Limitní věty teorie pravděpodobnosti

Jiří Neubauer

Katedra ekonometrie, FVL, UO Brno
kancelář 69a, tel. 973 442029
email:Jiri.Neubauer@unob.cz

Zákon velkých čísel

- Jestliže opakujeme nezávisle nějaký pokus, můžeme z pozorovaných hodnot sestavit rozdělení relativních četností a informaci o tomto rozdělení shrnout do charakteristik.
- Toto rozdělení resp. charakteristiky nazveme **empirickým rozdělením** resp. **empirickými charakteristikami**.
- Při dodržování jistých podmínek můžeme očekávat, že empirické rozdělení (resp. charakteristiky) se bude blížit teoretickému rozdělení (resp. charakteristikám) a to tím více, čím větší bude počet realizovaných pokusů.

Zákon velkých čísel

- Jestliže opakujeme nezávisle nějaký pokus, můžeme z pozorovaných hodnot sestavit rozdělení relativních četností a informaci o tomto rozdělení shrnout do charakteristik.
- Toto rozdělení resp. charakteristiky nazveme **empirickým rozdělením** resp. **empirickými charakteristikami**.
- Při dodržování jistých podmínek můžeme očekávat, že empirické rozdělení (resp. charakteristiky) se bude blížit teoretickému rozdělení (resp. charakteristikám) a to tím více, čím větší bude počet realizovaných pokusů.

Zákon velkých čísel

- Jestliže opakujeme nezávisle nějaký pokus, můžeme z pozorovaných hodnot sestavit rozdělení relativních četností a informaci o tomto rozdělení shrnout do charakteristik.
- Toto rozdělení resp. charakteristiky nazveme **empirickým rozdělením** resp. **empirickými charakteristikami**.
- Při dodržování jistých podmínek můžeme očekávat, že empirické rozdělení (resp. charakteristiky) se bude blížit teoretickému rozdělení (resp. charakteristikám) a to tím více, čím větší bude počet realizovaných pokusů.

Zákon velkých čísel

- Je třeba si uvědomit, že přibližování empirických hodnot k hodnotám teoretickým nemá charakter matematické konvergence, ale konvergence pravděpodobnostní.
- Pravděpodobnostní konvergencí rozumíme skutečnost, že při vzrůstajícím počtu pokusů se pravděpodobnost větších odchylek empirických hodnot od teoretických stále zmenšuje.

Zákon velkých čísel

- Je třeba si uvědomit, že přibližování empirických hodnot k hodnotám teoretickým nemá charakter matematické konvergence, ale konvergence pravděpodobnostní.
- Pravděpodobnostní konvergencí rozumíme skutečnost, že při vzrůstajícím počtu pokusů se pravděpodobnost větších odchylek empirických hodnot od teoretických stále zmenšuje.

Konvergence podle pravděpodobnosti

Definice

Jestliže pro posloupnost náhodných veličin $X_1, X_2, \ldots, X_n, \ldots$ platí vztah

$$\lim_{n\to\infty} P(|X_n - c| < \epsilon) = 1, \epsilon > 0,$$

říkáme, že posloupnost $\{X_n\}$ **konverguje podle pravděpodobnosti** ke konstantě c. Pravděpodobnostní konvergence se označuje

$$X_n \xrightarrow{P} c.$$

Čebyševova nerovnost

Věta

Pro libovolnou náhodnou veličinu X se střední hodnotou $E(X)$, konečným rozptylem $D(X)$ a pro každé $\epsilon > 0$ platí

$$P(|X - E(X)| < \epsilon) \geq 1 - \frac{D(X)}{\epsilon^2}.$$

Čebyševova nerovnost se uplatňuje především v oblasti teorie, je možno ji však použít i pro odhad určitých pravděpodobností u náhodných veličin, jejichž rozdělení jsou neznámá.

Bernoulliova věta

Věta

Jestliže náhodná veličina X je počet výskytů jevu v n nezávislých pokusech a π je stálá pravděpodobnost, že tento jev nastane v jednom pokuse, potom pro každé $\epsilon > 0$ platí

$$\lim_{n \to \infty} P\left(\left|\frac{X}{n} - \pi\right| < \epsilon\right) = 1.$$

Moivre-Laplaceova věta

Věta

Nechť náhodná veličina X má binomické rozdělení s parametry n a π, tj. $X \sim B(n, \pi)$ [a], střední hodnotu $E(X) = n\pi$ a rozptyl $D(X) = n\pi(1-\pi)$. Pro normovanou náhodnou veličinu

$$U = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}}$$

platí limitní vztah

$$\lim_{n \to \infty} P(U \leq u) = \Phi(u),$$

kde $\Phi(u)$ je distribuční funkce rozdělení $N(0,1)$.

