

POZVÁNKA NA KONFERENCI AMSE 2025 INVITATION TO THE AMSE 2025 CONFERENCE

Redakce

The **Applications of Mathematics and Statistics in Economics** conference aims to acquaint its participants with the latest mathematical and statistical methods that can be used in solving theoretical and practical problems and challenges of economics. The conference gives a unique chance to present research results achieved in the area where mathematics and statistics border with economics.

Conference date: 27. – 31. August 2025

Conference venue: Hotel U královny Elišky, Malé náměstí 117, Hradec Králové, The Czech Republic, <https://hotelukralovnyelisky.cz/en>

Website of the conference series: <http://www.amse-conference.eu>

Important dates:

- 11th May 2025 – Abstract and registration of participants (including co-authors).
- 30th May 2025 – Notification of abstract acceptance/rejection.
- 13th June 2025 – Conference fee payment.

Papers should be thematically related to applications of mathematics and statistics in economics and should fall in one of the following sections.

- Macroeconomics, Public Economics and Methodological Issues of Economics,
- Social Economics, Economic Sustainability and Demographic Economics,
- Microeconomic Issues,
- Financial Markets, Risk Measurement and Insurance,
- Multidimensional Statistics in Economics.

Types of participation in the conference. There are three kinds of participation at the conference:

- active participation with a presentation of the paper that will be included in one of three journals,
- active participation with a presentation of the paper that will not be included in journal,
- passive participation without a presentation.

Contact details:

- Organizing committee: amse2025@vse.cz
- Administrative issues: zivalova@vse.cz

ZVOLENA NOVÁ PŘEDSEDKYNĚ ČSTS NEW CHAIRWOMAN ELECTED

Výbor společnosti

Vážené kolegyně, vážení kolegové!

Účastníci členské schůze České statistické společnosti konané dne 30. ledna 2025 v Praze zvolili Martinu Litschmannovou novou předsedkyní ČStS.

Srdečně gratulujeme a přejeme hodně zdaru!

Výbor společnosti



*Ondřej Vencálek gratuluje ke zvolení
nové předsedkyni ČStS Martině Litschmannové*

ČASOVÝ ROZVOJ EPIDEMIE A MATEMATIKA THE DEVELOPMENT IN TIME OF AN EPIDEMIC AND MATHEMATICS

Zdeněk Půlpán

Adresa: Univerzita Pardubice, Dopravní fakulta Jana Pernera, Studentská 95, 532 10 Pardubice 2; Na Brně 1952, 500 09 Hradec Králové 9

E-mail: zdenek.pulpan@upce.cz, zdenek.pulpan@post.cz

Abstrakt: Vývoj epidemie v čase lze popsat matematicky mnoha způsoby. Každý popis má svou hlavní ideu. Příspěvek poukazuje na matematická zpracování některých základních předpokladů popisu šíření pandemie v modelech s diskretním a spojitým časem.

Klíčová slova: modely šíření pandemie v čase, logistická funkce.

Abstract: The development of an epidemic over time can be described mathematically in many ways. Each description has its main idea. The paper points to the mathematical processing of some basic assumptions of the description of the spread of the pandemic in discrete and continuous time models.

Keywords: pandemic spread models over time, logistic function.

MSC: A40, E20, B50, F10.

1. Úvod

Pandemií se nazývá podle [10] rychle se šířící epidemie po všech světadílech. Někdy se ještě zdůrazňuje takové rozšíření onemocnění, že nestačí léčit nemocné, ale k zabrzdění šíření onemocnění je třeba organizovat ochranu i zdravých jedinců. Světová zdravotnická organizace v dubnu 2008 tuto definici zeslabila a jako pandemií označuje každou infekci, která se šíří alespoň ve dvou státech jednoho regionu WHO a současně aspoň v jednom státě jiného regionu; regiony WHO zhruba odpovídají kontinentům, Asie je rozdělena na více regionů.

Význam matematiky si společnost opět mohla uvědomit v období pandemie Covid 19. Státy celého světa pochopily, že je třeba se v rozhodování opřít o *matematicky formulovanou* prognózu vývoje pandemie. Důležitá jsou při tom data, získaná od zdravotníků a znalosti a zkušenosti z předchozích pandemií. Matematika na podobné situace byla sice teoreticky připravená, na základě mnoha rozličných matematických pandemických modelů byl pokusně sestavován očekávaný vývoj v čase (hlavně odhad počtu nakažených jedinců)

měnit. Pokud určitou změnu hodnoty parametru v průběhu pandemie nezaregistrujeme včas, bude špatným odhadem ovlivněna další předpověď, i když představa o mechanismu šíření pandemie bude správná. Hodnoty parametrů, získané na počátku pandemie, prognózuji zde celý její vývoj (což neodpovídá realitě).

Chtěli jsme především ukázat, jak mnohostranné znalosti (a to nejen matematické) potřebuje matematik, který se věnuje aplikacím.

Poděkování: Autor děkuje ing. Pavlu Střížovi, Ph.D., za technické dopracování tohoto příspěvku.

Citovaná a doporučená literatura

- [1] Berec, L.: Když se matematika potká s biologií: matematická epidemiologie, *Pokroky matematiky, fyziky a astronomie*, roč. 63, č. 2 (2018).
- [2] Bonita, R., Beaglehole, R., Kjellström, T.: *Basic epidemiology*, 2nd edition, World Health Organization 2006, ISBN 92-4-154707-3, 205 s.
- [3] Wassertheil-Smoller, S.: *Biostatistics and Epidemiology: A Primer for Health and Biomedical Professionals*, Springer, 2004. ISBN 0-387-402-926.
- [4] Anděl, J.: *Statistické metody*, MatfyzPress 2019, 300 s.
- [5] Jekel, J., Katz, D., Elmore, J.: *Epidemiology*, Biostatistics and Preventive Medicine, Philadelphia, Penn: Saunders 2001.
- [6] Cipra, T.: *Statistická analýza časových řad s aplikacemi v ekonomii*, Praha, SNTL/Alfa 1986. *cit. 16*
- [7] Anděl, J.: *Statistická analýza časových řad*, Praha, SNTL 1976. *cit. 16*
- [8] Kermack, W. O., Mc Kendrick, A. G.: A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character*, 115(772), 700–721. *cit. 15*
- [9] Holčík, J.: *Modelování a simulace biologických systémů*, Praha, ČVUT, 2006, ISBN 80-01-03470-4. *cit. 16*
- [10] *Malý encyklopedický slovník A–Ž, heslo pandemie*, NČSAV Praha 1972. *cit. 4*
- [11] Wei-Chau Xie: *Differential Equations for Engineers*, Cambridge University Press, 2010, ISBN 978-511-77622-9. *cit. 18*
- [12] Hurt, J.: *Teorie spolehlivosti*, Praha, SPN 1984.

