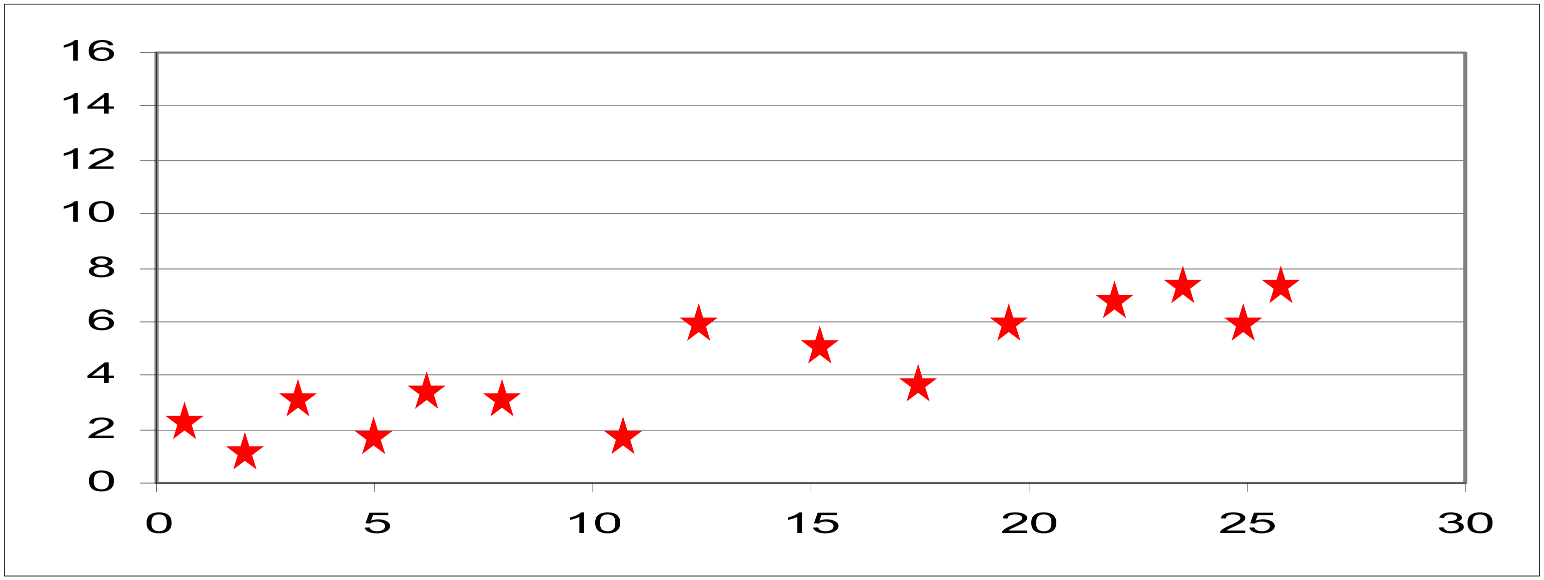
Statistical issues at online surveillance

Marianne Frisén
Statistical Research Unit
Göteborg University
Sweden

Outline

- I Inferential framework
- II Demonstration of computer program
- III Complicated problems - examples

Statistical methods
to separate important changes from stochastic variation.



Enough information for decision?

Continual observation of a time series,

with the goal of detecting an important change in the underlying process

as soon as possible after it has occurred.

- Monitoring
- Surveillance
- Change-point analysis

- SPC
- Control charts
- Early warnings
- Just in time

Monitoring of health

INDIVIDUALS:

- natural family planning
  - Hormone cycles
- regular health controls
  - pregnancy
- Intensive care
  - fetal heart rate
- surveillance after intervention
  - kidney transplant

POPULATIONS:

- control of epidemic diseases
- surveillance of known risk factors
- detection of new environmental risks