[a] X je součtem n nezávislých náhodných veličin $X_1, X_2, \dots, X_n$, z nichž každá má alternativní rozdělení $A(\pi)$

Moivre-Laplaceova věta

Moivreova-Laplaceova věta tedy říká, že při dostatečně velkém počtu nezávislých pokusů konverguje binomické rozdělení k rozdělení normálnímu. Aproximace je vhodná, jestliže

$$n\pi(1-\pi) > 9 \quad \text{a} \quad \frac{1}{n+1} < \pi < \frac{n}{n+1}.$$

Moivre-Laplaceova věta pro podíl

Věta

Nechť náhodná veličina X je součet n nezávislých alternativních veličin $X_1, X_2, \ldots, X_n$, které mají stejný parametr π, $0 < \pi < 1$, a náhodná veličina $\frac{X}{n}$ má střední hodnotu $E\left(\frac{X}{n}\right) = \pi$ a rozptyl $D\left(\frac{X}{n}\right) = \frac{\pi(1-\pi)}{n}$. Potom pro normovanou náhodnou veličinu

$$U = \frac{\frac{X}{n} - \pi}{\sqrt{\pi(1-\pi)}}\sqrt{n} \tag{1}$$

platí limitní vztah

$$\lim_{n\to\infty} P(U \leq u) = \Phi(u),$$

kde $\Phi(u)$ je distribuční funkce normovaného normálního rozdělení $N(0,1)$.

Lévy-Lindebergova věta

Věta

Nechť náhodná veličina X je součtem n nezávislých náhodných veličin $X_1, X_2, \ldots, X_n$, se stejným rozdělením s konečnou střední hodnotou $E(X_i) = \mu$ a konečným rozptylem $D(X_i) = \sigma^2$, $i = 1, 2, \ldots, n$. [a] Pro normovanou náhodnou veličinu

$$U = \frac{X - n\mu}{\sqrt{n\sigma^2}}$$

platí limitní vztah

$$\lim_{n \to \infty} P(U \leq u) = \Phi(u),$$

kde $\Phi(u)$ je distribuční funkce rozdělení $N(0, 1)$.

[a] Střední hodnota a rozptyl náhodné veličiny X jsou vzhledem k nezávislosti a stejnému rozdělení veličin X_i rovny $E(X) = n\mu$ a $D(X) = n\sigma^2$

Lévy-Lindebergova věta pro průměr

Věta

Nechť náhodná veličina $\overline{X}$ je průměr n nezávislých náhodných veličin $X_1, X_2, \ldots, X_n$, které mají libovolný identický zákon rozdělení s konečnou střední hodnotou $E(X_i) = \mu$ a konečným rozptylem $D(X_i) = \sigma^2$. Potom pro střední hodnotu a rozptyl náhodné veličiny $\overline{X}$ platí

$$E(\overline{X}) = \mu \text{ a } D(\overline{X}) = \frac{\sigma^2}{n}$$

a pro normovanou náhodnou veličinu

$$U = \frac{\overline{X} - \mu}{\sigma}\sqrt{n} \tag{2}$$

platí limitní vztah

$$\lim_{n\to\infty} P(U \leq u) = \Phi(u),$$

kde $\Phi(u)$ je distribuční funkce normovaného normálního rozdělení $N(0,1)$.

Lévy-Lindebergova věta

Pro výběrový úhrn M platí:

- $M = \sum_{i=1}^{n} X_i \sim as.N(n\mu, n\sigma^2), E(M) = n\mu, D(M) = n\sigma^2$
- $U = \frac{M - E(M)}{\sqrt{D(M)}} = \frac{M - n\mu}{\sqrt{n\sigma^2}} \sim as.N(0,1)$
- $P(M \leq m) = F(m) \approx \Phi\left(\frac{m - n\mu}{\sqrt{n\sigma^2}}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{m - n\mu}{\sqrt{n\sigma^2}} < u_{1-\alpha/2}\right) = 1 - \alpha$

Lévy-Lindebergova věta

Pro výběrový úhrn M platí:

- $M = \sum_{i=1}^{n} X_i \sim as.N(n\mu, n\sigma^2), E(M) = n\mu, D(M) = n\sigma^2$
- $U = \frac{M-E(M)}{\sqrt{D(M)}} = \frac{M-n\mu}{\sqrt{n\sigma^2}} \sim as.N(0,1)$
- $P(M \leq m) = F(m) \approx \Phi\left(\frac{m-n\mu}{\sqrt{n\sigma^2}}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{m-n\mu}{\sqrt{n\sigma^2}} < u_{1-\alpha/2}\right) = 1 - \alpha$

Lévy-Lindebergova věta

Pro výběrový úhrn M platí:

- $M = \sum_{i=1}^{n} X_i \sim as.N(n\mu, n\sigma^2), E(M) = n\mu, D(M) = n\sigma^2$
- $U = \frac{M - E(M)}{\sqrt{D(M)}} = \frac{M - n\mu}{\sqrt{n\sigma^2}} \sim as.N(0,1)$
- $P(M \leq m) = F(m) \approx \Phi\left(\frac{m - n\mu}{\sqrt{n\sigma^2}}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{m - n\mu}{\sqrt{n\sigma^2}} < u_{1-\alpha/2}\right) = 1 - \alpha$

Lévy-Lindebergova věta

Pro výběrový úhrn M platí:

- $M = \sum_{i=1}^{n} X_i \sim as.N(n\mu, n\sigma^2), E(M) = n\mu, D(M) = n\sigma^2$
- $U = \frac{M-E(M)}{\sqrt{D(M)}} = \frac{M-n\mu}{\sqrt{n\sigma^2}} \sim as.N(0,1)$
- $P(M \leq m) = F(m) \approx \Phi\left(\frac{m-n\mu}{\sqrt{n\sigma^2}}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{m-n\mu}{\sqrt{n\sigma^2}} < u_{1-\alpha/2}\right) = 1-\alpha$

Lévy-Lindebergova věta

Pro výběrový průměr $\overline{X}$ platí:

- $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim as.N(\mu, \frac{\sigma^2}{n}), E(\overline{X}) = \mu, D(\overline{X}) = \frac{\sigma^2}{n}$
- $U = \frac{\overline{X}-E(\overline{X})}{\sqrt{D(\overline{X})}} = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \sim as.N(0,1)$
- $P(\overline{X} \leq \overline{x}) = F(\overline{x}) \approx \Phi\left(\frac{\overline{x}-\mu}{\sigma}\sqrt{n}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{\overline{x}-\mu}{\sigma}\sqrt{n} < u_{1-\alpha/2}\right) = 1-\alpha$

Lévy-Lindebergova věta

Pro výběrový průměr $\overline{X}$ platí:

- $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim as.N(\mu, \frac{\sigma^2}{n}), E(\overline{X}) = \mu, D(\overline{X}) = \frac{\sigma^2}{n}$
- $U = \frac{\overline{X}-E(\overline{X})}{\sqrt{D(\overline{X})}} = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \sim as.N(0,1)$
- $P(\overline{X} \leq \overline{x}) = F(\overline{x}) \approx \Phi\left(\frac{\overline{x}-\mu}{\sigma}\sqrt{n}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{\overline{x}-\mu}{\sigma}\sqrt{n} < u_{1-\alpha/2}\right) = 1-\alpha$

Lévy-Lindebergova věta

Pro výběrový průměr $\overline{X}$ platí:

- $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i \sim as.N(\mu, \frac{\sigma^2}{n}), E(\overline{X}) = \mu, D(\overline{X}) = \frac{\sigma^2}{n}$
- $U = \frac{\overline{X} - E(\overline{X})}{\sqrt{D(\overline{X})}} = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \sim as.N(0,1)$
- $P(\overline{X} \leq \overline{x}) = F(\overline{x}) \approx \Phi\left(\frac{\overline{x}-\mu}{\sigma}\sqrt{n}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{\overline{x}-\mu}{\sigma}\sqrt{n} < u_{1-\alpha/2}\right) = 1 - \alpha$

Lévy-Lindebergova věta

Pro výběrový průměr $\overline{X}$ platí:

- $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i \sim as.N(\mu, \frac{\sigma^2}{n}), E(\overline{X}) = \mu, D(\overline{X}) = \frac{\sigma^2}{n}$
- $U = \frac{\overline{X}-E(\overline{X})}{\sqrt{D(\overline{X})}} = \frac{\overline{X}-\mu}{\sigma}\sqrt{n} \sim as.N(0,1)$
- $P(\overline{X} \leq \overline{x}) = F(\overline{x}) \approx \Phi\left(\frac{\overline{x}-\mu}{\sigma}\sqrt{n}\right)$
- $P\left(-u_{1-\alpha/2} < \frac{\overline{x}-\mu}{\sigma}\sqrt{n} < u_{1-\alpha/2}\right) = 1-\alpha$

