Outbreak detection

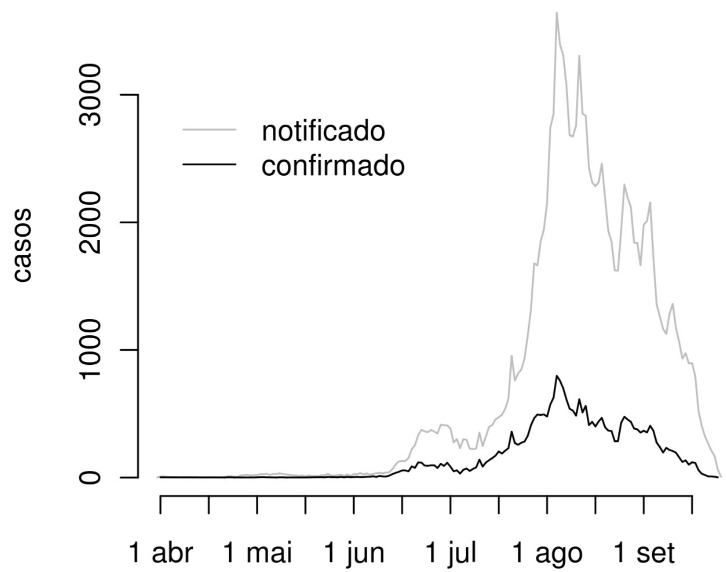
author: Marcelo F C Gomes date: autosize: true

Outbreak/epidemic activity detection

What defines the period of epidemic outbreak/activity? When does it start e how long it lasts? - Exponential growth? - Re > 1? - Number of cases?

On a noise time series, how to separate fluctuations and isolated clusters from sustained outbreak?

H1N1pdm09 in Brazil



***

Not knowing the operational details leading to the "first burst" in mid-June, should we classify it as season's starting date/outbreak?

What if it had happened in May?

Activity thresholds

Paradigm change:

Increase rate → incidence volume

The epidemic threshold

Paradigm change:

Increase rate → incidence volume

Incidence level above which there is a clear separation between baseline and sustained increase of case counts (*threshold*).

This turning point can be referred to as *(pre-)epidemic threshold*.

The epidemic threshold

How to estimate it?

- Sensitivity *vs* specificity

The epidemic threshold

How to estimate it?

- Sensitivity *vs* specificity

Do we want it to be triggered at any burst ou only when there is a clear pattern of sustained increase?

