Statistical Measures
for
Evaluation of Methods
for
Syndromic Surveillance

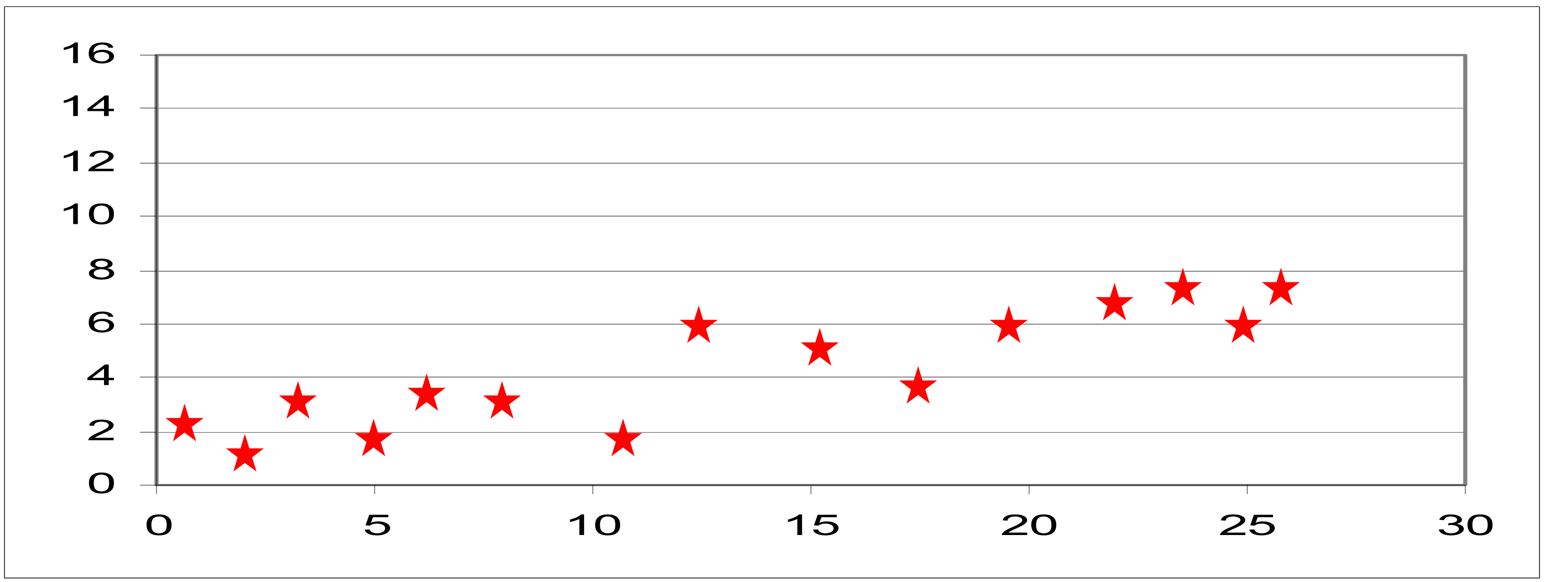
Marianne Frisén

Statistical Research Unit

Göteborg University

Sweden

Statistical methods
to separate important changes from stochastic variation.



Enough information for decision?

Surveillance

- **Repeated measurements**
- **Repeated decisions**
- **No fix hypothesis**
- **Time important**

Time important

- **Methods**

- **Evaluations**

  – Frisén, M. (2003), Statistical Surveillance. Optimality and Methods., *International Statistical Review,*71, 403-434.

