

Editor
MARIANNE FRISÉN

Financial Surveillance



 **WILEY**

STATISTICS IN PRACTICE

Financial Surveillance

Edited by

Marianne Frisé

Göteborg University, Sweden



John Wiley & Sons, Ltd

Copyright © 2008

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk

Visit our Home Page on www.wileyeurope.com or www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, Ontario, L5R 4J3, Canada

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Library of Congress Cataloging-in-Publication Data

Financial surveillance / edited by Marianne Frisén.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-06188-6 (cloth : acid free paper)

1. Econometric models. 2. Mathematical optimization. I. Frisén, Marianne.

HB141.F75 2007

332.01'519 – dc22

2007045546

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-470-06188-6 (HB)

Typeset in 10.5/13pt Times by Laserwords Private Limited, Chennai, India

Printed and bound in Great Britain by TJ International, Padstow, Cornwall

This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

List of Contributors

EVA ANDERSSON Statistical Research Unit, Department of Economics, Göteborg University, PO Box 640, SE 405 30 Göteborg, Sweden; and Department of Occupational and Environmental Medicine, Sahlgrenska University Hospital, PO Box 414, SE 405 30 Göteborg, Sweden

DAVID BOCK Statistical Research Unit, Department of Economics, Göteborg University, PO Box 640, SE 405 30 Göteborg, Sweden

MARIANNE FRISÉN Statistical Research Unit, Department of Economics, Göteborg University, PO Box 640, SE 405 30 Göteborg, Sweden

VASYL GOLOSNOY Institute of Statistics and Econometrics University of Kiel, Olshausenstraße 40, D-24118 Kiel, Germany

IRYNA OKHRIN Department of Statistics, Europa-Universität Viadrina, Grosse Scharrnstr. 59, D-15230 Frankfurt (Oder), Germany

YAREMA OKHRIN Department of Statistics, Europa-Universität Viadrina, Grosse Scharrnstr. 59, D-15230 Frankfurt (Oder), Germany

WOLFGANG SCHMID Department of Statistics, Europa-Universität Viadrina, Grosse Scharrnstr. 59, D-15230 Frankfurt (Oder), Germany

HELGI TÓMASSON Faculty of Economics and Business Administrations, University of Iceland, Oddi v/Sturlugotu, IS-101 Reykjavik, Iceland

Contents

List of Contributors	vii
1 Introduction to financial surveillance	1
<i>Marianne Frisé</i>	
2 Statistical models in finance	31
<i>Helgi Tómasson</i>	
3 The relation between statistical surveillance and technical analysis in finance	69
<i>David Bock, Eva Andersson and Marianne Frisé</i>	
4 Evaluations of likelihood-based surveillance of volatility	93
<i>David Bock</i>	
5 Surveillance of univariate and multivariate linear time series	115
<i>Yarema Okhrin and Wolfgang Schmid</i>	
6 Surveillance of univariate and multivariate nonlinear time series	153
<i>Yarema Okhrin and Wolfgang Schmid</i>	
7 Sequential monitoring of optimal portfolio weights	179
<i>Vasyl Golosnoy, Wolfgang Schmid and Iryna Okhrin</i>	
8 Likelihood-based surveillance for continuous-time processes	211
<i>Helgi Tómasson</i>	
9 Conclusions and future directions	235
<i>Marianne Frisé</i>	
Bibliography	239
Index	257

1

Introduction to financial surveillance

Marianne Frisé

Statistical Research Unit, School of Business, Economics and Law, Göteborg University, PO Box 660, SE 405 30 Göteborg, Sweden

1.1 What is financial surveillance?

In financial surveillance the aim is to signal at the optimal trading time. A systematic decision strategy is used. The information available at each possible decision time is evaluated in order to judge whether or not there is enough information for a decision about an action or if more information is necessary so that the decision should be postponed. Financial surveillance gives timely decisions.