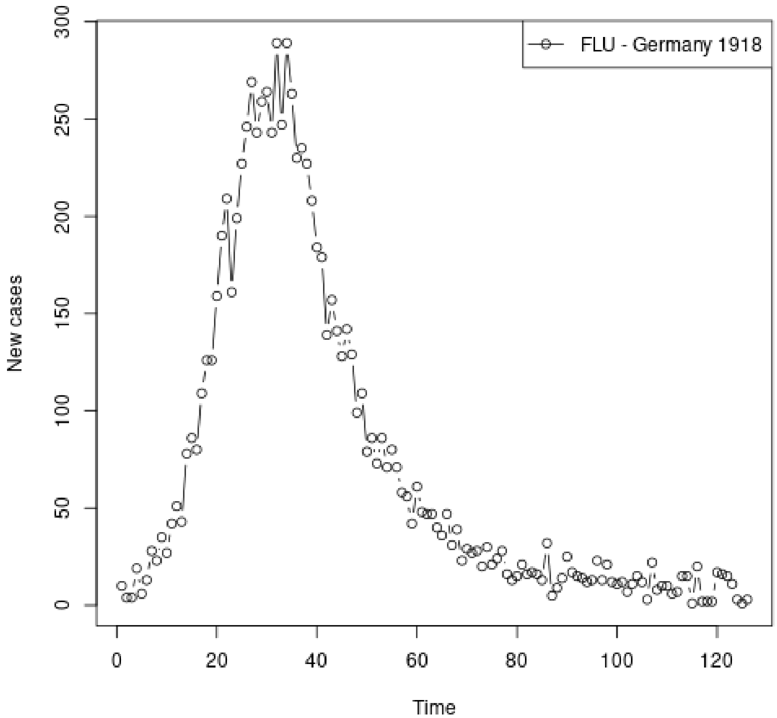
The epidemic threshold

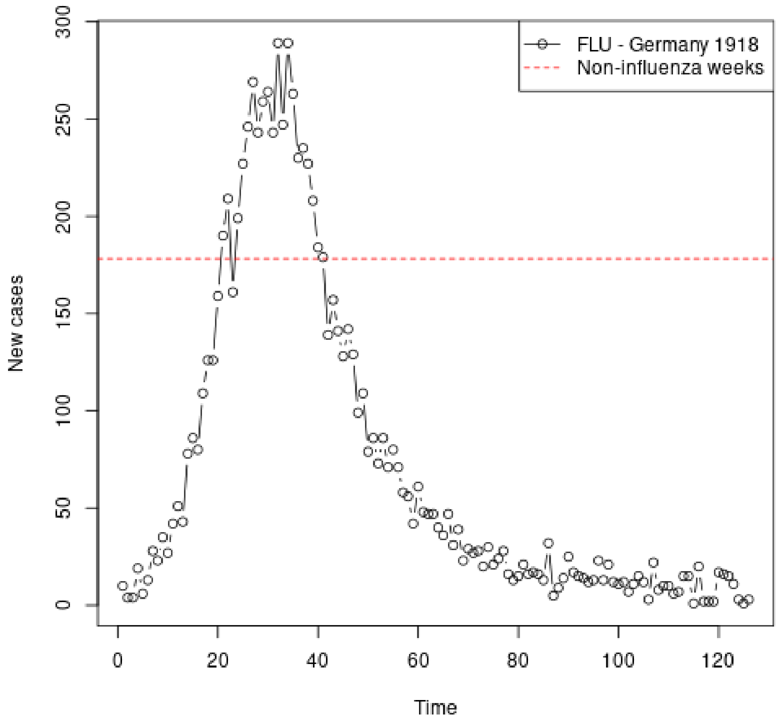
Methodology used by the USA CDC for Influenza-like illness (ILI): - The baseline is developed by calculating the mean percentage of patient visits for ILI during non-influenza weeks for the most recent two seasons and adding two standard deviations (2021-2022 and 2022-2023). A non-influenza time period (e.g., a "non-influenza week") is defined as two or more consecutive weeks in which each week accounted for less than 2% of the season's total number of specimens that tested positive for influenza in public health laboratories. Region-specific baselines are calculated using the same methodology. Due to the wide variability in regional level data, it is not appropriate to apply the national baseline to regional data.

http://www.cdc.gov/flu/weekly/overview.htm

The epidemic threshold – CDC

- A non-influenza time period (e.g., a "non-influenza week") is defined as two or more consecutive weeks in which each week accounted for less than 2% of the season's total number of specimens that tested positive for influenza in public health laboratories.