Evaluations

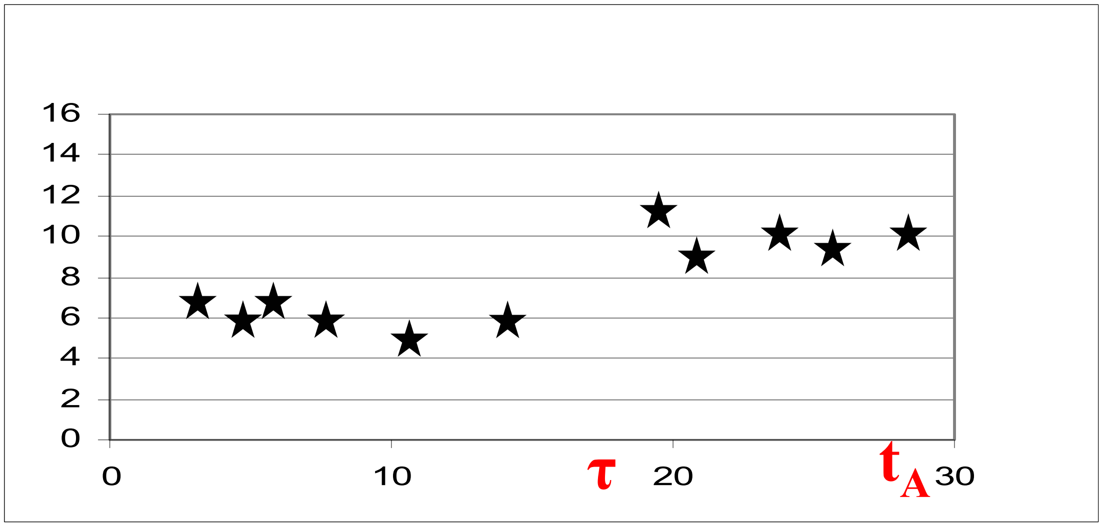
- Quick detection
- Few false alarms

- Frisén, M. (1992). Evaluations of methods for statistical surveillance. *Statistics in Medicine,* 11, 1489 - 1502**.**

Change in distribution

The First $(\tau-1)$ observations $x_{\tau-1} = x(1), ..., x(\tau-1)$ have density f^D

The following observations have density f^C



Alarm

Marianne Frisén
Syndromic Surveillance NY 2003

False alarms

- The Average Run Length at no change, $ARL^0 = E(t_A | D)$
- The false alarm probability $P(t_A < \tau)$

Delay

- $ED(t) = E[\max(0, \mathbf{t_A} - \mathbf{t}) \mid \tau = t]$
- The expected delay depends on the time τ of the change
- ED = Expected Delay = $E_\tau[ED(\tau)]$
- ARL[1] = ED(1) The Average Run Length until detection of a change (that occurred at the same time as the inspection started). $E(t_A \mid \tau = 1)$
- **P**robability of **S**uccessful **D**etection