Pochopitelně toto je jen ukázkový příklad. Pro seriózní odhady bychom potřebovali mít precizní šetření na větším počtu pacientů. Teprve potom by bylo možné příslušným odhadům dát odpovídající interpretaci.

Zkusme do vztahu pro $\lambda(t)$ za odhad hustoty vložit hustotu exponenciálního rozdělení $f(t) = \alpha \cdot e^{-\alpha t}$. Protože je pro toto rozdělení $ET = \alpha^{-1} \sim \bar{T} = 22,3$ a dosazením příslušných vztahů máme

$$\lambda(t) = \frac{\alpha \cdot e^{-\alpha t}}{1 - (1 - e^{-\alpha t})} = \alpha,$$

dostáváme hodnotu nezávislou na čase a je pak $\hat{\lambda} \sim \frac{1}{\bar{T}} \sim 0,045$.

Tabulka 5: V tabulce jsou v levé části posledního sloupce uvedeny odhady funkce $\lambda(t)$ pro fiktivní onemocnění šesti pacientů z empirických dat a v pravé části je odhad, předpokládající $f(t)$ jako odhad hustoty normálního rozdělení pro sledovaná vstupní data.

Dny	$\hat{F}(t)$	$\Delta \hat{F} / \Delta t$	$\hat{\lambda}(t)$	
10	0	0	0,00	0,010
15	1/6	1/30	0,04	0,049
17	2/6	1/12	0,17	0,078
21	3/6	1/24	0,08	0,161
24	4/6	1/18	0,17	0,241
27	5/6	1/18	0,33	0,335
30	1	1/18	∞	0,435

Předpokládejme ještě, že náhodná veličina T má přibližně normální rozdělení $N(\mu; \sigma^2)$. Pak v našem případě je $\mu \sim 22,3$ a $\sigma^2 \sim 33,5$. Příslušné odhady hodnot funkce $\lambda(t)$ jsou v pravé části posledního sloupce Tabulky 5.

5. Závěr

Ukázali jsme některé zjednodušené matematické modely, které se snaží o prognózu vývoje počtu nakažených jedinců v čase. Jejich parametry jsou především funkcí příslušného druhu epidemie, ale také funkcemi primární imunity populace i pandemické situace, spočívající například v rozdílných kontaktech osob infikovaných s neinfikovanými a jsou také funkcemi času (v závislosti na tom, jak postupně populace získává imunitu) i roční době, rozsahu testování osob a podobně. Parametry modelů se proto během rozvoje pandemie mohou

za určitých předpokladů šíření neznámého Covidu 19, přesto bylo zpočátku obtížné konkrétní model šíření této nákazy sestavit vzhledem k neznalosti vstupních parametrů modelu v naší populaci. (Pandemie se ve světě objevovaly od starověku až do dnešní doby a dá se očekávat, že nás budou provázet i v budoucnosti přes velký pokrok současné medicíny. Známý jsou pandemie černého kašle, spalniček, příušnic, eboly, HIV/AIDS, SARS, dýmějového moru, žloutenky atd. Jak se ukázalo později, v různých populacích se pandemie šířila s rozdílnými parametry vzhledem k rozdílné odolnosti vůči ní v rozdílných oblastech; například v Africe se šíří daleko pomaleji než v Číně, Americe či Evropě).

Použitelné matematické modely šíření pandemie jsou sice formálně složitější, jejich parametry musí zachycovat vlastnosti populace ke konkrétní variantě viru, ale i prostředí, ve kterém působí; základní typy závislosti počtu nakažených na čase u různých virových onemocnění jsou však formálně velmi podobné. Šíření nemoci závisí ale nejen na jejím druhu, ale i na pořadí působení jednotlivých druhů virů, roční době a způsobu ochrany lidí. Avšak teprve na základě dlouhodobější analýzy registrovaného průběhu šíření určité infekce lze dobrý model pro konkrétní pandemii vytvořit. Důležité je předem stanovit, ze kterých vstupních dat při modelování můžeme vyjít a jaké mohou být apriorní představy o přenosu konkrétní studované infekce, tedy jaké jsou znalosti a zkušenosti lékařů s dřívějšími vývoji pandemií a odlišnosti pandemie studované vzhledem k předchozím pandemiím.

A tu byl v poslední době velký problém s informacemi, protože žádné zkušenosti s šířením tohoto druhu Covid 19 u nás ani ve světě nebyly. Protože tvorba přiléhavého modelu je možná až po určitém rozvinutí pandemie, nelze v krátkém čase od vypuknutí pandemie použít některé matematické postupy, které předpokládají rozsáhlý soubor časově závislých dat. Zejména se nelze spoléhat na extrapolace interpolací nebo regresní modely s mnoha parametry. K dispozici byla nejprve pouze data z počátečního rozvoje pandemie. To je velké riziko možnosti včasné prognózy vývoje pandemie. Statistici spolu s lékaři v takových případech bojují o čas, musí se včas spojovat znalosti medicínské s biostatistickými k vytvoření aspoň přibližného matematického modelu předpokládaného vývoje s vhodnými parametry (to znamená například dobře odhadnutelnými) a teprve potom matematickou formu upřesňovat a model postupně ověřovat.

V následujících příkladech naznačíme, jaké jsou možnosti volby matematického modelu na základě znalosti jen relativně malého počtu počátečních dat z celkového časového průběhu epidemie. Je pochopitelné, že čím delší je pozorování a čím více specifických vlastností šíření epidemie zachytíme, tím přesněji jsme schopni situaci popsat.

