# The Relationship Between the Recurrence Interval and Time-to-Signal Properties of Surveillance Schemes

SHANNON FRAKER
AND BILL WOODALL

QUALITY AND PRODUCTIVITY RESEARCH CONFERENCE

JUNE 2007

# Contents




- Background
- The recurrence interval
- Time-to-signal measures
- The recurrence interval and non-resetting EWMA charts
- The recurrence interval and non-resetting CUSUM charts
- The recurrence interval and Markov-dependent signaling processes
- Conclusions

# Background



- Public health surveillance is the collection, analysis, and interpretation of data essential to disease monitoring.
- Health surveillance methods not only need to detect that there is an active cluster occurring but they also need to do so in a timely manner.

# Background



- Industrial statistical process control (SPC) is a set of monitoring methods that have been widely used in industry for decades.

- The use of control charts such as the EWMA chart and the CUSUM chart are common practice and their properties have been studied in detail.

# Background



- Public health officials have also begun to use control charts in disease monitoring.
- Often, adjustments to the traditional control charts need to be made.
- Although control charts are becoming more accepted in disease monitoring, scan methods are the most popular method of outbreak detection.

# Issue



- There has not been adequate evaluation of the disease monitoring schemes. Often times, a new technique is developed and is evaluated in terms of a real-life dataset that is often extremely complicated.

# Recurrence Interval



- The recurrence interval is widely used in public health surveillance.

- It is commonly applied to Center for Disease Control and Prevention's nationwide monitoring system *BioSense* (www.cdc.gov/biosense/).

- Kleinman (2005, *Spatial & Syndromic Surveillance*) stated that by using the recurrence interval, one can use any retrospective spatial clustering test repeatedly.

- Kleinman (2005, *Spatial & Syndromic Surveillance*) stated that the prospective monitoring case of Kulldorff (2001, JRSS-A) was not necessary.

# Recurrence Interval



- The recurrence interval is defined to be the fixed number of time periods required for the expected number of false alarms in a monitoring process to be one.
- Defined for any point of monitoring.
- Often used to evaluate scan methods.

# Recurrence Interval



- Count data are available at regular time intervals for a specified number of regions, $S$.

- The recurrence interval is defined to be

$$RI = (p\text{-value} \; * \; S)^{-1}$$

- The $p$-value is calculated as the probability of getting as many occurrences as observed in that region given the predicted model.

- A signal is produced if the recurrence interval is too big.

# Recurrence Interval



- A surveillance scheme can be developed for the sums of the counts in overlapping windows of up to $r$ recent time periods. In this situation, the recurrence interval is defined to be

$$RI = (p\text{-value} * r * S)^{-1}$$

# Recurrence Interval



The recurrence interval is limited in its application by:

1. The process must be in-control.

2. The marginal probabilities of a signal must be constant for all time periods.

# Marginal Probability of a Signal



- Let $$E_i = \begin{cases} 1 & \text{if time } i \text{ results in a signal} \\ 0 & \text{otherwise} \end{cases}$$

- The marginal probability of a signal at time *i* is then $P(E_i = 1)$.