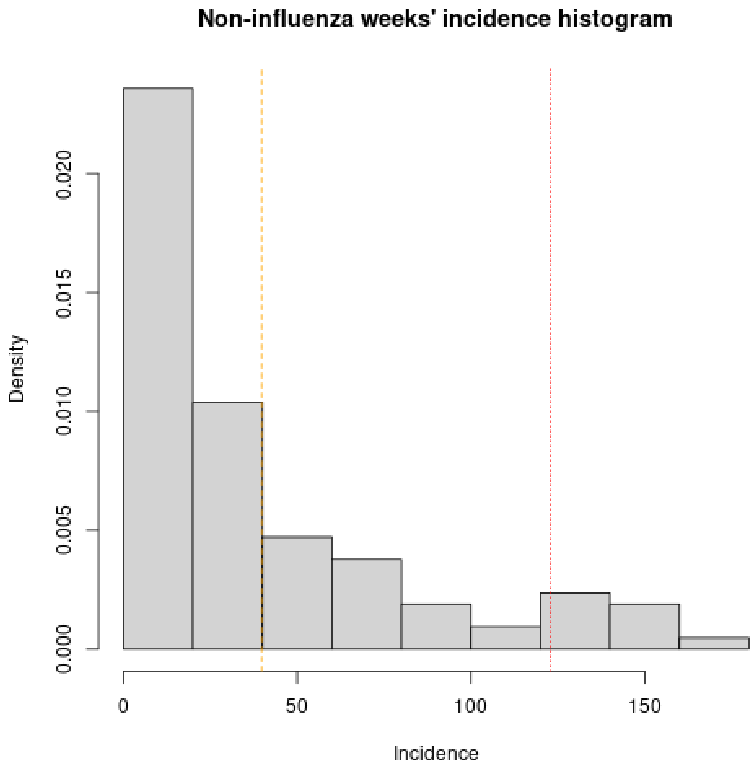


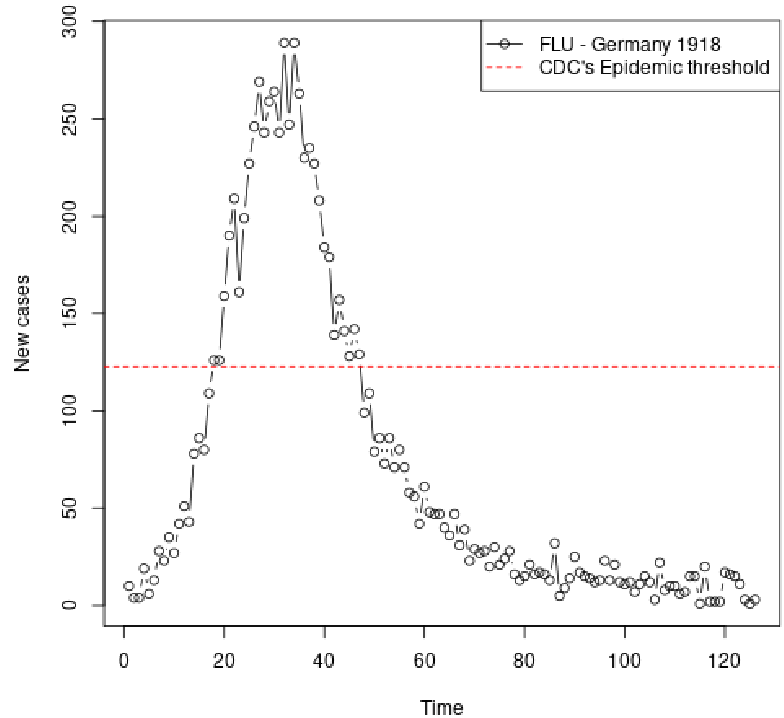


***

The epidemic threshold – CDC

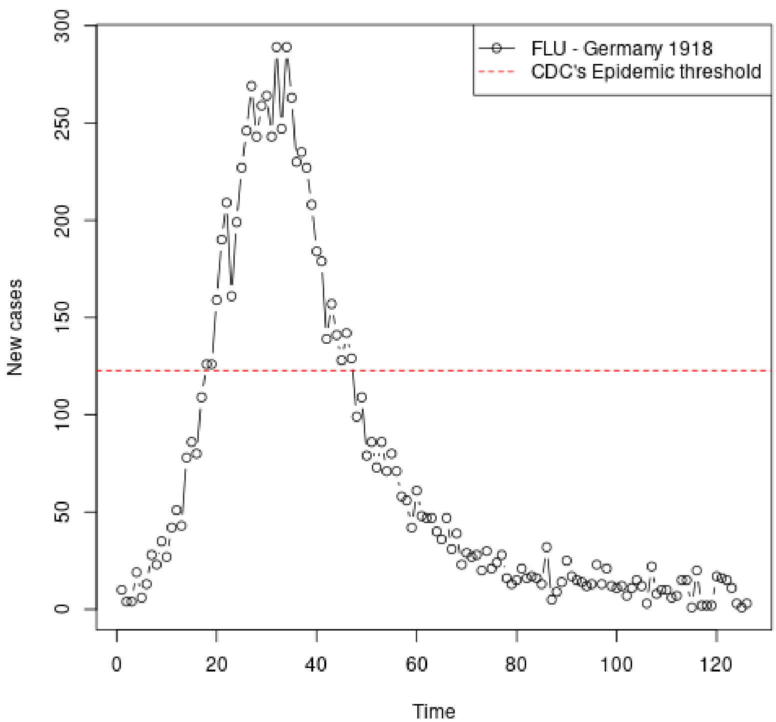
- The baseline is developed by calculating the mean percentage of patient visits for ILI during non-influenza weeks for the most recent two seasons and adding two standard deviations.

