GÖTEBORG UNIVERSITY

Department of Statistics

RESEARCH REPORT 1995:2
ISSN 0349-8034

DETECTION OF GRADUAL CHANGES. STATISTICAL METHODS IN POST MARKETING SURVEILLANCE

by

Alecka Sveréus

Statistiska institutionen
Göteborgs Universitet
Viktoriagatan 13
S-411 25 Göteborg
Sweden

DETECTION OF GRADUAL CHANGES.
STATISTICAL METHODS IN POST MARKETING SURVEILLANCE

By Alecka Sveréus
Department of Statistics
University of Göteborg, S-41125 Göteborg, Sweden

ABSTRACT

Surveillance can be viewed as continual observation in time where the goal is to detect a change in the underlying process as soon as possible after it has occurred. In many applications, such as post marketing surveillance, it is of special interest to detect gradual changes in the underlying process. When a drug has been marketed one needs a continuous surveillance of adverse drug reactions. This is now done by different statistical methods suggested by the Food and Drug Administration (FDA), among others.

The ability to detect a linear increase by different methods of surveillance is analysed . This is compared to the case where a sudden change to a constant level occurs. Often, as in the FDA recommendations, repeated significance tests are made and this technique of surveillance is identical to the Shewhart method. Different significance tests correspond to different transformations of the observations to the variable to be used in the Shewhart test. This method is evaluated by different measures of goodness as the false alarm probability, the probability of successful detection and the predicted value, for different cases of the critical event. Evaluations of the transformations suggested by the FDA are done in the case of the critical event being a sudden shift to a constant level and for the case of a linear increase. Considerable differences are demonstrated.

A method which is optimal to detect a linear increase is derived. It takes into consideration all the data up to the decision time. A linear approximation of this method is derived and the weights in this approximation are studied. A comparison with the linear approximation of the method which is optimal to detect a sudden shift, is made. For the cases studied the Shewhart method approximates the optimal method better when the critical event is a linear change than when it is a shift.

Keywords: Surveillance, Gradual changes, Linear increase, Post Marketing, Adverse Reactions, False alarm probability, Successful Detection, Predictive Value, Likelihood Ratio.

TABLE OF CONTENTS

1. INTRODUCTION

The aims of this work is to contribute with some new ideas and methods in the area of post marketing surveillance and to develop methods that can detect gradual changes in the surveillance. The description of practical aspects of post marketing surveillance of adverse drug reactions and how these are handled today was obtained through studies of the medical literature in the field, interviews and discussions with WHO staff. This is described in Section 2.

In Section 3 the common praxis in this area is described with the problems of data quality in Section 3.1. The statistical methods proposed for post marketing surveillance of adverse drug reactions by the FDA, are given in Section 3.2. In Section 3.3 the methods suggested by Praus (1993) et al, by Moussa (1978), and by Mandel et al (1976) are described.

In Section 4 the characteristics of surveillance in general are given.

One other important aim of this work is to develop methods to detect gradual changes through surveillance and also to evaluate these methods. A stochastic model used for describing the surveillance situation at hand is formulated in Section 5. In this model I assume that the change in the process occurs at the stochastic point τ.

Different measures of performance and optimality criteria are described in Section 6. The measures of goodness chosen in this work are the Average Run Length (ARL) in Section 6.1, the probability of false alarms in Section 6.2, the probability of successful detection in Section 6.3 and the predictive value in Section 6.4.

The performance of the type of methods suggested by the Food and Drug Administration (FDA) has been evaluated in Section 7 by the ARL (see Section 7.1), the probability for false alarm (see Section 7.2), the probability for successful detection (see Section 7.3) and the predicted value of an alarm (see Section 7.4).

An optimal method for post marketing surveillance is derived in Section 8.1, applying the Likelihood Ratio technique when the critical event is a linear increase and where all changes before the decision time point are of interest. This technique which was suggested by Frisén and De Maré (1991) was in that paper applied to the case where the critical event is a shift to some level μ^1. A linear approximation of the method is given in Section 8.2.

