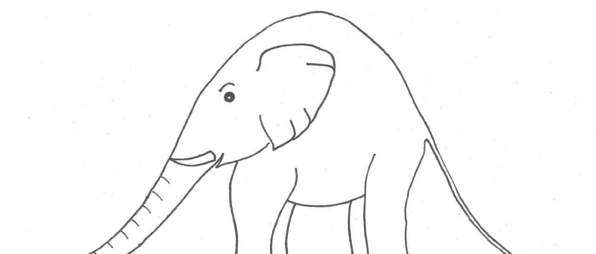
Pravděpodobnost III. – Návrat velkých čísel

Lékařství dělá lidi nemocnými, matematika mrzutými a teologie hříšnými.
Martin Luther

Milý příteli,
vítej u třetího dílu seriálu o pravděpodobnosti! Zatímco v předchozím díle jsme Ti chtěli ukázat, jak se pomocí pravděpodobnosti dají řešit všemožné matematické úlohy, zde bychom Ti především rádi naznačili, proč je vlastně pravděpodobnost tak užitečným nástrojem i mimo svět matematiky. Naším hlavním cílem bude pomocí náhodných veličin vysvětlit princip, kterému se říká *zákon velkých čísel.*

Nejprve si však zopakujme, co jsme si řekli v předchozím díle. Představili jsme si pravděpodobnostní metodu, jež umožňuje pomocí pravděpodobnosti řešit úlohy, které s ní zdánlivě vůbec nesouvisí. Poté jsme si zavedli náhodné veličiny (značené velkými písmeny, nejčastěji X), což jsou v podstatě proměnné nabývající různých hodnot s různou pravděpodobností. Definovali jsme je jako funkce z pravděpodobnostního prostoru Ω do reálných čísel. Díky této definici můžeme s různými náhodnými veličinami definovanými na stejném pravděpodobnostním prostoru provádět různé operace a stále dostaneme náhodnou veličinu: jsou-li například X, Y definované na Ω, pak náhodná veličina $2X + Y^2$ přiřadí každému ω z Ω číslo $2X(\omega) + Y(\omega)^2$. Po stručném shrnutí sumační notace jsme ji hned aplikovali při zavedení střední hodnoty náhodné veličiny $\mathrm{E}(X)$, což je prostě průměrná hodnota X. Nakonec jsme si pro libovolné náhodné veličiny X, Y a číslo c dokázali vztahy $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$ a $\mathrm{E}(c \cdot X) = c \cdot \mathrm{E}(X)$, které se dohromady nazývají linearita střední hodnoty. V řadě případů nám tyto vlastnosti velmi usnadňují její výpočet. Všechny nově nabyté znalosti jsme pak použili k vylepšení pravděpodobnostní metody, té se však ve třetím díle věnovat vůbec *nebudeme.*



Začátek tohoto dílu bude věnován především intuici za již zmíněným zákonem velkých čísel, který může sice na první pohled působit poněkud složitě, ale ve skutečnosti pouze potvrzuje správnost našich představ o pravděpodobnosti. Dále prozkoumáme několik užitečných myšlenek, jako je nezávislost náhodných veličin (což je zobecnění nezávislosti *jevů* z prvního dílu) nebo rozptyl.

Zákon velkých čísel

Zákony jsou jako párky, lepší nevědět, jak vznikají.

John Godfrey Saxe

Vraťme se opět k našemu oblíbenému příkladu s házením mincí. Předpokládejme, že si n-krát hodíme mincí a spočítáme, kolik nám padlo panen. Tomu odpovídá monstrózní pravděpodobnostní prostor obsahující 2^n elementárních jevů – každý odpovídá jedné z možných posloupností hodnot. V minulém díle jsme si ukázali, jak takovéto pokusy popisovat pomocí náhodných veličin. Pro hod s pořadovým číslem i můžeme vytvořit náhodnou proměnnou A_i, která bude říkat, zda v i-tém hodu padla panna (v tomto případě bude rovná jedné), či nikoli (bude rovna nule). Takovým proměnným také říkáme indikátory, neboť indikují, zda nastal daný jev (padla panna). Střední hodnota této proměnné $\mathrm{E}(A_i)$ bude rovna $\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}$. Použitím linearity střední hodnoty dostáváme, že pro střední počet padlých panen platí $\mathrm{E}(A_1 + A_2 + \cdots + A_n) = \mathrm{E}(A_1) + \mathrm{E}(A_2) + \cdots + \mathrm{E}(A_n) = n/2$.

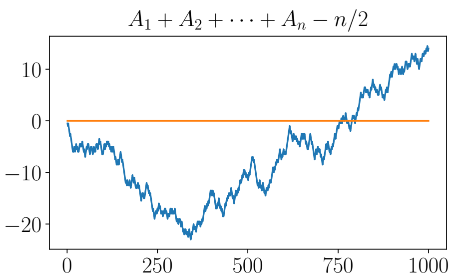
Všimni si, že ačkoli v předchozím příkladu byly jednotlivé hody mincí nezávislé, tak to pro použití linearity střední hodnoty nebylo nutné, protože platí i obecně. Můžeme třeba uvážit jinou náhodnou veličinu B odpovídající tomu, že si nejprve náhodně s poloviční pravděpodobností vybereme mezi mincí, která má na obou stranách pannu, a mincí, která má na obou stranách orla, a poté si touto mincí n-krát hodíme. Střední hodnota B je $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot n = n/2$ (mohou nastat jen dvě možnosti). Zároveň ale můžeme totéž spočítat pomocí indikátorů: Zavedeme-li B_i jako indikátor toho, zda je i-té slovo na papíře „panna“, bude opět platit, že B_i je rovno jedné s pravděpodobností $\frac{1}{2}$, a podle stejného vzorečku jako výše dostaneme, že $\mathrm{E}(B) = \mathrm{E}(B_1) + \mathrm{E}(B_2) + \cdots + \mathrm{E}(B_n) = n/2$. Tentokrát už ale o nezávislosti mluvit nemůžeme: Nastane-li jev $B_1 = 1$, víme, že už jistě nastanou i všechny ostatní jevy $B_i = 1$.

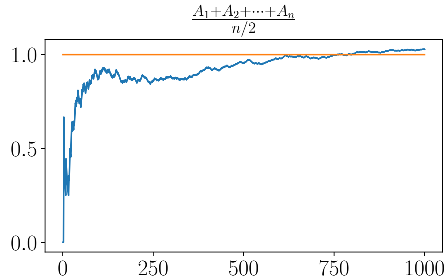
Veličiny A a B nás budou provázet po zbytek dílu a budeme se držet jejich značení (náhodné veličiny jinak obvykle značíme písmeny z konce abecedy).

První příklad s veličinou A jsme dokonce použili v prvním díle jako motivaci k zavedení pravděpodobnosti: hodíme-li si n-krát férovou mincí, očekáváme, že přibližně v půlce případů dostaneme pannu. První a druhý příklad se v tomto ohledu zásadně liší. Ačkoli v druhém příkladu také platí $\mathrm{E}(B) = n/2$, nemá smysl očekávat, že budou hodnoty B blízko průměru, poněvadž veličina nabývá pouze hodnot 0 a n.

Co přesně tedy máme na mysli, když v prvním případě očekáváme přibližně $n/2$ panen? Na následujících dvou obrázcích vidíš výsledek pokusu, kdy jsme na počítači postupně generovali 1000 náhodných veličin, přičemž každá nabývá hodnoty 0, nebo 1 s pravděpodobností $\frac{1}{2}$. To odpovídá házení mincí a sledování počtu panen, takže dále budeme o pokusu mluvit tak, jako kdybychom si doopravdy hodili tisíckrát mincí.

Na prvním obrázku jsme postupně vyznačili to, jak se počet padlých panen liší od střední hodnoty, neboli pro n od 1 do 1000 jsme vyznačili, kolik je $A_1 + A_2 + \cdots + A_n - n/2$. Na druhém obrázku je vyznačen podíl $\frac{A_1 + A_2 + \cdots + A_n}{n/2}$. Oba obrázky odpovídají stejnému pokusu, takže například platí, že hodnota nalevo je větší než nula právě tehdy, když je hodnota napravo větší než jedna.





Když se podíváš na první obrázek, vidíš, že celkový počet panen nebyl 500 (doopravdy to bylo 514). Dá se spočítat, že pravděpodobnost, že padne přesně 500 panen, se pohybuje kolem pouhých 2,5 %. Graf na levém obrázku ani nevypadá, že by se snažil zůstávat v okolí nuly. Naopak se zdá, že se rozdíl počtů panen a orlů s rostoucím počtem hodů postupně zvětšuje.

Tento fakt není tak překvapivý, pokud se smíříme s tím, že mince „nemá paměť" – to je to, co myslíme, když říkáme, že jednotlivé hody jsou nezávislé. V prvních šestnácti hodech padly jen dvě panny[1] a graf se odchýlil od nuly k hodnotě -7. To ale vůbec neznamená, že by se v dalších hodech mince musela snažit, aby padalo více panen (a tím se graf vrátil k nule). Panna a orel jsou v každém dalším hodu stejně pravděpodobní, a to nezávisle na historii.

Lidé si to ovšem občas neuvědomují, čímž se dopouštějí tzv. gamblerova omylu. V roce 1913 třeba v jistém kasinu v Monte Carlu hráči prohráli rekordní částky, když v tamní ruletě dvacetšestkrát po sobě skončila kulička na černém políčku. Hráči samozřejmě sázeli na to, že v příštím kole už přece kulička musí padnout na červené políčko. Pozorování, že málokdy padne tolik černých po sobě, je sice správné; když ale tímto nepravděpodobným jevem podmíníme (tedy když už víme, že nastal), je v dalším roztočení rulety černá stejně pravděpodobná jako červená. To, že jsme v minulosti měli smůlu, nás bohužel nijak nechrání před dalším neštěstím.[2]

Následující úloha ukazuje, že ačkoli je po $2n$ hodech nejpravděpodobnější počet padlých panen roven n, pravděpodobnost tohoto jevu se postupně zmenšuje (a jak jsme si řekli, pro $n = 1000$ je to jen 2,5 %).

