Rozhledy matematicko-fyzikální

Pavel Pokorný

Souběh narozenin

Rozhledy matematicko-fyzikální, Vol. 96 (2021), No. 4, 18–29

Persistent URL: http://dml.cz/dmlcz/149339

Souběh narozenin

Pavel Pokorný, VŠCHT Praha

„Jé, ty máš narozeniny stejný den v roce jako já? To je ale vzácná náhoda!“

Je to opravdu vzácná náhoda? Mezi dvěma osobami je pravděpodobnost souběhu narozenin ve stejném dni v roce opravdu malá, jen 1/365, tedy přibližně 0,3 %. A mezi více osobami? Jaká je např. pravděpodobnost, že mezi 22 fotbalisty na hřišti dojde k alespoň jednomu souběhu narozenin? A k souběhu alespoň trojích narozenin v jednom dni v roce?

1. Souběh dvojích narozenin

Budeme uvažovat nepřestupný rok, tedy rok, který má 365 dní. Budeme předpokládat, že každý den v roce je na narozeniny stejně pravděpodobný a že mezi uvažovanými osobami nejsou žádné vazby, např. že se nejedná o dvojčata. Tedy že výskyty narozenin různých osob jsou nezávislé jevy. Zajímáme se jen o dny v roce, zatímco rok nás nezajímá.

1.1. Dvě osoby

Uvažujme nejdříve pouze 2 osoby. Každá z nich může mít narozeniny libovolný den v roce se stejnou pravděpodobností. To je celkem

$$c = 365^2$$

možných případů. Aby nedošlo k souběhu narozenin, může mít první osoba narozeniny kterýkoliv z 365 dnů v roce, ale druhá osoba může mít narozeniny pouze jeden ze zbývajících 364 dnů v roce. Takže celkový počet možností, kdy nedojde k souběhu, protože každý den má narozeniny nejvýše jeden člověk, je

$$c_1 = 365 \cdot 364.$$

Tedy pravděpodobnost nesouběhu bude

$$n = \frac{c_1}{c} = \frac{365 \cdot 364}{365^2}.$$

Tento zlomek by šlo krátit, ale my ho ponecháme v tomto tvaru, protože to bude výhodné později.

A pravděpodobnost souběhu je

$$1 - n = \frac{1}{365},$$

jak jsme uvedli v úvodu.

1.2. Tři osoby

Pro 3 osoby máme celkem

$$c = 365^3$$

možností, protože každá ze tří osob může mít narozeniny libovolný den v roce. Aby nedošlo k žádnému souběhu narozenin, může mít první osoba narozeniny libovolný den v roce, tedy v 365 případech. Druhá osoba již musí mít narozeniny pouze libovolný den ve zbývajících 364 dnech a třetí osoba ve zbývajících 363 dnech. Tedy celkový počet možností bez souběhu, kdy má každý den narozeniny nejvýše jeden člověk, je

$$c_1 = 365 \cdot 364 \cdot 363.$$

Tedy pravděpodobnost nesouběhu bude

$$n = \frac{c_1}{c} = \frac{365 \cdot 364 \cdot 363}{365^3}.$$

1.2.1. Značení

Připravíme si užitečný způsob zápisu takovýchto součinů. Podobně jako součet čísel lze zapsat pomocí velkého řeckého Sigma, např.

$$a_1 + a_2 + a_3 = \sum_{i=1}^{3} a_i,$$

tak pro součin používáme velké řecké Pi, např.