The authors of this book hope that it will serve two purposes. First, we hope that it will stimulate an increased use of surveillance in finance by providing methods which have not been available before. Second, we hope that the statistical community will use the book as a spur to further development of techniques and to research on questions unanswered by the following chapters. Financial surveillance is a new area, and some open problems are discussed in Chapter 9.

Financial decision strategies are based, in one way or another, on continuous observation and analysis of information. This is financial surveillance. Statistical surveillance uses decision theory and statistical inference in order to derive timely decision strategies. Hopefully, this book will serve as a bridge between finance and statistical surveillance.

This book is written by statisticians with an interest in finance. Textbooks describing financial problems and statistical methods are for example Föllmer and Schied (2002), Härdle, Kleinow and Stahl (2002), Gouriéroux and Jasiak (2002), Franke, Härdle and Hafner (2004), Cizek, Härdle and Weron (2005) and Scherer and Martin (2005). Many and various statistical techniques are described in these books.

In Section 1.2 statistical methods which are useful for financial decisions will be discussed. In Section 1.3 the area of statistical surveillance is described, and the characteristics of surveillance are compared to other areas in statistics. Evaluations in surveillance are described in Section 1.4. This is an important area, since the choice of evaluation measures will decide which methods are considered appropriate. General methods for aggregating information over time are described in Section 1.5. Special aspects of surveillance for financial decisions are discussed in Section 1.6.

The content of the book will be described in Section 1.7, which deals with the relation between the chapters. This section also provides reading guidelines.

1.2 Statistical methods for financial decision strategies

Statistical methods use observations of financial data to give information about the financial process which produces the data. This is in contrast to probability theory, where assumptions about the financial process are used to derive which observations will be generated.

1.2.1 Transaction strategies based on financial data

In finance, the relation between observations and decisions is often informal. Statisticians have taken on the role of presenting statistical summaries of quantitative data. In many areas, including finance, this means providing point and interval estimates for the quantities of interest. Methods for providing such summaries are highly formalized and constantly evolving. The discipline of statistics uses observations to make deductions about the real world. It has its own set of axioms and theorems besides those of probability theory. While decision making is the incentive for much statistical analysis, the process that transforms statistical summaries into decisions usually remains informal and ad hoc.

In finance the timeliness of transactions is important to yield a large return and a low risk. The concept of an efficient arbitrage-free market is of great interest. One central question is whether the history of the price of an asset contains information which can be used to increase the future return. A natural aim

is to maximize the return. The theory of stochastic finance has been based on an assumption of an efficient market where the financial markets are arbitrage-free and there is no point in trying to increase the return. Even though this view is generally accepted today, there are some doubts that it is generally applicable. When the information about the process is incomplete, as for example when a change point could occur, there may be an arbitrage opportunity, as demonstrated by Shiryaev (2002). In Chapter 3 it will be discussed how technical analysis relies on the possibility of using history to increase future returns. The support for the efficient market hypothesis depends on the knowledge about the model, as is discussed below in Section 1.2.2.

1.2.2 Modelling

In finance, advanced stochastic models are necessary to capture all empirical features. The expected value could depend on time in a complicated nonlinear way. Parameters other than the expected value are often of great interest, and the risk (measured by variance) is often of great concern. Complicated dependencies are common, which means that complicated measures of variance are necessary. Multivariate data streams are of interest for example when choosing portfolio. The models may be described in continuous or discrete time. Chapter 2 gives an overview of models of interest in finance. Statistical methods for estimation and model choice will also be briefly described in the chapter. The use of the models should be robust to errors in the model specification.

1.2.2.1 Stochastic model assumed known

When the stochastic model is assumed to be completely known there is no expected return to be gained. We will have an arbitrage-free market. We can use probability theory to calculate the optimal transaction conditions. Important contributions are found in the book by Shiryaev (1999) or in articles in the scientific journal *Finance and Stochastics*. Also the proceedings of the conference *Stochastic Finance* in 2004 and 2007 are informative on how to handle financial decisions when the model is completely known.