2. Modely šíření pandemie v čase

Ukážeme některé velmi jednoduché modely vývoje pandemie v čase s použitím určitých vstupních dat a představ o jejím šíření. Předpokládejme, že získáváme data z určité oblasti. Nejprve popíšeme tvorbu diskrétního modelu šíření epidemie, pak se budeme věnovat modelům se spojitým časem. Modely s představou diskrétně odhadovaného času většinou vycházejí z určité představy o mechanismu šíření epidemie mezi jedinci. Modely se spojitým časem se opírají o histogram zjištěného počtu nakažených v čase a snaží se to odhadnout funkcí $n(t)$ vyjadřující počet nakažených v určité oblasti v okamžiku t .

a) Diskrétní model šíření

Začneme úplně jednoduchým modelem. Zavedeme veličiny

n_t ... počet nakažených za čas t ,

k_1 ... střední (očekávaný) počet osob, které nakazí infikovaná osoba za jednotku času.

Začneme předpokladem, že v počátečním čase $t = 0$ je v populaci jediný nakažený pacient. Pak v čase $t = 1$ bude v populaci $n_1 = 1 + k_1$ nakažených osob, v čase $t = 2$ bude v populaci $n_2 = n_1 k_1 = (1 + k_1)k_1$ nakažených jedinců, v čase $t = 3$ to bude $n_3 = n_2 k_1 = (1 + k_1)k_1^2$ nakažených osob atd. V obecném čase t bude za předpokladu, že žádná z nich se dosud neuzdravila, roven $n_t = (1 + k_1)k_1^{t-1}$.

Parametr k_1 je velice obtížné, pokud vůbec, měřitelný. Proto budeme používat parametry jiné:

p ... pravděpodobnost, že nakažená osoba při jednom kontaktu nakazí osobu zdravou,

k ... střední počet kontaktů jedné osoby za jednotku času.

Pokud je v populaci velice málo (zanedbatelně mnoho) nakažených osob, můžeme psát $k_1 = pk$. Při tomto označení dostaneme výsledný model počtu nakažených jedinců v populaci ve tvaru $n_t = (1 + pk) \cdot (pk)^{t-1}$ kde t je počet časových jednotek po prvním nakaženém.

Příklad 1

Předpokládejme, že $p = 0,7$; $k = 5$; pak je $pk = 3,5$; $n_t = 4,5 \cdot 3,5^t$ a tedy podle předchozího vzorce dostaneme (zaokrouhlené hodnoty na jednotky jsou uvedeny v Tabulce 1):

$$\begin{aligned} n_1 &= 4,5; n_2 = 15,75; n_3 = 55,125; n_4 = 192,9375; n_5 = 675,28125; \\ n_6 &= 2363,484375; n_7 = 8272,1953125; \dots \end{aligned}$$

$$\begin{aligned} \lambda(t) &= \lim_{h \rightarrow 0^+} \frac{P(t < T \leq t+h | T \geq t)}{h} = \lim_{h \rightarrow 0^+} \frac{P(t < T \leq t+h)}{P(T \geq t)} \cdot \frac{1}{h} = \\ &= \lim_{h \rightarrow 0^+} \frac{F(t+h) - F(t)}{h} \cdot \frac{1}{P(T \geq t)} = \frac{f(t)}{1 - F(t)} = \\ &= \frac{f(t)}{R(t)} = \frac{-R'(t)}{R(t)} = -\frac{d \ln R(t)}{dt}. \end{aligned}$$

Z posledního vztahu pro $\lambda(t)$ vyplývá

$$\ln R(t) = - \int_0^t \lambda(s) ds = -\Lambda(t).$$

Takže máme pro $R(t)$ a $F(t)$ výraz

$$R(t) = e^{-\Lambda(t)} = 1 - F(t).$$

Derivováním předchozího vztahu dostaneme pro hustotu pravděpodobnosti $f(t)$ vztah

$$f(t) = \lambda(t) \cdot e^{-\Lambda(t)},$$

který vyjadřuje hustotu $f(t)$ pomocí funkce $\lambda(t)$.

V našem případě bychom mohli funkci $\lambda(t)$ interpretovat jako intenzitu obranyschopnosti organismu v čase t se vypořádat s chorobou. Otázkou je, jaký tvar funkce $f(t)$, která je klíčová pro popis rozvoje choroby, předpokládat. Opět si připomeneme technické obory, které zde předpokládají nejčastěji exponenciální nebo Weibullovo rozdělení. V epidemiologických šetřeních však funkce $f(t)$ může mít jiný tvar. Je však možné postupovat klasickým způsobem a vyjít z experimentu, distribuční funkci $F(t)$ nahradit empirickou distribuční funkcí a využít její empirické derivace.

Příklad 5

Bylo sledováno 6 pacientů, kteří onemocněli stejnou chorobou, byli stejně staří a před vypuknutím choroby byli zcela zdraví. Doba jejich nemoci sledovanou chorobou byla měřena ve dnech. Jednotliví pacienti byli nemocni postupně v počtu 15, 21, 17, 30, 24, 27 dní. Tomu odpovídající empirická distribuční funkce $\hat{F}(t)$ je uvedena v Tabulce 5. V téže tabulce jsou pak ještě uvedeny relativní difference $\Delta \hat{F} / \Delta t$ jako odhady derivací a pak odhady hodnot funkce λ (v levé části sloupce) ve tvaru

$$\hat{\lambda}(t) = \frac{\Delta \hat{F} / \Delta t}{1 - \hat{F}(t)}.$$

teprve v čase $t \rightarrow +\infty$ by se počet nakažených ustálil na hodnotě blízké $\frac{a}{b}$ (viz vztah (B)), a to je stav stabilní.

Poznámka: Porovnáme-li výraz (6) s výrazem (A), označení y odpovídá četnosti n , konstanta r odpovídá konstantě a , konstanta K odpovídá podílu $\frac{a}{b}$. Veličina n je tedy diskrétní, zatím co veličina y je spojitá. V naznačených aplikacích, kdy veličinu y chápeme jako odhad relativní četnosti měření na rozsáhlém souboru, to nemusí být problém.

Dalším příkladem autonomní rovnice je rovnice s rozdílem dvou funkcí proměnné y na pravé straně ve tvaru

$$\frac{dy}{dt} = g(y) - h(y).$$

Pro stacionární bod y_0 této diferenciální rovnice platí, že $g(y_0) = h(y_0)$. Bod y_0 je stabilní, když $\frac{dg}{dy} < \frac{dh}{dy}$, nestabilní, když $\frac{dg}{dy} > \frac{dh}{dy}$. Pro grafy funkcí $g(y)$ a $h(y)$ to znamená, že v bodě y_0 v případě stability se setkávají obě funkce tak, že vlevo od bodu y_0 je funkce h níže než funkce g (vpravo je funkce h výše než funkce g). V případě nestability je to obráceně.

