Níže naleznete řešení některých příkladů z 12. cvičení z předmětu MB 103/203, podzim 2018. Číslo v hranaté závorce odkazuje na řešený příklad s ohledem na číslování pdf souboru ke cvičení, který naleznete ve studijních materiálech předmětu (soubor mlllc-2018-12.pdf). Řešení k úlohám by mělo být možné nalézt ve velice podobné formě ve skriptech k tomuto předmětu, odkud jsem je převzal a pouze mírně uzpůsobil s ohledem na notaci a teorii, kterou jsme používali.

Kromě jiných výsledků plynoucích z teorie budeme při řešení příkladů potřebovat následující tvrzení.

Věta 1 (Čebyševova nerovnost). *Nechť X je náhodná veličina s konečným rozptylem DX a nechť $\epsilon > 0$ je libovolná, fixně zvolená konstanta. Pak platí*

$$P(|X - EX| \geq \epsilon) \leq \frac{DX}{\epsilon^2}$$

Věta 2 (Markovova nerovnost). *Nechť X je nezáporná náhodná veličina[1], pro kterou známe střední hodnotu EX a nechť $\epsilon > 0$ je libovolná, fixně zvolená konstanta. Pak platí*

$$P(X \geq \epsilon) \leq \frac{EX}{\epsilon}$$

Věta 3 (O střední hodnotě, rozptylu a aproximaci binomického rozdělení). *Nechť X je náhodná veličina s binomickým rozdělením $X \sim Bi(n,p)$. Pak pro střední hodnotu EX a rozptyl DX platí*

$$\begin{aligned} EX &= np \\ DX &= np(1-p)\ . \end{aligned}$$

Navíc lze binomické rozdělení aproximovat normálním rozdělením s parametry $\mu = np$ a $\sigma^2 = np(1-p)$, tj. $X \approx N(np, np(1-p))$.

Řešení. [2. úlohy]
X je náhodná veličina s $EX = \mu$. Chceme určit $P(X \geq 3\mu)$. Podle Markovovy nerovnosti z věty 2 máme

$$P(X \geq 3\mu) \leq \frac{\mu}{3\mu} = \frac{1}{3}\ .$$

Jestliže navíc víme, že náhodná veličina má exponenciální rozdělení s parametrem μ, tj. $X \sim Ex(\frac{1}{\mu})$, pak platí

$$P(X \geq 3\mu) = 1 - P(X < 3\mu) = 1 - P(X \leq 3\mu) = 1 - F(3\mu)\ ,$$

kde $F(x) = 1 - e^{-\frac{x}{\mu}}$ je distribuční funkce uvažovaného rozdělení. Proto

$$P(X \geq 3\mu) = 1 - (1 - e^{-3}) = \frac{1}{e^3}\ . \qquad \diamond$$

[1] Tj. $X(\omega) \geq 0\ \forall \omega \in \Omega$

Řešení. [3. úlohy]
Čebyšev: Náhodná veličina X, která popisuje počet padlých šestek, má binomické rozdělení $X \sim \mathrm{Bi}(1200, \frac{1}{6})$. Podle věty 3 to v našem případě znamená, že střední hodnota pro X je

$$EX = 1200 \cdot \frac{1}{6} = 200$$

a rozptyl je

$$DX = 1200 \cdot \frac{1}{6}(1 - \frac{1}{6}) = 200 \cdot \frac{5}{6} = \frac{500}{3} \ .$$

Pravděpodobnost, že v 1200 pokusech padně alespoň 150-krát a nejvýše 250-krát hodnota 6 se zapíše jako $P(150 \leq X \leq 250)$. Přepíšeme tuto nerovnost pomocí absolutní hodnoty (uvnitř vyjde $|X - EX|$), využijeme pravděpodobnosti doplňku a použijeme Čebyševovu nerovnost z věty 1.

$$\begin{aligned} P(150 \leq X \leq 250) &= P(|X - 200| \leq 50) \\ &= 1 - P(|X - 200| \geq 51) \\ &\geq 1 - \frac{500}{3 \cdot 51^2} \\ &\approx 0.94 \ . \end{aligned}$$