$$a_1 \cdot a_2 \cdot a_3 = \prod_{i=1}^{3} a_i.$$

S tímto značením tedy můžeme psát

$$c_1 = 365 \cdot 364 \cdot 363 = \prod_{i=363}^{365} i.$$

Pro značení i pro výpočty je užitečné zavést operaci faktoriál. Značí se vykřičníkem za číslem a faktoriál přirozeného čísla je součin všech přirozených čísel od jedné do daného čísla, tedy např. $3! = 1 \cdot 2 \cdot 3 = 6$, obecně

$$k! = \prod_{i=1}^{k} i.$$

Protože

$$k! = \frac{(k+1)!}{k+1},$$

je přirozené dodefinovat $0! = 1$. S použitím faktoriálu můžeme psát

$$c_1 = 365 \cdot 364 \cdot 363 = \frac{365!}{362!}.$$

1.3. Obecný počet osob

Abychom naše výsledky mohli použít i pro řešení jiných úloh, než je souběh narozenin v roce, zobecníme zadání a zavedeme toto značení: h počet osob (hráčů) a d počet dnů. Pak bude celkový počet možných případů

$$c(h, d) = d^h.$$

Počet možností, kdy nedochází k souběhu narozenin, tedy kdy má každý den narozeniny nejvýše jedna osoba, bude (za předpokladu $h \leq d$) součin h čísel

$$c_1(h, d) = d(d-1)\dots(d-h+1) = \frac{d!}{(d-h)!}.$$

Pravděpodobnost nesouběhu bude

$$n(h, d) = \frac{c_1(h, d)}{c(h, d)}$$

a pravděpodobnost souběhu alespoň dvojích narozenin bude

$$s_2(h, d) = 1 - n(h, d).$$

Pro $h = 2$ hráče a pro $d = 365$ dnů dostáváme pravděpodobnost souběhu alespoň dvojích narozenin

$$s_2(2{,}365) \doteq 0{,}00274,$$

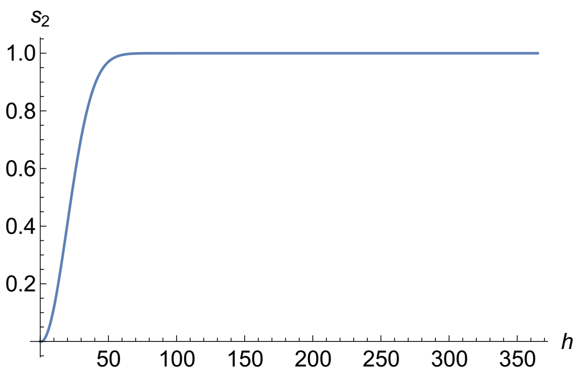
tedy méně než 0,3 %, takže souběh narozenin v páru je výjimečný. Pro $h = 22$ hráčů je pravděpodobnost souběhu

$$s_2(22, 365) \doteq 0{,}476,$$

tedy přibližně 48 %, takže souběh narozenin mezi 22 hráči na fotbalovém hřišti je běžná záležitost. A to ani nepočítáme rozhodčí a trenéry.

No a kdybychom zahrnuli i diváky, tak pokud bude celkový počet osob větší než počet dnů v roce, obecně pokud $h > d$, tak je souběh jistý. Toto samozřejmé tvrzení se nazývá Dirichletův princip, někdy označované také jako zásuvkový princip nebo holubníkový princip. Jedna jeho podoba zní: Když máme stůl s d zásuvkami a v nich uloženo h tužek, tak pokud máme více tužek než zásuvek, tedy pokud $h > d$, tak budou v alespoň jedné zásuvce alespoň dvě tužky.

Závislost pravděpodobnosti souběhu narozenin na počtu hráčů ukazuje obr. 1. Pro malý počet hráčů je pravděpodobnost téměř nulová, pro velký počet hráčů je rovna jedné, tedy souběh je jistý.



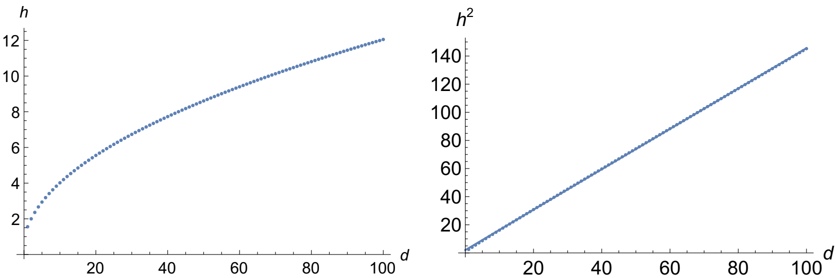
Obr. 1: Závislost pravděpodobnosti, že dojde k souběhu narozenin mezi h hráči na počtu hráčů pro $d = 365$ dnů. Např. pro $h = 57$ je $s_2 = 0.9901$.