An evaluation of the transformations of the data to achieve a test statistic suggested by the FDA is made in Section 9. The expected values are calculated in Section 9.1. To examine if successive values of the transformed variables are independent the covariances are calculated in Section 9.2.

Concluding remarks are given in Section 10.

2. THE NEED FOR POST MARKETING MONITORING

New drugs are registered on the basis of data from clinical trials usually involving less than 3000 patients. It is almost impossible to include enough patients in clinical trials to detect uncommon adverse reactions. Such clinical trials consist of certain types of exposed population and exclude some groups of patients, e.g. the very young, pregnant women and patients with severe diseases.

Adverse drug reactions which occur a long time after the start of the use of the drug or which are very rare can often go undetected until after the drug is on the market. Further surveillance is therefore required once marketing approval is obtained. A system of post marketing surveillance (PMS) is therefore essential to determine the safety profile of medicines when used in clinical practice.

This need is recognized by the World Health Organization (WHO) with a Collaborative Research Centre at the Drug Regulatory Authority of Sweden in Uppsala. The aim of this WHO program is to develop the existing signal generation potential and to ensure that analysis and investigation of all important safety signals proceeds consistently.

3. COMMON PRAXIS OF POST MARKETING SURVEILLANCE

The surveillance for adverse drug reactions (ADR) in Europe rests almost exclusively on a spontaneous reporting system for practicing physicians established according to the recommendations of the WHO. Spontaneous reports of ADR's are systematically collected by the pharmaceutical company and reported to regulatory authorities or are directly reported to regulatory authorities by the physician. At present 40 countries are participating and the national centres send their ADR reports to the WHO centre. The information on an individual case transmitted to the centre is the case identification, the patient data, the description of the adverse reaction and information about administered drugs. Finally, the background data and comments by the national centre are given.

The WHO centre, at present, apply a procedure whereby after every two ADR's a signal is given. To date no extensive statistical evaluation of the method used in the WHO centre has been made.

The 1992 revised Draft Guideline for Post marketing Reporting of Adverse Drug Reactions was published by the Office of Epidemiology and Biostatistics, Centre for Drug Evaluation and Research, Food and Drug Administration (FDA) in the USA. This guideline includes the definition of "reporting interval", "comparison interval" and the statistical procedure for determination of whether a significant increase in the reporting of ADR's has occurred.

What we require is to quantify the number of ADR's and the population at risk. In the literature much attention is paid to the measure of the number of ADR's. How do we count ADR's? Norwood and Sampson (1988) propose that "ADR's be counted in the time interval in which they were reported rather than the time interval in which they occurred if these periods differ. This avoids the bias that would occur in the current period because of ADR's that have actually occurred but have not yet been reported. In other words, in doing so, we do not bias the results by a comparison of the current period rate, with a necessarily short follow-up period, to a historical rate inflated by a longer follow-up."

Several aspects of drug usage have potential importance in the determination of ADR incidence. Some adverse reactions occur primarily upon initial use of a drug; some occur only on subsequent drug usage; while others only occur after long-term use. The appropriate measure of drug usage to monitor a particular ADR depends on knowledge of its relationship to drug use.

In most cases we have no direct measure of any of these parameters of drug usage, although a variety of indirect measures are available. Manufacturers' sales volume is often used to monitor ADR's. The use of estimated numbers of new prescriptions available from those ADR's that occur with first time use is one other measure. Attainment of suitable indirect measures of subsequent usage of long-term use poses even greater difficulties.

It is important to classify the ADR's according to labelled categories and that a historical database exists with which one can compare current rates.

3.1. Data Quality

The data we need for the surveillance are the number of ADR's and the population at risk. The biases which can influence the number of ADR's through the spontaneous reporting system according to Sachs and Bortnichak (1986) are:

- Length of time a drug has been on the market. In general an increase of ADR reporting can be observed in the first few years during a drug has been on the market. Thereafter a decrease is detectable regardless of the sales volume.