Pokračující zobecnění našich úvah by spočívalo v aplikaci *autonomního systému diferenciálních rovnic* na popis časového vývoje epidemie. Jsou to soustavy diferenciálních rovnic prvního řádu, jejichž pravé strany nezávisí na čase. [11]

4. Využití teorie spolehlivosti

Zatím co v klasické teorii spolehlivosti se uvažuje čas T jako náhodná veličina, označující dobu bezporuchové činnosti nebo v biologii dobu života určitého druhu živočichů, zde se jako nezáporná náhodná veličina T označuje doba trvání určité nemoci. Distribuční funkce této náhodné veličiny pak je $F(t) = P(T \leq t)$. Distribuční funkce v našem případě označuje pravděpodobnost, že doba nemoci nepřekročí t . Častěji se ale používá pravděpodobnost $R(t) = 1 - F(t) = P(T > t)$, že po čas t se pacient nevyléčí (v teorii spolehlivosti označuje $R(t)$ míru spolehlivosti – spolehlivost v čase t). Uvažujeme-li spojitost funkce $R(t)$ pro všechna nezáporná t , můžeme předpokládat existenci hustoty pravděpodobnosti náhodné veličiny T , tj. nezáporné funkce $f(t)$ pro $t \geq 0$, že platí $\frac{dF}{dt} = f(t)$ a také $\int_0^\infty f(t) \cdot dt = 1$. Pro funkci $R(t)$ z toho vyplývá, že $R(t) = 1 - \int_0^t f(s) \cdot ds$. Pak také střední doba nemoci ET se dá vyjádřit integrálem $ET = \int_0^\infty t \cdot f(t) dt$.

V teorii spolehlivosti se zavádí ještě funkce $\lambda(t)$, která vyjadřuje intenzitu poruch (nebezpečí poruchy) vztahy

Tabulka 1: Výsledky z Příkladu 1, zaokrouhlené na jednotky.

t	n_t
1	4
2	16
3	55
4	193
5	675
6	2363

Počet nakažených se tedy podle tohoto modelu vyvíjí jako geometrická posloupnost (indexovaná časem) s kladným kvocientem pk . Podle učiněného předpokladu může tento model adekvátně popisovat šíření choroby v populaci jen po čas, který je dostatečně krátký na to, aby se žádná z nakažených osob neuzdravila a celkový počet nakažených byl zanedbatelně malý vzhledem k celkové velikosti populace. Z geometrického (tj. neohraničeného) růstu počtu nakažených je také zřejmé, že takto jednoduchý model může popisovat pouze počáteční fázi epidemie. Přesto z něho můžeme učinit jeden rozumný závěr: pokud $pk > 1$, choroba se začne v uvažované oblasti šířit.

Jestliže se po střední době l nemocný vyléčí s pravděpodobností q , je aktuální střední počet nakažených v čase t roven n_t^*

$$\begin{aligned} n_t^* &= n_t & \text{pro } 1 \leq t < l, \\ n_t^* &= n_{t-1}^* \cdot (pk) - n_{t-l}^* \cdot q & \text{pro } t \geq l. \end{aligned}$$

Z předchozího vztahu je možné explicitní vyjádření pro n_t^* určit (je to lineární diferenční rovnice l -tého stupně).

Předpokládáme zde skokové získávání dat a konstantnost parametrů k, p a q . „Správnost“ vztahu je však nutné experimenty ověřit. (Slovo „správnost“ dáváme do uvozovek, protože jeho význam je v tomto případě relativní. Je věcí našeho názoru, kdy budeme model považovat za správný. „Správnost“ modelu totiž nemusí potvrzovat ani numerický souhlas výpočtu podle odvozeného vztahu s právě naměřenými daty, například jiný experimentální výsledek nemusí odpovídat dřívější předpovědi.)

Je ovšem ještě možné, že pravděpodobnosti p nebo q se během epidemie mění (třeba v závislosti na počtu vyléčených), jsou tedy funkcemi času i způsobu, jak se populace začíná bránit šíření. Odhad parametru k modelu se může měnit i se změnou podmínek kontaktu osob (závisí na jeho formě

i délce). V době pandemie je vysoká individuální variabilita důsledků kontaktů.

Jak uvidíme v následujícím příkladu, podle našeho modelu střední doba léčby l má významný vliv na vývoj počtu nakažených.

Příklad 2

Volme $p = 0,7$; $k = 5$; $l = 3$; $q = 0,8$, pak je $pk = 3,5$. Dosazením do předchozích vzorců dostaneme hodnoty, uvedené v Tabulce 2, které jsou zaokrouhlené na jednotky.

Tabulka 2: Výsledky z Příkladu 2, zaokrouhlené na jednotky.

t	n_t	n_t^*
1	4	4
2	16	16
3	55	55
4	675	189
5	2363	648
6	8272	2225

Předpokládejme ještě, že když r je poměrný počet osob s uvažovanou chorobou zemřelých za jednotku času, pak $1 - r$ je pravděpodobnost, že nakažená osoba přežije časový interval jednotkové délky. Počet nakažených n_t^* po t časových jednotkách musí splňovat následující vztahy

$$\begin{aligned} n_0^{**} &= n_0 = 1, \\ n_1^{**} &= (1 + pk) \cdot (1 - r) = n_1 \cdot (1 - r), \\ n_2^{**} &= (1 + pk) \cdot (pk) \cdot (1 - r)^2 = n_2 \cdot (1 - r), \\ &\dots \\ n_t^{**} &= n_t \cdot (1 - r) \quad \text{pro } 1 \leq t < l, \\ n_t^{**} &= n_{t-1}^{**} (pk) \cdot (1 - r) - n_{t-l}^{**} \cdot q \quad \text{pro } t \geq l. \end{aligned}$$

Ještě připomínáme, že bychom model mohli ještě doplnit možností *testování* na infekci, protože to se ukázalo během pandemie jako důležitý činitel, zpomalující šíření nemoci nebo vliv na podíl zemřelých z počtu dříve nakažených atd. Pro přenos infekce je také důležitá délka nebo i forma kontaktu.

ního počtu nakažených v epidemii velmi často odpovídá jistému typu řešení diferenciální rovnice. Příkladem mohou být *autonomní diferenciální rovnice*. To jsou diferenciální rovnice tvaru

$$\frac{dy}{dt} = f(y), \quad \text{kde } y = y(t); \quad t \in \langle 0; t_1 \rangle. \quad (4)$$

Předchozí rovnice je separovatelná a je ji možno řešit integrací (za předpokladu $f(y) \neq 0$) výrazů na levé i pravé straně rovnice

$$\frac{dy}{f(y)} = dt.$$

Jestliže $f(y_0) = 0$, je konstantní funkce $y(t) = y_0$ řešením rovnice (4).