Moivre-Laplace: Jelikož přímé vyčíslení hodnoty $P(150 \leq X \leq 250)$ pomocí distribuční funkce je (bez užití počítače) komplikované, můžeme využít aproximace pomocí normálního rozdělení. Moivre-Laplaceova věta je důsledkem centrální limitní věty pro náhodné veličiny s binomickým rozdělením. Konkrétně platí, že je-li X náhodná veličina s binomickým rozdělením $\mathrm{Bi}(n, p)$, pak náhodná veličina S_n dána vztahem

$$S_n = \frac{X - np}{\sqrt{np(1-p)}}$$

bude pro dostatečně velká n konvergovat[2] k náhodné veličině se standardizovaným normálním rozdělením[3] $N(0, 1)$ a distribuční funkcí Φ. Pro nás to znamená, že náhodnou veličinu X nahradíme náhodnou veličinou

$$Z = \frac{X - 200}{\sqrt{\frac{500}{3}}} = \frac{\sqrt{3}(X - 200)}{10\sqrt{5}} \ ,$$

pro jejíž distribuční funkci platí $F_Z \approx \Phi$. Přepíšeme podle výše uvedeného vztahu pro Z

[2] aproximace se považuje za dobrou, platí-li $np(1-p) > 9$, což je v našem případě se značnou rezervou splněno

[3] tj. normální rozdělení se střední hodnotou 0 a rozptylem 1.

nerovnost v pravděpodobnosti $P(150 \leq X \leq 250)$ a upravíme

$$\begin{aligned}
P(150 \leq X \leq 250) &= P(150 - 200 \leq X - 200 \leq 250 - 200) \\
&= P(\frac{\sqrt{3}(150-200)}{10\sqrt{5}} \leq Z \leq \frac{\sqrt{3}(250-200)}{10\sqrt{5}}) \\
&= P(\frac{-50\sqrt{3}}{10\sqrt{5}} \leq Z \leq \frac{50\sqrt{3}}{10\sqrt{5}}) \\
&= P(-\sqrt{15} \leq Z \leq \sqrt{15}) \\
&\approx \Phi(\sqrt{15}) - \Phi(-\sqrt{15}) \\
&= \Phi(\sqrt{15}) - (1 - \Phi(\sqrt{15})) \\
&= 2\Phi(\sqrt{15}) - 1 \ .
\end{aligned}$$

Dále musíme postupovat podle tabulky normálního rozdělení (viz. tabulka standardizovaného normálního rozdělení např. ve skriptech, na anglické wikipedii apod.). [4]. Platí $\sqrt{15} \approx 3.87$ a v tabulce této hodnotě odpovídá číslo 0.49995, které je rovno obsahu plochy pod grafem funkce hustoty mezi hodnotou 0 a 3.87. Abychom získali $\Phi(\sqrt{15})$, musíme tedy k této hodnotě ještě příčíst 0.5. Tj. $\Phi(\sqrt{15}) \approx 0.99995$ a dohromady získáváme

$$P(150 \leq X \leq 250) \approx 2\Phi(\sqrt{15}) - 1 \approx 2 \cdot 0.99995 - 1 \approx 0.9999 = 99.99\% \ . \qquad \diamond$$

Řešení. [4. úlohy]
Jako náhodnou veličinu X máme dle zadání označit rychlost větru a chceme zjistit $P(X \leq 60)$. Průměrná rychlost je 20 km/hod, tj. $EX = 20$ a podle Markovovy nerovnosti pak[5]

$$P(X \leq 60) = 1 - P(X \geq 60) = 1 - \frac{20}{60} = \frac{2}{3} \ .$$

V druhé části úkolu navíc víme, že směrodatná odchylka $\sigma = \sqrt{DX} = 1$ km/hod a úloha vede na aplikaci Čebyševovy nerovnoti. Chceme zjistit, pro jaké x platí $0.9 \leq P(|X - 20| < x)$. Přitom platí $P(|X - 20| < x) = 1 - P(|X - 20| \geq x)$ a podle Čebyševovy nerovnosti máme $P(|X - 20| \geq x) \leq \frac{1}{x^2}$, tedy $P(|X - 20| < x) \leq 1 - \frac{1}{x^2}$. Dohromady

$$0.9 \leq 1 - \frac{1}{x^2} \Rightarrow x \geq \sqrt{10} \approx 3.2 \ .$$

Jelikož jsme zjistili, jak velký výkyv může být kolem průměrné hodnoty větru, víme, že interval rychlostí, kterých vítr s pravděpodobností alespoň 0.9 nabývá, je $I \approx (20 - 3.2, 20 + 3.2) = (16.8, 23.2)$. $\diamond$

Řešení. [5. úlohy]
V této úloze X je náhodná veličina, která udává počet studentů s prospěchem do 1.2 ze skupiny n náhodně vybraných studentů. X má tedy binomické rozdělení $X \sim Bi(n, 0.1)$, jelikož pouze 10% studentů z celé fakulty má požadovaný prospěch. Úloha se nás ptá, pro jaké nejmenší n bude splněna rovnost $0.95 = P(0.08n \leq X \leq 0.012n)$. Nerovnost uvnitř

[4] POZOR: Podívejte se, jak tabulka funguje, aby vás při písemce nepřekvapila. Tabulka totiž nevyjadřuje přímo $\Phi(x)$, ale $\Phi(x) - 0.5$.