1.2.2.2 Incomplete knowledge about the stochastic model

When the model is not completely known, the efficient and arbitrage-free market assumptions are violated. Changes at unknown times are possible. One has to evaluate the information continuously to decide whether a transaction at that time is profitable. Statistical inference is needed for the decision (Shiryaev 2002).

1.2.3 Evaluation of information

Statistical inference theory gives guidelines on how to draw conclusions about the real world from data. Statistical hypothesis testing is suitable for testing a single hypothesis but not a decision strategy including repeated decisions, as will be further described in Section 1.4.1. Statistical surveillance is an important branch of inference. The relatively new area of statistical surveillance deals with the sequential evaluation of the amount of information at hand. It provides a theory for deciding at what time the amount of information is enough to make a decision and take action. This bridges the gap between statistical analysis and decisions.

In this book, we concentrate on the methodology of statistical surveillance. This methodology is of special interest for financial decision strategies, but it is also relatively new in finance. The ambition here is to give a comprehensive description of such aspects of statistical surveillance that may be of interest in finance. Thus the next sections of this chapter will give a short review on statistical surveillance.

1.3 What is statistical surveillance?

1.3.1 General description

Statistical surveillance means that a time series is observed with the aim of detecting an important change in the underlying process as soon as possible after the change has occurred. Statistical methods are necessary to separate important changes in the process from stochastic variation. The inferential problems involved are important for the applications and interesting from a theoretical viewpoint, since they bring different areas of statistical theory together.

Broad surveys and bibliographies on statistical surveillance are given by Lai, who concentrates on the minimax properties of stopping rules, by Woodall and Montgomery (1999) and Ryan (2000), who concentrate on control charts, and by Frisé (2003), who concentrates on the optimality properties of various methods.

The theory of statistical surveillance has developed independently in different statistical subcultures. Thus, the terminology is diverse. Different terms are used to refer to ‘statistical surveillance’ as described here. However, there are some differences in how the terms are used. ‘Optimal stopping rules’ are most often used in probability theory, especially in connection with financial problems. However, this does not always include the statistical inference from the observations to the model. Literature on ‘change-point problems’ does not always treat the case of continuous observations but often considers the case

of a retrospective analysis of a fixed number of observations. The term ‘early warning system’ is most often used in the economic literature. ‘Monitoring’ is most often used in medical literature and as a nonspecific term. Timeliness, which is important in surveillance, is considered in the vast literature on quality control charts, and here also the simplicity of procedures is stressed. The notations ‘statistical process control’ and ‘quality control’ are used in the literature on industrial production and sometimes also include other aspects than the statistical ones.

The statistical methods suitable for surveillance differ from the standard hypothesis testing methods. In the prospective surveillance situation, data accumulated over time is analysed repeatedly. A decision concerning whether, for example, the variance of the price of a stock has increased or not has to be made sequentially, based on the data collected so far. Each new possibility demands a new decision. Thus, there is no fixed data set but an increasing number of observations. In sequential analysis we have repeated decisions, but the hypotheses are fixed. In contrast, there are no fixed hypotheses in surveillance. The statistics derived for a fixed sample may be of great value also in the case of surveillance, but there are great differences between the systems for decision. The difference between hypotheses and on-line surveillance is best seen by studying the difference in evaluation measures (see Section 1.4.1).

In complicated surveillance problems, a stepwise reduction of the problem may be useful. Then, the statistics derived to be optimal for the fixed sample problem can be a component in the construction of the prospective surveillance system. This applies, for example, to the multivariate problems described in Section 1.6.7 in this chapter and in Chapters 5–7.

1.3.2 History

The first modern control charts were developed in the 1920s, by Walter A. Shewhart and coworkers at Bell Telephone Laboratories. In 1931 the famous book *Economic Control of Quality of Manufactured Product* (Shewhart 1931) was published. The same year Shewhart gave a presentation of the new technique to the Royal Statistical Society. This stimulated interest in the UK. The technique was used extensively during World War II both in the UK and in the US. In the 1950s, W. E. Deming introduced the technique in Japan. The success in Japan spurred the interest in the West, and further development started.

