

# Počáteční šíření epidemie

*Keywords: kombinatorika, pravděpodobnost a statistika, statistika, zpracování dat, aritmetický průměr, geometrický průměr, regresní analýza, GeoGebra*

Nejznámějším případem epidemie z poslední doby je pandemie virové choroby COVID-19, která je způsobena koronavirem SARS-CoV-2. Tato epidemie vypukla v prosinci 2019 ve městě Wu-chan v provincii Chu-pej v centrální Číně.

Začátek epidemie COVID-19 v Evropě se datuje na leden 2020. První potvrzené případy byly hlášeny ve Francii dne 24. ledna 2020. Jednalo se o tři pacienty, kteří se nedávno vrátili z Číny, kde byla epidemie již v plném proudu.

Do České republiky se pak nemoc rozšířila před 1. březnem 2020, kdy byly potvrzeny první tři případy. K 18. březnu 2020 bylo u nás potvrzeno již 464 případů.

Šíření nemoci mívá na začátku epidemie exponenciální charakter. Až později se obvykle toto šíření zpomalí a dále už má jiný průběh (lineární, logistický apod.).

## Typický průběh epidemie

Na základě získaných dat zkusíme namodelovat počet nakažených osob v závislosti na dnech od začátku epidemie. V tabulce jsou data popisující počet nakažených v závislosti na dnech od začátku epidemie.

| Den ( $n$ ) | Počet nakažených ( $a_n$ ) |
|-------------|----------------------------|
| 1           | 3                          |
| 2           | 3                          |
| 3           | 5                          |
| 4           | 6                          |
| 5           | 9                          |
| 6           | 20                         |
| 7           | 27                         |
| 8           | 33                         |
| 9           | 39                         |
| 10          | 64                         |
| 11          | 95                         |
| 12          | 117                        |
| 13          | 142                        |
| 14          | 190                        |
| 15          | 299                        |
| 16          | 384                        |

*Poznámka.* Jedná se o reálná data z České republiky od 1. března 2020.

**Úloha 1.** Vypočítejte poměr počtu nakažených osob v jednom dni vzhledem k předchozímu dni.

*Řešení.* Označíme-li  $q_n$  poměr nakažených v  $n$ -tém dnu ku počtu nakažených v  $n-1$ -ním dnu, dostaneme

pro výpočet vztah

$$q_n = \frac{a_n}{a_{n-1}}, \quad n \geq 2.$$

Můžeme tak vypočítat všechny údaje a doplnit tabulku.

| Den ( $n$ ) | Počet nakažených ( $a_n$ ) | Poměr počtu nakažených |
|-------------|----------------------------|------------------------|
| 1           | 3                          | —                      |
| 2           | 3                          | 1,000                  |
| 3           | 5                          | 1,667                  |
| 4           | 6                          | 1,200                  |
| 5           | 9                          | 1,500                  |
| 6           | 20                         | 2,222                  |
| 7           | 27                         | 1,350                  |
| 8           | 33                         | 1,222                  |
| 9           | 39                         | 1,182                  |
| 10          | 64                         | 1,641                  |
| 11          | 95                         | 1,484                  |
| 12          | 117                        | 1,232                  |
| 13          | 142                        | 1,214                  |
| 14          | 190                        | 1,338                  |
| 15          | 299                        | 1,574                  |
| 16          | 384                        | 1,284                  |

Vypočtenou hodnotu poměru počtu nakažených v jednom dni vzhledem k předchozímu dni můžeme interpretovat jako rychlost šíření nemoci. Pro přesně exponenciální růst nebo pokles by tento poměr vycházel konstantní. My pracujeme pouze s “naměřenými” daty, takže nám hodnota poměru vychází pouze přibližně. Pojďme se proto na tento poměr podívat blíže.

**Úloha 2.** Vypočítejte aritmetický a geometrický průměr poměru počtu nakažených. Který z nich je v tomto případě vhodnější?

**Řešení.** Aritmetický průměr  $\bar{q}$  je dán vzorcem

$$\bar{q} = \frac{q_2 + \dots + q_{16}}{15} = 1,40731.$$

Podobně pro geometrický průměr  $G$  platí

$$G = \sqrt[15]{q_2 \cdot \dots \cdot q_{16}} = 1,38191.$$

Lepší je v tomto případě použít geometrický průměr.

Aritmetický průměr může být ovlivněn extrémními hodnotami (pro nás například hodnota  $q_6$ ), což může zkreslit interpretaci průměrného růstu nebo poklesu. Geometrický průměr naopak tento vliv zmírňuje.