Položme si otázku: Pro které h tato závislost překročí jednu polovinu v závislosti na počtu dnů d? Tedy řešíme rovnici

$$s_2(h, d) = \frac{1}{2}.$$

Pokud neexistuje celočíselné řešení této rovnice, můžeme se ptát, pro které nejmenší přirozené h je $s_2(h,d) \geq \frac{1}{2}$. K otázce neceločíselnosti h se ještě vrátíme.

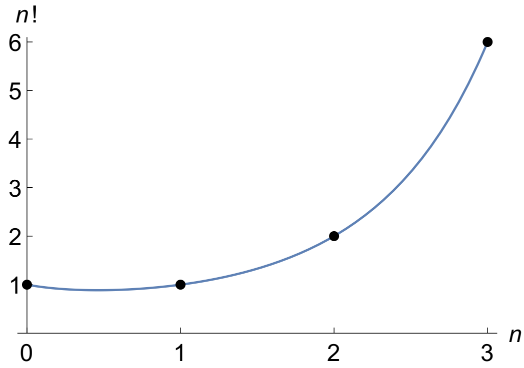
Odpověď na tuto otázku lze hledat dvojím způsobem. Můžeme zvolit numerický experiment. Pro různé hodnoty d např. mezi 1 a 100 můžeme postupně zvyšovat h, počítat s_2 a zaznamenat si takovou hodnotu h, při které s_2 překročí $\frac{1}{2}$. Obr. 2 vlevo ukazuje závislost h na d za podmínky $s_2(h,d) = \frac{1}{2}$. Graf připomíná odmocninu. V pravé části tohoto obrázku vidíme stejná data, ale na svislé ose je vynesena hodnota h^2. Body se srovnaly podél přímky. Když těmito body proložíme přímku (např. metodou nejmenších čtverců) a vykreslíme ji také do grafu, vidíme, že body leží velice blízko této přímky. Závěr je, že je-li počet hráčů přibližně odmocnina z počtu dnů, tak je pravděpodobnost souběhu přibližně jedna polovina, tedy běžný jev.



Obr. 2: Vlevo: závislost počtu hráčů h na počtu dnů d, aby byla pravděpodobnost souběhu narozenin jedna polovina. Tato závislost připomíná graf funkce odmocnina. Vpravo: stejná data jako na obrázku v levé části, ale na svislé ose je vynesena veličina h^2. Body leží podél přímky, což ukazuje, že závislost h na d je přibližně odmocninová.

Pozorný čtenář si všimne, že h na obr. 2 nabývá i neceločíselných hodnot. To je dáno tím, že obr. 2 jsme vytvořili pomocí matematického software *Mathematica* tak, že jsme pro jednotlivé hodnoty d řešili rovnici $s_2(h,d) = \frac{1}{2}$ a tento software používá zobecnění faktoriálu i pro neceločíselné argumenty. Pro zvídavého čtenáře: Toto zobecnění se (až na posunutí o jedničku) nazývá Gama funkce, viz obr. 3.

Druhý způsob, jak dojít k tomuto závěru, je nahradit faktoriály v rovnici $s_2(h,d) = \frac{1}{2}$ pomocí Stirlingovy formule. Potom za předpokladu, že d je velké, bychom dospěli opět k závěru, že h je přibližně rovno odmocnině z d.



