bio bi

$$n \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \right).$$

Tu vrijednost možemo lako izračunati koristeći se aproksimacijom

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \approx \ln n + \gamma + \frac{1}{2n},$$

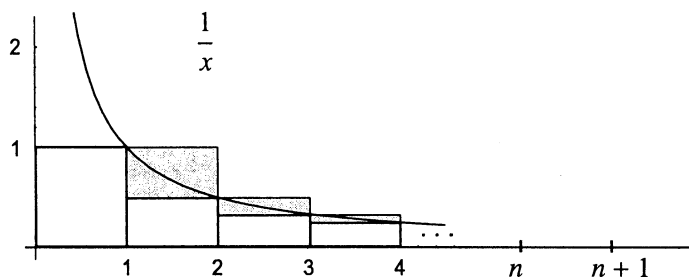
gdje je γ slavna Eulerova konstanta, $\gamma = 0.577218 \dots$

Na primjer, za $n = 5$ nalazimo

$$5 \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} \right) \approx 5 \ln 5 + 5\gamma + \frac{1}{2} \approx 11.4$$

što je, na jednu decimalu, ista vrijednost koju dobijemo i direktnim računanjem vrijednosti $5(1 + 1/2 + \dots + 1/5)$.

Objasnit ćemo kako možemo doći do te aproksimacije, i do same konstante γ . Krenimo od iznosa površina na sljedećoj slici:



Ukupna površina nezatamnjenih pravokutnika, čije se osnovice protežu od 0 do n , ima iznos

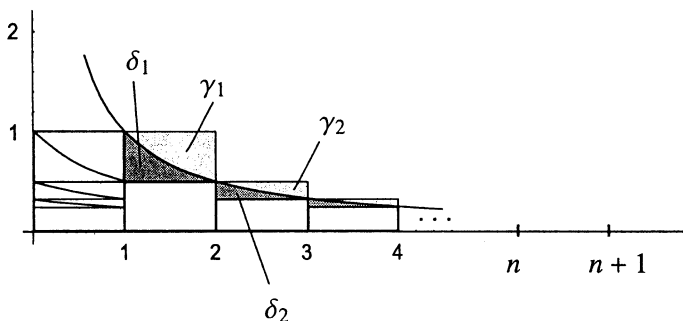
$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}.$$

Isti iznos ima i ukupna površina pravokutnika sa zatamnjenim vrhovima, čije se osnovice protežu od 1 do $n+1$. Površina od osi x do grafa $y = 1/x$, nad intervalom $[1, n]$, iznosi $\ln n$. Odavde, uz pokrate $\bar{\delta}_n = \delta_1 + \delta_2 + \dots + \delta_{n-1}$ i $\bar{\gamma}_n = \gamma_1 + \gamma_2 + \dots + \gamma_{n-1}$, lako slijedi (vidi prethodnu sliku)

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = \ln(n+1) - (\delta_1 + \delta_2 + \dots + \delta_{n-1}) = \ln(n+1) - \bar{\delta}_n, \quad (1)$$

$$H_n = 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} = \ln(n+1) + (\gamma_1 + \gamma_2 + \dots + \gamma_n) = \ln(n+1) + \bar{\gamma}_{n+1}. \quad (2)$$

Prikažemo li sve površine $\gamma_1, \gamma_2, \dots; \delta_1, \delta_2, \dots$ u 1×1 pravokutniku (vidi sljedeću sliku) lako ćemo se uvjeriti da vrijede sljedeće tvrdnje:



$$\sum_{n=1}^{\infty} \gamma_n = \lim_{n \rightarrow \infty} \bar{\gamma}_n = \gamma > \frac{1}{2}, \quad \sum_{n=1}^{\infty} \delta_n = \lim_{n \rightarrow \infty} \bar{\delta}_n = \delta < \frac{1}{2}, \quad \gamma + \delta = 1 \quad (3)$$

$$(\gamma - \bar{\gamma}_n) + (\delta - \bar{\delta}_n) = \frac{1}{n}, \quad \delta - \bar{\delta}_n \approx \frac{1}{2n} \approx \gamma - \gamma_n \quad (4)$$

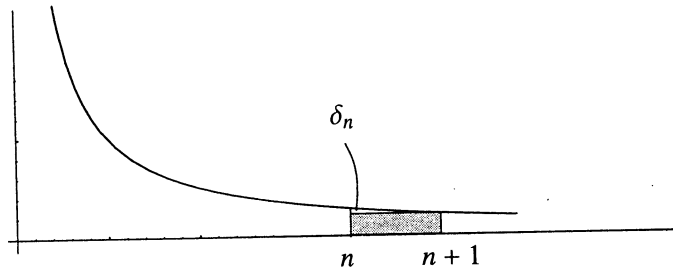
Iz (1) i (4) sada lako slijedi

$$H_n = 1 + \ln n - \bar{\delta}_n = 1 + \ln n + (\delta - \delta_n) - \delta = \ln n + \gamma + (\delta - \delta_n) \approx \ln n + \gamma + \frac{1}{2n},$$