Navíc v exponenciálním případě růstu (pokud se omezíme na základní typ tvaru  $k \cdot a^x$ ) se násobí konstantním faktorem (tj. podíl dvou naměřených hodnot po pevně zvoleném intervalu je konstantní). Geometrický průměr přímo reprezentuje tuto násobnou povahu změn. Pokud například máme dvě po sobě jdoucí měření s poměry

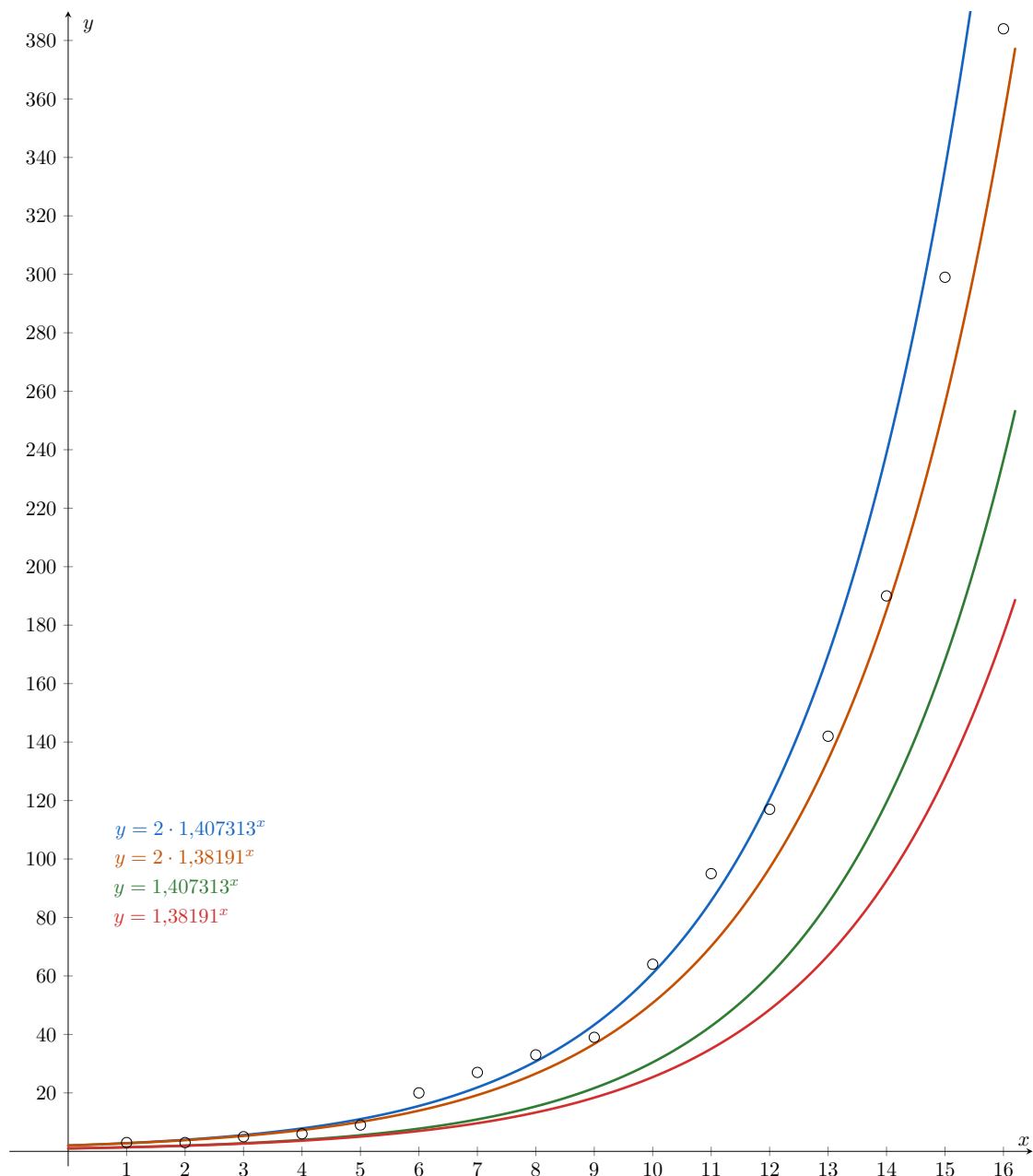
$$q_1 = 2, \quad q_2 = 0,5,$$

pak aritmetický průměr  $\bar{q} = 1,25$  by naznačoval stabilní růst, zatímco geometrický průměr  $G = 1$  správně reflektuje nulový čistý růst.

**Úloha 3.** Navrhňte funkci, která by aproximovala počet nakažených v jednotlivých dnech. Vytvořte graf v Geogebře (případně jiném programu).