Úloha 1. Vzpomeň si, že pravděpodobnost, že v n hodech padne přesně k panen, je $\binom{n}{k} \cdot 2^{-n}$.

(1) Dokaž, že pro sudé n a nezáporné $k \leq n$ platí $\binom{n}{n/2} \geq \binom{n}{k}$, neboli nejpravděpodobnější počet padlých panen je $n/2$.

(2) Dokaž, že se s rostoucím $2n$ zmenšuje pravděpodobnost, že při $2n$ hodech padne přesně polovina panen.

Pohlédni však nyní na druhý obrázek, kde udáváme podíl počtu padlých panen a střední hodnoty $n/2$. Ačkoli se tento graf chová pro prvních přibližně 30 hodů docela nevyzpytatelně, poměr se následně rychle přiblíží hodnotě jedna, přesně jak napovídá intuice.

Rozdílné chování obrázků ukazuje i následující úloha. První podúloha se týká obrázku nalevo a její výsledek je přímým důsledkem toho, že mince nemá paměť. Druhá podúloha se týká obrázku napravo a naznačuje, že graf napravo má tendenci vracet se k hodnotě 1, jestliže se od ní vychýlil (prvních několik hodů je totiž stále méně a méně důležitých).

Úloha 2. Předpokládejme, že v prvních deseti hodech padly samé panny. Pro $n \geq 10$ označme $R_n = 10 + A_{11} + \cdots + A_n - n/2$, tedy rozdíl počtu panen a střední hodnoty po n hodech. Obdobně podíl po n hodech označme $P_n = \frac{10+A_{11}+\cdots+A_n}{n/2}$.

(1) Spočítej, že pro všechna $n \geq 10$ platí $\mathrm{E}(R_n) = \mathrm{E}(R_{10}) = 5$.

[1] Opravdu! Samotné nás to překvapilo.

[2] Jak by Ti Job jistě ochotně potvrdil.

(2) Spočítej, kolik je $\mathrm{E}(P_n)$, a ověř, že pro $m \geq n \geq 10$ je $\mathrm{E}(P_m) \leq \mathrm{E}(P_n)$.

S podobným jevem se mimochodem můžeš setkat i za jiných okolností a obvykle se označuje jako regrese k průměru. Nad jeho jinými variantami se můžeš zamyslet v následující úloze.

Úloha 3. (na zamyšlení, nemusíš nic počítat)

(1) Jistá soutěž probíhá ve dvou soutěžních dnech, přičemž každý den účastníci řeší tři úlohy. Předpokládejme, že každý soutěžící má pro každou úlohu nezávisle na ostatních poloviční pravděpodobnost, že ji vyřeší. Rozmysli si, že ti, kteří první den vyřešili alespoň dvě úlohy, budou po druhém dni pravděpodobně umístěni hůře než po tom prvním.

(2) Jistý učitel zpozoroval, že pozitivní motivace nefunguje tak dobře jako ta negativní. Pokud se žákovi povedl test výjimečně dobře, učitel jej pochválil. I přesto se příští test obvykle nepovedl tak dobře. Pokud se test žákovi vůbec nepovedl, učitel ho seřval. Příští test se žákovi obvykle povedl lépe. Máš pro to nějaké jiné vysvětlení?

(3) Bylo vypozorováno, že v roce 2017 se v jistých zatáčkách odehrálo hodně bouraček. Do zatáček byly nainstalovány bezpečnostní kamery a v příštím roce v těchto zatáčkách již k tolika haváriím nedošlo. Znamená to nutně, že bezpečnostní kamery zde přispěly k bezpečnějšímu provozu?

Vraťme se k obrázku podílu napravo. Fakt, že se graf na obrázku (s velkou pravděpodobností) drží poblíž hodnoty jedna, se nazývá **zákon velkých čísel**. Tento zákon doopravdy neplatí jen pro férové mince, ale třeba i pro házení falešnou mincí nebo kostkou – a obecně pro nezávislá opakování jakéhokoliv pokusu. My si dokážeme jeho jistou verzi, ale nejprve si povězme, proč je tento fakt vůbec zajímavý.

Zákon velkých čísel potvrzuje naši intuici, že se pravděpodobnost chová předvídatelně, pokud daný pokus opakujeme mnohokrát (jestliže bychom si hodili mincí jen třikrát, nic zajímavého bychom asi nevypozorovali). Ve druhém díle jsme si například jako součást naší představy o tom, co je střední hodnota, pověděli, že se vyplatí hrát hry, ve kterých je střední hodnota výdělku kladná. Ne vždy je to ale tak jednoduché – hrál(a) bys třeba hru, ve které vyhraješ 6 000 Kč s pravděpodobností 90 % a prohraješ 50 000 Kč s pravděpodobností 10 %? I když je střední hodnota výdělku kladná, zdá se zbytečné riskovat své týdenní kapesné kvůli almužně 6 000 Kč. Úplně jiná situace je, pokud bys měl(a) odehrát pět tisíc her, přičemž v každé vyhraješ 1,2 Kč s pravděpodobností 90 % a prohraješ 10 Kč s pravděpodobností 10 %. Pro dostatečné množství pokusů už se prostě můžeme spolehnout na to, že vyhrajeme přibližně 90 % her a vyděláme si. To proto, že platí zákon velkých čísel.

Rozmysleme si, jak formulovat zákon velkých čísel o něco přesněji než „graf na obrázku bude blízko jedné“. Zabýváme se náhodnou veličinou $A = A_1 + A_2 + \cdots + A_n$ a speciálně nás zajímá poměr $\frac{A}{n/2}$. Pravděpodobnost, že bude tento podíl přesně 1, je mizivá, ale můžeme si zvolit interval, řekněme od 0,9 do 1,1, a pak se ptát, jaká je pravděpodobnost jevu $0{,}9 < \frac{A}{n/2} < 1{,}1$. Jestliže dokážeme, že pravděpodobnost tohoto jevu je velká, řekněme alespoň 90 %, můžeme si gratulovat.

Problém je, že tyto podmínky obecně neplatí. Třeba pro $n = 5$ může veličina $\frac{A_1+A_2+\cdots+A_5}{5/2}$ nabývat hodnot $0, \frac{2}{5}, \frac{4}{5}, \frac{6}{5}, \frac{8}{5}$ a $\frac{10}{5}$. Jev $0{,}9 < \frac{A_1+A_2+\cdots+A_5}{5/2} < 1{,}1$ tak nenastane nikdy! Buď je tento podíl nejvýše 0,8, nebo alespoň 1,2. Problém je v tom, že zákon velkých čísel je zákon pro velká čísla, ne pro všechna čísla. Můžeme tak doufat, že podmínky budou s velkou pravděpodobností splněny, pokud si mincí hodíme aspoň, řekněme, tisíckrát, ale je mylné očekávat, že poměr $\frac{A}{n/2}$ bude s velkou pravděpodobností blízko jedné i pro malá n.

Později si opravdu dokážeme, že pro alespoň tisíc pokusů bude poměr $\frac{A}{n/2}$ v intervalu mezi 0,9 a 1,1 s pravděpodobností alespoň 90 %. Důkaz bude obecný a sám (sama) si pak budeš moci rozmyslet, že když zvětšíme počet pokusů, můžeme třeba zúžit interval nebo zvýšit pravděpodobnost, že $\frac{A}{n/2}$ bude ležet po n pokusech v daném intervalu.

Jak vidíš, už jenom formulace principu, který vidíme na obrázku napravo, je nesnadný úkol, proto Ti doporučujeme si ještě jednou přečíst poslední úsek této kapitoly.

Markovova nerovnost

Abychom nějakou verzi zákona velkých čísel dokázali, musíme si nejprve ukázat pravděpodobnostní nerovnost, kterou k důkazu použijeme. Zatím jedinou pravděpodobnostní nerovností, kterou známe, je fakt, že pravděpodobnost sjednocení jevů je nejvýš součet jednotlivých pravděpodobností. S tou si již ale nyní nevystačíme.

Ukážeme si nerovnost, která říká, že pokud nějaká náhodná veličina nabývá nezáporných hodnot, tak jen s malou pravděpodobností může nabývat hodnot podstatně větších než její střední hodnota. To zní docela složitě, ale myšlenka je jednoduchá. Představ si, že na dlouhém drátu napnutém mezi dvěma sloupy sedí hejno špačků. Řekněme, že průměrná vzdálenost špačka od levého sloupu je 10 metrů. Kolik špačků může sedět ve vzdálenosti aspoň 20 metrů? Nejvýše polovina! Pokud by totiž více než polovina špačků seděla ve vzdálenosti alespoň 20 metrů, jen díky této polovině by průměrná vzdálenost byla větší než 10 metrů. Lepší odhad než jedna polovina však dostat nemůžeme, neboť se může stát, že polovina špačků sedí ve vzdálenosti 0 od levého sloupu, zatímco druhá polovina ve vzdálenosti 20 metrů. Podobnou úvahu můžeme udělat i pro náhodné veličiny. Získaná nerovnost se jmenuje podle Andreje Markova[3].

Tvrzení. (Markovova nerovnost) *Nechť X je náhodná veličina definovaná na konečném pravděpodobnostním prostoru Ω nabývající pouze nezáporných hodnot. Potom pro libovolné $c > 0$ platí*

$$P\big(X \geq c \cdot E(X)\big) \leq \frac{1}{c}.$$

Tedy stejně jako maximálně polovina špačků může sedět dále než dvojnásobek průměrné vzdálenosti, může nezáporná náhodná veličina s nanejvýš třetinovou pravděpodobností nabývat hodnot větších než trojnásobek střední hodnoty. Povšimni si, že se v nerovnosti vyskytuje jednou „$\geq$“ a jednou „$\leq$“. Nejprve se „$\geq$“ vyskytuje uvnitř pravděpodobnosti, kde definuje jev „X nabývá velké hodnoty“. Následně „$\leq$“ říká, že pravděpodobnost tohoto jevu je malá.