što je naša aproksimacija. Osim toga, iz (2) i (4) slijedi:

$$H_n = \ln(n+1) + \bar{\gamma}_{n+1} = 1 + \ln(n+1) + \bar{\gamma}_{n+1} - \gamma + \gamma \approx \ln(n+1) + \gamma - \frac{1}{2(n+1)} \approx \ln n + \frac{1}{n+1} + \gamma - \frac{1}{2(n+1)} = \ln n + \gamma - \frac{1}{2(n+1)}$$

Naime, $\ln(n+1) - \ln n = 1/(n+1) + \delta_n \approx 1/(n+1)$, što se lako vidi:



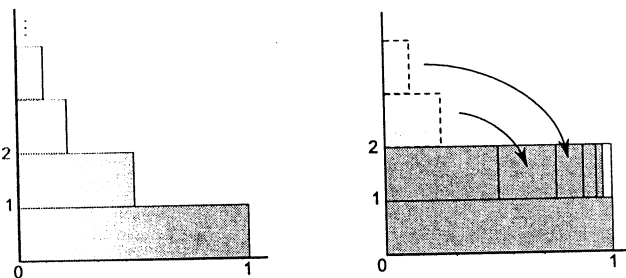
Sve u svemu, imamo sljedeću procjenu:

$$\ln n + \gamma - \frac{1}{2(n+1)} < H_n < \ln n + \gamma + \frac{1}{2n}$$

3. Vjerojatnost pisma je $1/2$. Vjerojatnost 2 pisma u nizu je $1/4$. Vjerojatnost 3 pisma u nizu je $1/8$. Vjerojatnost 4 pisma u nizu je $1/16$. I tako dalje. Dakle, vjerojatnost bloka duljine n je $1/2^n$. Slijedi da je prosječna duljina bloka

$$D = \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \dots + \frac{1}{2^n} \cdot n + \dots$$

Ovaj beskonačni zbroj možemo izračunati na razne načine. Pokazat ćemo kako ga je (u sasvim drugom kontekstu) izračunao Nikola Oresme u 14. st. On je konstruirao sljedeću površinu koja predstavlja tu sumu (lijeva slika):



Zatim je tu površinu presložio kao na prethodnoj slici (desno). Dakle, $D = 2$, tj. prosječna duljina bloka pisama je 2.

I ovdje možemo zaobići beskonačna zbrajanja. Zapitajmo se na koliko promjena (iz glave u pismo ili iz pisma u glavu) prosječno nailazimo u nizu od n bacanja novčića. Budući da je vjerojatnost promjene $1/2$, prosječno nailazimo na $n/2$ promjena. To znači da je prosječna duljina bloka 2. Uistinu jednostavnije rješenje.

4. Većina ljudi nekako osjeća da je vjerojatnost ista u svim trima slučajima. (Prosječan broj šestica na šest bačenih kocki je 1, na 12 je 2, na 18 je 3, a $1 : 6 = 2 : 12 = 3 : 18$.) Nažalost, to je jedan od čestih slučajeva kad nas vjerojatnosni osjećaj vara. (Prosjeci i vjerojatnost nisu isto). Zaboravimo zato osjećaje i krenimo računati. Vjerojatnost od x šestica u n bacanja je

$$\binom{n}{x} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{n-x} \quad x = 0, 1, \dots, n$$

Kada bacimo $6n$ kocaka vjerojatnost od n ili više šestica je

$$\sum_{x=n}^{6n} \binom{6n}{x} \left(\frac{1}{6}\right)^x \left(\frac{5}{6}\right)^{6n-x}$$

Newton je izračunao ove vrijednosti "ručno", a mi se danas možemo poslužiti tablicama kumulativne binomne distribucije, ili računalnim paketima koji ih sadrže, i dobit ćemo sljedeće vrijednosti:

n	$6n$	$P(\geq n \text{ šestica})$
1	6	0.665
2	12	0.619
3	18	0.597
4	24	0.584
5	30	0.576
\vdots	\vdots	\vdots
100	600	0.517

Dakle, najvjerojatnija je barem 1 šestica u 6 bacanja, a najmanje vjerojatne su barem 3 šestice u 18 bacanja. To je bilo i Newtonovo rješenje. Samuel Peppys je doznao na što se treba kladiti.