Řešení rovnice (4) vykazuje, když $\frac{df(y_0)}{dy} < 0$, tzv. asymptotickou stabilitu, kdy malá výchylka od y_0 vede na řešení, které konverguje opět k tomuto bodu y_0 . Naopak, když $\frac{df(y_0)}{dy} > 0$, je řešení nestabilní.

Příkladem diferenciální rovnice (4) může být rovnice ve tvaru

$$\frac{dy}{dt} = -ky, \quad \text{kde } k > 0. \quad (5)$$

Její řešení $y = y(t) = Ce^{-kt}$ se blíží k $y = 0$, protože v tomto případě je $\frac{df}{dy} < 0$. Bod $y = 0$ odpovídá tzv. stacionárnímu řešení (pro $k < 0$ by řešení rostlo neomezeně).

Když by se relativní výskyt počtu nakažených řídil předchozím vztahem (5), znamenalo by to, že pro čas $t \rightarrow +\infty$ by byl výskyt epidemie nulový, a to je stabilní stav.

Oběcný tvar *logistické diferenciální rovnice* s konstantními koeficienty je tvaru

$$\frac{dy}{dt} = f(y) = ry \left(1 - \frac{y}{K} \right), \quad r > 0, \quad K > 0. \quad (6)$$

Tato diferenciální rovnice má dva stacionární body, které se získají řešením rovnice

$$f(y) = ry \left(1 - \frac{y}{K} \right) = 0.$$

Dvě řešení $y = 0$ a $y = K$ odpovídají nestabilnímu a stabilnímu bodu. Pro první kořen je $\frac{df}{dy} > 0$ (to odpovídá nestabilnímu bodu), pro druhý je $\frac{df}{dy} < 0$ (to odpovídá bodu stabilnímu).

Když by se počet nakažených $n(t)$ řídil vztahem (6), mohl by být v určitém okamžiku počet nakažených hodně malý, ale tento stav by nebyl stabilní,

Vztah vyplynul z přímé úměrnosti úbytku nemocných v závislosti na čase a celkovým počtem infekčních jedinců. To je vidět z prvního vztahu pro derivace S, I, R podle času, ze kterého dostaneme pro $\frac{dR}{dt}$ podmínku, vyplývající z předchozích vztahů v jednoduchém tvaru $\frac{dR}{dt} = a \cdot I$. Z modelu SIR vyplývá, že epidemie propukne, když $\frac{rS}{a} > 1$. Viz např. [9]. Pro posouzení časového průběhu epidemie je nejdůležitější závislost $I = I(t)$.

Dlouhodobější vývoj epidemie popisuje model $SIRS$, který předpokládá časovou ztrátu odolnosti vůči infekci (ze stavu R je možné přejít zase na začátek do stavu S , tento přechod je charakterizován mírou ztráty odolnosti). Někdy je nutné zařadit do modelu ještě i skupinu infikovaných, kteří ještě nemají žádný příznak onemocnění. Pro ně použijeme znak E (exposed) a model pak předpokládá přechod ze stavu S do stavu E a z něho teprve do stavu I a pak R (model se označuje $SEIR$). Takový model předpokládá ale dobré trasování.

Pro rychlé posouzení předpokládaného vývoje infekce se používá ukazatel R_0 , nazývaný *základní reprodukční číslo*. Charakterizuje, kolik k nákaze náchylných osob se přibližně nakazí od infekční osoby za předpokladu běžného chodu života. Ukazatel R_t se nazývá *efektivní reprodukční číslo* a počítá se za ovlivnění běžného chodu života protiepidemickými opatřeními, je tedy časově závislý. Pro poněkud objektivizovaný odhad efektivního reprodukčního čísla stanovil Institut Roberta Kocha přesnější výpočet, který zohledňuje období nakažlivosti jedince a zpožděnou diagnostiku onemocnění, ve tvaru odpovídajícím určitému typu infekce. Pro infekci Covid 19 to byl podíl dvou sedmidenních součtů nově nakažených, posunutých vzájemně o 5 dní. Za významnou hodnotu reprodukčního čísla se bere hodnota 1. Je-li reprodukční číslo větší než 1, epidemie se pravděpodobně šíří, je-li menší než 1, dochází pravděpodobně k útlumu šíření. Vždy je však třeba počkat s jistějším tvrzením několik dní, reprodukční číslo je funkcí času. Více než reprodukční číslo charakterizuje vývoj pandemie graf, zachycující časový průběh počtu nakažených. Z něho příslušnou tendenci šíření nebo útlumu pandemie odhadneme s větší objektivitou na základě použití matematické teorie časových řad. [6, 7]

3. Autonomní rovnice

Znalosti o obvyklém průběhu závislosti u časově omezené řady měření (například počtu nakažených v populaci určitého rozsahu v jistém počátečním období) u zkoumané epidemie nám umožňují předpovědět přibližný průběh vývoje choroby i v časech, kdy neznáme experimentální data. K odhadu matematické formy hledaného vztahu pro předpověď slouží i znalosti řešení určitých typů diferenciálních rovnic. Empiricky získaná časová závislost relativ-

Proto lze očekávat v hodnotách p velké individuální rozdíly (tedy velkou variabilitu odhadu).

Příklad 3

Volme následující hodnoty parametrů matematického modelu šíření epidemie $p = 0,7$; $k = 5$; $pk = 3,5$; $l = 3$; $q = 0,8$; $r = 0,02$. Pak dosazením do předchozího vzorce dostaneme hodnoty (zaokrouhlené na celá čísla), uvedené v Tabulce 3.