Důkaz. Pokud by tomu tak nebylo, měl by jev $X \geq c\mathrm{E}(X)$ pravděpodobnost ostře větší než $\frac{1}{c}$. Jenže pokud nezáporná náhodná veličina nabývá hodnoty alespoň $c\mathrm{E}(X)$ s pravděpodobností vyšší než $\frac{1}{c}$, musí už její střední hodnota $\mathrm{E}(X)$ být vyšší než $c\mathrm{E}(X) \cdot \frac{1}{c} = \mathrm{E}(X)$, což je nesmysl.

Úloha 4. Rozmysli si:

(1) Kdy v Markovově nerovnosti nastává rovnost?
(2) Najdi příklad náhodné veličiny nabývající záporných hodnot, pro kterou nerovnost neplatí.

Úloha 5. Mediánem $m(X)$ náhodné veličiny X myslíme libovolné číslo, pro které platí, že $\mathrm{P}(X \geq m(X)) \geq \frac{1}{2}$ a $\mathrm{P}(X \leq m(X)) \geq \frac{1}{2}$. Například pro počet panen při jednom hodu mincí je mediánem libovolné číslo od 0 do 1 a pro hod férovou devítistěnnou kostkou (nabývá hodnot od jedné do devíti) je mediánem pouze číslo 5.

(1) Rozmysli si, že pro každou veličinu X existuje vždy aspoň jeden medián.
(2) Dokaž, že pro libovolný medián platí $m(X) \leq 2\mathrm{E}(X)$.

Úloha 6. Ve Wonkově továrně na čokoládu je na výběr ze 320 příchutí čokolády. Při procházce továrnou si každé z pěti dětí smělo ochutnat 160 příchutí. Protože bylo na spěch, každé dítě si

---

[3]Andrej Markov (1978–) je ruský hokejista, který v roce 2018 dopomohl svému týmu Ak Bars Kazaň k vítězství Gagarinova poháru. Jeho jmenovec Andrej Markov (1856–1922), žák P. Čebyševa, je v teorii pravděpodobnosti znám zejména díky tzv. Markovovým řetězcům, pomocí kterých se modelují nejrůznější pravděpodobnostní procesy.

vybralo náhodných 160 příchutí nezávisle na ostatních dětech. Dokaž, že příchutí, které si žádné dítě nevybralo, je alespoň dvacet s pravděpodobností nejvýše jedna polovina.

Vraťme se nyní k zákonu velkých čísel. Zajímá nás náhodná veličina $A = A_1 + A_2 + \cdots + A_n$, kde každé A_i odpovídá jednomu hodu mincí. Speciálně bychom rádi odhadli pravděpodobnost jevu $\frac{A}{n/2} \geq 1{,}1$, který také můžeme zapsat jako $A \geq 0{,}55n$. Díky symetrii panen a orlů je pravděpodobnost, že $A \leq 0{,}45n$, stejná, takže stačí spočítat, že pravděpodobnost jevu $A \geq 0{,}55n$ je malá. Pravděpodobnost, že A neleží v intervalu mezi $0{,}45n$ a $0{,}55n$, pak bude dvojnásobná (a tedy pořád malá), neboť se jedná o disjunktní jevy.

Pravděpodobnost jevu $A \geq 0{,}55n$ odhadneme pomocí Markovovy nerovnosti. Máme $\mathrm{E}(A) = \frac{n}{2}$, takže Markovova nerovnost říká, že

$$\mathrm{P}(A \geq 0{,}55n) = \mathrm{P}(A \geq 1{,}1 \cdot \mathrm{E}(A)) \leq \frac{1}{1{,}1} \doteq 0{,}9.$$

Pravděpodobnost, že A neleží v intervalu mezi $0{,}45n$ a $0{,}55n$, je tak nejvýše $2 \cdot \frac{1}{1{,}1} \doteq 1{,}8$. Ouha, aplikace Markovovy nerovnosti tedy skončila naprostým fiaskem. Chtěli jsme dokázat, že pravděpodobnost tohoto jevu malá, ale pouze jsme zjistili, že je menší než 1,8. Kde se stala chyba?

Problém je v tom, že kdybychom takto jednoduše pomocí Markovovy nerovnosti dokázali něco pro veličinu A odpovídající opakovanému házení mincí, dokázali bychom to samé i pro veličinu B, která nabývá hodnoty 0 a n se stejnou pravděpodobností. Pro ni ale nic takového neplatí – průměrná hodnota je $n/2$, ale pravděpodobnost, že při n hodech získáme hodnotu mezi $0{,}45n$ a $0{,}55n$, je nulová. V následující kapitole se proto zamyslíme nad nezávislostí, protože důležitý rozdíl mezi A a B je právě ten, že jednotlivé hody mincí, pomocí kterých definujeme A, jsou nezávislé. Následně si pro náhodné veličiny zavedeme další parametr, tzv. rozptyl, který už veličiny A a B spolehlivě rozliší.

Nezávislost náhodných veličin

Už v prvním díle jsme si pověděli o nezávislosti jevů. Protože ale nyní pracujeme spíše s veličinami než s konkrétními jevy, hodí se rozšířit si i naši definici nezávislosti. Co znamená, že na sobě dvě veličiny X, Y – tedy dva pokusy – nezávisí? Intuitivně by to mělo znamenat, že známe-li výsledek jednoho pokusu, neovlivní to možné výsledky toho druhého. Řečeno jazykem podmíněné pravděpodobnosti, pro libovolné x a y by mělo platit $\mathrm{P}(Y = y \mid X = x) = \mathrm{P}(Y = y)$. Z prvního dílu víme, že tento výraz můžeme přepsat jako $\mathrm{P}(X = x \cap Y = y) = \mathrm{P}(X = x) \cdot \mathrm{P}(Y = y)$, tedy jevy $X = x$ a $Y = y$ jsou nezávislé. Přesně tak se definuje nezávislost veličin.

Definice. Mějme pravděpodobnostní prostor Ω a na něm dvě náhodné veličiny X a Y takové, že X nabývá hodnot $x_1, x_2, \ldots, x_m$ a Y nabývá hodnot $y_1, y_2, \ldots, y_n$. Řekneme, že X a Y jsou nezávislé, pokud jsou pro všechna $1 \leq i \leq m$ a $1 \leq j \leq n$ jevy $X = x_i$ a $Y = y_j$ nezávislé.

Příkladem je třeba hod dvěma férovými kostkami. V takovém případě pracujeme s pravděpodobnostním prostorem obsahujícím 36 možných výsledků, přičemž každý nastane se stejnou pravděpodobností. Můžeme se nyní ptát na to, kolik ok padlo na první a kolik na druhé kostce, a zavést k tomuto účelu náhodné veličiny X a Y.

Spočítáme, že pro libovolné $1 \leq i \leq 6$ platí

$$\mathrm{P}(X = i) = \mathrm{P}(X = i \cap Y = 1) + \cdots + \mathrm{P}(X = i \cap Y = 6) = 6 \cdot \frac{1}{36} = \frac{1}{6},$$

neboť to, že nastal jev $X = i$, odpovídá šestici elementárních jevů. Stejně tak pro libovolné $1 \leq j \leq 6$ dostaneme $\mathrm{P}(Y = j) = \frac{1}{6}$. Pro libovolné i a j proto platí $\mathrm{P}(X = i \cap Y = j) = \frac{1}{36}$.

Zároveň $\mathrm{P}(X = i) \cdot \mathrm{P}(Y = j) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$. Všechny dvojice jevů $X = i$ a $Y = j$ jsou tedy nezávislé, a můžeme tak směle tvrdit, že i náhodné veličiny X a Y jsou nezávislé.

Úloha 7. Na stole leží dvě modré, jedna červená a jedna zelená propiska. Vybereme si dvě náhodně. Nechť X je náhodná veličina značící počet vybraných modrých propisek a Y značí počet vybraných červených propisek. Jsou X a Y nezávislé?

Řešení. Pravděpodobnostní prostor odpovídá šesti možným výsledkům. Dáme-li si tu práci, můžeme si nakreslit tabulku a vyplnit do ní pravděpodobnosti všech jevů $X = i \cap Y = j$.

Y \ X	0	1	2
1	$\frac{1}{6}$	$\frac{2}{6}$	0
0	0	$\frac{2}{6}$	$\frac{1}{6}$

Pravděpodobnosti jevů $X = i$ dostaneme jako součty po sloupcích (vyjde $\frac{1}{6}, \frac{4}{6}, \frac{1}{6}$) a pravděpodobnosti $Y = j$ jako součty po řádcích (vyjde $\frac{1}{2}, \frac{1}{2}$). Aby byly dané veličiny nezávislé, musela by hodnota na každém políčku být rovna součinu součtu odpovídajícího řádku a součtu odpovídajícího sloupce. To ale neplatí, třeba pro jev $X = 0 \cap Y = 0$ máme $0 \neq \frac{1}{6} \cdot \frac{1}{2}$.

Ačkoli nám trvalo tři díly seriálu, než jsme si ukázali definici nezávislých náhodných veličin, je to jeden z konceptů, který často bereme jako samozřejmý. Když si hodíme kostkou, zatímco kamarád si lízne kartu ze zamíchaného balíčku, výsledky na sobě samozřejmě „nezávisí“. Abychom tento koncept mohli pořádně popsat, zavedli jsme si postupně pravděpodobnostní prostory, které obsahují všechny možné výsledky pokusu (to jsou elementární jevy, v tomto případě dvojice stěna kostky a karta). Následně jsme definovali nezávislost, ale jen pro jevy. Poté jsme zavedli náhodné veličiny, které odpovídají jednotlivým pokusům, a teprve nyní jsme všechno spojili dohromady. Plodem našeho snažení je to, že nyní naprosto přesně víme, co to znamená, že jsou dané pokusy nezávislé, a můžeme proto s nimi snáze pracovat. Pro všechny nezávislé veličiny kupříkladu platí, že součin jejich středních hodnot je střední hodnotou jejich součinu.