- Country specific differences. Spontaneous reporting of ADR's is affected by differences in the reporting procedures and differences in the compliance to a reporting system.

- Reporting environment. Several extraneous factors including scientific and other publications concerning a certain drug with ADR lead to an increased level of publicity

and an increase in the spontaneous report of ADR's. Furthermore, changes in the reporting system itself can affect the number of ADR reports.

- Extent of interaction between physicians and manufacturer's sales force. The field sales force comprises the manufacturer's representatives who have closest contact with the physicians. and a high percentage of ADR's are directly reported to them. The amount of time they spend with physicians or in giving special training may lead to an increase in ADR reports.

Spontaneous reporting schemes have significant limitations as a consequence of "underreporting" of suspected reactions. Therefore the spontaneous reporting system should not be used to estimate the number of ADR's but used to detect rare or delayed adverse effects. It is best suited as a signal generator i.e. an early warning system for ADR. However if we want, for some reason, to estimate the number of ADR, we need to estimate the amount of "underreporting". This issue is handled by Haramburu (1993) in his article "Estimation of underreporting".

In measuring the population at risk it is assumed that the patients, have actually been treated with the drug. The bias we will get if we use the manufacturer's sales volume or number of prescriptions as the measure of population at risk is that all drugs sold or prescribed have not necessarily been consumed. This may lead to an over-optimistic measure.

One way to solve the problem of "underreporting" and to obtain better measures for the population at risk is suggested by Inman (1981) by Prescription Event Monitoring (PEM) system. The aims of this system may be summarized according to Waller (1991) as follows:

- To monitor all new chemical entities that are used in general practice in the United Kingdom.
- To study a cohort of at least 10 000 patients.
- To generate hypotheses about adverse reactions.
- To test specific hypotheses.

- To measure the frequency of adverse events after exposure to a new drug.
- To complement the existing spontaneous reporting system.

When a pharmacist dispenses a drug for a National Health Service patient, he or she must send the prescription to the Prescription Pricing Authority if he or she wishes to be reimbursed the cost. All prescriptions for a drug are photocopied and sent to the Drug Safety Research Unit (DSRU). This is done until a cohort of sufficient size has been identified. The copies are processed so as to identify the patient, doctor and the fact that drug exposure has occurred. The next step is to find out what happened after the drug was prescribed. This is done by means of a standard questionnaire that is mailed to the prescribing doctor. The PEM system is limited in time which implies that the adverse reactions that occurs a long time after the use of the drug will never be discovered. Since it is a cohort we could use hypothesis and standard statistical procedures. The situation described in Section 4 is intended for surveillance and is not intended for this system.

The spontaneous reporting system is the one most used today. In the future one may expect higher quality in the data collected for the purpose of PMS since considerable effort is placed on this issue among medical researchers. It is therefore important to develop statistical methods with nice properties that can generate signals for ADR's.

3.2. FDA Guidelines

In the FDA guidelines we can find the definition of "reporting interval", "comparison interval" and a statistical procedure for determination of whether a significant increase of ADE's (ADR) has occurred. For drugs marketed for less than three years, the "reporting interval" is the most recent quarter of marketing and the "comparison interval" is the interval from initial marketing until the day preceding the "reporting interval". For drugs marketed for three years or longer, the "reporting interval" is the most recent year of marketing and the "comparison interval" is the year preceding the "reporting interval". This distinction is made due to the fact that the practicing physicians are more alert during the first years the drug is released on the market and as a consequence more ADR's are reported. When the drug is marketed for three years or longer it is difficult with the recommended procedure to detect an increase of ADR's when there is a small increase annually, which is discussed in Section 9.

Let

$Y(i)$= the random variable of the number of ADR's reported in quarter or year *i*.

$f(i)$= the estimate of sales volume in quarter or year *i*, concidered as non-random.