Obr. 3: Faktoriál $n! = 1 \cdot 2 \cdot \dots \cdot n$ pro přirozená n lze dodefinovat pro nulu hodnotou $0! = 1$ a také pro kladná neceločíselná n pomocí gamma funkce $n! = \Gamma(n+1)$

1.3.1. Použití v kryptografii

Zatím jsme naše výpočty podávali jako matematické řešení úlohy o pravděpodobnosti souběhu narozenin fotbalových hráčů, tedy bez praktického použití. Tyto úvahy mají však použití mimo jiné v kryptografii. To je část matematiky, která se zabývá metodami bezpečného přenosu zpráv a digitálním podpisem dokumentu. Pro tyto účely se používají mimo jiné také hešovací funkce (anglicky hash function) [1]. Vstupem pro takovou funkci je posloupnost znaků libovolné konečné délky a výstupem je celé číslo mezi 0 a jistou maximální hodnotou. Hešovací funkce se navrhují tak, aby bylo snadné spočítat výsledek ze vstupních dat, ale aby bylo velice obtížné nalézt vzor ze znalosti funkční hodnoty, to znamená vstup dávající daný výsledek. Nebo nalézt jiný vzor, který dá stejnou funkční hodnotu. Je-li počet možných vstupů větší než počet výstupních hodnot, pak taková funkce nemůže být prostá a je jisté, že existují různé vstupní hodnoty, které dají stejný výsledek. Tedy že dochází k souběhu. A zde se nabízí srovnání s naší úlohou o souběhu narozenin. Funkce, kterou uvažujeme, přiřazuje každému člověku den narozenin (z d možných). My poté vybíráme ze všech lidí h hráčů.

Volme nyní náhodně h vstupních hodnot hešovací funkce, která má d možných funkčních hodnot. Pokud jich zvolíme dostatek, konkrétně je-li $h > \sqrt{d}$, je velice pravděpodobné, že najdeme dva různé vstupy, které dají stejný výsledek. Je-li dokonce $h > d$, je souběh jistý.

Proto je důležité navrhovat hešovací funkce tak, aby počet možných výsledků byl veliký. Např. kryptoměna bitcoin používá hešovací funkci SHA-256, která má funkční hodnoty dlouhé 256 bitů, tedy nabývají jedné z 2^{256} hodnot. V našem značení $d = 2^{256}$. Abychom s velkou pravděpodobností našli souběh, tedy dva různé argumenty, které dají stejný výsledek, měli bychom vyzkoušet

$$h = \sqrt{d} = \sqrt{2^{256}} = 2^{128} \doteq 10^{38}$$

vstupních hodnot. Jak by to asi dlouho trvalo? Dnešní počítače pracují s hodinovým kmitočtem v řádu GHz, tedy za jednu sekundu provedou přibližně 10^9 operací. I kdybychom zapojili 10 miliard procesorů najednou (to je trochu více, než je současný počet obyvatel naší planety), tak jsme schopni provést 10^{19} operací za vteřinu. Takže provést 10^{38} operací by nám trvalo $10^{38}/10^{19} = 10^{19}$ s. To je déle, než je odhadované stáří vesmíru, což je přibližně 14 miliard let $= 14 \cdot 10^9 \cdot 365 \cdot 24 \cdot 3600 \doteq 10^{17}$ s. Tedy, u této hešovací funkce nehrozí, že by někdo pouhým zkoušením našel dva různé vzory, které dají stejný výsledek.

2. Souběh trojích narozenin

A jaká je pravděpodobnost, že tři hráči budou mít narozeniny ve stejný den v roce? Očekáváme, že to bude řidší jev než souběh dvojích narozenin. Podívejme se nejprve na krajní případy. Bude-li počet hráčů $h < 3$, pak k souběhu tří narozenin nikdy nedojde. Bude-li naopak $h > 2d$, pak souběh trojích narozenin je jistý. A jak to bude mezi těmito dvěma krajními případy? Počet případů, kdy dochází k souběhu alespoň trojích narozenin v jednom dni, dostaneme tak, že od celkového počtu c případů odečteme počet c_1 případů, kdy nedochází k souběhu, a počet c_2, kdy dochází k souběhu dvou (ale ne více) narozenin.

Ukážeme si dvě možné cesty k řešení: útok hrubou silou a diferenční rovnici.