5. Većina ljudi ima osjećaj da je u skupini veličine školskog razreda vjerojatnost podudarnog rođendana puno manja od 50%. I to je jedan od čestih slučajeva lošeg osjećaja za vjerojatnost. Naime, u skupini te veličine podudarnost je oko

o). Dakle, $D = 2$,

se na koliko pro-
zimo u nizu od n
osječno nailazimo
Jistinu jednostav-

a slučajima. (Pro-
18 je 3, a $1 : 6 =$
kad nas vjerojat-
vimo zato osječajje

70%. Okrenimo se opet računu. Neka je $N = 365$ broj jednako vjerojatnih dana (zanemarit ćemo prijestupne godine), a s broj ljudi u promatranj skupini. Izračunajmo vjerojatnost svih da svih s ljudi ima međusobno različite rođendane.

Rođendan prvoga od s ljudi može biti bilo kojeg od N dana. Drugi mora biti različit, pa zato može tek biti bilo koji od preostalih $N - 1$ dana. Treći može biti bilo koji od preostalih $N - 2$, itd., do preostalih $N - s + 1$ dana. Dakle, ukupan broj načina na koji se mogu realizirati svi međusobno različiti rođendani je

$$N(N - 1)(N - 2) \dots (N - s + 1) = \frac{N!}{(N - s)!}$$

S druge strane, ukupan broj načina na koji se mogu realizirati svi mogući rođendani je $N \cdot N \cdot N \cdot \dots \cdot N = N^s$.

Oдавде slijedi da vjerojatnost svih međusobno različitih rođendana iznosi

$$\frac{N}{N} \frac{N - 1}{N} \frac{N - 2}{N} \dots \frac{N - s + 1}{N} = \frac{N!}{(N - s)!N^s}$$

Komplementarna vjerojatnost,

$$P_s = 1 - \frac{N!}{(N - s)!N^s},$$

možemo poslužiti
paketima koji ih

je vjerojatnost da skupina od s ljudi ima barem jedan podudarni rođendan. U sljedećoj tablici dani su iznosi te vjerojatnosti za $N = 365$ i neke s .

s	5	10	20	23	30	40	60
P_s	0.027	0.117	0.411	0.507	0.706	0.891	0.994

Vidimo da je vjerojatnost podudarnog rođendana veća od 50% već u skupini od 23 čovjeka. U skupini od 60 ljudi podudaran rođendan gotovo je siguran.

Najteži problem u gornjemu izvodu je računanje vrijednosti P_s . Naime, $365!$, $(365 - 23)!$ i 365^{23} ogromni su brojevi koji zahtijevaju ogromne račune. Zato te račune ovdje nismo ni uveli, nego smo (u tablici) dali samo njihove konačne rezultate.

anje vjerojatne su
e. Samuel Peppys

Puno je lakše provesti približne račune, poput sljedećeg koji se koristi dobro poznatom aproksimacijom:

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \frac{x^4}{4!} - \dots \approx 1 - x.$$

zreda vjerojatnost
d čestih slučajeva
odudarnost je oko

Ta je aproksimacija dobra za male vrijednosti x i obično se koristi tako da desnom stranom $1 - x$ aproksimiramo lijevu stranu e^{-x} . Mi ćemo je koristiti u obrnutom smjeru.

$$\begin{aligned} \frac{N!}{(N-s)!N^s} &= \frac{N}{N} \frac{N-1}{N} \frac{N-2}{N} \cdots \frac{N-s+1}{N} = \\ &= \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \cdots \left(1 - \frac{s-1}{N}\right) \approx \\ &\approx e^{-\frac{1}{N}} e^{-\frac{2}{N}} \cdots e^{-\frac{s-1}{N}} = e^{-\frac{1+2+\dots+(s-1)}{N}} = \\ &= e^{-\frac{s(s-1)}{2N}} \end{aligned}$$

Uočimo da je ta aproksimacija (zbog $e^{-x} > 1 - x$) veća od tražene vrijednosti. Sada lako zaključujemo da je

$$P_s = 1 - \frac{N!}{(N-s)!N^s} \approx 1 - e^{-\frac{s(s-1)}{2N}}$$

i da je ta aproksimacija manja od tražene vrijednosti. Želimo li naći s za koji je $P_s > 0.5$, valja nam riješiti jednadžbu

$$1 - e^{-\frac{s(s-1)}{2 \cdot 830}} = 0.5, \quad e^{-\frac{s(s-1)}{830}} = 0.5, \quad \frac{-s(s-1)}{830} = \ln 0.5,$$

$$s^2 - s + 830 \ln 0.5 = 0.$$

Ta kvadratna jednadžba ima pozitivno rješenje $24 < s < 25$. Dakle, za $s = 25$ sigurno vrijedi $P_s > 0.5$. S obzirom na to da su naše aproksimacije bile manje od tražene vrijednosti, možda je $P_s > 0.5$ i za $s = 23$ ili čak $s = 22$. To ovaj račun s e^{-x} ne može odlučiti. Međutim, točniji rezultati pokazuju da je $s = 23$ prvi s za koji je $P_s > 0.5$.