Tvrzení. *Nechť X a Y jsou nezávislé náhodné veličiny. Potom platí, že $E(X \cdot Y) = E(X) \cdot E(Y)$.*

Než se pustíme do důkazu, ukažme si použití na následujícím příkladu. Uvažme pravděpodobnostní prostor o velikosti 24 odpovídající denním hodinám a veličiny X, Y, kde X pro každý interval o délce jedné hodiny říká, kolik vlaků včera projelo jihlavským nádražím během tohoto intervalu. Dále Y říká, kolik průměrně vagónů měly tyto vlaky. Pokud tedy mezi půlnocí a jednou hodinou ranní projel jeden vlak se dvěma a jeden vlak se čtyřmi vagóny, bude pro odpovídající elementární jev našeho prostoru nabývat X hodnoty 2 a Y hodnoty 3.

Veličina XY říká, kolik vagónů projelo nádražím v daný časový interval. Víme-li, že $\mathrm{E}(X) = 20$ a $\mathrm{E}(Y) = 3$, dává smysl, že $\mathrm{E}(XY) = 60$. To ale platí jen za předpokladu, že jsou veličiny X a Y nezávislé!

Jestliže každou z dvanácti nočních hodin projede stanicí jeden vlak s jedním vagónem, zatímco každou denní hodinu to bude 39 vlaků po pěti vagónech, bude platit $\mathrm{E}(X) = 20$ a $\mathrm{E}(Y) = 3$. Jenže průměrný počet vagónů bude $\frac{12 \cdot 1 \cdot 1 + 12 \cdot 39 \cdot 5}{24} = 98$, což je mnohem více než 60. Rozmysli si, že v tomto případě nejsou X a Y nezávislé.

Důkaz. Nazvěme $H(X)$ a $H(Y)$ obory hodnot X a Y. Na jednu stranu pro veličinu $X \cdot Y$ platí

$$\mathrm{E}(X \cdot Y) = \sum_{x \in H(X)} \sum_{y \in H(Y)} \mathrm{P}(X = x \cap Y = y) xy,$$

protože vždy s pravděpodobností $\mathrm{P}(X = x \cap Y = y)$ bude $X = x$ a $Y = y$, pročež $XY = xy$. Na druhou stranu zatneme zuby a uvědomíme si, že

$$\mathrm{E}(X)\mathrm{E}(Y) = \left(\sum_{x \in H(X)} \mathrm{P}(X = x)x\right) \cdot \left(\sum_{y \in H(Y)} \mathrm{P}(Y = y)y\right) =$$
$$= \sum_{x \in H(X)} \sum_{y \in H(Y)} \mathrm{P}(X = x)x \cdot \mathrm{P}(Y = y)y,$$

kde jsme roznásobili dvě závorky. Tyto dva výrazy jsou ale stejné, protože z nezávislosti máme $\mathrm{P}(X = x \cap Y = y) = \mathrm{P}(X = x) \cdot \mathrm{P}(Y = y)$ pro všechna x a y.

Úloha 8. Kuba dluží Honzovi obnos, jehož hodnotu již oba dávno zapomněli, a tak se rozhodli, že mu ho Kuba splatí následujícím způsobem.

(1) Kuba si dvakrát hodí kostkou. Počet ok při prvním hodu postupně určí jednu z šesti hodnot $1, 2, 5, 10, 20, 50$. Počet ok při druhém hodu pak určí, kolik mincí této hodnoty dá Kuba Honzovi. Jaká je střední hodnota zaplacené částky?

(2) Honzu napadlo, že aby se zbytečně neošoupala kostka, bude stačit, když si Kuba hodí jen jednou, a tato hodnota určí částku i počet mincí (tři oka znamenají třikrát pět korun). Jaká je střední hodnota teď?

Úloha 9.

(1) Náhodné veličiny $X_1, X_2, \ldots, X_n$ jsou nezávislé, pokud pro všechna $x_1, x_2, \ldots, x_n$ platí, že n-tice jevů $X_1 = x_1$, $X_2 = x_2, \ldots, X_n = x_n$ je nezávislá. Ověř, že pro nezávislé veličiny platí $\mathrm{E}(X_1 \cdot X_2 \cdots X_n) = \mathrm{E}(X_1) \cdot \mathrm{E}(X_2) \cdots \mathrm{E}(X_n)$.

(2) David se rozhodl podvádět v jisté hře, ve které se desetkrát hodí kostkou (jednotlivé hody jsou nezávislé) a následně se spočte a) součet b) součin jednotlivých počtů ok. Součet ok na protilehlých stranách kostky je vždy sedm. David se naučil házet kostkou tak, že si vždy může vybrat dvojici protějších stěn a hodit tak, že tyto stěny zaručeně nepadnou (ostatní stěny padnou s pravděpodobností $\frac{1}{4}$). Jak má házet kostkou, aby maximalizoval střední hodnotu a) součtu b) součinu ok?

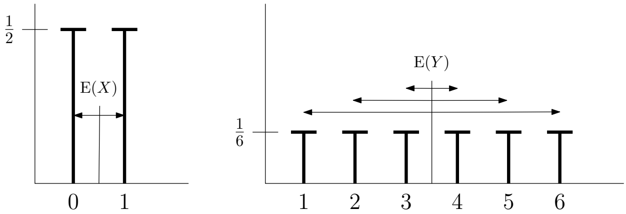
Rozptyl

Tragické fiasko s aplikací Markovovy nerovnosti na veličinu A odpovídající hodu n mincemi ukázalo, že znát střední hodnotu dané veličiny nestačí k tomu, abychom rozlišili A od veličiny B, která nabývá hodnoty 0 a n se stejnou pravděpodobností. V této kapitole si povíme o **rozptylu**, což je parametr náhodné veličiny, který nám říká, jak moc jsou hodnoty dané veličiny vzdálené od průměru, a umožní nám tak tyto veličiny rozlišit.

Definice. Rozptylem náhodné veličiny X myslíme střední hodnotu veličiny $(X - \mathrm{E}(X))^2$.

Rozptyl se značí $\mathrm{Var}(X)$ (značení pochází z anglického *variance*), takže můžeme psát $\mathrm{Var}(X) = \mathrm{E}((X - \mathrm{E}(X))^2)$. Ačkoli tento výraz vypadá kuriózně, dává dobrý smysl. Předně, $(X - \mathrm{E}(X))^2$ je zase náhodná veličina. Známe-li střední hodnotu původní veličiny X (na obrázku vyznačená svislou čarou), pro výpočet rozptylu si nejdříve představíme, že od všech možných hodnot X odečteme $\mathrm{E}(X)$ (tím posuneme celý histogram o $\mathrm{E}(X)$). Dostaneme veličinu $X - \mathrm{E}(X)$, jejíž střední hodnota je nula. Nakonec možné hodnoty této veličiny umocníme na druhou a spočítáme střední hodnotu.

Číslo, které dostaneme, nám v jistém smyslu říká, jaká je průměrná vzdálenost hodnot veličiny X od její střední hodnoty. Nicméně se nejedná o průměr vzdáleností (tomu by odpovídala veličina $|X - \mathrm{E}(X)|$), ale o průměr druhých mocnin vzdáleností. Druhá mocnina zařídí, že může-li veličina nabývat hodnot hodně vzdálených od průměru, bude její rozptyl obrovský. Proto platí, že rozptyl pro hod mincí vyjde jen $\frac{1}{2} \cdot (0{,}5^2 + 0{,}5^2) = \frac{1}{4}$, protože výsledek je vždy blízko průměru. Naproti tomu rozptyl hodu kostkou je $\frac{1}{6}(2{,}5^2+1{,}5^2+0{,}5^2+0{,}5^2+1{,}5^2+2{,}5^2) = \frac{35}{12}$, neboť výsledky mohou být od průměru dále.



Naším cílem je nyní spočítat rozptyly dvou veličin A a B ze začátku tohoto dílu. Začněme s B. Abychom se při výpočtu rozptylu moc nenadřeli, rozmysli si nejprve dvě obecné vlastnosti rozptylu.

Úloha 10. Dokaž, že pro libovolná čísla a a b platí $\mathrm{Var}(X+a) = \mathrm{Var}(X)$ a $\mathrm{Var}(bX) = b^2\mathrm{Var}(X)$.

Jinými slovy, přičtením jedničky ke všem hodnotám náhodné veličiny se její rozptyl nezmění, a pokud hodnoty vynásobíme deseti, rozptyl se zvětší stokrát.

Z druhé vlastnosti plyne, že rozptyl B je n^2-násobek rozptylu veličiny indikující, zda v jednom hodu mincí padla panna. Už jsme spočítali, že rozptyl této veličiny je $\frac{1}{4}$; dostáváme $\mathrm{Var}(B) = \frac{1}{4} \cdot n^2$.

Následující tvrzení ukazuje jiný způsob, jak počítat rozptyl.

Tvrzení. *Pro náhodnou veličinu X platí $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$.*

Důkaz. Po rozepsání rozptylu z definice a roznásobení dostáváme rovnost

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2 - 2X\mathrm{E}(X) + \mathrm{E}(X)^2\right).$$

Na ni můžeme aplikovat linearitu střední hodnoty:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}\left(2X\mathrm{E}(X)\right) + \mathrm{E}\left(\mathrm{E}(X)^2\right).$$

Jakkoli je to na první pohled zvláštní, i $2X\mathrm{E}(X)$ a $\mathrm{E}(X)^2$ jsou náhodné veličiny definované na stejném pravděpodobnostním prostoru jako X: první z nich přiřadí elementárnímu jevu ω číslo $2\mathrm{E}(X) \cdot X(\omega)$, zatímco druhá prostě přiřadí všem jevům stejné číslo $\mathrm{E}(X)^2$.

Protože $2\mathrm{E}(X)$ je číslo a ne veličina, platí $\mathrm{E}\left(2X\mathrm{E}(X)\right) = \mathrm{E}\left(2\mathrm{E}(X) \cdot X\right) = 2\mathrm{E}(X) \cdot \mathrm{E}(X) = 2\mathrm{E}(X)^2$, a dále $\mathrm{E}\left(\mathrm{E}(X)^2\right) = \mathrm{E}(X)^2$. Dohromady tedy dostáváme

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - 2\mathrm{E}(X)^2 + \mathrm{E}(X)^2 = \mathrm{E}(X^2) - \mathrm{E}(X)^2.$$

Úloha 11. Spočítej $\mathrm{Var}(B)$ pomocí předchozího tvrzení.