In the Shewhart method each observation is judged separately. The next important step was taken when Page suggested the CUSUM method for aggregating information over time. Shortly afterwards, Roberts (1959) suggested another method for aggregating information – the EWMA method. A method

based on likelihood which fulfils important optimality conditions was suggested by Shiryaev (1963).

In recent years there have been a growing number of papers in economics, medicine, environmental control and other areas, dealing with the need of methods for surveillance. The threat of bioterrorism and new contagious diseases has been an important reason behind the increased research activity in the theory of surveillance. Hopefully, the time is now ripe for finance to benefit from all these results.

1.3.3 Specifications of the statistical surveillance problem

The general situation of a change in distribution at a certain change-point time τ will now be specified. The variable under surveillance could be the observation itself or an estimator of a variance or some other derived statistic, depending on the specific situation. We denote the process by $X = \{X(t) : t = 1, 2, \dots\}$, where $X(t)$ is the observation made at time t . The purpose of the monitoring is to detect a possible change. The time for the change is denoted by τ (see Figure 1.1). This can be regarded either as a random variable or as a deterministic but unknown value, depending on what is most suitable for the application.

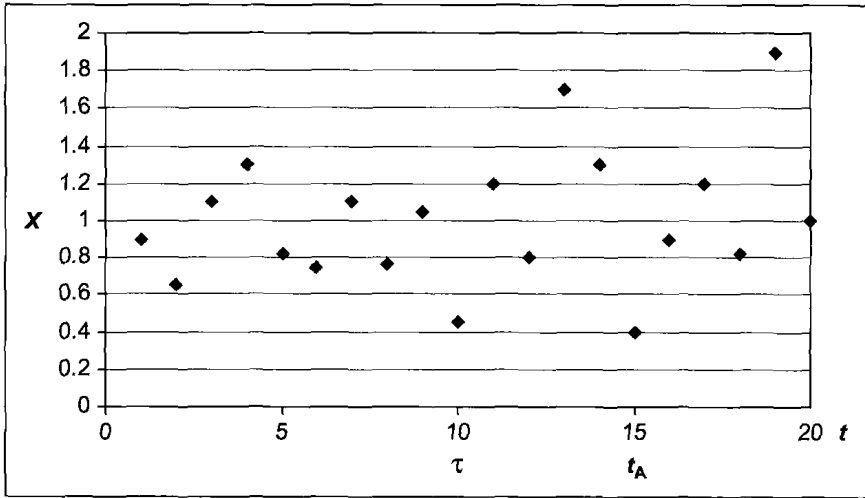


Figure 1.1 The first $\tau - 1$ observations $X_{\tau-1} = \{X(t) : t \leq \tau - 1\}$ are ‘in-control’ with a small variance. The subsequent observations (from $t = \tau$ (here 10) and onwards) have a larger variance. The alarm time is t_A , which happens to be 15. Thus the delay is $t_A - \tau = 5$.

The properties of the process change at time τ . In many cases we can describe this as

$$X(t) = \begin{cases} Y(t), & t < \tau \\ Y(t) + \Delta, & t \geq \tau \end{cases} \quad (1.1)$$

where Y is the ‘in-control’ or ‘target’ process and Δ denotes the change. More generally we can denote the ‘in-control’ state by D and the state which we want to detect by C . The (possibly random) process that determines the state of the system is here denoted by $\mu(t)$. This could be an expected value, a variance or some other time-dependent characteristic of the distribution. Different types of states between which the process changes are of interest for different applications. Descriptions of the details of the states are made within each chapter.

