

Univariate and Multivariate Surveillance of Outbreaks

Övervakning av utbrott

Linus Schiöler

Doktorsavhandling i statistik vid Göteborgs universitet.

Inom många områden görs fortlöpande observationer i syfte att snabbt upptäcka förändringar. Inom tillverkningsindustrin började enkla metoder användas redan 1930. Teorin för statistisk övervakning ger möjlighet att konstruera optimala metoder. Behovet av övervakning inom finans, miljö och medicin har medfört att teorin utvecklats snabbt de senaste åren.

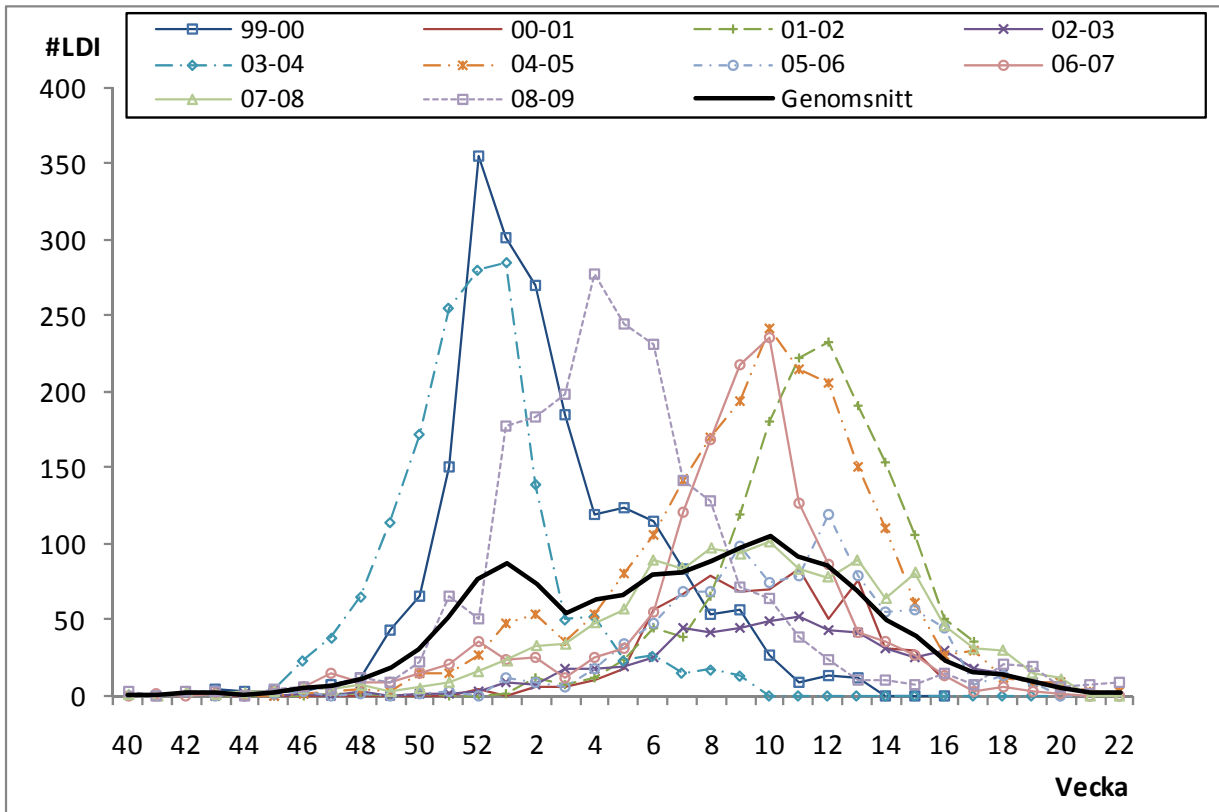
Det har blivit allt mer påtagligt att det finns ett behov att snabbt upptäcka utbrott av epidemiska sjukdomar. Utbrott av nya virus som svin- och fågelinfluensa visade att det fanns ett behov av beredskap vid krissituationer. Genom att snabbt upptäcka nya utbrott kan sjukvård planeras och preventiva åtgärder vidtas. Detta kan avsevärt minska förlusten av människoliv samt samhällsekonomiska utgifter för sjukskrivningar och sjukvård. Detsamma gäller det årliga influensautbrottet under vintern, som varje år leder till tusentals dödsfall och kostar samhället stora belopp. Det är således inte förvånansvärt att intresset av statistisk övervakning är större än någonsin.