Poznámka k opravě na spojitost

Používáme-li normální rozdělení jako aproximaci pro rozdělení diskrétní náhodné veličiny, doporučuje se, zejména u silněji asymetrických rozdělení, provést tzv. **opravu na spojitost**, která tuto aproximaci zlepšuje. Při výpočtu pravděpodobností $P(X \leq x)$ resp. $P(X \geq x)$ pomocí normální aproximace jsou totiž výsledky podhodnocené. Naopak při výpočtu pravděpodobností $P(X < x)$ resp. $P(X > x)$ pomocí normální aproximace jsou výsledky nadhodnocené.

Poznámka k opravě na spojitost

Podstata výpočtů pravděpodobností pomocí opravy na spojitost spočívá v tom, že výpočet pravděpodobnosti pro argument x provedeme přibližně pomocí argumentu $x + 0{,}5$ resp. $x - 0{,}5$. Příklady provedených korekcí ukazuje tabulka:

před opravou	$x < 3$	$x \leq 3$	$x = 5$	$x \geq 7$	$x > 7$
po opravě	$x < 2{,}5$	$x < 3{,}5$	$4{,}5 < x < 5{,}5$	$x > 6{,}5$	$x > 7{,}5$

Příklad 1

Pravděpodobnost zásahu cíle při jednom výstřelu je 0,8. Jaká je pravděpodobnost, že se počet zásahů při 200 výstřelech nebude lišit od střední hodnoty počtu zásahů o více než 10 zásahů?

Příklad 1

Binomické rozdělení:

$$E(X) = n\pi = 200 \cdot 0{,}8 = 160$$

$$D(X) = n\pi(1-\pi) = 200 \cdot 0{,}8 \cdot (1 - 0{,}8) = 32$$

$$\begin{aligned} P(150 \leq X \leq 170) &= p(150) + p(151) + \cdots + p(170) = \\ &= \binom{200}{150} 0{,}8^{150} \cdot 0{,}2^{50} + \binom{200}{151} 0{,}8^{151} \cdot 0{,}2^{49} + \cdots + \\ &\quad + \binom{200}{170} 0{,}8^{170} \cdot 0{,}2^{30} = 0{,}937 \end{aligned}$$

Příklad 1

Binomické rozdělení:

$$E(X) = n\pi = 200 \cdot 0{,}8 = 160$$

$$D(X) = n\pi(1-\pi) = 200 \cdot 0{,}8 \cdot (1-0{,}8) = 32$$

$$\begin{aligned} P(150 \leq X \leq 170) &= p(150) + p(151) + \cdots + p(170) = \\ &= \binom{200}{150} 0{,}8^{150} \cdot 0{,}2^{50} + \binom{200}{151} 0{,}8^{151} \cdot 0{,}2^{49} + \cdots + \\ &\quad + \binom{200}{170} 0{,}8^{170} \cdot 0{,}2^{30} = 0{,}937 \end{aligned}$$

Příklad 1

Pomocí Moivre-Laplaceovy věty:

$$F(x) \approx \Phi\left(\frac{x - n\pi}{\sqrt{n\pi(1-\pi)}}\right)$$

$$P(150 \leq X \leq 170) = F(170) - F(149) \approx \Phi\left(\tfrac{170-160}{\sqrt{32}}\right) - \\ -\Phi\left(\tfrac{149-160}{\sqrt{32}}\right) = 0{,}936$$

Příklad 1

Pomocí Moivre-Laplaceovy věty (s opravou na spojitost):

$$F(x) \approx \Phi\left(\frac{x - n\pi}{\sqrt{n\pi(1-\pi)}}\right)$$

$$P(150 \leq X \leq 170) \approx P(149{,}5 < X < 170{,}5) = F(170{,}5) - F(149{,}5) =$$
$$= \Phi\left(\frac{170{,}5-160}{\sqrt{32}}\right) - \Phi\left(\frac{149{,}5-160}{\sqrt{32}}\right) = 0{,}937$$