# Time-to-signal Measures



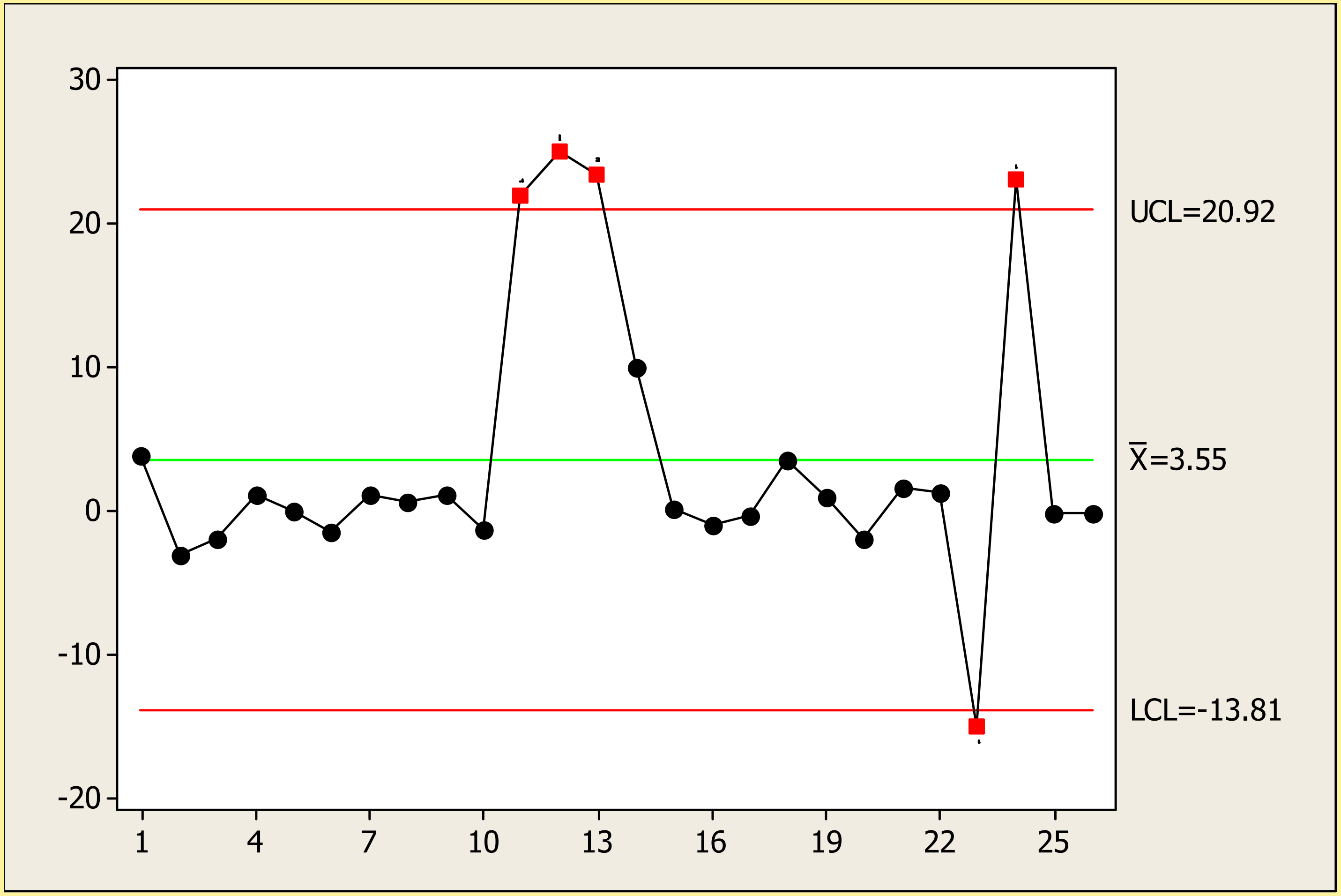
- The in-control average time-to-signal (ATS) is defined to be the average number of time periods until the first signal.

- The in-control average signaling event length (ASEL) is defined to be the average length of consecutive time periods that result in signals on the same side of the centerline for a two-sided chart.

- The in-control average time-between-signaling events (ATBSE) is defined to be the average number of time periods between signaling events.

# Time-to-signal Measures





# Non-resetting EWMA Chart



- It is assumed that observations, $X_1$, $X_2$, $X_3$,…, collected over time, are the result of an in-control process.
- These observations are distributed i.i.d. N(0,1)
- In the simulation, 2,500 control charts containing 10,000 time periods of observations are examined.