Surveillance

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

Scources of knowledge

Quality control

Stopping rules in probability theory

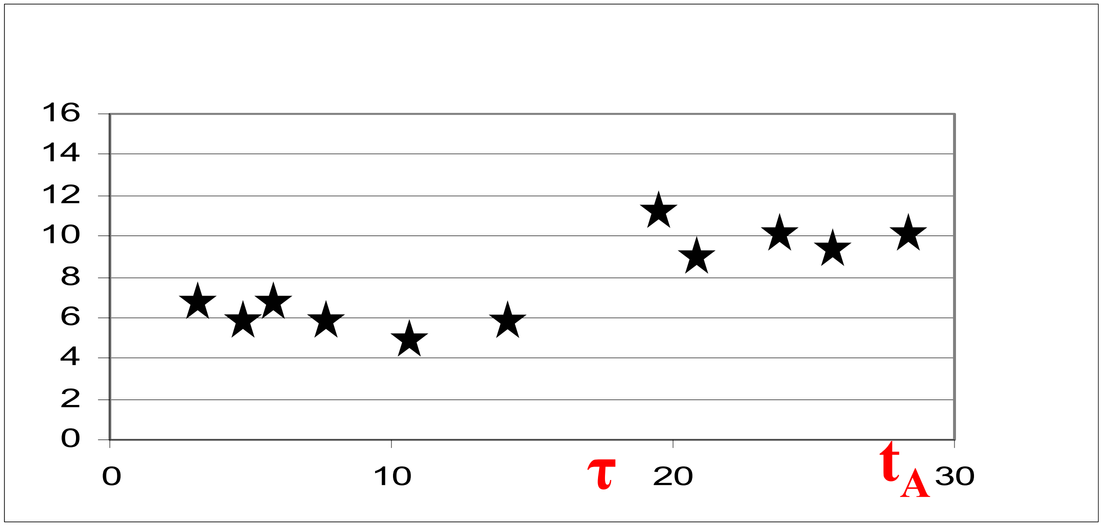
Inference

Medicine

Change in distribution

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

The following observations have density f^C



Alarm

Timely detection

of a change in a process

from state D to state C

Evaluations

- Quick detection
- Few false alarms

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

False alarms

- The Average Run Length at no change,

  $ARL^0 = E(t_A | D)$

- The false alarm probability

  $P(t_A < \tau)$.

Motivated alarms

- ARL^1 The Average Run Length until detection of a change (that occurred at the same time as the inspection started) $E(t_A|\tau=1)$.

- $ED(t) = E[\max(0, t_A - t) \mid \tau = t]$
    - $ARL^1 = ED(1)$
    - $CED(t) = E[t_A - t \mid \tau = t, t_A \geq t]$
- $ED = E_\tau[ED(\tau)]$
- Probability of Successful Detection $PSD(\tau, d) = P(t_A - \tau \leq d \mid t_A \geq \tau)$.

Predictive value

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

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

Optimality

- ARL-optimality
- ED-optimality
- Minimax-optimality

- Frisén, M. and de Maré, J. (1991). Optimal surveillance. *Biometrika*, 78, 271-80.
- Frisén, M. (in press), Statistical Surveillance. Optimality and Methods., *International Statistical Review*.
- Frisén, M. and Sonesson, C. (2003): Optimal surveillance by exponentially moving average mehtods. Submitted.

ARL Optimality

- Minimal ARL^1 for fixed ARL^0
- Observe that $\tau=1$
- Consequences demonstrated in
  - Frisén, M. (in press), Statistical Surveillance. Optimality and Methods., *International Statistical Review*.
  - Frisén, M. and Sonesson, C. (2003): Optimal surveillance by exponentially moving average mehtods. Submitted.
- Use only with care!