[5] v nerovnostech využijeme, že X je spojitá, tj. $P(X > 60) = P(X \geq 60)$

argumentu P můžeme transformovat podobně jako ve 3. úloze a získat aproximaci pomocí standardizovaného normálního rozdělení

$$\begin{aligned} 0.95 &= P(0.08n \leq X \leq 0.12n) \\ &= P(\frac{-0.02n}{0.3\sqrt{n}} \leq \frac{X - 0.1n}{0.3\sqrt{n}} \leq \frac{0.02n}{0.3\sqrt{n}}) \\ &\approx 2\Phi(\frac{0.02}{0.3}\sqrt{n}) - 1 \\ &\approx 2\Phi(0.0666\sqrt{n}) - 1 \end{aligned}$$

Tedy máme

$$0.95 \approx 2\Phi(\frac{2}{30}\sqrt{n}) - 1 \Rightarrow \Phi(\frac{2}{30}\sqrt{n}) \approx 0.975 \ .$$

V tabulce proto hledáme[6] *vzor* hodnoty $0.975 - 0.5 = 0.475$. Zjistíme, že vzor hodnoty 0.475 je číslo 1.96, což znamená $P(Z \leq 1.96) = \Phi(1.96) = 0.975 \approx \Phi(\frac{2}{30}\sqrt{n})$. Konečně vyvozujeme $\frac{2}{30}\sqrt{n} \approx 1.96 \Rightarrow n \approx (1.96 \cdot 15)^2 = 864.36$. Potřebujeme tedy skupinu 865 studentů, abychom dosáhli požadovaných parametrů. ◇

Řešení. [6. úlohy]
Zde volíme náhodnou veličinu X jako počet stromů, které se ujmou. Máme proto $X \sim Bi(500, 0.8)$, $EX = 400, DX = 80$ a analogicky jako v předchozích úlohách můžeme aproximovat pomocí standardizovaného rozdělení náhodné veličiny $Z = \frac{X-400}{\sqrt{80}}$. Chceme zjistit $P(X \geq 380)$

$$\begin{aligned} P(X \geq 380) &= P(Z \geq \frac{380 - 400}{\sqrt{80}} \\ &= P(Z \geq -\sqrt{5}) \\ &\approx 1 - \Phi(-\sqrt{5}) \\ &= \Phi(\sqrt{5}) \\ &\approx 0.987 = 98.7\% \end{aligned}$$

kde rovnost $1 - \Phi(-\sqrt{5}) = \Phi(\sqrt{5})$ plyne z vlastností standardizovaného normálního rozdělení (funke hustoty je symetrická podle osy y). ◇

Řešení. [6. úlohy]
Nechť X je náhodná veličina určující počet padlých jedniček při 600 hodech *ideální* kostkou, tj. $X \sim Bi(600, \frac{1}{6})$. Podle věty 3 můžeme X aproximovat pomocí normálního rozdělení $N(100, \frac{250}{3})$. Situaci, ve které padla jednička přesně 45-krát, můžeme uvažovat jako jednoprvkový náhodný výběr. Chceme zjistit, zda-li na hladině spolehlivosti $\alpha = 0.01$ může uvažovaná kostka být ideální kostkou. V ideální případě padne jednička v 600 hodech 100 krát. Chceme tedy určit oboustranný interval spolehlivosti pro střední hodnotu počtu padlých jedniček naší kostky a zjistit, zda-li v něm *ideální střední hodnota* 100 leží. Při určování oboustranného intervalu spolehlivosti pro parametr μ vycházíme ze vzorce

$$1 - \alpha = P(\bar{X} - \frac{\sigma}{\sqrt{n}}z(\frac{\alpha}{2}) < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}}z(\frac{\alpha}{2})) \ ,$$

[6] opět zdůrazňuji, mrkněte se na tabulku pro standardizované normální rozdělení a jak se s ní pracuje

kde číslo $z(x)$ odpovídá vzoru hodnoty x v tabulce standardizovaného normálního rozdělení. V našem případě máme $\alpha = 0.01, \bar{X} = 45, \frac{\sigma}{\sqrt{n}} = \frac{250}{3}$ a dostáváme proto, že interval spolehlivosti na hladině $\alpha = 0.01$ pro střední hodnotu μ je $(21, 69)$. Jelikož hodnota 100 v tomto intervalu neleží, nedá se s pravděpodobností 99% tvrdit, že naše kostka je ideální.

$\diamond$