*Řešení.* Metodou pokus-omyl se pro tvorbu exponenciální funkce, která vhodně reprezentuje naše data, nabízí například využití vypočítaných průměrů. Pro zápis funkce budeme využívat obvyklé značení, které je použité i v Geogebře a na obrázku níže. V grafu  $x$  značí čas ve dnech a  $y$  počet nakažených.

S aritmetickým průměrem získáme funkci  $y = 1,40731^x$ , s geometrickým průměrem  $y = 1,38191^x$ . Tyto funkce jsou od vyznačených bodů celkem daleko, proto je mírně upravíme:  $y = 2 \cdot 1,40731^x$ ,  $y = 2 \cdot 1,38191^x$ . Grafy navržených funkcí jsou zobrazeny v obrázku. Samozřejmě můžete získat ještě přesnější exponenciální funkce.



Obrázek 1: Navržené exponenciální funkce metodou pokus omyl

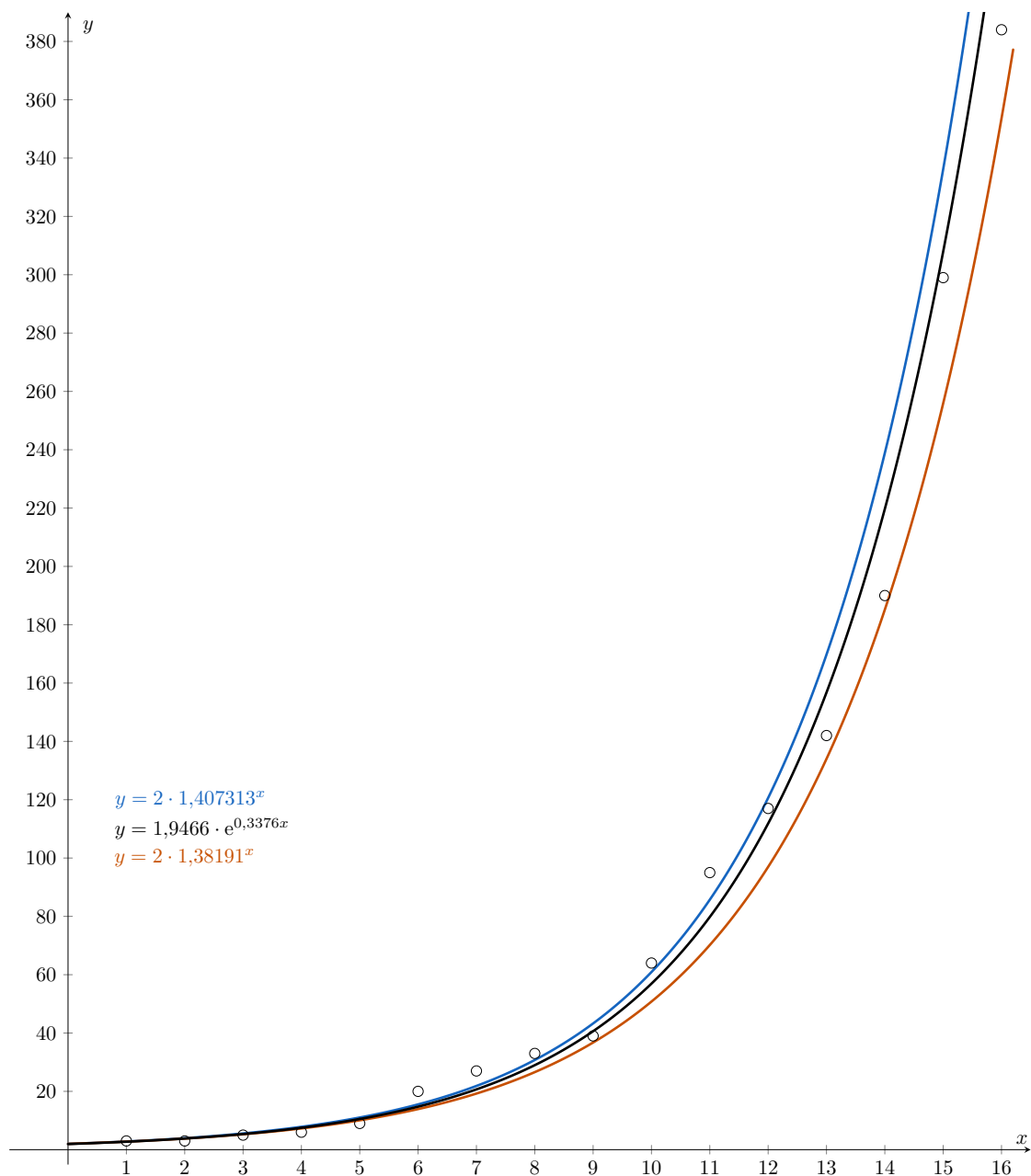
**Úloha 4.** Po využití regresní analýzy dat lze získat vhodnější funkci popisující chování počtu nakažených, která je tvaru  $y = 1,9466 \cdot e^{0,3376x}$ . Porovnejte v Geogebře (nebo v jiném vhodném softwaru) vaši funkci s touto funkcí. Vypočítejte hodnoty navržených funkcí pro 14. až 16. den, výsledek zaokrouhlete na celé číslo. Porovnejte je s hodnotami v tabulce.