Úloha 12. Dokaž, že pro libovolnou náhodnou veličinu X platí $\mathrm{E}(X^2) \geq \mathrm{E}(X)^2$. Kdy v této nerovnosti nastává rovnost?

Úloha 13. Mějme náhodnou veličinu X, která nabývá pouze hodnot z množiny $\{1, 2, \ldots, n\}$. Pro jakou největší hodnotu n je každá taková X jednoznačně určena (neboli pravděpodobnost každého jevu $X = i$ je pevně dána):

(1) svojí střední hodnotou?

(2) svojí střední hodnotou a rozptylem?

(3) svým rozptylem?

Pokusme se nyní spočítat rozptyl veličiny A. Na první pohled vůbec není jasné, jak na to. Pro počítání střední hodnoty přijde vhod její linearita, tedy fakt, že střední hodnota součtu je součet středních hodnot. Platí něco takového i pro rozptyl? Obecně ne, což si můžeš rozmyslet v následující úloze.

Úloha 14. Uvažme jeden hod férovou mincí (odpovídající pravděpodobnostní prostor má dva prvky) a veličinu P indikující, zda padla panna (tedy je rovná jedné, padla-li panna, a jinak nule), a veličinu O indikující, zda padl orel. Spočítej, že $\mathrm{Var}(O+O) > \mathrm{Var}(O) + \mathrm{Var}(O)$ a $\mathrm{Var}(O+P) < \mathrm{Var}(O) + \mathrm{Var}(P)$.

Klíčovou vlastností rozptylu je nicméně to, že pro nezávislé veličiny je součet jejich rozptylů opravdu rozptylem jejich součtu. To plyne přímo z toho, že pro nezávislé veličiny je $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$:

Tvrzení. *Jsou-li X a Y nezávislé náhodné veličiny, platí $Var(X+Y) = Var(X) + Var(Y)$.*

Důkaz. Rozepišme výraz $\mathrm{Var}(X+Y)$ s pomocí linearity střední hodnoty.

$$\begin{aligned}\mathrm{Var}(X+Y) &= \mathrm{E}\left((X+Y)^2\right) - \mathrm{E}(X+Y)^2 = \\ &= \mathrm{E}(X^2) + \mathrm{E}(Y^2) + 2\mathrm{E}(XY) - \left(\mathrm{E}(X)^2 + 2\mathrm{E}(X)\mathrm{E}(Y) + \mathrm{E}(Y)^2\right).\end{aligned}$$

Pro nezávislé náhodné veličiny platí $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$, takže se tyto členy odečtou a zbude nám rovnost $\mathrm{Var}(X+Y) = \mathrm{E}(X^2) + \mathrm{E}(Y^2) - \mathrm{E}(X)^2 - \mathrm{E}(Y)^2 = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Úloha 15. Zobecni předchozí tvrzení pro libovolný počet náhodných veličin, tedy dokaž, že máme-li náhodné veličiny $X_1, \ldots, X_n$ takové, že každá dvojice X_i, X_j $(i \neq j)$ je nezávislá, pak platí $\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \cdots + \mathrm{Var}(X_n)$.

Spočítat rozptyl náhodné veličiny A je nyní snadné. Máme

$$\mathrm{Var}(A) = \mathrm{Var}(A_1) + \mathrm{Var}(A_2) + \cdots + \mathrm{Var}(A_n) = n \cdot \mathrm{Var}(A_1) = \frac{1}{4}n.$$

Porovnáním s $\mathrm{Var}(B) = \frac{1}{4}n^2$ vidíme, že $\mathrm{Var}(A)$ je pro velké n mnohem menší než $\mathrm{Var}(B)$. To je přesně výsledek, který jsme čekali! Rozptyl veličiny B je obrovský, protože hodnoty, kterých B nabývá, jsou vždy daleko od průměru.

Jdeme do finále. Markovova nerovnost, kterou jsme již potkali, nám totiž překvapivě často o náhodné veličině X řekne více, pokud ji aplikujeme nikoliv přímo na X, nýbrž na veličinu $(X - \mathrm{E}(X))^2$, která měří „vzdálenosti hodnot X od průměru“. Dosazení tohoto výrazu do Markovovy nerovnosti je tak užitečným trikem, že výsledná nerovnost má své vlastní jméno po P. Čebyševovi[4].

Tvrzení. (Čebyševova nerovnost) *Pro libovolnou náhodnou veličinu X a $a > 0$ platí*

$$P\Big((X - E(X))^2 \geq a^2 Var(X)\Big) \leq \frac{1}{a^2}.$$

Důkaz. Vzhledem k tomu, že $(X - \mathrm{E}(X))^2$ nabývá pouze nezáporných hodnot a $\mathrm{Var}(X) = \mathrm{E}((X - \mathrm{E}(X))^2)$, můžeme aplikovat Markovovu nerovnost na veličinu $(X - \mathrm{E}(X))^2$ a a^2, čímž dostaneme dokazovanou nerovnost.

[4]Ruský matematik Pafnutij Lvovič Čebyšev (1821–1894) je kromě své nerovnosti znám také díky svým příspěvkům k teorii čísel: dokázal kupříkladu tzv. Bertrandův postulát, který tvrdí, že pro libovolné přirozené n se mezi n a $2n$ nachází alespoň jedno prvočíslo. V cizojazyčné literatuře vystupuje pod přezdívkami Chebyshev, Chebysheff, Chebychov, Chebyshov, Tchebychev, Tchebycheff, Tschebyschev, Tschebyschef, Tschebyscheff či Chebychev.

Ekvivalentně můžeme psát $\mathrm{P}\left(|X - \mathrm{E}(X)| \geq a\sqrt{\mathrm{Var}(X)}\right) \leq \frac{1}{a^2}$. Odmocnina z rozptylu se nazývá **směrodatná odchylka** a obvykle se značí σ.[5] Stejně jako rozptyl říká, jak moc jsou hodnoty veličiny vzdáleny od průměru. Dost možná ses s ní již setkal(a) ve fyzice. Pokud pětkrát změříš hustotu neznámého předmětu, pokaždé Ti vyjde trochu jiný výsledek kvůli různým nepřesnostem – měření proto můžeme popisovat náhodnou veličinou. Směrodatnou odchylku této veličiny pak obvykle odhadneš z naměřených dat pomocí jistého vzorečku z fyzikálních tabulek. Odchylka je představitelnější než rozptyl: pokud náhodnou veličinu X měříš v metrech, tak rozptyl má jednotku metry čtvereční, zatímco směrodatná odchylka je opět v metrech.

Podobně jako Markovova nerovnost tvrdí, že náhodná veličina nabývá hodnot vyšších než a-násobek střední hodnoty s pravděpodobností nejvýše $1/a$, Čebyševova nerovnost prohlašuje, že pravděpodobnost, že náhodná veličina je vzdálena od své střední hodnoty o a-násobek směrodatné odchylky, je nejvýše $1/a^2$. Proto směrodatná odchylka dává představu, jak moc se jednotlivá měření obvykle liší od průměru (tedy střední hodnoty).

Úloha 16. Dokaž, že pro libovolnou náhodnou veličinu X nabývající nezáporných hodnot platí $\mathrm{P}(X = 0) \leq \frac{\mathrm{Var}(X)}{\mathrm{E}(X)^2}$.

Jako velké finále se nyní naposled vraťme k veličině A popisující hod n mincemi. Máme $\mathrm{E}(A) = \frac{n}{2}$ a $\mathrm{Var}(A) = n/4$. Čebyševova nerovnost tak po odmocnění podmínky dává

$$\mathrm{P}\left(|A - \mathrm{E}(A)| \geq a\sqrt{n/4}\right) \leq \frac{1}{a^2}.$$

Dosadíme-li z estetických důvodů $a = 2c\sqrt{n}$, dostáváme

$$\mathrm{P}\left(|A - \mathrm{E}(A)| \geq cn\right) \leq \frac{1}{4c^2 n}.$$

Jestliže za c nyní dosadíme 0,05 a zvolíme $n = 1000$, dostaneme, že počet panen se bude od průměru lišit o více než $0{,}05n = 0{,}1 \cdot \frac{n}{2}$ s pravděpodobností nejvýše $\frac{1}{4 \cdot 0{,}05^2 \cdot 1000} = \frac{400}{4 \cdot 1000} = \frac{1}{10}$, přesně jak jsme slibovali v první kapitole. S rostoucím n bude tato pravděpodobnost dále klesat. Můžeme si vybírat různá c a zkoušet různé přesnosti, pro každé c nicméně výraz $\frac{1}{4c^2 n}$ bude klesat a nakonec bude zanedbatelně malý.

Postup, který jsme aplikovali pro házení férovou mincí, samozřejmě funguje i pro házení falešnou mincí, kostkou, měření výšky lidí, které potkáš na ulici, … Vždy platí, že se zvětšujícím se počtem pokusů se pravděpodobnost chová předvídatelněji a ukázněněji. S pomocí nově nabytých poznatků dokonce můžeme zformulovat obecnější verzi našeho tvrzení:

Tvrzení. (Verze zákona velkých čísel) *Nechť $X_1, \ldots, X_n$ jsou po dvou nezávislé náhodné veličiny, které mají všechny stejnou střední hodnotu i rozptyl. Potom pro jejich součet $Y = X_1 + X_2 + \cdots + X_n$ a libovolné kladné c platí nerovnost*

$$P\left(|Y - E(Y)| \geq cn\right) \leq \frac{Var(X_1)}{c^2 n}.$$

Ačkoli toto tvrzení kvůli mnoha písmenkům může působit složitě, nejedná se o nic jiného než o výpočet, který jsme udělali pro házení mincí (vzoreček nahoře dostaneme, pokud si vzpomeneme, že rozptyl jednoho hodu mincí je $\frac{1}{4}$).