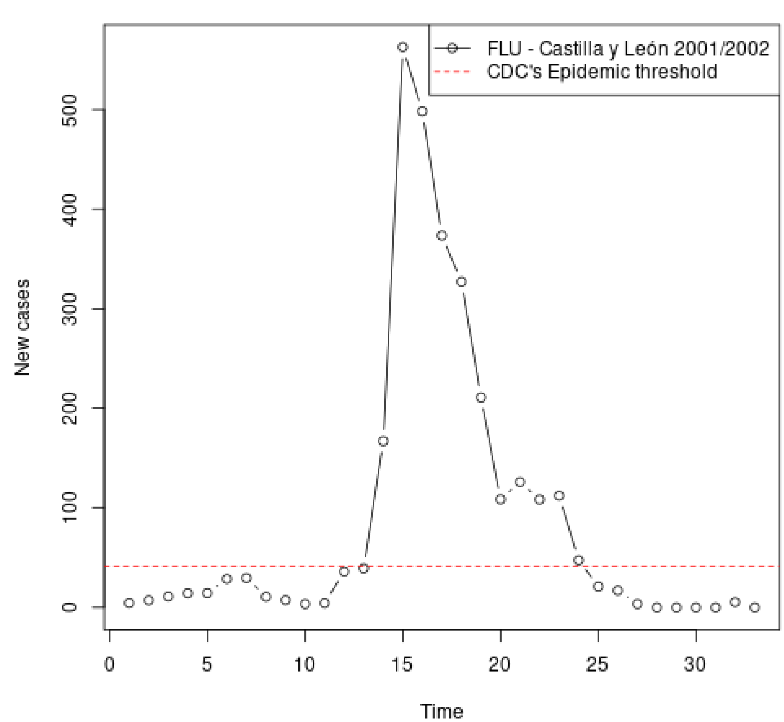




***

The epidemic threshold – CDC





***

Moving Epidemics Method (MEM)

The MEM r-package (*Moving Epidemics Method*), uses a similar idea, but with a different methodology. The package offers several methods, all of them based on the Maximum **Accumulated Percentage** (MAP) to define "active phase". This package also offers methods for other activity thresholds and seasonal trends (endemic channels).

Bibliography: - Vega T., Lozano J.E. et al. (2004) Modelling influenza epidemic - can we detect the beginning and predict the intensity and duration? International Congress Series 1263 (2004) 281-283. - Vega T., Lozano J.E. et al. (2012) Influenza surveillance in Europe: establishing epidemic thresholds by the Moving Epidemic Method. Influenza and Other Respiratory Viruses, DOI:10.1111/j.1750-2659.2012.00422.x. - Vega T., Lozano J.E. et al. (2015) Influenza surveillance in Europe: comparing intensity levels calculated using the moving epidemic method. Influenza and Other

Respiratory Viruses, DOI:10.1111/irv.12330.

========================= MAP Curve

For a season with S time units (weeks), calculate the maximum fraction of accumulated news cases on windows of size $r = 1, 2, ..., S$.

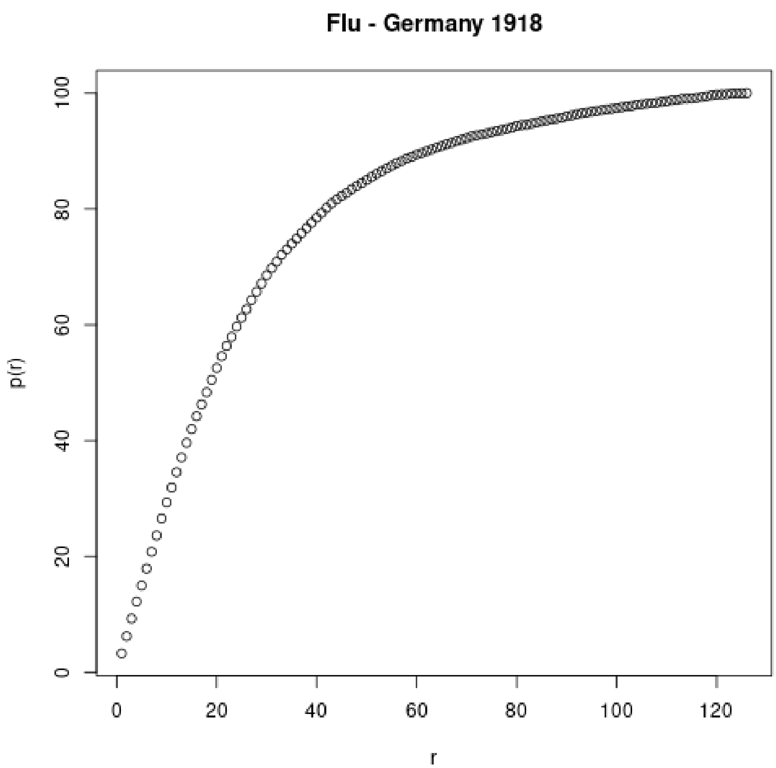
$$C(r) = \max_{i=1,S-r+1} \sum_{t=i}^{t+r-1} I_t$$