# Non-resetting EWMA Chart



$$M_i = \lambda X_i + (1-\lambda) M_{i-1}$$

$$\pm 3\left(\frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2i}\right]\right)^{1/2}$$

$i$ = 1, 2, 3, ... λ = 0.05, 0.10, 0.20, 0.30, or 0.40

# Non-resetting EWMA Chart



| λ | Estimated ATS | Estimated ATBSE | Estimated ASEL | Recurrence Interval |
|---|---|---|---|---|
| 0.05 | 1332.1 | 746.6 | 2.48 | 370.4 |
| 0.10 | 814.9 | 603.0 | 1.82 | 370.4 |
| 0.20 | 537.5 | 482.6 | 1.38 | 370.4 |
| 0.30 | 463.8 | 423.1 | 1.20 | 370.4 |
| 0.40 | 403.1 | 396.0 | 1.11 | 370.4 |

# Non-resetting EWMA Chart



- A second study simulated 10,000 charts under the same assumptions as the previous simulation. The number of signals that appear in the first 370 consecutive time periods was calculated for each of the 10,000 charts.

# Non-resetting EWMA Chart



| λ | Proportion with no false alarms | Proportion with one false alarm | Proportion with more than one false alarm |
|---|---|---|---|
| 0.05 | 0.7547 | 0.0796 | 0.1657 |
| 0.10 | 0.6369 | 0.1466 | 0.2165 |
| 0.20 | 0.5142 | 0.2391 | 0.2467 |
| 0.30 | 0.4487 | 0.2933 | 0.2580 |
| 0.40 | 0.4083 | 0.3282 | 0.2635 |
| 1.00 | 0.3678 | 0.3684 | 0.2639 |

# Non-resetting CUSUM Chart



- It is assumed that observations, $X_1$, $X_2$, $X_3$,…, collected over time, are the result of an in-control process.
- These observations are distributed i.i.d. N(0,1).