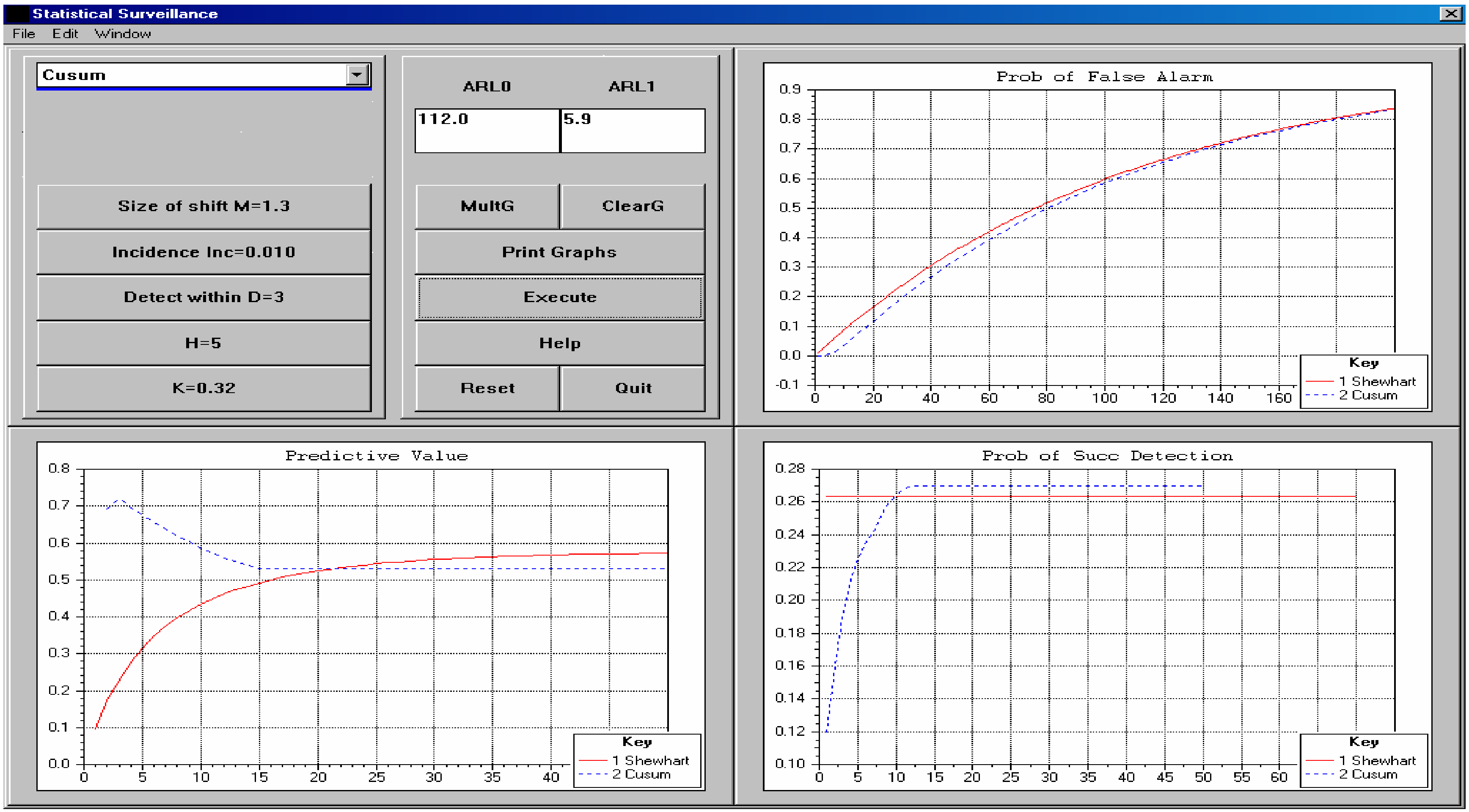
$$PSD(\tau, d) = P(t_A - \tau \le d \mid t_A \ge \tau).$$

Predictive value

$$\Pr(\tau \leq t \mid t_A = t)$$

The predictive value reflects the trust you should have in an alarm.

Computer program



Optimality

- ARL-optimality
- Expected Delay-optimality
- Minimax-optimality

- Frisén, M. and de Maré, J. (1991). Optimal surveillance. *Biometrika*, 78, 271-80.
- Frisén, M. (2003). Statistical Surveillance. Optimality and Methods., *International Statistical Review,* 71, 403-434.

ARL Optimality

- Minimal ARL^1 for fixed ARL^0
- Useless methods are ARL optimal
  - Frisén, M. (2003), Statistical Surveillance. Optimality and Methods. *International Statistical Review*, 71,403-434.
- Use only with care!

Minimal expected delay

- For a fixed false alarm probability

- The full likelihood ratio method LR is optimal
- EWMA Approximately optimal
  - for a certain value of the parameter
  - Frisén, M. and Sonesson, C. (2003): Optimal surveillance by exponentially moving average mehtods. Submitted.

Minimax Optimality

- Minimal expected delay
  for the worst value of τ
  and for the worst history of observations before τ
  - **CUSUM is minimax optimal**
  - Moustakides, G. V. (1986), "Optimal Stopping Times for Detecting Changes in Distributions," *The Annals of Statistics*, 14, 1379-1387.