The change to be detected differs depending on the application. Most studies in literature concern a step change, where a parameter changes from one constant level, say, $\mu(t) = \mu^0$ to another constant level, $\mu(t) = \mu^1$. The case $\mu > 0$ is described here. We have $\mu(t) = \mu^0$ for $t = 1, \dots, \tau-1$ and $\mu(t) = \mu^1$ for $t = \tau, \tau + 1$.

Even though autocorrelated time series are studied for example by Schmid and Schöne (1997), Petzold, Sonesson, Bergman and Kieler (2004), and in Chapters 5–7, processes which are independent given τ are the most studied and used also in Chapters 3 and 4. This simple situation will be used to introduce general concepts of evaluations, optimality and standard methods.

Some cases of special interest in financial surveillance are discussed in Section 1.6.

1.4 Evaluations

Quick detection and few false alarms are desired properties of methods for surveillance. Knowledge about the properties of the method in question is important. If a method calls an alarm, it is important to know whether this alarm is a strong indication of a change or just a weak indication. The same methods can be derived by Bayesian or frequentistic inference. However, evaluations differ. Here we present measures suitable for frequentistic inference.

1.4.1 The difference between evaluations for hypothesis testing and on-line surveillance

Measures for a fixed sample situation can be adopted for surveillance, but some important differences will be pointed out. In Table 1.1 the measures

Table 1.1 Evaluation measures for hypothesis testing and the corresponding measures for on-line surveillance.

	<i>Test</i>	<i>Surveillance</i>
<i>False alarms</i>	Size α , Specificity	ARL ⁰ , MRL ⁰ , PFA
<i>Detection ability</i>	Power, Sensitivity	ARL ¹ , MRL ¹ , CED, ED, maxCED, PSD, SADT

conventionally used in hypothesis testing and some measures for surveillance are given. These measures will be described and discussed below.

Different error rates and their implications for a decision system were discussed by Friséen and de Maré (1991). Using a constant probability of exceeding the alarm limit for each decision time means that we have a system of repeated significance tests. This may work well also as a system of surveillance and is often used. The Shewhart method described in Section 1.5.2 has this property. This is probably also the motive for using the limits with the exact variance in the EWMA method described in Section 1.5.4.

Evaluation by significance level, power, specificity and sensitivity, which is useful for a fixed sample, is not appropriate without modification in a surveillance situation since these measures do not have unique values in a surveillance system. One problem with evaluation measures originally suggested for the study of a fixed sample of, say, n observations is that the measures depend on n . For example, the specificity will tend to zero for most methods and the size of the test will tend to one when n increases.

Chu, Stinchcombe and White (1996) and others have suggested methods with a size less than one:

$$\lim_{n \rightarrow \infty} P(t_A \leq n | D) < 1.$$

This is convenient since ordinary statements of hypothesis testing can be made. However, Friséen (2003) demonstrated that the detection ability of methods with this property declines rapidly with the value of time τ of the change. Important consequences were illustrated by Bock (2008).

The performance of a method for surveillance depends on the time τ of the change. Generally, the sensitivity will not be the same for early changes as for late ones. It also depends on the length of time for which the evaluation is made. Thus, there is not one unique sensitivity value in surveillance, but other measures may be more useful. Accordingly, conventional measures for fixed samples should be supplemented by other measures designed for statistical surveillance, as will be discussed in the following.

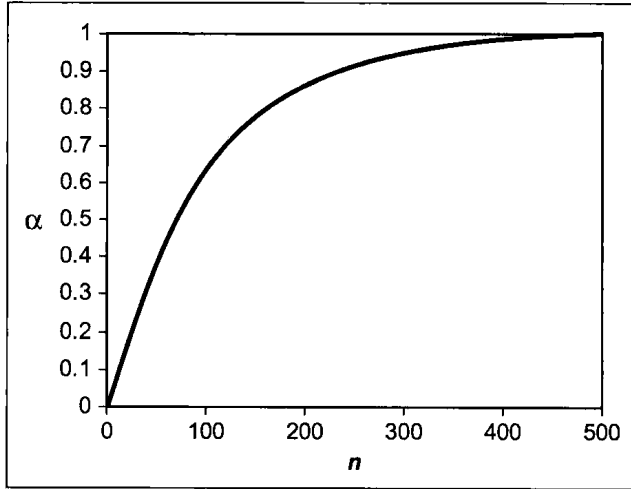


Figure 1.2 The size, α , of a surveillance system which is pursued for n time units, when the probability of a false alarm is 1% at each time point.