# Non-resetting CUSUM Chart



$$S_i = \max(0, S_{i-1} + X_i - k)$$

A signal is given as soon as $S_i > h$

$i = 1, 2, 3, \ldots$

$S_0 = 0$

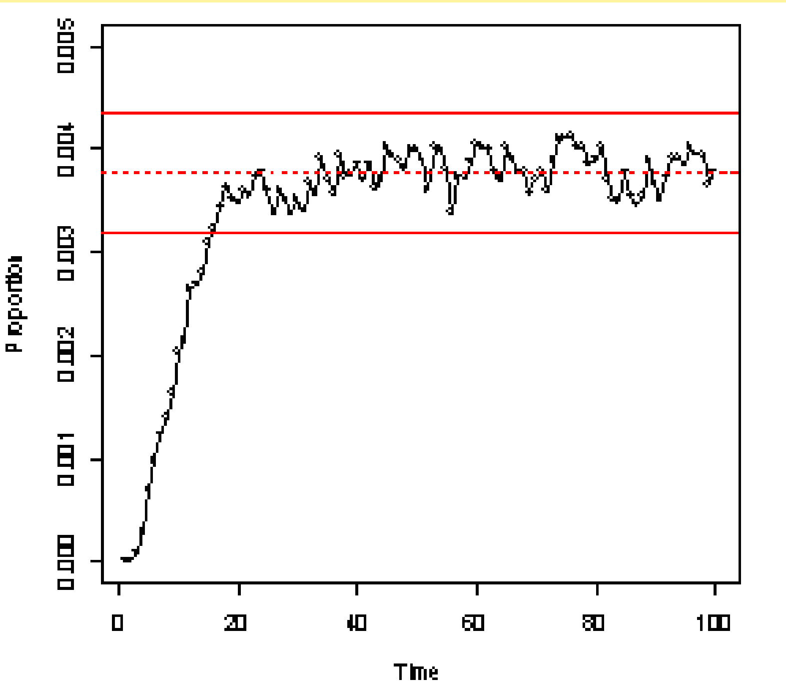
h = 4.0 or 5.0

k = 0.5 or 1.0

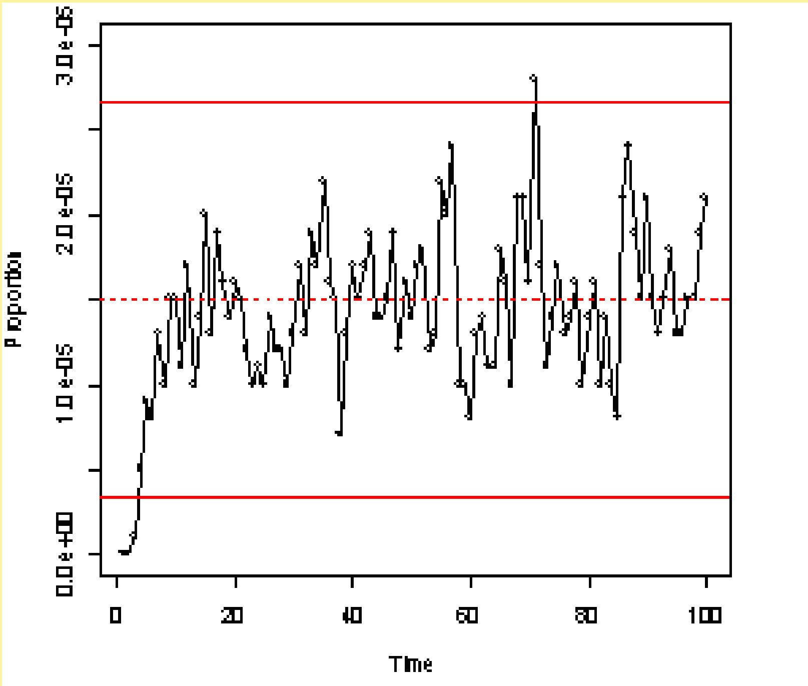
# Non-resetting CUSUM Chart



**Estimated Marginal Probabilities of a Signal**
***h* = 5.0    *k* = 0.5**




**Estimated Marginal Probabilities of a Signal**
***h* = 5.0    *k* = 1.0**



# Steady-state



- To calculate the recurrence interval one might wish to consider a steady-state value.
- For the CUSUM charts, the time to reach steady-state depends on the parameters.

# Markov-dependent signaling process



Let $a$ = P(signal after a non-signal)

Let $b$ = P(non-signal after a signal)