*Poznámka.* Exponenciální funkci získanou regresní analýzou lze vytvořit v tabulkovém procesoru nebo Geogebře. V Geogebře je nutný příkaz, v němž zadáme body získané z tabulky: `RegreseExponencialni({(1,3),`

Results matter!

$(2,3), (3,5), \dots, (16,384)\}$ .

**Řešení.** Všechny tři funkce jsou v obrázku níže. Pouhým pohledem můžeme vidět, že všechny funkce jsou si na začátku podobné, ale pro další hodnoty funkce  $y = 1,9466 \cdot e^{0,3376x}$  vypadá jako nejvhodnější pro popis zdaných dat.



Obrázek 2: Srovnání navržených funkcí s funkcí dle regresní analýzy

Toto pozorování můžeme i kvantifikovat pomocí údajů z tabulky.

| Den ( $n$ ) | Počet nakažených ( $a_n$ ) | Počet dle funkce $y = 2 \cdot 1,40731^x$ | Počet dle funkce $y = 2 \cdot 1,38191^x$ | Počet dle funkce $y = 1,9466 \cdot e^{0,3376x}$ |
|-------------|----------------------------|------------------------------------------|------------------------------------------|-------------------------------------------------|
| 14          | 190                        | 239                                      | 185                                      | 220                                             |
| 15          | 299                        | 336                                      | 256                                      | 308                                             |
| 16          | 384                        | 473                                      | 354                                      | 432                                             |

Jako ukazatel přesnosti s jakou naše funkce reprezentují daná data můžeme uvážit, jak moc se liší jednotlivé hodnoty od sebe. Postupně tak dostaneme

$$|239 - 190| + |336 - 299| + |473 - 384| = 175$$

$$|185 - 190| + |256 - 299| + |354 - 384| = 92$$

$$|220 - 190| + |308 - 299| + |432 - 384| = 87.$$

Obvykle se ale pro posouzení přesnosti používají druhé mocniny (čtverce) uvedených rozdílů. A to proto, že malé rozdíly (menší než 1) se umocněním zmenšují a velké rozdíly se ještě zvětšují.

V tomto případě dostáváme

$$(239 - 190)^2 + (336 - 299)^2 + (473 - 384)^2 = 11691$$

$$(185 - 190)^2 + (256 - 299)^2 + (354 - 384)^2 = 4174$$

$$(220 - 190)^2 + (308 - 299)^2 + (432 - 384)^2 = 3285.$$

Tedy i v tomto případě se jako nejlepší ukázala třetí křivka. Toto není žádná náhoda, jelikož křivka vytvořená pomocí regrese je křivka, která má tuto chybu nejmenší.

**Úloha 5.** Bude se epidemie šířit neustále podle navržené funkce? Co může mít vliv na její další chování?

**Řešení.** Pro zachování exponenciálního průběhu epidemie by se nesměly měnit podmínky šíření nemoci. V praxi se lidé začnou chránit ochrannými prostředky, minimalizovat kontakt s ostatními nebo očkovat, čímž se nárůst počtu nakažených sníží. Epidemie se dostane do svého vrcholu a počty nakažených se začnou snižovat. Pro trvalý exponenciální růst by také muselo existovat neomezené množství jedinců, kteří se mohou danou nemocí nakazit.

## Literatura

- European Centre for Disease Prevention and Control *Novel Coronavirus: three cases reported in France*. Dostupné z [https://www.flickr.com/photos/ecdc\\_eu/50321985653/in/dateposted/](https://www.flickr.com/photos/ecdc_eu/50321985653/in/dateposted/) [cit. 08.08.2024]
- Ministerstvo zdravotnictví ČR. *Onemocnění aktuálně - COVID 19*. Dostupné <https://onemocneni-aktualne.mzcr.cz/api/v2/covid-19> [cit. 30.8.2023]