1.4.2 Measures of the false alarm rate

The false alarm tendency is more complicated to control in surveillance than in hypothesis testing, as was seen above (for example in Figure 1.2). There are special measures of the false alarm properties which are suitable for surveillance. The most commonly used measure is the Average Run Length when there is no change in the system under surveillance, $ARL^0 = E(t_A|D)$. A variant of the ARL is the Median Run Length, MRL.

A measure commonly used in theoretical work is the false alarm probability, $PFA = P(t_A < \tau)$. This is the probability that the alarm occurs before the change.

1.4.3 Delay of the alarm

The delay time of the detection of a change should be as short as possible. The most commonly used measure of the delay is the average run length until the detection of a true change (that occurred at the same time as the surveillance started), which is denoted by ARL^1 . The part of the definition within parentheses is seldom spelled out but generally used in the literature (see, for example, Page 1954 and Ryan 2000). Instead of the average, Gan (1993) advocates that the median run length should be used on the grounds that it may be more easily interpreted. However, also here only a change occurring at the same time as the surveillance started is considered.

In most practical situations it is important to minimize the expected delay of detection whenever the change occurs. Shiryaev (1963) suggested measures of the expected value of the delay. The expected delay from the time of change, $\tau = t$, to the time of alarm, t_A , is denoted by

$$ED(t) = E[\max(0, t_A - t) | \tau = t].$$

Note that $ARL^1 = ED(1) + 1$. The $ED(t)$ will typically tend to zero as t increases. Thus, it is easier to evaluate the conditional expected delay

$$CED(t) = E[t_A - \tau | t_A \geq \tau = t] = ED(t)/P(t_A \geq t).$$

$CED(\tau)$ is the expected delay for a specific change point τ . The expected delay is generally not the same for early changes as for late ones. For most methods, the CED will converge to a constant value. This value is sometimes named the 'steady state average delay time' or SADT. It is, in a sense, the opposite to ARL^1 since only a very large value of τ is considered. SADT has been advocated for example by Srivastava and Wu (1993), Srivastava (1994) and Knoth (2006).

For some situations and methods the properties are about the same regardless of when the change occurs. However, this is not always true, as illustrated by Frisén and Wessman (1999). Then, it is important to consider more and other cases than just $\tau = 1$. The values of CED can be summarized in different ways. One is the maximal value over τ . Another approach is to regard τ as a random variable with the probabilities $\pi(t) = P(\tau = t)$. These probabilities can also be regarded as priors. The intensity of a change is defined as $\nu(t) = P(\tau = t | \tau \geq t)$, which is usually assumed to be constant over time. Shiryaev (1963) suggested a summarized measure of the expected delay

$$ED = E[ED(\tau)].$$

Sometimes the time available for action is limited. The Probability of Successful Detection suggested by Frisén (1992) measures the probability of detection with a delay time no longer than d

$$PSD(d, t) = P(t_A - \tau < d | t_A \geq \tau = t).$$

This measure is a function of both the time of the change and the length of the interval in which the detection is defined as successful. Also when there is no absolute limit to the detection time it is often useful to describe the ability to detect the change within a certain time. In such cases it may be useful to calculate the PSD for different time limits d . This has been done by Marshall, Best, Bottle and Aylin (2004). The ability to make a very quick detection

(small d) is important in surveillance of sudden major changes, while the long-term detection ability (large d) is more important for ongoing surveillance where smaller changes are expected.

