# 8 LIMITNÍ VĚTY

**Limitní věty** jsou tvrzení, která jsou důležitá pro popis pravděpodobnostních modelů v případě rostoucího počtu náhodných pokusů..

Pro orientaci v této problematice jsme se seznámili s několika novými pojmy:

Jestliže posloupnost náhodných veličin $X_n$ **konverguje podle pravděpodobnosti k náhodné veličině X**, pak:

$$\forall \varepsilon > 0: \quad \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$$

Jestliže posloupnost náhodných veličin $\{X_n\}$ **konverguje k náhodné veličině X v distribuci**, pak:

$$\lim_{n\to\infty} F_n(x) = F(x),$$

F(x) v tomto případě nazýváme **asymptotickou distribuční funkcí**.

Velice hrubý odhad pravděpodobnosti odchylky náhodné veličiny X od její střední hodnoty nám umožňuje **Čebyševova nerovnost**:

$$\forall\, \varepsilon > 0: \quad P(|X - EX| \geq \varepsilon) \leq \frac{DX}{\varepsilon^2}$$

Chceme-li odhadnout pravděpodobnost, že odchylka náhodné veličiny X od její střední hodnoty je k násobkem směrodatné odchylky $(k \cdot \sigma)$, pak použijeme upravenou verzi Čebyševovy nerovnosti, kdy za ε dosadíme k.σ:

$$\forall\, \sigma, k > 0: \quad P(|X - EX| \geq k \cdot \sigma) \leq \frac{1}{k^2}$$

Konvergenci průměru v posloupnosti nezávislých veličin se zabývá **zákon velkých čísel**, který říká, že průměr nezávislých náhodných veličin se stejnými středními hodnotami konverguje podle pravděpodobnosti k jejich střední hodnotě.

Důsledkem zákona velkých čísel je **Bernoulliho věta**. Bernoulliho věta říká, že relativní četnost sledovaného jevu konverguje podle pravděpodobnosti k jeho pravděpodobnosti. To nám umožňuje experimentálně odhadovat neznámou pravděpodobnost pomocí pozorované relativní četnosti (viz. klasická definice pravděpodobnosti).

Konvergencí rozdělení k normálnímu rozdělení se zabývá **centrální limitní věta**, která má dvě dílčí formulace: Lindebergovu-Lévyho větu a Moivreovo-Laplaceovu větu.

**Lindebergova-Lévyho věta** říká, že pro dosti velké n má součet i průměr náhodných veličin se stejným rozdělením, stejným průměrem a stejným rozptylem přibližně normální rozdělení:

Jestliže $X_1, X_2, \ldots, X_n$ jsou nezávislé náhodné veličiny se stejným (libovolným) rozdělením, stejnými středními hodnotami $EX_1 = EX_2 = \ldots = EX_n = \mu$ a se stejnými (konečnými) rozptyly $DX_1 = DX_2 = \ldots = DX_n = \sigma^2$, pak platí:

$$X = \sum_{i=1}^{n} X_i \Rightarrow X \to N\left(n\mu; n\sigma^2\right)$$

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \to N\left(\mu; \frac{\sigma^2}{n}\right)$$

**Moivreova-Laplaceova věta** vyjadřuje konvergenci binomického rozdělení k normálnímu rozdělení:

Nechť $X \to Bi(n; p)$, pak: pro dostatečně velká n: $X \to N(np; np(1-p))$

Přičemž poměrně dobré výsledky dává tato aproximace v případě, že:

$$np(1-p) > 9 \quad \text{nebo} \quad \min\{np; n(1-p)\} > 5$$

Mezi důležité aplikace centrální limitní věty pak patří možnost **aproximace výběrové relativní četnosti normálním rozdělením**:

$$p \to N\left(\pi; \frac{\pi \cdot (1-\pi)}{n}\right),$$

možnost **aproximace Poissonova rozdělení rozdělením normálním**:

Nechť $X \to Po(\lambda t)$, pak pro dostatečně velké t můžeme X aproximovat normálním rozdělením s parametry: $X \to N(\lambda t, \lambda t)$

a možnost **aproximace průměrného počtu události za časovou jednotku normálním rozdělením**:

Nechť Y je průměrný počet výskytu události za časovou jednotku, pak: $Y \to N\left(\lambda, \frac{\lambda}{t}\right)$

Na závěr zbývá připomenout, že chceme-li dostat co nejlepší výsledky při aproximaci diskrétního rozdělení rozdělením spojitým, nezapomeneme při výpočtech na **opravu na spojitost**.

**8.1. Odhadněte pravděpodobnost, že náhodná veličina je od své střední hodnoty vzdálená o více než 3σ.**

**Řešení:**

$$\forall\, \sigma > 0: \quad P(|X - EX| \geq 3 \cdot \sigma) \leq \frac{1}{3^2} (= 0,\bar{1})$$

Hledaná pravděpodobnost nepřekračuje 11%. (Je to opravdu hrubý odhad, srovnejte si s pravidlem 6 sigma platným pro normální rozdělení.)

**8.2. Pravděpodobnost vyrobení zmetku je 0,5. Odhadněte pravděpodobnost, že při vyrobení 1000 výrobků bude 400 – 600 zmetků.**

**Řešení:**

X ... počet zmetků v 1000 výrobcích, proto:

$$X \to Bi(1000;0{,}5) \Rightarrow EX = n.p = 500; \quad DX = n.p.(1-p) = 250; \quad \sigma_X = \sqrt{DX} = \sqrt{250}$$

Pravděpodobnost, že počet zmetků bude v rozmezí 400 až 600 lze vyjádřit ve tvaru:

$$P(400 < X < 600) = P(|X - 500| < 100) = 1 - P(|X - 500| \geq 100)$$

Vyjádříme-li si povolenou odchylku od střední hodnoty (100) jako násobek směrodatné odchylky ( $\sqrt{250}$ ), můžeme z Čebyševovy nerovnosti zjistit, že:

$$P(|X - 500| \geq 100) = P\left(|X - 500| \geq \frac{100}{\sqrt{250}} \cdot \sqrt{250}\right) \leq \frac{1}{\left(\frac{100}{\sqrt{250}}\right)^2} = \frac{250}{10000} = 0{,}025\,,$$

z čehož lze jednoduše odvodit, že:

$$\underline{\underline{P(400 < X < 600) > 0{,}975}}$$

(Je-li $P(|X - 500| \geq 100) \leq 0{,}025$, pak $1 - P(|X - 500| \geq 100) > 0{,}975\ (= 1 - 0{,}025)$ ). Znovu si připomeňme, že jde o velice hrubý odhad.

---

**8.3. Dlouhodobým průzkumem bylo zjištěno, že doba potřebná k objevení a odstranění poruchy stroje má střední hodnotu 40 minut a směrodatnou odchylku 30 minut. Jaká je pravděpodobnost, že doba potřebná k objevení a opravení 100 poruch nepřekročí 70 hodin?**

**Řešení:**

$X_i$ ... doba potřebná k objevení a odstranění i-té poruchy

Víme, že: $EX_i = 40$ minut a $DX_i = 30^2$ minut$^2$, rozdělení náhodné veličiny $X_i$ neznáme

X ... celková doba do objevení sté poruchy

$$X = \sum_{i=1}^{100} X_i$$

Na základě Lindebergovy-Lévyho věty víme, že součet n náhodných veličin se stejným rozdělením (nemusíme vědět jakým), stejnými středními hodnotami a stejnými rozptyly můžeme aproximovat normálním rozdělením s parametry: $\mu = n \cdot EX_i$, $\sigma^2 = n \cdot DX_i$, proto:

$$X \to N\left(100 \cdot 40; 100 \cdot 30^2\right)$$

*Nyní již není problém určit hledanou pravděpodobnost (nesmíme jen zapomenout na užívání stejných jednotek, v našem případě minut (70h = 4200 minut):*

$$\underline{\underline{P(X < 4200)}} = F(4200) = \Phi\left(\frac{4200 - 4000}{\sqrt{90000}}\right) = \Phi(0,67) \underline{\underline{= 0,749}}$$