Utility

- **The loss of a false alarm is a function of the the time between the alarm and the change point.**
- **The gain of an alarm is a linear function of the same difference.**

$$u(t_A, \tau) = \begin{cases} h\left(t_A - \tau\right) & , t_A < \tau \\ a_1 \cdot \left(t_A - \tau\right) + a_2, & t_A \geq \tau \end{cases}$$

Shiryaev, A. N. (1963), "On Optimum Methods in Quickest Detection Problems," *Theory of Probability and its Applications*, 8, 22-46

ED Optimality

Minimal expected delay — ED

for a fixed false alarm probability — $\mathrm{P}\left[t_A < \tau\right]$

Maximizes the utility by Shiryaev

Minimax Optimality

- Minimal expected delay
  for the worst value of τ
  and for the worst history of observations before τ

  - Pollak, M. (1985), "Optimal Detection of a Change in Distribution," *The Annals of Statistics*, 13, 206-227
  - Lai, T. L. (1995), "Sequential Changepoint Detection in Quality-Control and Dynamical-Systems," *Journal of the Royal Statistical Society Ser. B*, 57, 613-658.

Methods

- LR
  - Shiryaev-Roberts
- Shewhart
- EWMA
  - Moving average
- CUSUM

Partial likelihood ratio

– Detection of $\tau=t$

- $C=\{\tau=t\}$ $D=\{\tau>s\}$
- $L(s, t) = f_{X_s}(x_s \mid \tau=t) / f_{X_s}(x_s \mid \tau>s)$

LR

- Full likelihood ratio
  - $LR(s) = f_{X_s}(x_s \mid C) / f_{X_s}(x_s \mid D)$
  - $C=\{\tau \leq s\}$ $D=\{\tau > s\}$
  - $LR(s)= \sum_{t=1}^{s} w(s, t)L(s, t) > G_s$

LR

- Fulfills several optimality criteria e.g.
  - Maximum expected utility

- Frisén, M. and de Maré, J. (1991). Optimal surveillance. *Biometrika*, 78, 271-80.

LR

- Alarmrule equivalent to rule with constant limit for the **posterior probability**
  - if only two states C and D.
  - Frisén, M. and de Maré, J. (1991). Optimal surveillance. *Biometrika*, 78, 271-80.
- ”The Bayes method”
- Frequentistic inference possible

- Comparison: **Hidden Markov Modeling** and LR
  - Andersson, E., Bock, D. and Frisén, M. (2002) Statistical surveillance of cyclical processes with application to turns in business cycles. *Submitted*.

Shirayev Roberts

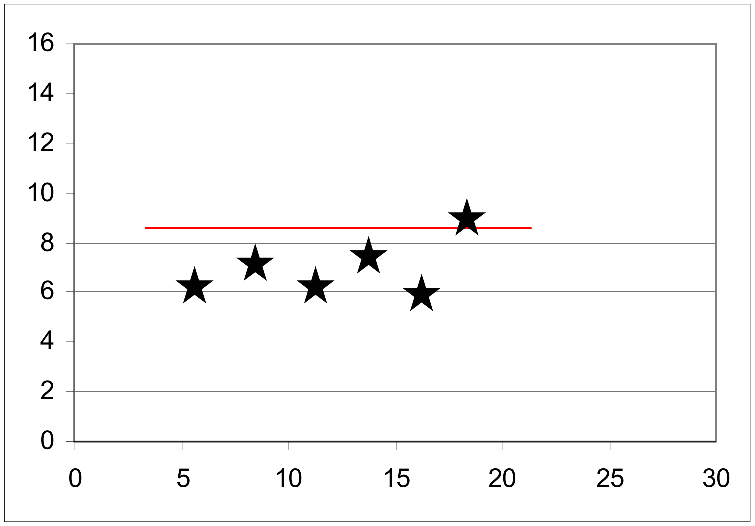
- The LR method with a non-informative prior.
- The limit of the LR method when the intensity ν tends to zero.
- Can often be used as an approximation of LR for rather large values of ν

Frisén, M., and Wessman, P. (1999), "Evaluations of Likelihood Ratio Methods for Surveillance. Differences and Robustness.," *Communications in Statistics. Simulations and Computations*, 28, 597-622.