Tabulka 3: Hodnoty vypočtené v Příkladu 3, zaokrouhlené na jednotky.

t	n_t	n_t^*	n_t^{**}
1	4	4	4
2	16	16	15
3	55	55	51
4	675	189	173
5	2363	648	581
6	8272	2225	1951

Předchozí modely s diskrétním časem se pro popis reálných situací používají velmi málo. Obtížně se z pozorování stanovují jejich parametry $k, p, q, \dots$. Jsou založeny na představě o šíření epidemie, která může mít během rozvoje pandemie rozdílné podoby.

Pokusme se nyní řešit problém počtu nakažených za předpokladu spojitěho časového záznamu průběhu epidemie.

b) Snahy o popis průběhu epidemie se spojitým časem

Myslím, že nebude na škodu uvést nejprve i příklady, které náš problém vhodné formulace časové závislosti počtu nakažených jedinců *neřeší*. Snahou je ukázat, že sice matematika nabízí řadu aproximačních metod, ale jen některé vedou k uspokojivé prognóze časového rozvoje určité pandemie. Zkušenost říká, že je nutné využít co nejvíce znalostí z předchozích pandemií. Vyjdeme-li u jednotlivých modelů ze stejných vstupních dat, ale protože každý z modelů je jiný, prognózuje jiné hodnoty $n(t)$. Chceme naznačit, že otázka vhodnosti modelu není tolik otázkou aktuálního souhlasu předpovědi s naměřenými daty, ale otázkou nalezení co nejobecnějšího vztahu, který by vedl

k jeho fungování jako nástroje postačující předpovědi i v poněkud rozdílných situacích.

Označme $n(t)$ počet jedinců nakažených v čase t . Přitom víme, že počet nakažených vždy nejprve roste (v závislosti na podmínkách), později se šíření zpomaluje, až se nakonec zastaví; zde nyní budeme předpokládat, že v určitém časovém intervalu představuje $n(t)$ polynomicke závislost (polynomicke závislosti se nejčastěji uvažují jako prvotní zkušební modely) počtu nakažených na čase.

Již jsme se zmínili o tom, že z histogramů časové závislosti počtu nakažených vyplývá, že počáteční růst počtu nakažených je po určité době vystřídán jeho poklesem, proto se dá předpokládat, že i spojitý průběh závislosti $n(t)$ počtu nakažených na čase bude mít podobný tvar. To se dá matematicky vyjádřit derivací funkce

$$\frac{dn}{dt} = at - bt^2,$$

kde $a, b > 0$ jsou konstanty, vztažené ke konkrétní epidemiologické situaci.

Z předchozího vztahu vidíme, že extrém funkce $n(t)$ nastává pro $t = 0$ a $t = \frac{a}{b}$. $t = \frac{a}{b}$ v intervalu $(0; \frac{a}{b})$ je derivace kladná, tj. funkce $n(t)$ je rostoucí, v intervalu $(\frac{a}{b}; +\infty)$ je záporná, tj. funkce $n(t)$ je klesající. (To ovšem bezprostředně plyne z toho, že konstanty a, b jsou kladné.)

Obecné řešení předchozí diferenciální rovnice je zřejmě neurčitý integrál z její pravé strany, tedy

$$n = n(t) = \frac{1}{2}at^2 - \frac{1}{3}bt^3 + C,$$

kde C je integrační konstanta.

Použijeme jen tu část polynomu $n(t)$, pro něž je $t \in \langle 0; t_m \rangle$ a kde t_m je jeho největší kladný kořen.

Volíme-li počáteční podmínky $n(0) = n_0$, $n(1) = n_1$, $n(2) = n_2$, dostaneme předchozí řešení s parametry

$$C = n_0; a = -0,5n_2 + 4n_1 - 3,5n_0; b = -0,75n_2 + 3n_1 - 2,25n_0.$$

Pro konkrétní vyjádření této závislosti použijme data (pro počáteční časové okamžiky rozvoje pandemie) z předchozího příkladu. Dosadíme-li $n_0 = 1$, $n_1 = 4$, $n_2 = 15$, dostaneme $a = 5$, $b = -1,5$, což neodpovídá předpokladu $b > 0$. Navíc pro $n(t)$ vyjádření $n(t) = 2,5t^2 + 0,5t^3 + 1$ již na první pohled nevyhovuje představě o možném časovém průběhu pandemie (pro uvedená data jsou všechny koeficienty nezáporné, znamená to neomezený růst počtu

který se může nakazit „je ve stavu S “, jedinec, který se nakazil v důsledku infekce „je ve stavu I “, atd. Celková velikost populace pak je $N = S + I$. Každý jedinec může přejít ze stavu S do stavu I . Model používá jediný koeficient, který označuje míru nakažlivosti $r > 0$ ve vztazích pro časové změny veličin S a I . Předpokládá pro každého jedince možnost přechodu ze stavu S do stavu I daný nám již známými podmínkami $\frac{dS}{dt} = -r \cdot S \cdot I$; $\frac{dI}{dt} = r \cdot S \cdot I = r \cdot (N - I)I = r \cdot N \cdot I - rI^2$. Porovnáním s předchozím modelem (A) je zde označeno $n = I$; $a = rN$; $b = r$. Jak jsme již viděli, řešení pro $n = I$ vede k logistické funkci.

S modelem SI souvisí model SIS , kde se uvažuje ještě o přechodu ze stavu I do stavu S na základě uzdravení dříve infekčního jedince (je to charakterizováno mírou uzdravení). Vývoj populace modelujeme v dostatečně krátkém časovém úseku, takže zanedbáváme případná narození nových jedinců nebo úmrtí starých. To je vidět i z toho, že pro derivaci celkové velikosti populace $N = S + I$ platí $\frac{dS}{dt} + \frac{dI}{dt} = 0$, velikost populace zůstává konstantní. Pro popis tedy stačí jediná rovnice $\frac{dS}{dt} = -rSI + aI$, kde $r > 0$ je mírou nákazy a parametr $a > 0$ je mírou uzdravení. Model je určen těmito dvěma kladnými parametry r, a . Nepředpokládá tedy žádnou imunitu, dříve nakažený se může nakazit znovu. Dosadíme-li do základní rovnice za I odpovídající vztah $I = N - S$, dostaneme základní rovnici tohoto modelu ve tvaru $\frac{dS}{dt} = rSN - rS^2$ jejíž řešení je také logistická funkce typu (B).