---

**8.4. Životnost elektrického holícího strojku Adam má exponenciální rozdělení se střední hodnotou 2 roky. Určete pravděpodobnost, že průměrná životnost 150-ti prodaných strojků Adam bude vyšší než 27 měsíců.**

**Řešení:**

$X_i$ ... životnost i-tého holícího strojku Adam

$$X_i \to Exp\left(\frac{1}{2}\right) \Rightarrow EX_i = \frac{1}{\lambda} = 2\ roky \Rightarrow \lambda = \frac{1}{2} rok^{-1} \Rightarrow DX_i = \frac{1}{\lambda^2} = 4\ rok^2$$

$\overline{X}$ ... průměrná životnost 150-ti strojků Adam

$$\overline{X} = \frac{\sum_{i=1}^{150} X_i}{150} = \frac{1}{150}\sum_{i=1}^{150} X_i$$

Z Lindebergovy-Lévyho věty víme, že: $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \to N\left(\mu; \frac{\sigma^2}{n}\right)$

V našem případě: $\overline{X} = \frac{1}{150}\sum_{i=1}^{150} X_i \to N\left(2; \frac{4}{150}\right)$

Nyní, když známe rozdělení průměrné životnosti 150-ti strojků Adam, můžeme řešení dokončit (27 měsíců = 2,25 roků):

$$\underline{\underline{P(\overline{X} > 2,25)}} = 1 - F(2,25) = 1 - \Phi\left(\frac{2,25 - 2}{\sqrt{\frac{4}{150}}}\right) = 1 - \Phi(1,53) = 1 - 0,937 \underline{\underline{= 0,063}}$$

---

**8.5. Na telefonní ústřednu je napojeno 3000 účastníků. Každý z nich bude volat telefonní ústřednu během hodiny s pravděpodobností 10%. Jaká je pravděpodobnost, že během následující hodiny zavolá ústřednu:**

**a) právě 300 účastníků?**
**b) více než 310 účastníků?**
**c) mezi 200 a 450 účastníky(včetně)?**

**Řešení:**

X … počet účastníků volajících ústřednu během následující hodiny (z 3000)

Z definice náhodné veličiny X je zřejmé, že X má binomické rozdělení: $X \to Bi(3000;0{,}1)$, jehož pravděpodobnostní funkce je:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

ada) $P(X = 300) = \binom{3000}{300} \cdot (0{,}1)^{300} \cdot (1-0{,}1)^{3000-300} = \dfrac{3000!}{2700! \cdot 300!} \cdot (0{,}1)^{300} \cdot (0{,}9)^{2700}$

Zde narážíme na problém. S pomocí kalkulačky nedokážeme určit žádný z výše uvedených faktoriálů. Proto v tomto případě provedeme alespoň přibližný výpočet (aproximaci).

Z Moivreovy-Laplaceovy věty víme, že binomické rozdělení můžeme aproximovat rozdělením normálním:

**Moivreova-Laplaceova věta:**

Nechť $X \to Bi(n;p)$; $\mathrm{EX} = \mathrm{np}$; $\mathrm{DX} = \mathrm{np}(1-\mathrm{p})$, pak dostatečně velká n: $X \to N(np;np(1-p))$

**V našem případě:**

$X \to Bi(3000;0{,}1)$; $\mathrm{EX} = 3000 \cdot 0{,}1 = 300$; $\mathrm{DX} = 3000 \cdot 0{,}1 \cdot 0{,}9 = 270$,

proto X mohu aproximovat normálním rozdělením s parametry $\mu=300$; $\sigma^2=270$

$$X \to N(300;270)$$

Nyní musíme vyřešit ještě jednu komplikaci. Při aproximaci diskrétní náhodné veličiny spojitou dochází k tomu, že výpočet pravděpodobnostní funkce nelze jednoduše provést (pravděpodobnostní funkce spojité náhodné veličiny je nulová).

$$P(X = 300) = 0$$

Proto se provádí tzv. **oprava na spojitost**.