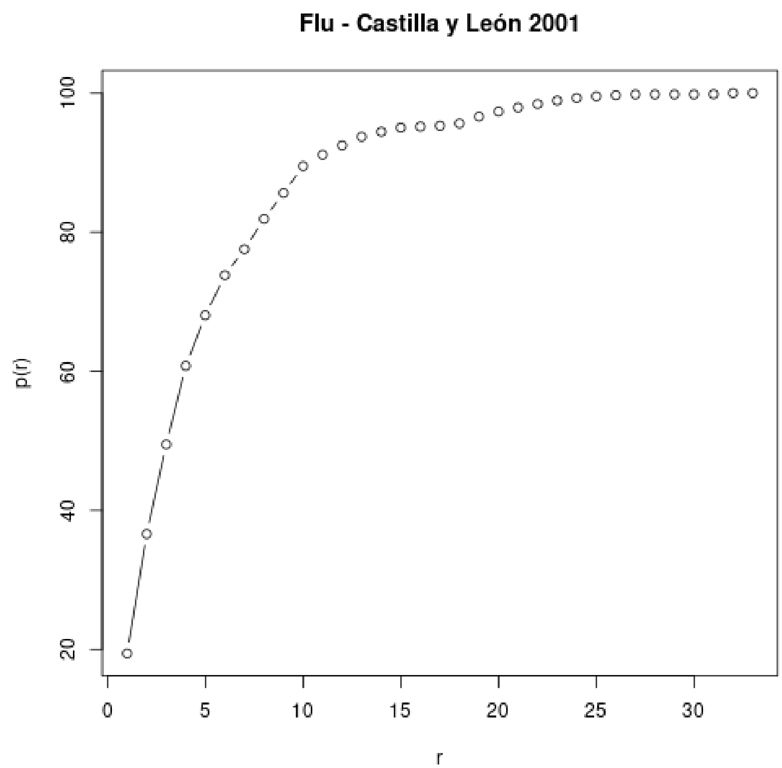
$$p(r) = \frac{C(r)}{C(S)}$$

$p(r)$: MAP curve

From this curve, different methods can be applied to define "non-influenza weeks" and then estimate the (pre-)epidemic threshold based on the corresponding case distribution. For instance, based on a percentile of the distribution, such as 90, for instance. The higher the percentile, the higher the threshold, lowering sensitivity but increasing specificity. And vice-versa by lowering the target percentile.

MAP Curve





***

MAP Curve optimal window - Method 1

The optimal window is defined by minimizing the sum of normalized r and the normalized derivative of p as an approximation for inflection point.

$$\Delta p = p^{r+1} - p^r$$

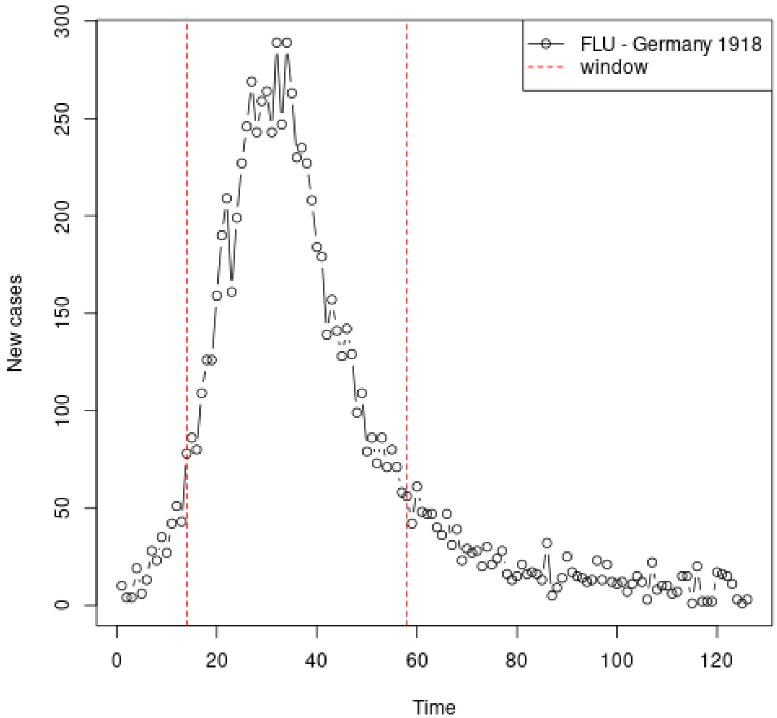
$$\hat{r} = \frac{r - r_{min}}{r_{max} - r_{min}}$$