- Shewhart Optimal for recent large change

Predictive value

A **constant predicted** value
makes the same kind of action appropriate
both for early and late alarms.

- Shewhart
  - Many early alarms.
  - These alarms are often false.
  - The predicted value of early alarms is very low.
  - Different actions motivated for early and late alarms

Syndromic surveillance

- Multivariate
- If the different symptoms change at the same time (or with a known time lag)
  - the multivariate situation is easily reduced to a univariate one
  - Wessman, P. (1998) Some Principles for surveillance adopted for multivariate processes with a common change point. Communications in Statistics. Theory and Methods, 27, 1143-1161.

Outbreak detection

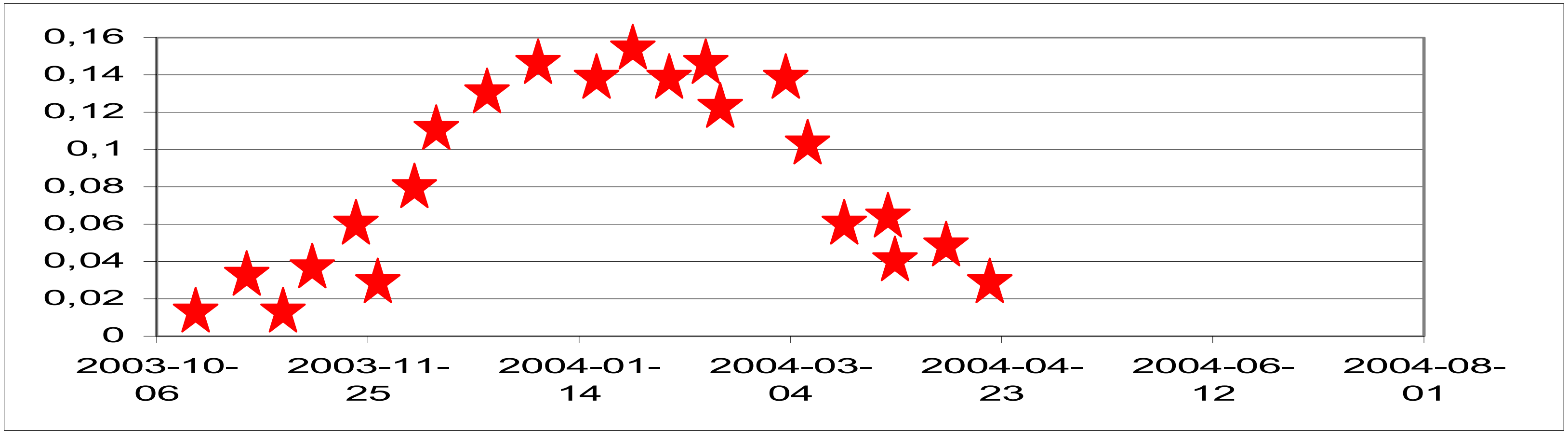
- Non-normal distribution
- Multivariate
  - Syndrome
  - Spatial
- Complicated changes

  - Frisén, M. (2003), Statistical Surveillance. Optimality and Methods., *International Statistical Review,* 71, 403-434.

Type of change

- Change from unknown baseline
- Gradual change
- Change in monotonicity

Outbreak



- Rise
- Decline

New robust method for detection of rise or decline

Maximum likelihood ratio

$$\left\{ x_s : \quad \frac{\max f(x_s | C)}{\max f(x_s | D)} \geq k' \right\}$$

- Frisén, M. (1986) Unimodal regression. *The Statistician,* **35,** 479-485.
- Andersson, E., Bock, D. and Frisén, M. (2002) Statistical surveillance of cyclical processes with application to turns in business cycles. *Submitted*.

Concluding remarks

- Robustness is important
- Knowledge of properties of a system is useful
- This involves many aspects
- Statistical measures which take care of time