Nyní si můžeme konečně s úlevou oddechnout, neboť jsme potvrdili, že intuitivní představa o povaze pravděpodobnosti je správná! Cesta k tomuto důkazu byla sice poměrně dlouhá a náročná, ale zase jsme díky ní narazili na spoustu užitečných a zajímavých konceptů. Dříve než si budeš moci

[5] Uvedený symbol je malé řecké sigma; jeho velkou variantu už znáš jako sumační symbol.

vychutnat zasloužený odpočinek, rádi bychom Ti ještě v hvězdičkových kapitolách představili pár zajímavostí, které souvisí s tím, čím jsme se v tomto díle zabývali.

Zvěřinec náhodných veličin*

Jeden z pojmů, který jsme doteď nezaváděli, je rozdělení náhodné veličiny. Rozdělení je histogram, kterým si popisujeme, jak často daná veličina nabývá jednotlivých hodnot. Jedná se tak o jakýsi prototyp náhodných veličin, u kterých už nás vlastně vůbec nezajímá původní pravděpodobnostní prostor. Zde jsme pro Tebe uspořádali katalog několika důležitých rozdělení. Přidali jsme k nim také některé jejich vlastnosti, přičemž s většinou z nich ses už někde v seriálu potkal(a). Pokud ne, zkus si je sám (sama) odvodit!

Alternativní (Bernoulliho) rozdělení – náhodná veličina nabývající s pravděpodobností p hodnoty 1 a jinak nabývající hodnoty nula. Veličina odpovídá hodu falešnou mincí a my ji z druhého dílu známe jako tzv. indikátor. Pro alternativní náhodnou veličinu X platí $\mathrm{E}(X) = p$ a $\mathrm{Var}(X) = p - p^2$.

Binomické rozdělení – náhodná veličina značící počet panen při hodu n mincemi, přičemž na každé z nich padne panna s pravděpodobností p. Platí $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Dále máme $\mathrm{E}(X) = n \cdot p$ (z linearity střední hodnoty) a $\mathrm{Var}(X) = n(p - p^2)$ (X je součet nezávislých veličin).

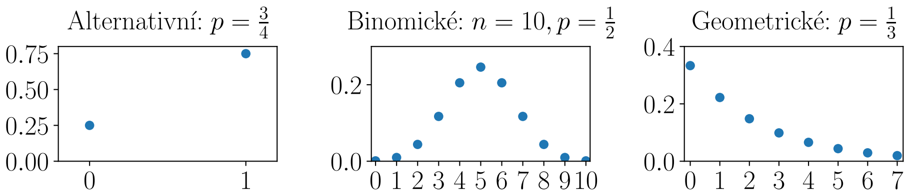
Pro $p = \frac{1}{2}$ jsou hodnoty binomického rozdělení přesně kombinační čísla $\binom{n}{k}$ vynásobená 2^{-n}. Proto binomické rozdělení úzce souvisí s kombinačními čísly, tzv. Pascalovým trojúhelníkem nebo tzv. binomickou větou, se kterými ses již možná někde setkal(a). Toto rozdělení nás provázelo po celý tento díl a víme o něm již mnoho. Jestliže si hodíme n mincemi pro sudé n, nejpravděpodobnější výsledek je, že padne $n/2$ panen. Přesto je pravděpodobnost tohoto jevu čím dál menší, když zvětšujeme počet hodů (viz první úlohu). Čebyševova nerovnost říká, že pravděpodobnost, že binomické rozdělení bude nabývat hodnot vzdálených od $n/2$ o více než, řekněme, tři směrodatné odchylky (tedy $3 \cdot \sqrt{\frac{n}{4}}$), je jen $\frac{1}{9}$. Intuice, že počet panen se od střední hodnoty $n/2$ odchýlí o řádově odmocninu z n, vysvětluje rozdílné chování dvou grafů ze začátku tohoto dílu. Zatímco v prvním grafu se křivka postupně vzdaluje od nuly ($\sqrt{n}$ postupně roste), ve druhém grafu se postupně přibližuje k jedničce ($\sqrt{n}/n$ je fakt malé pro velká n).

Úloha 17. (těžká) Mějme sudé n. Nejprve si rozmysli, že $\binom{n}{n/2}$ je určitě menší než 2^n, ale větší než $2^n/(n+1)$. Předchozí úvaha o Čebyševově nerovnosti doopravdy dává lepší představu o tom, jak velké je toto kombinační číslo.

(1) Pomocí Čebyševovy nerovnosti dokaž, že $\binom{n}{n/2} \geq 2^n \cdot \frac{3}{(8\sqrt{n}+4)}$.

(2) Naopak využij toho, že platí $\binom{n}{n/2} \cdot 2^{-n} = (1 - \frac{1}{n})(1 - \frac{1}{n-2}) \cdots (1 - \frac{1}{2})$ a zároveň $(1 - \frac{1}{n+1})(1 - \frac{1}{n}) \cdots (1 - \frac{1}{3})(1 - \frac{1}{2}) = \frac{1}{n+1}$, abys dokázal(a), že $\binom{n}{n/2} \leq \frac{2^n}{\sqrt{n+1}}$.

Geometrické rozdělení – náhodná veličina odpovídající počtu hodů, než na falešné minci padne panna. Padne-li panna s pravděpodobností p, máme $\mathrm{P}(X = k) = (1-p)^{k-1} \cdot p$ a $\mathrm{E}(X) = \frac{1}{p}$ (viz minulý díl). Jedná se o jedno z mála rozdělení s vlastností $\mathrm{P}(X \geq a + b \,|\, X \geq a) = \mathrm{P}(X \geq b)$. Jinými slovy, mince nemá paměť (vzpomeň na gamblerův omyl).



Rovnoměrné rozdělení – náhodná veličina odpovídající tomu, že zlomíme hůl na náhodném místě tak, že pro libovolný úsek hole (třeba pro první třetinu hole) je pravděpodobnost, že zlom bude někde v tomto úseku, úměrná jeho délce.

Spojitých rozdělení kromě toho rovnoměrného existuje mnoho a obecně se popisují funkcí, které se říká hustota pravděpodobnosti. V případě hole délky 1 by hustotou byla funkce rovná jedné pro $0 \leq x \leq 1$ a jinak rovná nule. Pravděpodobnost, že si vybereš dané x z nějakého intervalu, pak odpovídá obsahu pod grafem této funkce, což je v případě hole to samé jako délka intervalu. Opravdová hůl se láme snáze poblíž prostředku, což můžeme popsat funkcí, která je od 0 do $\frac{1}{2}$ rostoucí a následně od $\frac{1}{2}$ do 1 klesající. Stejně dlouhé intervaly pak obecně odpovídají různým obsahům, a tedy i pravděpodobnostem. Detaily o hustotě pravděpodobnosti jsou nicméně daleko nad rámec seriálu.

Jako varovný příklad toho, že se spojité náhodné veličiny chovají složitěji, uvádíme úlohu, se kterou přišel na konci 19. století francouzský matematik Joseph Bertrand[6].

Úloha 18. (Bertrandův paradox) Uvažme kružnici k se středem S a do ní vepsaný rovnostranný trojúhelník se stranou délky 1. Jaká je pravděpodobnost, že bude délka náhodně zvolené tětivy kružnice větší než 1?

(1) Zvolit si náhodnou tětivu znamená zvolit rovnoměrně a nezávisle dva body na obvodu kružnice a spojit je. Na základě toho si rozmysli, že pravděpodobnost, že dostaneme tětivu delší než 1, je jedna třetina.

(2) Zvolit si náhodnou tětivu ale také můžeme tak, že si zvolíme náhodný bod A uvnitř kružnice a ten následně interpretujeme jako střed tětivy kolmé na úsečku SA. Rozmysli si, pro které body je daná tětiva delší než 1, a porovnáním obsahů z toho vyvoď, že daná pravděpodobnost je jedna čtvrtina.

(3) Ještě jiný způsob, jak zvolit bod A z předešlého odstavce, je náhodně si zvolit, pod jakým úhlem se vydáš ze středu kružnice S a v jaké vzdálenosti se zastavíš. První je náhodné číslo od 0 do 2π, druhé je náhodné číslo od 0 do poloměru kružnice. V takovém případě si rozmysli, že vyjde jedna polovina.

(4) Který z předchozích přístupů je správný? :-)

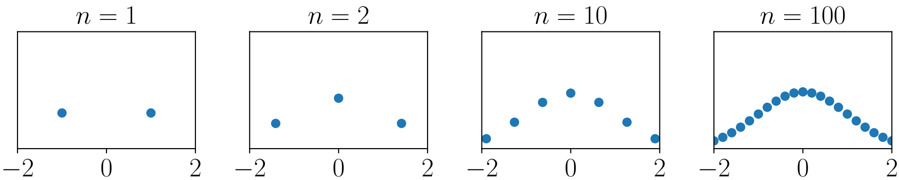
Centrální limitní věta*

Statistika nuda je, má však cenné údaje.
Zdeněk Svěrák, Jaroslav Uhlíř

To nejlepší nakonec. Podívej se na následující obrázky zobrazující binomické rozdělení s $p = \frac{1}{2}$ pro $n = 1, 2, 10$ a 100. Třeba druhý obrázek odpovídá třem možným počtům panen po hodech dvěma mincemi, které nastanou postupně s pravděpodobnostmi $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$. Obecně je vidět, že rozdělení se

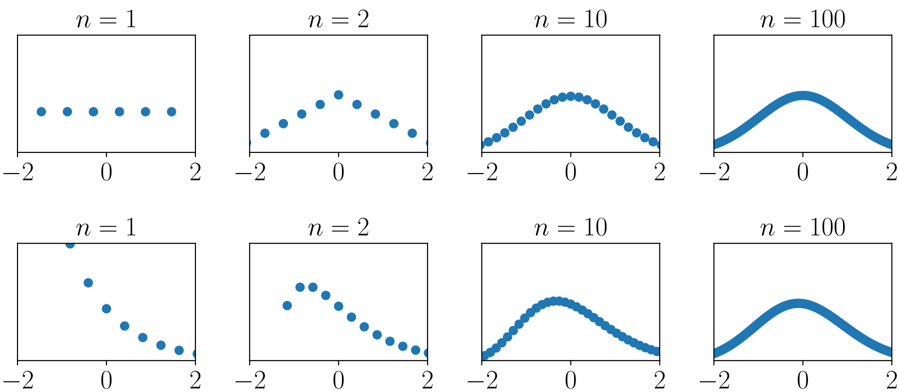
[6] Joseph Bertrand (1,98 m) je basketbalista, kterého můžeš znát díky jeho výkonům v týmu Drážďanští Titáni. Jeho jmenovec Joseph Bertrand (1822–1900) je znám zejména díky tzv. Bertrandovu postulátu, který dokázal až P. Čebyšev.