Uvažujme o něco komplikovanější model SIR . Takový model je uveden v článku [8] jako nejjednodušší speciální případ vytvořené obecné teorie šíření epidemií. Zde znakem R (recovered class, někdy také removed) označujeme jedince, kteří buď nemoc měli a vyléčili se nebo jedince, kteří nemoci podlehlí a zemřeli. Celkový počet jedinců v jednotlivých kategoriích je tedy $N = S + I + R$. Předpokládá se pohyb jedinců mezi stavy S, I, R v tomto pořadí. Přechod ze stavu S do stavu I je charakterizován *mírou nákazy* r . Ta souvisí s rychlostí šíření nákazy, která zase závisí na kvalitě preventivních opatření a informovanosti o nemoci a možnostech šíření. Pohyb ze stavu I do R je charakterizován *mírou uzdravení nebo úmrtí* v důsledku epidemie parametrem a . Model, popisující jen dva přechody ($S \rightarrow I$, $I \rightarrow R$), popsáné parametry r, a , se označuje jako SIR model. Předpokládá následující vztahy mezi proměnnými: ze vztahu $N = S + I + R$ derivováním podle času dostaneme $0 = \frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt}$; jestliže (například z pozorování nebo dokonce z důvodu zjednodušení modelu) předpokládáme přímou úměrnost mezi časovou změnou počtu náchylných k onemocnění a jejich současným počtem a počtem infekčních, dostaneme jednoduchý tvar závislosti $\frac{dS}{dt} = -r \cdot S \cdot I$, kde parametr $r > 0$ a podobně vztah pro $\frac{dI}{dt} = r \cdot S \cdot I - a \cdot I$, kde $a > 0$.

$$0, 3, 6, 9, \dots, k, \text{ (kde } k \text{ je dělitelné třemi),}$$

$$1, 4, 7, 10, \dots, (k+1),$$

$$2, 5, 8, 11, \dots, (k+2),$$

a určit tři rovnice pro tři neznámé $a, b, C = \ln k$. Tyto tři rovnice získáme jako součty v každé ze tří posloupností (časových řad) logaritmu $n(t_i)$:

$$\sum_{j=0}^{\frac{k}{3}} \ln n(t_{3j}) = \left(\frac{k}{3} + 1\right) \cdot (\ln a - \ln b + C) +$$

$$+ a \sum_{j=0}^{\frac{k}{3}} t_{3j} - \sum_{j=0}^{\frac{k}{3}} \ln (e^{at_{3j}+C} - 1), \quad (1)$$

$$\sum_{j=0}^{\frac{k}{3}} \ln n(t_{3j+1}) = \left(\frac{k}{3} + 1\right) \cdot (\ln a - \ln b + C) +$$

$$+ a \sum_{j=0}^{\frac{k}{3}} t_{3j+1} - \sum_{j=0}^{\frac{k}{3}} \ln (e^{at_{3j+1}+C} - 1), \quad (2)$$

$$\sum_{j=0}^{\frac{k}{3}} \ln n(t_{3j+2}) = \left(\frac{k}{3} + 1\right) \cdot (\ln a - \ln b + C) +$$

$$+ a \sum_{j=0}^{\frac{k}{3}} t_{3j+2} - \sum_{j=0}^{\frac{k}{3}} \ln (e^{at_{3j+2}+C} - 1). \quad (3)$$

Uvedená soustava je soustavou tří nelineárních rovnic. Její řešení je možné například iteracemi. V praxi uvažované modely mají ještě implementovány další parametry, pomocí kterých se dají odhadovat například vlivy různých organizačních opatření bránících „přirozenému“ rozvoji pandemie. Pak je ovšem potřeba podobným způsobem vytvořit odpovídající počet rovnic. Je jasné, že tím se numerická stránka popisu rozvoje pandemie komplikuje, protože potřebujeme více pozorování v různých časech. Přitom nám jde hlavně o včasnou predikci počtu nakažených.

d) Kompartimentové modely epidemie

Pro popis časového vývoje epidemie byly vyvinuty různě komplikované kompartimentové modely. Nejjednodušší z nich má název *model SI*. Označuje se v nich soubor těch, kteří se mohou nakazit S (susceptibles) a soubor nemocných jedinců v důsledku infekce I (infections). Říkáme při tom, že jedinec,

nakažených). Připomeňme, že model konstruujeme proto, abychom z několika počátečních hodnot počtu nakažených mohli stanovit prognózu počtu nakažených v budoucnosti.

Polynom $n(t) = a_0 + a_1 t + \dots + a_n t^n$ z předcházejícího příkladu je obecně interpolačním polynomem pro počáteční hodnoty (t_i, n_i) , $i = 0, 1, 2, \dots, n$. Platí pro něho interpolační podmínka $n(t_i) = n_i$, která představuje pro $t_i < t_j$, $i < j$ řešitelnou soustavu lineárních rovnic pro koeficienty $a_0, a_1, \dots, a_n$. Snadno ale již vytušíme, že například zvyšováním stupně polynomu (konstrukcí interpolačních polynomů) vhodné řešení tímto způsobem také nenajdeme. (Odhad interpolačním polynomem $n(t)$ se s růstem jeho stupně více „vlní“, a to také odporuje experimentálním zjištěním, „vlní“ se ale i derivace tohoto polynomu, proto žádné polynomiální řešení nevyhovuje.)

Jiný model (později bude označen jako *SI* model) je empiricky lépe ověřený, než je model předchozí a je dán základním vztahem

$$\frac{dn}{dt} = an - bn^2. \quad (A)$$

V předchozím vztahu je $n = n(t)$ počet nakažených (infekčních) v čase $t > 0$, a, b jsou kladné konstanty, které se určí (jak uvidíme dále) z některých známých hodnot funkce n .

Podle tohoto modelu roste počet nakažených jen do $n < \frac{a}{b}$, po dosažení maxima při $n \sim \frac{a}{b}$ začne počet nakažených v závislosti na čase klesat. Závislost $n = n(t)$ má tedy jediné maximum. I tvar grafu $n(t)$ je v souladu s představou o šíření pandemie např. jedné vlny epidemie Covid 19 v určité oblasti.

Náš vztah představuje opět separovatelnou diferenciální rovnici, kterou můžeme za daných předpokladů převést na tvar

$$\frac{dn}{n(a - bn)} = dt,$$

ze kterého integrací dostaneme vztah

$$\ln \frac{n}{n - \frac{a}{b}} = at + C,$$

kde C je integrační konstanta.