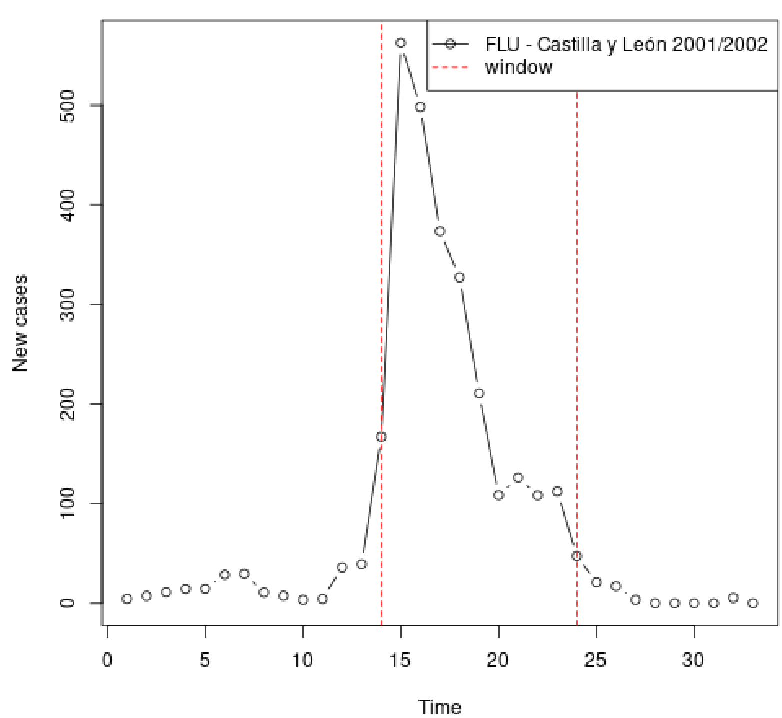
$$\hat{\Delta p} = \frac{\Delta p - \Delta p_{min}}{\Delta p_{max} - \Delta p_{min}}$$

Optimal window:

$r \mid \sqrt{\hat{r}^2 + \hat{\Delta p}^2 - (\hat{r} + \hat{\Delta p})^2/2}$ is minimum

MAP Curve - Method 1





***

MAP Curve - Method 2

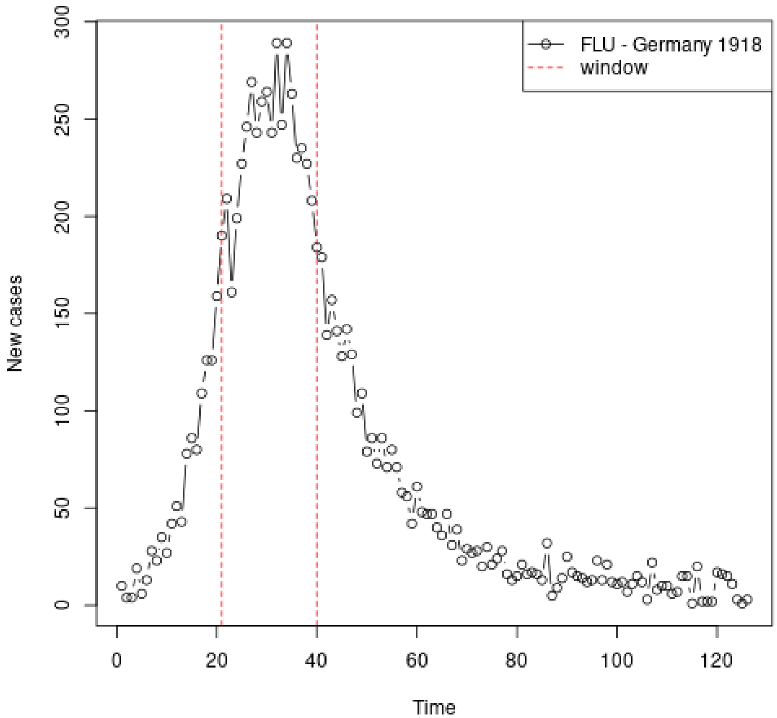
Define the optimal window as the lowest in which the derivative of p is less than a value defined based on sensitivity/specificity analisis. In general, a value betwenn 2 and 4 (empirical).