Shewhart

- Alarmstatistic
  $X(s)=L(s,s)$
- Alarmlimit
  constant (often 3σ)
- Alarmrule

$t_A = \min\{s: X(s) > 3\sigma\},$



EWMA

Alarmstatistic

$$Z_s = \lambda \sum_{j=0}^{s-1} (1-\lambda)^j x(s-j) = \lambda(1-\lambda)^s \sum_{t=1}^{s} (1-\lambda)^{-t} x(t) \propto \sum_{t=1}^{s} b^t x(t)$$

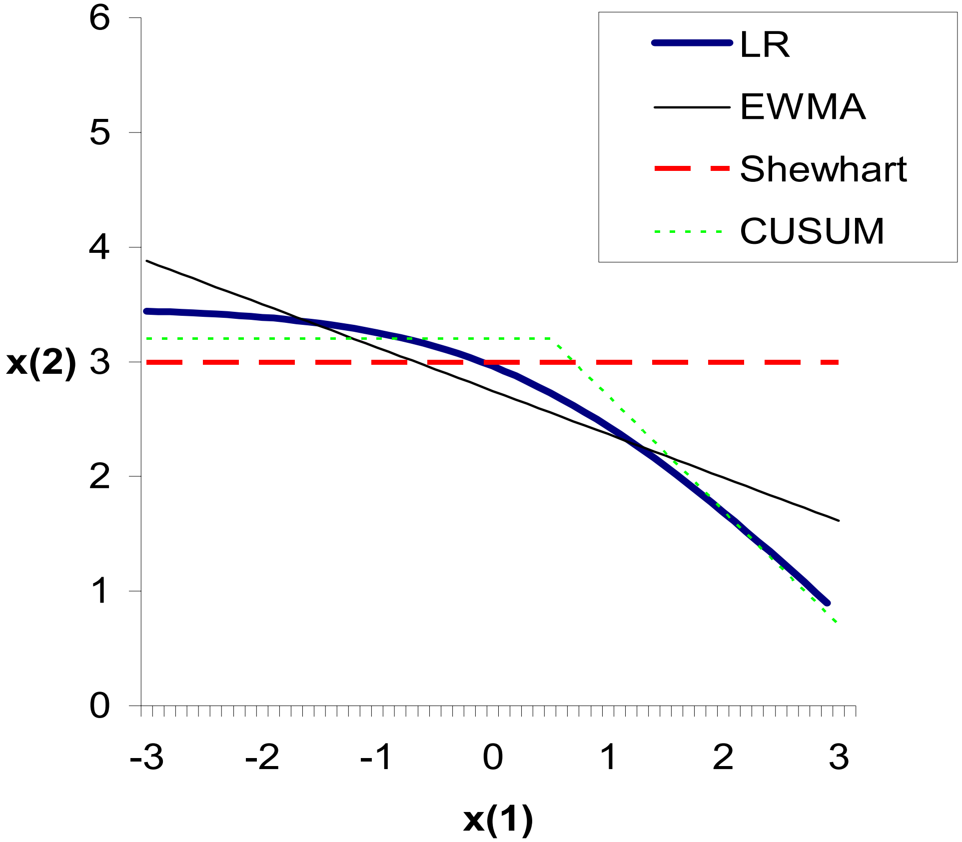
Approximates LR if $\lambda = 1 - \exp(-\mu^2/2)/(1-\nu)$

- Frisén, M. (in press), Statistical Surveillance. Optimality and Methods., *International Statistical Review*.
- Frisén, M. and Sonesson, C. (2003): Optimal surveillance by exponentially moving average mehtods. Submitted.

CUSUM

- Alarmrule
  - max(L(s, t); t=1, 2,.., s) > G
- Minimax optimality

Alarm limits at the second observation



Parameters for optimizing

The **Shewhart** method has **no** parameters

The **CUSUM** and the **Shiryaev-Roberts** methods have one parameter **M** to optimize for the size of the shift μ.

The **LR**-method has besides **M** also the parameter **V** to optimize for the intensity ν.

Similarity

The LR, **Shiryaev-Roberts** and the **CUSUM** methods tend to the Shewhart method when the parameter M tends to infinity.

This explains some earlier claims of similarities between some methods. These studies were made for very large values of M.

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 LR and the Shiryaev-Roberts methods have relatively constant predicted values.