

Metody stanovení epidemického prahu.


Jeden z důležitých matematických nástrojů používaných v epidemiologii je modelování časového trendu různých diagnóz. Tento trend pak slouží k stanovení epidemického prahu (sledované populace), tedy k identifikaci odlehklých hodnot. Zásadním problémem pro tvorbu takovýchto modelů je sezónní kolísání a dlouhodobý trend v souvislosti s výskytem neočekávaně velkých hodnot incidence v krátkých (špatně predikovatelných) časových intervalech - výskyt epidemií (například u chřipky).

V návrhu metodiky jsem se zaměřil na rozklad hodnot incidence na nalezení dlouhodobého trendu, sezónního kolísání, náhodný šum a nakonec na identifikaci vysokých hodnot. Dlouhodobý trend je obvykle modelován regresními metodami, protože tento trend často nelze popsat matematickou funkcí, je zde volen LOESS odhad a to ve tvaru multiplikativním (ovlivnění incidence má obvykle multiplikativní charakter). Sezónní trend je modelován pomocí Serflingova přístupu (goniometrických funkcí) a nově navrhaného přístupu využívajícího modelování pro jednotlivé kalendářní týdny (jako kategoriální veličinu) a následným vyhlazením (zde nazýván ANOVA model).

Podstatným problémem pro tyto odhady je eliminace případných odlehklých hodnot. K tomu jsou navrženy dvě odlišná řešení:

- Využití analýzy cenzurovaných dat – hodnota pozorovaná v době epidemie je navýšena o příspěvek této epidemie. Epidemie může být identifikována experty, nebo nalezena formálně z dat jako zvolené procento týdnů a hodnotami nejbližšími od modelu. Tento přístup sice eliminuje vysoké hodnoty v epidemiích ale vyžaduje předpoklad o tvaru rozložení incidencí.
- Použití regresních kvantilů – tedy pro zvolený model (Serflingův nebo ANOVA model) nalezení hodnot modelu, jejichž hodnota je pro zvolené procento týdnů nad nebo pod daty. Tento přístup eliminuje vliv epidemií a dokonce nevyžaduje předpoklad o tvaru rozložení incidencí.

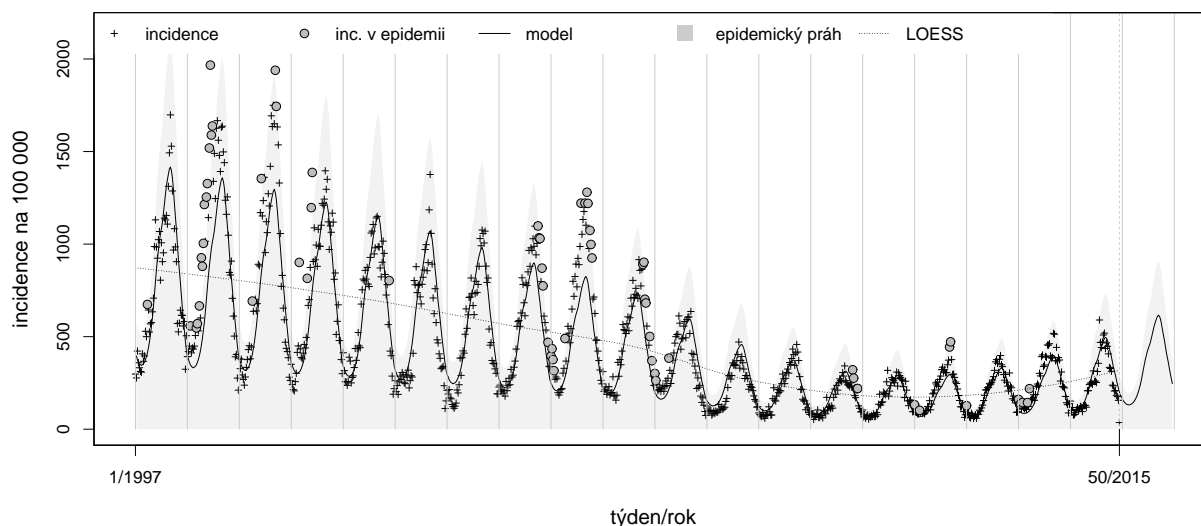
Dalším krokem je pak stanovení epidemického prahu. Při použití cenzurovaného přístupu je možno z předpokladu na tvar rozložení incidence odhadnout predikční mez přes, kterou s například 95% jistotou nepřekročí zvolené procento týdnů (např. 10%). Pro regresní kvantily je možno provést totéž dokonce bez předpokladu na tvar rozložení. rozdíl mezi oběma přístupy je v tom, že v prvním případě je epidemický práh násobkem baseline zatímco v druhém není. Druhý přístup připouští tedy i velmi různorodý tvar rozložení incidencí.

Vývoj a ověření těchto metod bylo provedeno v programu  (©2013 The R Foundation for Statistical Computing ISBN, 3-900051-07-0). V tomto programu je připraven i nástroj na rutinní výpočet modulů pro všechny infekční diagnózy surveillance systému EpiDat a ARI/ILI.

Tento model umožňuje i vytváření krátkodobých predikcí s výhradou že každá extrapolace má problematickou interpretaci. Je použitelný jak v infekční, tak i v neinfekční epidemiologii.

Příklad pro Salmonelózu – A02:

A02 ANOVA LOESS mult. model s 5% epidemií
podle data onemocn ní



Tři nejvýznamnější práce:

PROCHÁZKA B. KYNČL J.: Estimating of baseline and treshold for diagnoses with seasonal and longtime trend. Central European Journal of Public Health, (2015 December).

PROCHÁZKA, B. Regression quantiles and trimmed least-squares estimator in the nonlinear-regression model. Computational statistics & data analysis. 1988, vol. 6, no. 4, p. 385-391.

KYNČL, J., PROCHÁZKA, B., HAVLÍČKOVÁ, M., OTAVOVÁ, M., ČÁSTKOVÁ, J., KRÍŽ, B. Excess death attributable to influenza in the Czech Republic in 1982-2000. Journal of epidemiology and community health. 2004, vol. 58, Suppl. 1, p. A31-A32.

Za nejvýznamnější přínos mé práce považuji navrzení metodiky jak získat nejen odhad baseline neovlivněný odlehlými hodnotami, ale především odhad epidemického prahu, který je užitečný pro vyslovení podezření na epidemii v reálném čase.