När man försöker upptäcka en förändring måste man göra en avvägning mellan att göra upptäckten så snabbt som möjligt och säkerheten i bedömningen. Genom att använda statistisk övervakning kan denna avvägning göras på ett optimalt sätt. Man kan härigenom utveckla effektiva beslutsystem som grundar sig på vedertagen statistisk teori. En övervakningsmetods egenskaper kan utvärderas genom att man härleder dem från statistisk teori eller genom att använda sig av simuleringstudier, där man använder en statistisk modell för att efterlikna verkligheten. Detta gör att det på förhand går att ha en mycket god uppfattning av hur en statistisk metod fungerar när den tillämpas i praktiken. Det går alltså att uttala sig om hur säker man är på att en förändring verkligen har skett när man får ett larm från metoden. Man har också en uppfattning om hur lång fördröjning man kan förvänta sig från att en förändring sker tills att ett larm ges. Denna objektivitet är en stor fördel i praktiken. För att jämföra en objektiv statistisk metod med subjektiva bedömningar har ett experiment genomförts. Resultatet var att det objektiva systemet upptäcker förändringar snabbare och säkrare. Det visade sig även att de subjektiva bedömningarna är starkt beroende av vem som gjort dem.

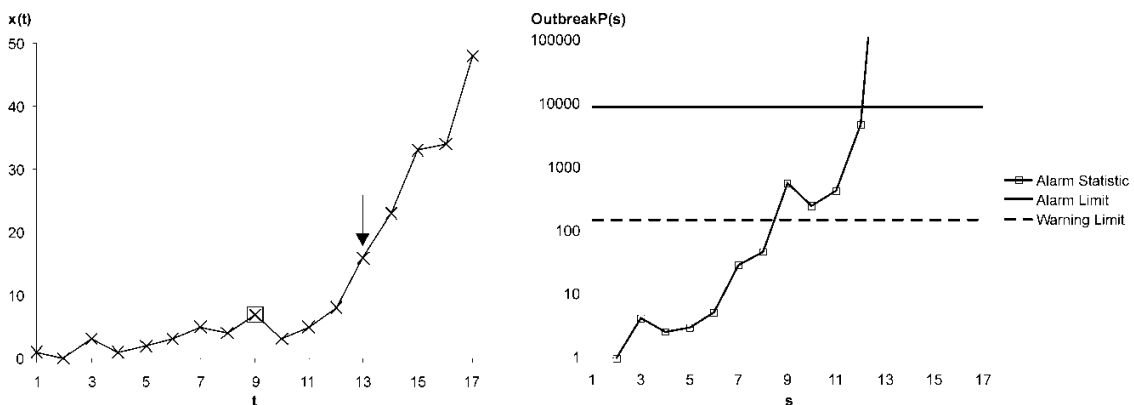
För att göra övervakningsmetoden enkel att använda har vi utvecklat användarvänliga programtillägg för Microsoft Excel. Dessa program kan fritt laddas ned på www.statistics.gu.se/surveillance. Metoden har även gjorts tillgänglig i Smittskyddsinstitutets övervakningssystem CASE, samt paketet "surveillance" i R.

Liksom vid all annan statistisk analys är man vid statistisk övervakning tvungen att göra vissa antagande rörande datamaterialet. Många metoder för upptäckt av influensa-utbrott baseras på avvikelser från en genomsnittskurva över åren. Detta kan bli ineffektivt eftersom det är stor variation mellan åren (se figur 1). När säsongens utbrott sker sent har dessa metoder svårt att upptäcka utbrottet.