$$1-b > a$$

ATS = $1/a$

ATBSE = $1/a$

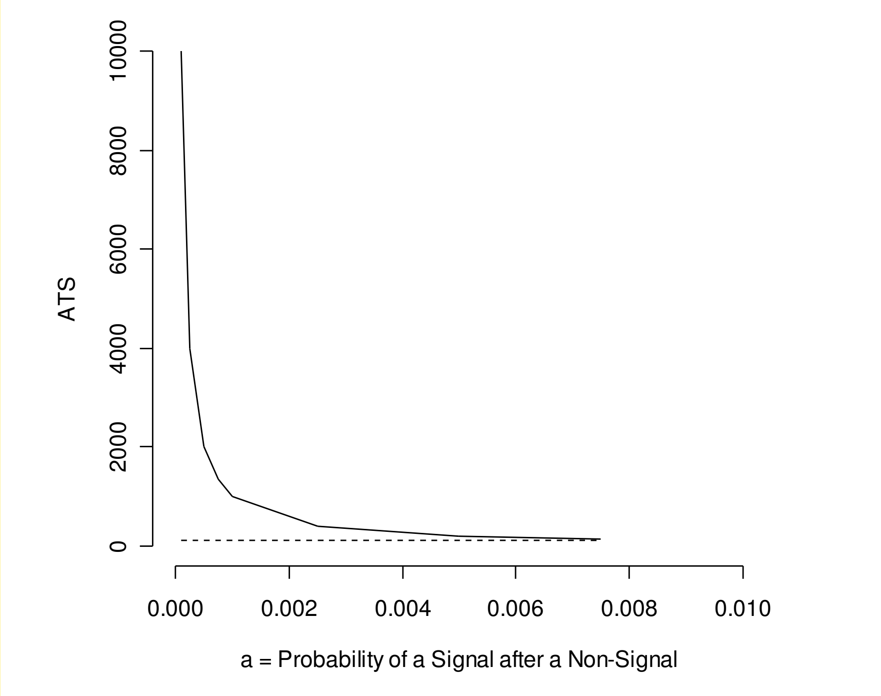
ASEL = $1/b$

SSRI = $1 + b/a$

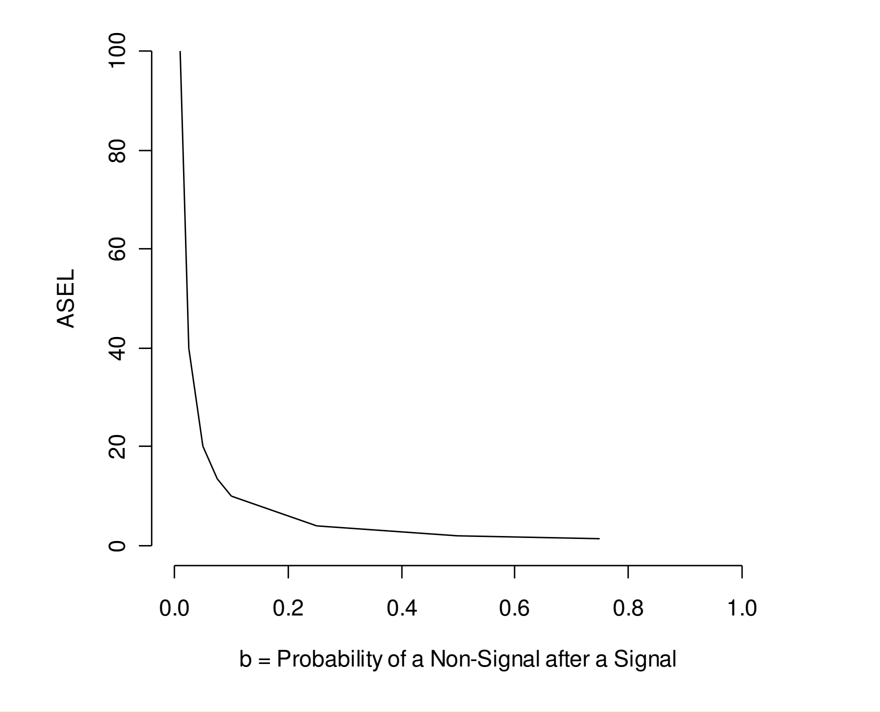
# Markov-dependent signaling process



**In-Control ATS and In-control ATBSE Values for a Constant Recurrence Interval of 101.00**




**In-Control ASEL Values for a Constant Recurrence Interval of 101.00**



# Conclusions



- To compare the recurrence interval and time-to-signal properties, the ATBSE and the ASEL were introduced.
- Processes with widely different in-control time-to-signal properties can have the same recurrence interval.
- The recurrence interval does not seem to adequately summarize a process and is limited in its applicability.
- Since the recurrence interval is only defined for the in-control case, the time-to-signal properties will always be needed to assess out of control performance.

# Publications




- Woodall, W.H., Marshall, J. B., Joner, M.D. Jr., Fraker, S.E., and Abdel-Salam, A.G. (2007) On the Use and Evaluation of Scan Methods in Prospective Public Health Surveillance. To appear in the *Journal of the Royal Statistical Society*, Series A.

- Fraker, S.E., Woodall, W. H., and Mousavi, S. (2007) The relationship between the recurrence interval and time-to-Signal properties of surveillance schemes, submitted for publication to *Quality Engineering* (special issue on SPC in health care).

# Acknowledgements




- I would like to think Olivia Grigg, Howard Burkom, and Shabnam Mousavi.

- I would like to thank the Natrella Scholarship Committee.

# The Relationship between the Recurrence Interval and Time-to-Signal Properties of Surveillance Schemes

SHANNON FRAKER

SFRAKER@VT.EDU