1.4.4 Predictive value

When an alarm is called, one needs to know whether to act as if the change is certain or just plausible. To obtain this, both the risk of false alarms and the risk of delay must be considered. If τ is regarded as a random variable this can be done by one summarizing measure. The probability that a change has occurred when the surveillance method signals was suggested by Frisén (1992) as a time-dependent predictive value

$$PV(t) = P(\tau \leq t_A | t_A = t).$$

When there is an alarm ($t_A = t$), PV indicates whether there is a large probability or not that the change has occurred ($t_A \geq \tau$). Some methods have a constant PV. Others have a low PV at early alarms but a higher one later. In such cases, the early alarms will not prompt the same serious action as later ones.

1.4.5 Optimality

1.4.5.1 Minimal expected delay

Shiryaev (1963) suggested a highly general utility function, in which the expected delay of an alarm plays an important role. Shiryaev treated the case where the gain of an alarm is a linear function of the value of the delay, $t_A - \tau$, and the intensity of the change is constant. The loss associated with a false alarm is a function of the same difference. This utility can be expressed as $U = E\{u(\tau, t_A)\}$, where

$$u(\tau, t_A) = \begin{cases} h(t_A - \tau) & \text{if } t_A < \tau \\ a_1(t_A - \tau) + a_2 & \text{else.} \end{cases}$$

The function $h(t_A - \tau)$ is usually a constant (say, b), since the false alarm causes the same cost of alerts and investigations irrespectively of how early the false alarm is given. In this case, we have

$$U = b P(t_A < \tau) + a_1 ED + a_2.$$

We would have a maximal utility if there is a minimal (a_1 is typically negative) expected delay from the change point for a fixed probability of a false alarm (see Section 4.3). This is termed the ED criterion. Variants of the utility function leading to different optimal weighting of the observations are suggested for example by Poor (1998) and Beibel (2000).

1.4.5.2 Minimax optimality

The minimum of the maximal expected delay after a change considers several possible change times, just like the ED criterion. However, instead of an expected value, which requires a distribution of the time of change, the least favourable value of $CED(t)$ is used. This criterion is used in Chapter 5.

Moustakides (1986) uses an even more pessimistic criterion, the ‘worst possible case’, by using not only the least favourable value of the change time, but also the least favourable outcome of $X_{\tau-1}$ before the change occurs. This criterion is very pessimistic. The CUSUM method, described in Section 1.5.3, provides a solution to the criterion proposed by Moustakides. The merits of the studies of this criterion have been thoroughly discussed for example by Yashchin (1993) and Lai (1995). Much theoretical research is based on this criterion.

1.4.5.3 ARL optimality

Optimality is often stated as a minimal ARL^1 for a fixed ARL^0 . ARL^1 is the expected value under the assumption that all observations belong to the ‘out-of-control’ distribution, whereas ARL^0 is the expected value given that all observations belong to the ‘in-control’ distribution. Efficient methods for surveillance (see Section 1.5) will put most weight on the most recent observations. Statistical inference with the aim of discriminating between the two alternatives that all observations come from either of the two specified distributions should, by the ancillarity principle, put the same weight on all observations. To use efficient methods and evaluate them by the ARL criterion is thus in conflict with this inference principle.

Pollak and Siegmund (1985) argue that for many methods, the maximal value of $CED(t)$ is equal to $CED(1)$, and with a minimax perspective this can be an argument for using ARL^1 since $CED(1) = ARL^1 - 1$. However, this argument is not relevant for all methods. In particular, it is demonstrated by Friséen and Sonesson (2006) that the maximal CED-value is not $CED(1)$ for the EWMA method in Section 1.5.5. In the case of this method, there is no similarity between the optimal parameter values according to the ARL criterion and the minimax criterion, while the optimal parameter values by the criterion of expected delay and the minimax criterion agree well.

The dominating position of the ARL criterion was questioned by Friséen (2003), since methods useless in practice are ARL optimal. The ARL can be used as a descriptive measure and gives a rough impression, but it is questionable as a formal optimality criterion.