When the drugs are marketed for less than three years the hypothesis is:

$$H_0: E\left(\frac{Y(t)}{f(t)}\right) = E\left(\frac{\sum_{i=1}^{t-1} Y(i)}{\sum_{i=1}^{t-1} f(i)}\right) \text{ and the alternative } H_1: E\left(\frac{Y(t)}{f(t)}\right) > E\left(\frac{\sum_{i=1}^{t-1} Y(i)}{\sum_{i=1}^{t-1} f(i)}\right)$$

The transformation suggested by this hypothesis is (see Section 9.1 for more details):

$$Z_A(t) = \frac{Y(t)}{f(t)} - \frac{\sum_{i=1}^{t-1} Y(i)}{\sum_{i=1}^{t-1} f(i)}, \text{ which implies that } H_0 \text{ is equivalent to } E[Z(t)]=0$$

A similar test statistic was suggested by Finney (1973).

When the drugs are marketed for three years or longer the hypothesis is: $H_0: E\left(\frac{Y(t)}{f(t)}\right) = E\left(\frac{Y(t-1)}{f(t-1)}\right)$ and the alternative $H_1: E\left(\frac{Y(t)}{f(t)}\right) > E\left(\frac{Y(t-1)}{f(t-1)}\right)$

The transformation suggested for this situation is (see Section 9.1 for more details):

$Z_B(t) = \frac{Y(t)}{f(t)} - \frac{Y(t-1)}{f(t-1)}$, which implies that H_0 is equivalent to $E[Z(t)]=0$

In the article by Tsong (1992) a description of 6 statistical procedures are given for testing the above hypothesis. One of these is the statistical procedure provided in the 1991 version of the FDA Draft Guideline for Reporting Adverse Drug Reactions. This procedure is based on the normal approximation test proposed by Norwood and Sampson (1988). The letter proposed the normal approximation of the binomial test through Fisher's Exact test under the assumption that the sales volume is sufficiently large. It can also be obtained as the upper 0.05 critical level of the regular normal approximation test for comparing two proportions.

By use of these methods we ignore the fact that repeated decisions have to be taken. When we take this into account, this is equivalent to the Shewhart method proposed by Shewhart (1931). In this method an alarm is triggered for the first time t for which $|X_t - \mu^0| > g$ where g is a constant and for the definition of μ^0 see Section 5.

It will be assumed that the variables Y(t) are independent and approximately normally distributed in accordance with other papers in this area.

3.3. Some Other Methods

Most papers on post marketing surveillance concentrate on data quality. However, some statistical articles have handled some suggestions on statistical methods within this area.

Mandel (1976) modified the method corresponding to the statistica $Z_A(t)$ to make it more sensitive to trends. This test statistic compares the last observation with the sum of the three preceding observations.

Moussa (1978) and Prauss et al (1993) suggests the use of the cumulative sum scores (CUSUM) technique introduced initially by Page (1954). The CUSUM method is often used in industrial process control and later proved to be a useful method for the monitoring of birth defects or other rare diseases. We assume that the background incidence level k_0 of reported ADR's per unit of sales volume is known, together with its standard deviation. Let $k_1 > k_0$ be an increased incidence level which is considered to be the alert or rejection level. The CUSUM score S_i for the period i is then defined in relation to its value at i-1 by

$$S_i = \max\left(0, S_{i-1} + \frac{y_i}{s_i} - k_r\right) \text{ for } i > 0, \text{ where } S_0 = 0$$

Here k_r, where $k_0 < k_r < k_1$, is called the reference level and is usually taken as the mean of k_0 and k_1. Whenever S_i exceeds a certain detection boundary h, an increased incidence is suspected. The equations defining this detection boundary h cannot be solved explicitly but tables for practical use are available.

Moussa (1978) also adresses the case of subpopulations. The method proposed issues a local signal in the jth subpopulation, if $S_{ij} > h_j$, and a general signal is triggered on a national level if $S_{ij} > h_j$ for all j=1, 2, ..,s.