Felaktiga antaganden kan få som konsekvens att övervakningssystemet blir ineffektivt. I (I) utvecklas robusta metoder som är särskilt lämpliga för att upptäcka utbrott. Ingen antagande görs angående en baselinenivå, i stället utnyttjas att influensautbrott kan karakteriseras av en ändring från en konstant incidens till en ökande. En larmstatistika som bygger på den generaliserade likelihoodkvoten mellan en konstant nivå och en monotont ökande konstrueras. För en-parameters exponentialfamiljen ges maximum-likelihood-skattningen av ickeparametriskt isoton regression. Metoden är alltså semiparametrisk. I figur 2 visas resultat från ett år för hela Sverige.



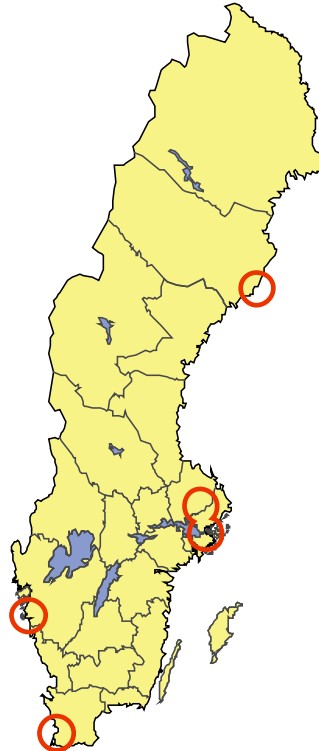
Figur 1. Antal laboratorie bekräftade influensa fall i Sverige under 9 säsonger samt en genomsnittskurva.



Figur 2. Vänster: Antal laboratorie bekräftade fall av influensa i Sverige 07-08 säsongen. Den första observationen är för vecka 40 2007. Varnings och larm signaler för metoden är markerade med fyrkant och oil respektive. Höger: Resultat från datorprogrammet för data till vänster.

Vid övervakning av sjukdomsutbrott har man ofta data tillgänglig från flera olika platser. I (IV) analyseras flera års svensk influensadata från ett tjugotal laboratorier. Inga tydliga tendenser av ett geografiskt

spridningsmönster visar sig. Slutsatsen är i stället att influensautbrottet i Sverige börjar cirka en vecka tidigare i storstäderna än i övriga landet (se figur 3). I (V) utnyttjar vi denna information för att konstruera ett mer effektivt övervakningssystem. Det rör sig om övervakning av flera variabler, så kallad multivariat övervakning. Vid multivariata problem är det en fördel att reducera komplexiteten så mycket som möjligt utan att gå miste om information. En uttömmande statistiska härleds för övervakning av ett hopp i medelvärde i (III) och vidareutvecklas för situationen med monotonitetsrestriktioner i (V). Här vidareutvecklas även metoderna från (I) för att utnyttja denna statistiska. Resultatet blir ett övervakningssystem som upptäcker utbrott snabbare än ett som behandlar hela landet som en enhet.



Figur 3. Storstadsområdena där influensautbrotten sker först

Avhandlingen baseras på:

- I. Frisé, M., Andersson, E. and Schiöler L. (2009)
Robust outbreak surveillance of epidemics in Sweden.
Statistics in Medicine, **28**, 476-493.
- II. Frisé, M., Andersson, E. and Schiöler L. (2010)
Evaluation of Multivariate Surveillance.
Journal of Applied Statistics, **37**, 2089-2100.
- III. Frisé, M., Andersson, E. and Schiöler L. (2011)
Sufficient Reduction in Multivariate Surveillance.
Communications in Statistics – Theory and Methods, **40**, 1821-1838.
- IV. Schiöler, L. (2011)
Characterisation of influenza outbreaks in Sweden.
Scandinavian Journal of Public Health, **39**, 427-436.

Schiöler Övervakning av utbrott.docx

- V. Schiöler, L. and Frisén M.
Multivariate outbreak detection.
Kommande i *Journal of Applied Statistics*.