2.1. Útok hrubou silou

Útok hrubou silou (anglicky brute force attack) znamená způsob řešení, kdy si napíšeme všechny možné případy a pak spočítáme, kolik z nich splňuje danou podmínku. To je možné jen pro malé hodnoty, které dají malý počet případů. Tužkou na papíře lze zvládnout desítky, nejvýše stovky případů, na počítači lze pracovat s miliony případů. Ani to ale často nestačí. Uvažme naši úlohu, která pro $h = 22$ hráčů a $d = 365$ dní má celkový počet možných případů $c(h,d) = d^h = 365^{22} \doteq 2 \cdot 10^{56}$. Zpracovat tak velký počet případů nelze stihnout za dobu menší než je stáří našeho vesmíru. Další vážné omezení této metody je kapacita paměti pro uchování informací o všech případech. To lze často obejít tím, že každý jednotlivý případ průběžně vytvoříme, zpracujeme a vymažeme z paměti. Ilustrujme si metodu hrubou silou tedy pro malé hodnoty parametrů, např. pro počet hráčů $h \in \{1,2,3,4,5\}$ a pro počet dnů $d \in \{1,2,3,4,5\}$.

Následující tabulka uvádí celkový počet možných případů $c(h,d) = d^h$.

h \ d	1	2	3	4	5
1	1	2	3	4	5
2	1	4	9	16	25
3	1	8	27	$\boxed{64}$	125
4	1	16	81	256	625
5	1	32	243	1024	3125

Pro ilustraci zvolme např. $h = 3$ hráče a $d = 4$ dny. Pak celkový počet případů je $c(3,4) = 4^3 = 64$. Tato hodnota je v tabulce zdůrazněna rámečkem. Těchto 64 případů lze rozdělit do tří skupin.

V první skupině jsou případy, kde nenastává souběh. Těchto případů je $c_1(3,4) = \frac{4\,!}{(4-3)\,!} = 4 \cdot 3 \cdot 2 = 24$. Tyto případy bez souběhu ukazuje následující tabulka. Buňky jsou seřazeny vzestupně, číslo vpravo má nejmenší váhu (podobně jako řadíme přirozená čísla).

1 2 3	1 2 4	1 3 2	1 3 4	1 4 2	1 4 3
2 1 3	2 1 4	2 3 1	2 3 4	2 4 1	2 4 3
3 1 2	3 1 4	3 2 1	3 2 4	3 4 1	3 4 2
4 1 2	4 1 3	4 2 1	4 2 3	4 3 1	4 3 2

Např. políčko $\boxed{2\ 1\ 3}$ na druhém řádku vlevo znamená případ, kdy první hráč má narozeniny druhý den, druhý hráč má narozeniny první den, třetí hráč má narozeniny třetí den a čtvrtý den už nemá z našich třech hráčů narozeniny žádný.

Ve druhé skupině jsou případy, kdy nastává souběh dvou (ale ne více) narozenin. Tyto případy ukazuje následující tabulka.

1 1 2	1 1 3	1 1 4	1 2 1	1 2 2	1 3 1	1 3 3	1 4 1	1 4 4
2 1 1	2 1 2	2 2 1	2 2 3	2 2 4	2 3 2	2 3 3	2 4 2	2 4 4
3 1 1	3 1 3	3 2 2	3 2 3	3 3 1	3 3 2	3 3 4	3 4 3	3 4 4
4 1 1	4 1 4	4 2 2	4 2 4	4 3 3	4 3 4	4 4 1	4 4 2	4 4 3

Např. políčko $\boxed{2\ 1\ 1}$ na druhém řádku vlevo znamená případ, kdy první hráč má narozeniny druhý den, zatímco druhý i třetí hráč mají souběžně narozeniny první den. Třetí a čtvrtý den už nemá narozeniny nikdo. Označme symbolem $c_2(h,d)$ počet případů, kdy dochází k souběhu dvou (ale ne více) narozenin. Tato tabulka má 36 políček, takže jsme hrubou silou zjistili, že $c_2(3,4) = 36$.