Je-li X definováno jako počet účastníků volajících během následující hodiny ústřednu, můžeme tvrdit, že pravděpodobnost, že příští hodinu bude volat 300 účastníků je stejná jako pravděpodobnost, že bude volat alespoň 299,5 a méně než 300,5 účastníků. (V intervalu $\langle 299{,}5;300{,}5)$ je pouze 300 účastníků.)

$$P(X=300)=P(299{,}5\le X<300{,}5)$$

$P(299{,}5\le X<300{,}5)$ již není pro spojitou náhodnou veličinu nulová a tak můžeme provést aproximační výpočet. Této úpravě se říká oprava na spojitost.

$$\underline{\underline{P(X=300)}}=P(299{,}5\le X<300{,}5)=P(X<300)-P(X<299{,}5)=$$
$$=F(300{,}5)-F(299{,}5)=\Phi\left(\frac{300{,}5-300}{\sqrt{270}}\right)-\Phi\left(\frac{299{,}5-300}{\sqrt{270}}\right)=$$
$$=\Phi(0{,}03)-\Phi(-0{,}03)=\Phi(0{,}03)-[1-\Phi(0{,}03)]=2\cdot\Phi(0{,}03)-1=$$
$$=2\cdot 0{,}512-1=\underline{\underline{0{,}024}}$$

adb) $P(X>310)=\sum_{k=321}^{3000}\binom{3000}{k}\cdot(0{,}1)^k\cdot(1-0{,}1)^{3000-k}=1-\sum_{k=0}^{310}\binom{3000}{k}\cdot(0{,}1)^k\cdot(1-0{,}1)^{3000-k}$

I zde nastává problém. Vidíme, že vyčíslení příslušných součtů (sum) by nám zabralo spoustu času (pokud bychom to vůbec s pomocí kalkulačky dokázali). Proto i v tomto případě přistoupíme k přibližnému výpočtu (aproximaci).

$$X\to N(300;270)$$

Nyní můžeme provést přibližný výpočet:

$$P(X>310)=1-P(X\le 310)$$

Znovu nastává problém způsobený aproximací diskrétního rozdělení spojitým a proto i zde přistoupíme k opravě na spojitost:

$$P(X>310)=1-P(X\le 310)=1-P(X<310{,}5)$$

$$\underline{\underline{P(X>310)}}=1-F(310{,}5)=1-\Phi\left(\frac{310{,}5-300}{\sqrt{270}}\right)=1-\Phi(0{,}64)=1-0{,}739=\underline{\underline{0{,}261}}$$

adc) $P(200\le X\le 450)=\sum_{k=200}^{450}\binom{3000}{k}\cdot(0{,}1)^k\cdot(1-0{,}1)^{3000-k}$

Opět máme ve výše uvedeném vztahu velký počet sčítanců a vysoké faktoriály, proto hledáme přibližný výsledek pomocí centrální limitní věty (Moivreovy-Laplaceovy věty). Zároveň i zde budeme provádět opravu na spojitost:

$$\underline{\underline{P(200 \le X \le 450)}} = P(X \le 450) - P(X < 200) = P(X < 450{,}5) - P(X < 200) =$$

$$= \mathrm{F}(450{,}5) - F(200) = \Phi\left(\frac{450{,}5 - 300}{\sqrt{270}}\right) - \Phi\left(\frac{200 - 300}{\sqrt{270}}\right) =$$

$$= \Phi(9{,}16) - \Phi(-6{,}09) = \Phi(9{,}16) - [1 - \Phi(6{,}09)] = \Phi(9{,}16) + \Phi(6{,}09) - 1 =$$

$$= 1 + 1 - 1 \underline{\underline{= 1}}$$

Použitá aproximace nám dává velmi dobré výsledky (velmi blízké skutečným), protože je splněna podmínka, že: $np(1-p) > 9$ $(27 > 9)$.