čím dál tím více podobají jisté křivce, která vypadá jako kopeček[7]. To není tak docela pravda, abychom dosáhli tohoto efektu, museli jsme postupně měnit měřítko na osách x a y. Po vhodném přeškálování ale opravdu platí, že se binomické rozdělení s rostoucím n blíží jisté křivce. Jmenuje se Gaussova křivka a je popsaná rovnicí $y = \frac{1}{\sqrt{2\pi}} \cdot \mathrm{e}^{-x^2/2}$, přičemž $\mathrm{e} \doteq 2{,}7$ je tzv. Eulerovo číslo.[8]



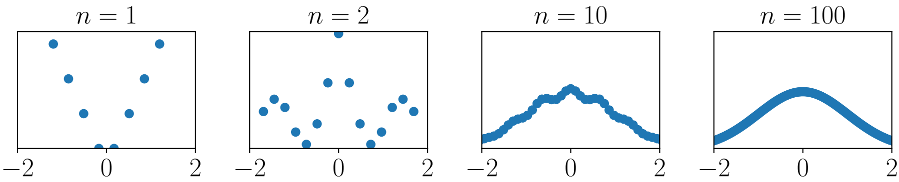
Co přesně jsme udělali s obrázkem? Nejprve jsme od veličiny odečetli střední hodnotu, tedy $n/2$. Tím jsme dosáhli toho, že vrcholy všech kopečků budou na pozici nula. To by samo o sobě ještě nestačilo. Jak víme, rozptyl součtu nezávislých veličin je roven součtu rozptylů, takže pro součet n veličin dostáváme rozptyl $n \cdot \mathrm{Var}(X_1) = n \cdot \frac{1}{4}$. Abychom dosáhli toho, že všechny veličiny budou mít stejný rozptyl, vydělili jsme jejich hodnoty výrazem $\sqrt{n \cdot \frac{1}{4}}$ (nezapomeň, že $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$). To odpovídá „zmáčknutí“ hodnot ve vodorovné ose. Následně jsme stejným poměrem „roztáhli“ hodnoty na svislé ose. To jsme udělali proto, že součet pravděpodobností se vždy musí nasčítat na 1, takže celkový „obsah pod křivkou“ musí být zachován (právě obsah totiž odpovídá naší intuitivní představě o pravděpodobnosti ve spojitých prostorech). Hodnoty na vodorovné ose odpovídají tomu, že se díváme jen do vzdálenosti dvou směrodatných odchylek od průměru (pro $n = 10$ se na obrázek vešlo jen 7 z 11 možných výsledků).

To je sice pěkné, ale hod férovou mincí je jen jednou z mnoha aktivit, která se dá opakovat. Co se stane, když budeme házet falešnou mincí, kostkou nebo si vymyslíme svoje vlastní potřeštěné rozdělení? Odpověď je šokující: stane se to samé! Na následujícím obrázku je nejprve stejný proces pro hod kostkou (rozmysli si, proč pro $n = 2$ vypadá obrázek jako střecha) a následně pro geometrické rozdělení, které vypadá jako skluzavka, a pro podivné rozdělení, které vypadá jako údolí. U všech rozdělení jejich součtu nakonec vypadá podezřele stejně!



[7] Případně jako hroznýš, který právě pozřel slona.

[8] Leonhard Euler (1707–1783) a Carl Friedrich Gauss (1777–1855) jsou všeobecně pokládáni za jedny z nejvlivnějších matematiků vůbec. Proto se objevují v poznámce pod čarou v každém druhém PraSečím seriálu.



Fakt, že pokud dost dlouho sčítáme nezávislé veličiny se stejným rozdělením, vypadá výsledek jako kopeček, se nazývá **centrální limitní věta**.[9] Nejen důkaz, ale i jen samotná formulace tohoto skvostu je nad rámec seriálu, neboť vyžaduje rozsáhlé znalosti z matematické analýzy. Přesto si o ní i o jejích důsledcích něco můžeme říci. Centrální se jí říká proto, že se jedná o jeden ze základních faktů pravděpodobnosti a statistiky, podobně jako fakt, že každé číslo lze jednoznačně rozložit na součin prvočísel, je základním kamenem teorie čísel. Limitní je proto, že v řeči vyšší matematiky je Gaussova křivka v jistém smyslu tzv. limitou.

K čemu je centrální limitní věta dobrá mimo svět matematiky? Například vysvětluje, proč je většina lidí téhož pohlaví podobně vysoká – na Tvoji výšku totiž má vliv mnoho okolností a speciálně hned několik různých genů. Zjednodušeně lze říct, že to, jak se Tvá výška liší od průměru, je součtem příspěvků všech těchto okolností. Protože na sobě různé okolnosti moc nezávisí a žádná není výrazně důležitější než ty zbylé, platí zde centrální limitní věta a rozdělení výšek vypadá jako kopeček (většina lidí je někde uprostřed).

Na druhou stranu barva očí záleží zejména na jednom genu, a proto je mnoho lidí s modrýma a hnědýma očima, ale není pravda, že by většina lidí měla oči v jakési hnědomodré barvě.

Zkus se sám (sama) zamyslet nad tím, kde všude můžeš Gaussovu křivku (kopeček) potkat. Závisí kupříkladu výběr auta nebo mobilního telefonu na mnoha podobně důležitých faktorech, nebo se lidé prostě dělí na dvě skupiny: ty, kdo chtějí nejlepší model za každou cenu, a na ty, kdo se spokojí se standardem stejným pro všechny? V prvním případě budou ceny různých modelů na trhu rozložené podobně jako binomické rozdělení, v druhém případě většina modelů bude mít přibližně jednu ze dvou cen.

I když je Gaussova křivka definovaná pro všechna reálná čísla, platí, že přibližně 99,7 % obsahu pod touto křivkou leží v intervalu od -3 do 3. Protože podle centrální limitní věty platí pro veličiny, které jsou součtem mnoha stejných nezávislých příspěvků, že $\frac{X-\mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}$ vypadá jako Gaussova křivka, můžeme pro tyto veličiny tvrdit, že s pravděpodobností přibližně 99,7 % nabývají hodnot vzdálených od střední hodnoty nejvýše o tři směrodatné odchylky. Podobné vlastnosti této křivky se často používají ve statistice. Ukažme si jednoduché použití na následující úloze z praxe.

Úloha 19. Na večírek přijde sto hostů. Víme, že každý host chce sníst nezávisle na ostatních nula chlebíčků s pravděpodobností 20 %, jeden s pravděpodobností 50 % a dva s pravděpodobností 30 %. Kolik máme nakoupit chlebíčků, aby s pravděpodobností aspoň 99 % na všechny zbylo?

Řešení. Neukážeme si přesné řešení, jen načrtneme, jak pomůže centrální limitní věta. Označme X_i počet chlebíčků, které snědl i-tý host. Zajímá nás veličina $X = X_1 + X_2 + \cdots + X_n$. Z centrální limitní věty víme, že histogram X (tedy její rozdělení) vypadá přibližně jako Gaussova křivka se středem v

$$\mathrm{E}(X) = 100 \cdot \mathrm{E}(X_1) = 100 \cdot (0{,}2 \cdot 0 + 0{,}5 \cdot 1 + 0{,}3 \cdot 2) = 110.$$

Dále vypočítáme rozptyl:

$$\mathrm{Var}(X) = 100 \cdot (\mathrm{E}({X_1}^2) - \mathrm{E}(X_1)^2) = 100 \cdot (0{,}2 \cdot 0^2 + 0{,}5 \cdot 1^2 + 0{,}3 \cdot 2^2 - 1{,}1^2) = 49 = 7^2.$$

Na internetu (nebo v tabulkách) můžeme najít, že 99 % obsahu pod křivkou leží v intervalu od $-\infty$ do 2,3. Hledaná odpověď je tak přibližně $110 + 2{,}3 \cdot 7 = 126{,}1$, takže musíme koupit aspoň 127 chlebíčků. Předchozí výpočet nicméně není žádný důkaz (doopravdy chlebíčků stačí 126).

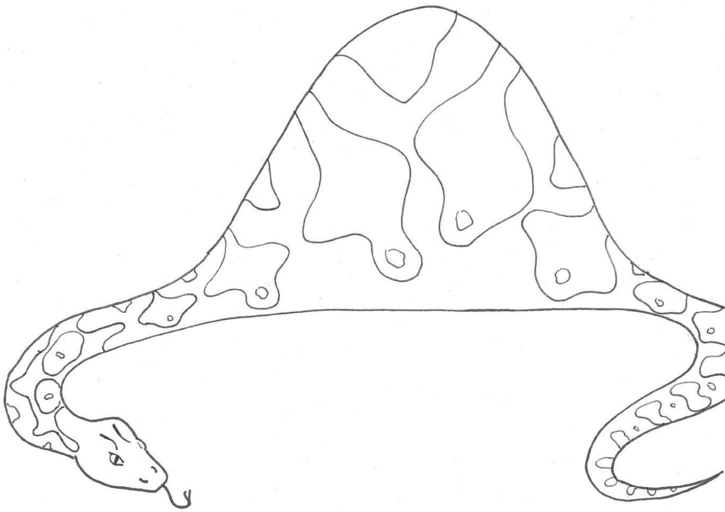
[9] Přesná matematická formulace nepoužívá slovo „kopeček“.