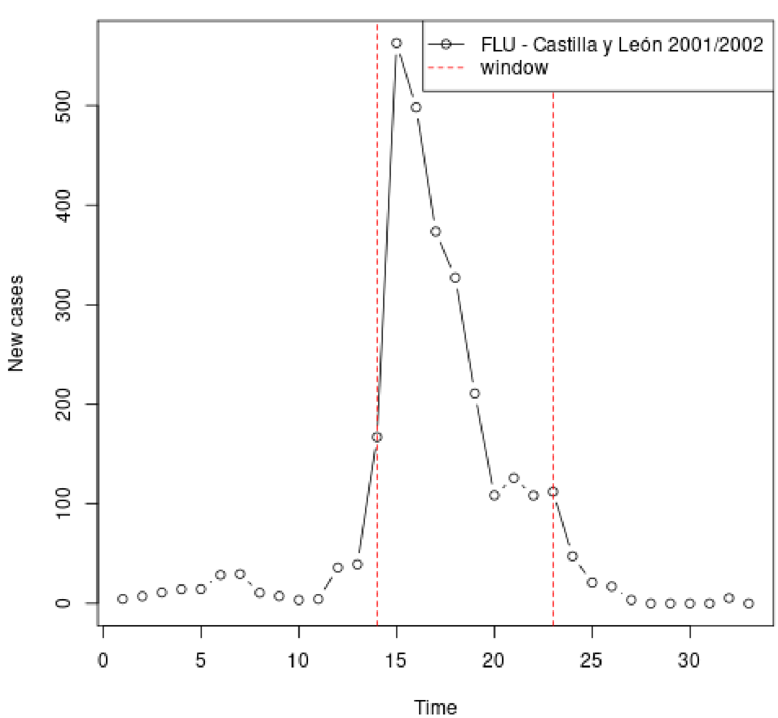
$$\Delta p = p^{r+1} - p^r$$

Window:

$$\min r \mid \Delta p < \epsilon$$

MAP Curve - Method 2





***

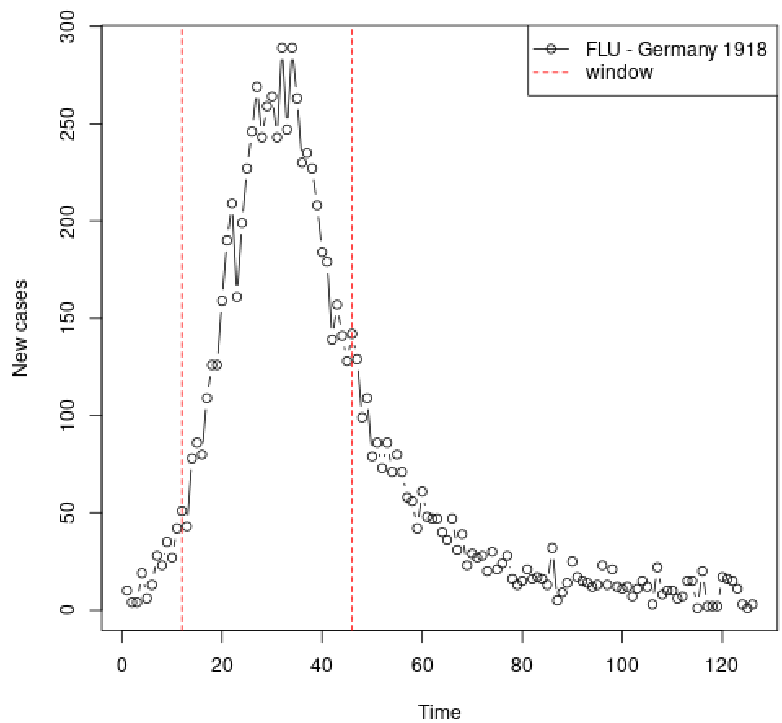
Inflection point

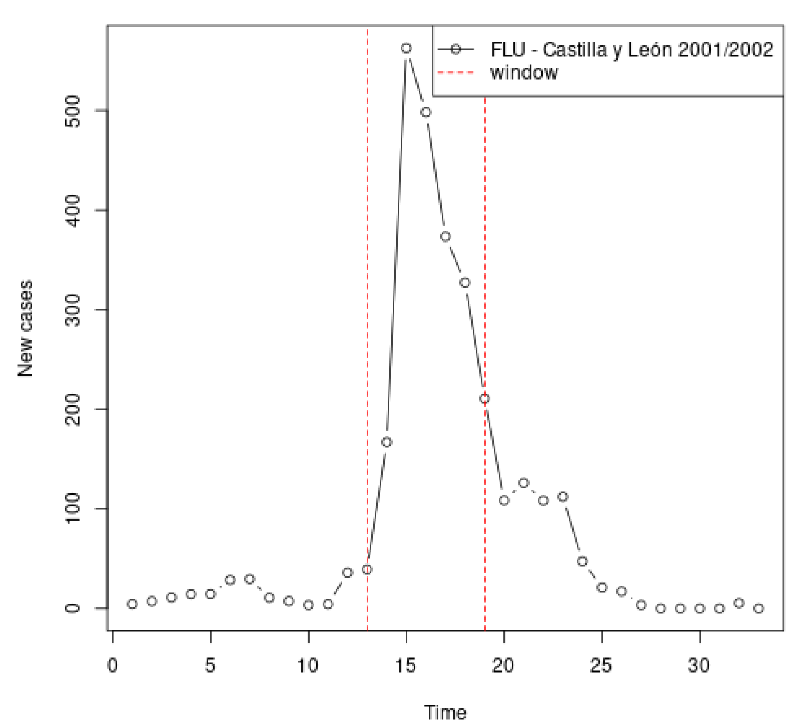
Time in which the second derivative of the incidence curve is null, thus changing the curve's concavity.

$$\Delta^2 I_t = I_{t+1} - 2I_t + I_{t-1} = 0$$

$$\Delta^2 I_{t-1} > 0 \; , \; \Delta^2 I_{t-1} < 0$$

Inflection point

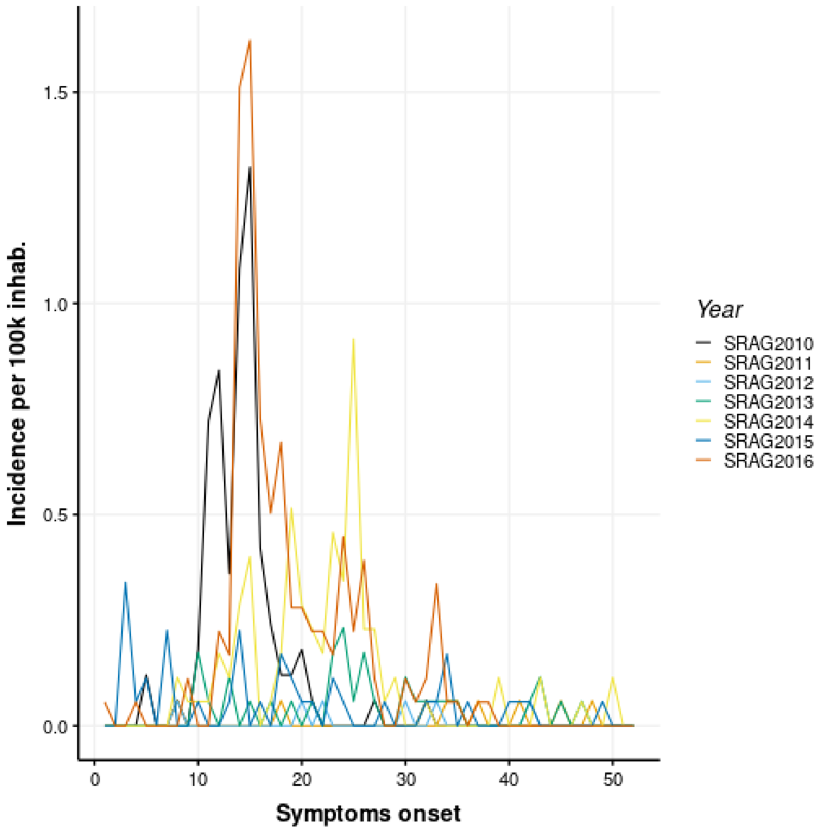




***

All nice and sweet, but...

SARI cases notified in Rondônia state (Brazil), from 2020-2016.



***