A ve třetí skupině budou případy, kdy nastává souběh tří (ale ne více) narozenin. Jejich počet označme $c_3(h,d)$. Tyto případy ukazuje následující tabulka.

1 1 1	2 2 2	3 3 3	4 4 4

Tato tabulka má 4 políčka, takže jsme hrubou silou zjistili, že $c_3(3,4) = 4$. Při třech hráčích nenastává souběh více než tří narozenin, takže náš výčet je úplný. Pro kontrolu sečteme

$$c_1(3,4) + c_2(3,4) + c_3(3,4) = 24 + 36 + 4 = 64,$$

což je přesně celkový počet případů

$$c = d^h = 4^3 = 64.$$

Podobným způsobem lze najít počet $c_2(h,d)$ případů se souběhem dvou (ale ne více) narozenin i pro jiné hodnoty h a d. Několik vybraných výsledků ukazuje následující tabulka.

h \ d	1	2	3	4	5
1	0	0	0	0	0
2	1	2	3	4	5
3	0	6	18	$\boxed{36}$	60
4	0	6	54	180	420
5	0	0	90	600	2100

Hodnota $c_2(3,4) = 36$, o které jsme již psali, je zdůrazněna rámečkem. Dvě hodnoty jsou v kroužku. To jsou hodnoty, které budeme později potřebovat pro rekurentní výpočet hodnoty v rámečku.

Pro výpočet můžeme s výhodou použít počítačový algebraický systém *Mathematica* a tyto příkazy

```
h=3;
d=4;
pripady=Tuples[Range[d],h];
maxsoubeh[x_]:=Tally[x][[All,2]]//Max;
soubehy = Map[maxsoubeh,pripady];
pocty = Tally[soubehy]

{{3, 4}, {2, 36}, {1, 24}}
```

Příkazem `Tuples` si připravíme všechny h-tice čísel od 1 do d, tedy všechny možné případy. Pak si připravíme funkci `maxsoubeh`, která pro každý případ najde, kolikrát se vyskytují nejčastější narozeniny. Zde jsme použili příkaz `Tally`, který ze vstupní posloupnosti najde, kolikrát se jednotlivé prvky v posloupnosti vyskytují, a vytvoří dvojice [prvek, počet jeho výskytů]. Dvojkou jsme vybrali pouze druhou položku dvojice, tedy počet výskytů. A příkazem `Max` dostaneme nejvyšší počet výskytů. Funkci `maxsoubeh` pustíme na každý případ a nakonec si zjistíme, kolikrát se jednotlivé typy souběhů vyskytují. Výsledek nám říká, že trojitý souběh nastává 4 krát, dvojitý souběh 36 krát a případů bez souběhu je 24. Ve shodě s výše uvedenými výsledky.

Tímto způsobem, tedy hrubou silou, nelze nalézt $c_2(22,365)$. Ukažme si tedy jiný způsob pomocí diferenční rovnice.

2.2. Diferenční rovnice

Odvodíme si vztah, který nám dovolí spočítat hodnotu $c_2(h,d)$ pomocí hodnot c_2 pro menší argumenty. Takový vztah se nazývá diferenční rovnice.

Příkladem jednoduché diferenční rovnice je

$$a_{n+1} = 2a_n.$$

Řešením je posloupnost, zde a_n, které se říká rekurentně zadaná posloupnost.

Samozřejmě, že bychom raději měli vztah, který by nám dovoloval spočítat hodnotu $c_2(h,d)$ jen ze znalosti h a d, ale když se nám takový vztah nepodaří najít, musíme se smířit alespoň s rekurentním vztahem.

Takže, kolik je možností, kdy nastane souběh dvou (ale ne více) narozenin? První hráč může mít narozeniny libovolný den. To je d možností. Pak:

- buď má v tento den narozeniny ještě některý ze zbylých $h-1$ hráčů a ostatní dny se vyskytují už buď bez souběhu (těch je $c_1(h-2,d-1)$) nebo nejvýše s dvojitým souběhem (těch je $c_2(h-2,d-1)$),
- anebo se tento den již neopakuje, pak musí narozeniny ostatních hráčů vykazovat právě dvojitý souběh (těch je $c_2(h-1,d-1)$).