4. SURVEILLANCE IN GENERAL

Surveillance can be viewed as continual observation in time where the goal is to detect a change in the underlying process as soon as possible after it has occured. In terms of statistical inference we have a situation with four characteristics:

- The number of observations is increasing.
- Decisions must be made gradually.
- The "catastrophe" under surveillance might occur at any time.
- The hypotheses undergo gradual change since at time t we are interested in the hypothesis that no change has occurred before t, while at time t+1 change before t+1 is of interest.

This situation arises in different areas of medicine. Each surveillance situation is unique and needs to be evaluated in the light of consequences of an alarm, the necessary timeliness, how data is measured, if an individual or a population is under surveillance etc. Examples of different situations are surveillance of fetal heart rate during labour, regular health controls, post marketing surveillance of drugs and surveillance of congenital malformations. When conducting the statistical analysis the fact that repeated decisions will be made and that no fixed hypothesis is of special interest is often ignored. Several methods which take into consideration the sequential structure of the surveillance situation are available e.g. CUSUM or Shewhart. Optimality in surveillance, as well as evaluation of the performance of surveillance methods has been discussed in the literature, cf. Frisén (1994).

5. STATISTICAL MODEL

The variable under surveillance is $X = \{X(t): t = 1,2,..\}$, where the observation at time t is X(t). X(t) is considered univariate, in the sence that only one adverse reaction at a time is considered. For surveillance in the multivariate case see Wessman (1995).

Let us assume that the sequence X(t) consists of independent and normally distributed observations with constant variance assumed to be one, without loss of generality,. If the adverse reaction is very rare then a discrete distribution would be more appropriate. A discrete distribution is studied by Arnkeldottir (1995). The variable X(t) at time point t could, for example, be based on a recursive residual such as

$$\frac{Y(t)}{f(t)} - \frac{\sum_{i=1}^{t-1} Y(i)}{\sum_{i=1}^{t-1} f(i)}, \text{ or } \frac{Y(t)}{f(t)} - \frac{Y(t-1)}{f(t-1)}$$

where Y(i) and f(i) are observed variables (see Section 3.2). The random process which determines the state of the system is denoted $\mu = \{\mu(t): t = 1,2...\}$. $E[X(t)] = \mu(t)$.

The time point of a change is denoted τ. The critical event of interest at decision time t is denoted C(t)= $\{\tau \le t\}$ and the alternative D(t)= $\{\tau > t\}$. Two types of critical events will be studied below. For the first type it is assumed that if a change in the process occurs, it is a linear change in time t. If the change occurs at time τ we have

$\{\mu(t) = (t - \tau + 1)b\}$ if $t \ge \tau$, $b > 0$

and

$\{\mu(t) = \mu^0\}$ if $t < \tau$.

The second type I have used as a comparison for the measures of performance is when the critical event C is a sudden change to another constant level. That is:

$\{\mu(t) = \mu^1\}$ if $t \ge \tau$

and

$\{\mu(t) = \mu^0\}$ if $t < \tau$.

It is assumed that the level μ^0 is known and without loss of generality it is assumed that this value is zero. When X(t) is a recursive residual then E[X(t)]=0, that is μ^0=0, see Section 9.1.

6. MEASURES OF PERFORMANCE IN POST MARKETING SURVEILLANCE

It is essential to evaluate different statistical methods in a way that takes into account the dependence on the length of the period of surveillance and the time point where the change occurs. The fact that the change can be gradual will also influence these measures of performance. This is not the case with the usual measures, such as the significance level and power of a test's performance that need to be generalized. Measures which take the above into account and chosen in this work are presented in Sections 6.1-6.4.

How a method of surveillance performs is dependent on the time τ between the start of the surveillance and the time of the change. We can express the performance as a function of τ, as in Frisén (1992). Sometimes we want to have a single criterion of optimality in order to get an index which is independent of τ. One way is to consider a probability distribution of τ and use summarizing measures over this distribution.

In quality control optimality is often defined as minimal ARL^1 / ARL^0, see the definition in the next section. One other suggested way is to find a method that minimizes ARL^1 for fixed ARL^0 . Drawbacks with these two criteria are discussed by Frisén (1994).