---

**8.6. Sekretářka Petra píše na stroji rychlosti 250 úhozů / min. Při této rychlosti udělá průměrně 3 chyby za 10 minut. Jaká je pravděpodobnost, že při 30-ti minutovém diktátu udělá více než 10 chyb?**

**Řešení:**

Definujme si náhodnou veličinu X jako počet chyb v diktátu (za 30 minut). Tato náhodná veličina (počet události v časovém intervalu) má Poissonovo rozdělení s parametrem $\lambda t = 9$ (průměrný počet chyb za 30 minut, = EX = DX).

$$X \to Po(9)$$

$$\underline{\underline{P(X > 10)}} = 1 - P(X \le 10) = 1 - \sum_{k=0}^{10} \frac{9^k}{k!} \cdot e^{-9} = 1 - e^{-9}\left(\frac{9^0}{0!} + \frac{9^1}{1!} + \ldots + \frac{9^{10}}{10!}\right) = 1 - 0{,}706 \underline{\underline{= 0{,}294}}$$

Tento výpočet byl poněkud pracný. Proveďme srovnávací výpočet pomocí centrální limitní věty:

Víme, že Poissonovu náhodnou veličinu s parametrem $\lambda t$ můžeme aproximovat **pro dostatečně velká $\boldsymbol{\lambda t}$** normálním rozdělením s parametry: $\mu = \lambda t$ , $\sigma^2 = \lambda t$.

$$X \to N(9;9)$$

*Pak:*

$$\underline{\underline{P(X > 10)}} = 1 - P(X \le 10) = 1 - P(X < 10{,}5) = 1 - F(10{,}5) = 1 - \Phi\left(\frac{10{,}5 - 9}{\sqrt{9}}\right) = 1 - \Phi(0{,}5) =$$

$$= 1 - 0{,}691 \underline{\underline{= 0{,}309}}$$

**Vyhodnocení aproximace pro tento případ:**

Aproximační postup byl mnohem rychlejší, výsledky obou postupů se nám liší o 1,5%, což je asi 5% -ní chyba (0,015/0,294).

---

**8.7. Výletní člun má nosnost 5000kg. Hmotnost cestujících je náhodná veličina se střední hodnotou 70kg a směrodatnou odchylkou 20kg. Kolik cestujících může člunem cestovat, aby pravděpodobnost přetížení člunu byla menší než 0,001?**

**Řešení:**

X … celková hmotnost všech cestujících

$$X = \sum_{i=1}^{n} X_i \Rightarrow X \to N\left(n\mu; n\sigma^2\right) \qquad \Rightarrow \qquad X \to N(n \cdot 70; n \cdot 400)$$

$$P(X > 5000) \qquad < 0{,}001$$

$$1 - F(5000) \qquad < 0{,}001$$

$$1 - \Phi\left(\frac{5000 - 70n}{\sqrt{400n}}\right) < 0{,}001$$

$$0{,}999 \qquad < \Phi\left(\frac{5000 - 70n}{\sqrt{400n}}\right)$$

$$3 \qquad < \frac{5000 - 70n}{\sqrt{400n}}$$

$$60\sqrt{n} \qquad < -70n + 5000$$

$$3600n \qquad < 4900n^2 - 700000n + 25000000$$

$$0 \qquad < 49n^2 - 7036n + 250000$$

$$n < 64{,}5 \quad \cup \quad n > 79 \quad \Rightarrow \underline{\underline{n_{\max} = 64}}$$

Člunem může cestovat maximálně 64 osob.