To dává rekurentní vztah

$$c_2(h,d) = d((h-1)(c_1(h-2,d-1)+c_2(h-2,d-1))+c_2(h-1,d-1)).$$

Např. pro $h=3$, $d=4$ dostaneme

$$c_2(3,4) = 4(2(c_1(1,3)+c_2(1,3))+c_2(2,3)) = 4(2(3+0)+3) = 36$$

v souladu s údaji v tabulce. Vidíme, že pro výpočet určité hodnoty $c_2(h,d)$ potřebujeme dvě hodnoty z této tabulky, a to hodnotu $c_2(h-1,d-1)$ (soused vlevo nahoře) a hodnotu $c_2(h-2,d-1)$ nad touto hodnotou. Tedy pro výpočet určité hodnoty $c_2(h,d)$ např. $c_2(22,365)$ potřebujeme znát hodnoty na prvních dvou řádcích tabulky, tedy $c_2(1,\cdot)$ a $c_2(2,\cdot)$, a hodnoty v levém sloupci, tedy $c_2(\cdot,1)$. Tyto hodnoty snadno dostaneme následující úvahou. Pro jednoho hráče, když je $h=1$, souběh dvojích narozenin nemůže nastat. Proto $c_2(1,d)=0$ pro všechny hodnoty d. Pro dva hráče může nastat souběh dvojích narozenin tolikrát, kolik je dní, tedy $c_2(2,d)=d$ pro všechny hodnoty d. A pokud uvažujeme pouze jeden den, tedy $d=1$, tak pro 3 a více hráčů nemůže nastat souběh pouze dvojích narozenin, ale právě h narozenin, tedy pro $h \geq 3$ je $c_2(h,1)=0$. Těmto podmínkách se říká okrajové podmínky.

Pro výpočet můžeme opět s výhodou použít počítačový algebraický systém *Mathematica*. Okrajové podmínky i diferenční rovnici můžeme zadat následujícími příkazy

```
c[h_,d_]:=d^h;
c1[h_,d_]:=d!/(d-h)!;
c2[2,1]:=1;
```

```
c2[1,d_]:=0;
c2[2,d_]:=d;
c2[h_,1]:=0;
c2[h_,d_]:=d*((h-1)*(c1[h-2,d-1]+c2[h-2,d-1])+c2[h-1,d-1]);
s2[h_,d_]:=(c[h,d]-c1[h,d])/c[h,d];
s3[h_,d_]:=(c[h,d]-c1[h,d]-c2[h,d])/c[h,d];
```

A pak se můžeme přesvědčit, jestli dostaneme stejný výsledek, jako jsme dostali hrubou silou

```
In[10]:= c2[3,4]
Out[10]= 36
```

Ano, to je stejný výsledek. A můžeme se přesvědčit, jaká je pravděpodobnost souběhu alespoň dvojích narozenin mezi 22 hráči

```
In[11]:= s2[22,365]//N
Out[11]= 0.475695
```

Funkci `N` jsme použili, abychom dostali přibližnou numerickou hodnotu, jinak bychom dostali zlomek, kde čitatel i jmenovatel jsou velká přirozená čísla.

A nakonec můžeme snadno spočítat pravděpodobnost souběhu alespoň trojích narozenin mezi 22 hráči

```
In[12]:= s3[22,365]//N
Out[12]= 0.0110777
```

Tedy přibližně jedno procento, což je výrazně méně než pravděpodobnost souběhu dvojích narozenin, jak jsme očekávali.

Poděkování. Za inspiraci k těmto úvahám děkuji svému příteli Lubošovi Rulíškovi.

Literatura

[1] Balková, Ľ., Legerský, J.: Hašovací funkce a kombinatorika na slovech. *Pokroky matematiky, fyziky a astronomie*, roč. 58 (2013), č. 4, s. 274–284.

[2] Birthday problem. `https://en.wikipedia.org/wiki/Birthday_problem`