Odhad v podobném duchu bychom dostali i s použitím Čebyševovy nerovnosti, která říká, že libovolná veličina se odchýlí o a směrodatných odchylek s pravděpodobností nejvýše $1/a^2$. Pointa je, že s předpokladem nezávislosti všech veličin (k použití Čebyševa by totiž stačila nezávislost po dvou) dostaneme mnohem přesnější odhad, k jehož použití nám kromě víry v centrální limitní větu stačí jen znalost střední hodnoty a rozptylu dané veličiny.

Závěr

Všudypřítomným a tajemným kopečkem skončila bonusová část posledního dílu. Vždy když si odteď koupíš kopeček zmrzliny, vzpomeň si na chladnou krásu střední hodnoty, roztaj nad definicí nezávislosti a vychutnej si sladkou příchuť zákona velkých čísel.

Doufáme, že sis čtení seriálu a řešení úloh užil(a) podobně, jako jsme si my užili psaní. Seriál pro Tebe psali Danil s Vaškem a obrázky nakreslila Hanka Pařízková. S psaním nám nesmírně pomohlo mnoho dalších PraSátek, za všechny děkujeme Hedvice, Kubovi K., Martinovi T., Michalovi, Mirkovi, Pepovi T., Radovi a Vikimu. A hlavně děkujeme Tobě, žes dočetl(a) až sem – měj se parádně!



Návody k úlohám

1. (1) Napiš si podíl po sobě jdoucích kombinačních čísel.
(2) Napiš si podíl pravděpodobností pro n a $n+1$.

2. V obou částech využij linearitu střední hodnoty.

3. (1) Jestliže soutěžící vyřešil alespoň dvě úlohy, skončí pravděpodobně po prvním dni mezi první polovinou soutěžících. Jenže druhý den jeho umístění bude náhodné, tedy pravděpodobně horší, než první den.
(2) Výjimečně dobrý výsledek nastává s malou pravděpodobností. Je-li příští test nezávislý na předchozím, s velkou pravděpodobností dopadne hůře.
(3) Počet bouraček je do značné míry náhodný, takže vybrané zatáčky v daný rok možná jen „měly smůlu“ a příští rok v nich pak pravděpodobně již k tolika haváriím nedojde.

4. Podobně jako v příkladu se špačky nastane rovnost jen tehdy, když daná veličina může nabývat pouze jedné nezáporné hodnoty.

Pro druhou část stačí uvážit veličinu, která nabývá pouze hodnot $-1, 0$, a k ní vhodnou hodnotu a.

5. (1) Postupuj od nejmenší možné hodnoty X po největší až do chvíle, kdy pravděpodobnost, že X je nejvýše x, přesáhne jednu polovinu. Dané x je pak medián.

(2) Použij Markovovu nerovnost pro $c = 2$.

9. (1) Postupuj stejně jako pro $n = 2$, neboj se představit si roznásobení n závorek.

(2) Pro součet použij linearitu střední hodnoty, pro součin předchozí bod.

10. Dosaď do vzorečku a vzpomeň, že $\mathrm{E}(X + a) = \mathrm{E}(X) + a$ a $\mathrm{E}(bX) = b\mathrm{E}(X)$.

13. Vyjde postupně 2, 3 a 1. Pokud $\mathrm{P}(X = i) = p_i$, musí ve všech podúlohách platit $\sum_{i=1}^n p_i = 1$. Známe-li $\mathrm{E}(X)$, respektive $\mathrm{Var}(X)$, známe hodnotu výrazů $\sum_{i=1}^n ip_i$, respektive $\sum_{i=1}^n i^2p_i - (\sum_{i=1}^n ip_i)^2$. Pro důkaz toho, že jsou čísla jednoznačně určena pro daná n, můžeme využít v prvních dvou částech jednoznačnosti řešení jisté soustavy lineárních rovnic. Pro důkaz maximality stačí najít dvě náhodné veličiny nabývající $n + 1$ hodnot, které mají stejný rozptyl či střední hodnotu.

15. Postupuj jako v předchozím tvrzení a neboj se roznásobit výraz $(X_1 + X_2 + \cdots + X_n)^2$.

16. Dosaď do Čebyševovy nerovnosti, zvol a tak, aby vyšla pravá strana.

17. Pro první část stačí použít, že $\binom{n}{n/2}$ je největší z $n+1$ čísel, která se dohromady sečtou na 2^n.

(1) Použij Čebyševovu nerovnost s $a = 2$ (ale fungují i jiné hodnoty), abys dokázal(a), že součet kombinačních čísel mezi $n/2 - \sqrt{n}$ a $n/2 + \sqrt{n}$ je velký. Poté použij, že $\binom{n}{n/2}$ je ze všech sčítanců ten největší.

(2) Umocni druhou nerovnost na druhou a porovnej výrazy.

Podrobné návody k úlohám

2. (1) Z linearity střední hodnoty plyne $\mathrm{E}(R_n) = 10 + \mathrm{E}(A_{11}) + \cdots + \mathrm{E}(A_n) - n/2 = 10 + (n - 10)/2 - n/2 = 5$.

(2) Z linearity střední hodnoty plyne $\mathrm{E}(P_n) = \frac{10+\mathrm{E}(A_{11})+\cdots+\mathrm{E}(A_n)}{n/2} = \frac{n+10}{n}$. Porovnáním po sobě jdoucích výrazů spočítáme, že $\mathrm{E}(P_n) - \mathrm{E}(P_{n+1}) > 0$, tedy $\mathrm{E}(P_n) > \mathrm{E}(P_{n+1})$.

6. Každá příchuť má poloviční pravděpodobnost, že si ji dané dítě vybere. Protože si děti vybírají nezávisle, má proto každá příchuť pravděpodobnost $\frac{1}{2^5} = \frac{1}{32}$, že si ji nikdo nevybere. Zavedeme si pro každou hodnotu indikátorovou proměnnou, která říká, zda si čokoládu nikdo nevybral. Z linearity střední hodnoty pak plyne, že střední počet nevybraných příchutí je $\frac{320}{32} = 10$. Dosazením $c = 2$ do Markovovy nerovnosti dostáváme, že nevybraných příchutí je alespoň dvacet s pravděpodobností nejvýše jedna polovina.

8. (1) Vzhledem k nezávislosti hodů to je součin obou středních hodnot, tedy $\frac{1+2+5+10+20+50}{6} \cdot \frac{1+2+3+4+5+6}{6} = 51 + \frac{1}{3}$ Kč.

(2) Veličiny již nyní nejsou nezávislé. Střední hodnota jejich součinu je $\frac{1}{6} \cdot (1 \cdot 1 + 2 \cdot 10 + \cdots + 6 \cdot 50) = 76 + \frac{2}{3}$ Kč.

11.

$$\mathrm{Var}(B) = \mathrm{E}(B^2) - \mathrm{E}(B)^2 = \frac{1}{2}(0 + n^2) - \left(\frac{1}{2}(0 + n)\right)^2 = \frac{1}{4}n^2.$$

12. Protože $\mathrm{Var}(X)$ je střední hodnotou nezáporné veličiny $(X - \mathrm{E}(X))^2$, je vždy nezáporný – proto $\mathrm{E}(X^2) - \mathrm{E}(X)^2 \geq 0$, což je požadovaná nerovnost. Aby byl rozptyl nulový, musí být všechny členy $(x - \mathrm{E}(X))^2$ nulové, tedy pro všechna x z oboru hodnot musí být $x = \mathrm{E}(X)$, tedy náhodná veličina může nabývat pouze jedné hodnoty (a není tak moc náhodná).

14. Máme $\mathrm{Var}(O) = \mathrm{Var}(P) = \mathrm{E}(O^2) - \mathrm{E}(O)^2 = \mathrm{E}(O) - \mathrm{E}(O)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$. Dále $\mathrm{Var}(O+O) = \mathrm{Var}(2O) = 4\mathrm{Var}(O) = 1$ a $O + P$ je veličina, která vždy nabývá stejné hodnoty, takže její rozptyl je nula.

18. (1) Nejprve náhodně zvolíme jeden z konců tětivy X a vepíšeme do k rovnostranný trojúhelník XAB, čímž kružnici rozdělíme na tři stejně dlouhé oblouky. Délka tětivy bude větší než jedna právě tehdy, když její druhý konec bude ležet na kratším oblouku AB. Tento oblouk tvoří třetinu obvodu kružnice, a proto je pravděpodobnost tohoto jevu $\frac{1}{3}$.

(2) Vepíšeme-li do kružnice rovnostranný trojúhelník ABC, bude střed kružnice S těžištěm trojúhelníku. Označme D střed úsečky AB. Úsečka DC je těžnice, kterou těžiště S dělí v poměru $1 : 2$. Vzdálenost DS je tedy rovna polovině poloměru kružnice.

Zvolíme-li náhodný bod tak, že jeho vzdálenost od S bude přesně polovinou poloměru kružnice, bude délka odpovídající tětivy přesně 1. Bude-li zvolený bod dál, bude tětiva kratší, a bude-li blíž, bude delší. Tětiva bude delší než jedna pro jevy odpovídající bodům v kruhu se středem v S a poloměrem polovičním oproti původní kružnici. Obsah tohoto obrazce je jednou čtvrtinou celkového obsahu.

(3) Na úhlu pochopitelně nezáleží. Tětiva bude delší než jedna právě tehdy, když zvolíme vzdálenost od středu větší než polovina poloměru (viz předchozí bod), což nastane v půlce případů.

(4) V jistém smyslu jsou správné všechny tři přístupy. Záleží na tom, co myslíme „náhodnou tětivou“. Rozdíl mezi přístupy se dá ilustrovat na druhé a třetí podúloze, kde vlastně vybíráme náhodný bod uvnitř kruhu a testujeme, zda je jeho vzdálenost od středu menší než polovina poloměru. Takové body leží v kružnici se stejným středem a polovičním poloměrem, která má proto čtvrtinový obsah.

Když ale vybíráme náhodný bod tak, že si náhodně zvolíme jeho vzdálenost od středu, bude tato vzdálenost menší než polovina poloměru s pravděpodobností $\frac{1}{2}$. Rozdíl mezi druhým a třetím způsobem tak je ten, že třetí v jistém smyslu preferuje body blízko středu nad body blízko okraje kruhu, zatímco ten druhý se chová ke všem stejně.