Příklad 1

Pomocí Čebyševovy nerovnosti:

$$P(|X - E(X)| < \epsilon) \geq 1 - \frac{D(X)}{\epsilon^2}$$

$$E(X) = n\pi = 200 \cdot 0{,}8 = 160$$

$$D(X) = n\pi(1 - \pi) = 200 \cdot 0{,}8 \cdot (1 - 0{,}8) = 32$$

$$P(|X - 160)| < 10) \geq 1 - \frac{32}{10^2} = 0{,}68$$

$$P(|X - 160)| < 11) \geq 1 - \frac{32}{11^2} = 0{,}736$$

Příklad 1

Pomocí Čebyševovy nerovnosti:

$$P(|X - E(X)| < \epsilon) \geq 1 - \frac{D(X)}{\epsilon^2}$$

$$E(X) = n\pi = 200 \cdot 0{,}8 = 160$$

$$D(X) = n\pi(1-\pi) = 200 \cdot 0{,}8 \cdot (1 - 0{,}8) = 32$$

$$P(|X - 160)| < 10) \geq 1 - \frac{32}{10^2} = 0{,}68$$

$$P(|X - 160)| < 11) \geq 1 - \frac{32}{11^2} = 0{,}736$$

Příklad 1

Pomocí Čebyševovy nerovnosti:

$$P(|X - E(X)| < \epsilon) \geq 1 - \frac{D(X)}{\epsilon^2}$$

$$E(X) = n\pi = 200 \cdot 0{,}8 = 160$$

$$D(X) = n\pi(1-\pi) = 200 \cdot 0{,}8 \cdot (1 - 0{,}8) = 32$$

$$P(|X - 160)| < 10) \geq 1 - \frac{32}{10^2} = 0{,}68$$

$$P(|X - 160)| < 11) \geq 1 - \frac{32}{11^2} = 0{,}736$$

Příklad 1

Pomocí Čebyševovy nerovnosti:

$$P(|X - E(X)| < \epsilon) \geq 1 - \frac{D(X)}{\epsilon^2}$$

$$E(X) = n\pi = 200 \cdot 0{,}8 = 160$$

$$D(X) = n\pi(1 - \pi) = 200 \cdot 0{,}8 \cdot (1 - 0{,}8) = 32$$

$$P(|X - 160)| < 10) \geq 1 - \frac{32}{10^2} = 0{,}68$$

$$P(|X - 160)| < 11) \geq 1 - \frac{32}{11^2} = 0{,}736$$

Příklad 2

Ve volbách dalo 52 % voličů hlas koaličním stranám. Jaká je pravděpodobnost, že při průzkumu veřejného mínění o rozsahu 2600 respondentů získala převahu opozice?

Příklad 2

$X \ldots$ počet respondentů volící opozici

$$X \sim B(2600; 0{,}48)$$

$$E(X) = n\pi = 2600 \cdot 0{,}48 = 1248$$

$$D(X) = n\pi(1-\pi) = 2600 \cdot 0{,}48 \cdot (1 - 0{,}48) = 648{,}96$$

Příklad 2

$X \ldots$ počet respondentů volící opozici

$$X \sim B(2600; 0{,}48)$$

$$E(X) = n\pi = 2600 \cdot 0{,}48 = 1248$$

$$D(X) = n\pi(1-\pi) = 2600 \cdot 0{,}48 \cdot (1-0{,}48) = 648{,}96$$

Příklad 2

$$\begin{aligned} P(X > 1300) &= 1 - P(X \leq 1300) = 1 - [p(0) + \cdots + p(1300)] = \\ &= 1 - \Big[\binom{2600}{0} \cdot 0{,}48^{0} \cdot 0{,}52^{2600} + \cdots + \\ &\quad + \binom{2600}{1300} \cdot 0{,}48^{1300} \cdot 0{,}52^{1300}\Big] = 1 - 0{,}98031 = 0{,}01969 \end{aligned}$$

Příklad 2

Pomocí Moivre-Laplaceovy věty:

$$F(x) \approx \Phi\left(\frac{x - n\pi}{\sqrt{n\pi(1-\pi)}}\right)$$

$$P(X > 1300) = 1 - P(X \leq 1300) = 1 - F(1300) \approx$$
$$\approx 1 - \Phi\left(\frac{1300-1248}{\sqrt{648{,}96}}\right) = 1 - 0{,}97939 = 0{,}02061$$

Příklad 2

Pomocí Moivre-Laplaceovy věty (s opravou na spojitost):

$$F(x) \approx \Phi\left(\frac{x - n\pi}{\sqrt{n\pi(1-\pi)}}\right)$$

$$\begin{aligned} P(X > 1300) = 1 - P(X \leq 1300) \approx 1 - P(X < 1300{,}5) = \\ = 1 - \Phi\left(\tfrac{1300{,}5 - 1248}{\sqrt{648{,}96}}\right) = \\ = 1 - 0{,}98034 = 0{,}01966 \end{aligned}$$