Snadnou úpravou z této rovnice získáme výraz pro časovou závislost počtu nakažených, tentokrát s integrační konstantou $k = e^C$ ve tvaru

$$n = n(t) = \frac{a}{b} \cdot \frac{k \cdot e^{at}}{k \cdot e^{at} - 1}. \quad (B)$$

Když $n(0) = n_0$, je $k = \frac{n_0}{n_0 - a/b}$, pro $n_0 = 1$ je pak $k = \frac{b}{b-a}$, vztah pro $n(t)$ je logistická funkce. (Jde o generický model omezeného růstu se saturací.)

Vztah (B) popisuje šíření pandemie od počátku do maxima. Zajímavé je, že předchozí vztah lze použít po malé úpravě i pro popis poklesu nakažených v čase; to využívají kompartmentové modely, jak uvidíme dále.

Statistický odhad parametrů ze vztahu (B) musí být konstruován pomocí výpočetní techniky, a to vždy pro větší počet členů časové řady. Pokusíme se na základě zjednodušujícího výpočtu ukázat (z malého počtu počátečních hodnot) numerickou obtížnost získání odhadu parametrů tohoto modelu.

Příklad 4

Upravme předchozí vztahy pro data. Dosadíme začátek časové řady měření (nezaokrouhlená data odpovídající Tabulce 3) do vztahu pro $n(t)$:

$$n_0 = \frac{a}{b} \frac{k}{k-1} = 1; \quad n_1 = \frac{a}{b} \frac{k \cdot e^a}{k \cdot e^a - 1} = 4,4; \quad n_2 = \frac{a}{b} \frac{k \cdot e^{2a}}{k \cdot e^{2a} - 1} = 15.$$

Z prvního vztahu získáme $\frac{a}{b} = \frac{k-1}{k}$, dosazením do druhého dostaneme $k = \frac{4,4 - e^a}{3,4e^a}$. Podobně dosazením za $\frac{a}{b}$ výrazu, obsahujícího jen k do třetího vztahu, získáme po jednoduchých úpravách $e^{2a} = \frac{15}{14k+1}$; tedy $e^a = \sqrt{\frac{15}{14k+1}}$.

Dosazením za e^a z předchozího vztahu do $k = \frac{4,4 - e^a}{3,4e^a}$, dostaneme výraz s neznámou k a můžeme tak řešit rovnici pro k ve tvaru

$$3,4 \cdot k \cdot \sqrt{\frac{15}{14k+1}} = 4,4 - \sqrt{\frac{15}{14k+1}}.$$

Dalšími úpravami (vynásobením výrazem $\sqrt{14k+1}$ a pak umocněním celé rovnice na druhou a zaokrouhlováním numerických odhadů) získáme kvadratickou rovnici pro k ve tvaru

$$173,40k^2 - 169,04k - 4,36 = 0.$$

První její odhadovaný (odhad označujeme znakem „ $\sim$ “) kořen $k_1 \sim 1,00$ nevyhovuje dost dobře zadání („hodně“ se numericky odchyluje při dosazení levá strana druhého daného vztahu od pravé), za bližší pro správné řešení proto považujeme odhad druhého kořene $k_2 \sim -0,025$. Pomocí něj dostaneme dosazením do prvního základního vztahu $\frac{a}{b} \sim \frac{-0,025-1}{-0,025} = 41,0$.

Dosazením odhadnutého kořene k_2 do vztahu pro e^a dostaneme $e^a = \sqrt{\frac{15}{14k+1}} = \sqrt{\frac{15}{14 \cdot (-0,025) + 1}} \sim 4,80$. Řešením naší úlohy je tedy (odhadovaný)

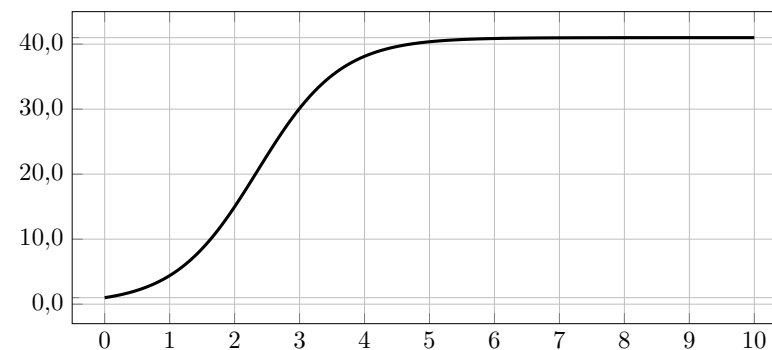
vztah

$$n = n(t) = \frac{a}{b} \frac{k \cdot e^{at}}{k \cdot e^{at} - 1} \sim 41,0 \cdot \frac{0,025 \cdot 4,80^t}{0,025 \cdot 4,80^t + 1}. \quad (C)$$

Závislosti $n = n(t)$ pro počáteční hodnoty t jsou uvedeny v Tabulce 4. I z tabulky je vidět, že pandemie má vrchol již po šestém kroku (rovnovážný stav je 41,0).

Tabulka 4: Hodnoty $n(t)$ se týkají Příkladu 4.

t	0	1	2	3	4	5	6	7	8	9	10	...
$n(t)$	1,0	4,4	15,0	30,1	38,1	40,4	40,9	41,0	41,0	41,0	41,0	41,0



Obrázek 1: Graf funkce $n = n(t)$ pro $t \in \langle 0; 10 \rangle$ podle vztahu (C), viz také Tabulka 4.

Použijeme-li více počátečních členů časové řady, předvedenou metodou, budeme muset řešit rovnice stále vyššího stupně. Je vidět, že takto numericky řešit pandemický problém nelze. Ukážeme sice vhodnější metodu odhadu koeficientů tohoto modelu, přesto se numerickým problémům nevyhneme.

c) Jak je také možné odhadnout parametry modelu (B) – logistické funkce

Když bychom znali „dostatečně“ dlouhou časovou posloupnost počtu nakažených $n(t)$ v diskrétních časových okamžicích t_i pro $i = 0, 1, 2, 3, \dots$, mohli bychom z ní vybrat 3 posloupnosti pro časové okamžiky