Two other important optimality criteria are:

a) Find the method which maximizes the detection probability for a fixed false alarm probability for each decision time.

b) Find the method which maximizes the expected utility. One important specification of utility is given by Girshick and Rubin (1952) and Shiryaev (1963). They treat the case where the gain of an alarm is a linear function of the difference $\tau - t_A$ between the time of the change and the time of an alarm. The loss of a false alarm is a function of the same difference.

The method suggested in Section 8 fulfills these criteria.

6.1. The Average Run Length

A measure often used in industrial process control is the average run length (ARL) until an alarm is given. The Run length (RL) is defined as the number of time points from the start of the surveillance until we get an alarm.. The ARL^0 is the average number of runs until an alarm when there is no change in the system under surveillance. The average run length for the alternative hypothesis, ARL^1, is the average number of decisions that must be taken to detect a true change in the system that occurred at the same time the inspection started, that is τ=0. These measures has the obvious disadvantages of being sensitive to skew run length distributions. One other disadvantage is that, when the true change in the system occurs at another time point than τ=0, which is most realistic in continual surveillance, we have little use of ARL^1 as defined above.

The average run length is the expected value of the run length (RL). That is:

$$ARL = E[RL] = \sum_{i=1}^{\infty} i \cdot P(RL = i)$$

In Section 7.1 this measure will be calculated for the Shewhart method in the case where the critical event, is a linear increase is compared to the case of a sudden shift to a constant level μ^1.

ARL is used here in order to calculate the coefficient b in the linear increase. Let us first decide a constant level μ^1 for the sudden shift and calculate ARL^0 and ARL^1. Then we want to find a value for the coefficient b in the case of a linear increase such that the ARL^0 and ARL^1 are the same as in the case of a shift.

6.2. The False Alarm Probability

One measure of performance is the probability of a false alarm at time point t, conditioned that it has not been an earlier alarm. This measure is important for post marketing

surveillance, since an alarm will indicate an increase of adverse reactions. If this signal is false considerable effort will be put on further research with high costs as a consequence.

Let us define the conditional probability of a false alarm at the time point t as

$\alpha(t) = P(RL = t | D(t), RL \geq t) = P(A(t) | D(t), A_{t-1}^c)$ where $A_t = A(1) \cup A(2) \cup ... \cup A(t)$ and A_t^c its complement. A(t) is the event: "alarm at decision time t".

The probability of a false alarm no later than at time t from the start is also of interest. A cumulative false alarm probability can then be defined by

$\alpha_t = P(RL \leq t | D(t)) = P(A_t | D_t) = 1 - P(A_t^c | D_t)$ where $D_t = D(1) \cap D(2) \cap ... \cap D(t)$

Then $\alpha_t = 1 - P(A^c(1) | D_t, A_0^c) ... P(A^c(t) | D_t, A_{t-1}^c)$

When $D(t) = \{\tau > t\}$ it can be simplified to

$\alpha_t = 1 - P(A^c(1) | D(1), A_0^c) \cdot ... \cdot P(A^c(t) | D(t), A_{t-1}^c)$

In Section 7.2 this measure will be calculated for the Shewhart method for t=1,...,100(see Figure 7.2.1) .

6.3. The Probability of Successful Detection

The distance between the change and the alarm is important to calculate for various methods. In the PMS area it is crucial to quantify this probability. If we have a serious adverse reaction we have to detect the change within some time period in order to avoid suffering or death. Thus it is important to calculate this probability for different time units d.

In these cases the probability to get an alarm within d time units from the change has occurred, given that there was no alarm before the change, that is the probability of successful detection (PSD), is of interest. The PSD is a function of the time distance d, the time of the change t and $\mu(t)$.

$PSD(d, t, \mu(t)) = P(RL \leq t + d | RL \geq t) = P(A_{t+d-1} | A_{t-1}^c, \tau = t)$

6.4. The Predictive Value

The predictive value (PV) is an important criterion of evaluation used by Frisén (1992), Frisén and Åkermo (1994) and Frisén and Cassel (1994). The probability that a change has occurred, given that there is an alarm, is even more important for the PMS area. The trust you should have in an alarm is important for the decision-maker in order to judge which decision is appropriate.

This criterion measures how strong an alarm is as an indication of a change. It can be interpreted as the relative frequency of motivated alarms among all alarms at t, and is defined as:

$$PV(t) = P(\tau \le t | RL = t) = P(C(t) | A(t), A_{t-1}^c) = \frac{PMA(t)}{PMA(t) + PFA(t)}$$

where PMA and PFA are defined below. The unconditional probability of a false alarm at t is

$$PFA(t) = P(A(t) \cap \tau > t) = (\alpha_t - \alpha_{t-1}) \cdot \prod_{j=1}^{t} (1 - \gamma_j)$$

where γ the incidence, is the probability that the stochastic time τ for the change takes the value t, given that there has been no change before t .

The probability of a motivated alarm at t is:

$$PMA(t) = P(A(t) \cap \tau \le t) = \sum_{j=1}^{t} \left(\prod_{i=1}^{j-1} (1 - \gamma_i) \right) \cdot \gamma_j \cdot P(A(t) \cap A_{t-1}^c | \tau = t)$$

Let us assume that this incidence is constant. Then

$$PFA(t) = (\alpha_t - \alpha_{t-1}) \cdot (1 - \gamma)^t \text{ and } PMA(t) = \sum_{j=1}^{t} (1 - \gamma)^{j-1} \cdot \gamma \cdot P(A(t) \cap A_{t-1}^c | \tau = t)$$

In order to get easy interpretations of alarms we could construct methods with a constant predicted value. This is studied by Åkemo, (1994).

7. RESULTS ON THE PERFORMANCE OF THE SHEWHART METHOD, GRADUAL CHANGES AND SUDDEN SHIFTS

As was discussed in Section 3.2, the Shewhart method is of special relevance for the methods used in post marketing surveillance, if we take into account the fact that repeated decisions have to be taken. In Section 9 these results will be used as a base for the discussions about the methods suggested by the FDA.

The fact that the change can be gradual will influence the measures of performance. I have calculated these measures under the assumption of a linear increase as described in Section 5. As a comparison I also calculated these measures for a sudden change to another constant level μ^1.

7.1. The Average Run Length

Two types of critical events for the statistical model are considered, described in Section 5. One is when the critical event C is a sudden shift and one is when we have a linear increase in time. The ARL^0 is the same for both these situations since no change has occurred. For the Shewart method, with the limit g, the run length (RL) follows a geometric distribution with the parameter $P(X(t) > g) = (1 - \Phi(g))$.

$$ARL^0 = \sum_{i=1}^{\infty} i \cdot P(RL = i) = \sum_{i=1}^{\infty} i \cdot (\Phi(g))^{i-1} \cdot P(1 - \Phi(g)) = \frac{1}{(1 - \Phi(g))}$$

where $\Phi(g)$ is the standardized normal probability distribution function.

For the average run length for the alternative hypothesis, ARL^1, we have for the case of a sudden shift to a constant level μ^1 and the parameter $P((X(t) - \mu^1) > g - \mu^1)$ $= (1 - \Phi(g - \mu^1))$ for the geometric distribution:

$$ARL^1 = \sum_{i=1}^{\infty} i \cdot P(RL = i) = \sum_{i=1}^{\infty} i \cdot (\Phi(g - \mu^1))^{i-1} \cdot P(1 - \Phi(g - \mu^1)) = \frac{1}{(1 - \Phi(g - \mu^1))}$$

For the case of a linear change in time we have different parameters for every i in the distribution for the RL. That is:

$$P(RL=i)=\prod_{k=1}^{i-1}\left(1-P\left(X(k)-b\cdot k>g-b\cdot k\right)\right)\cdot P\left(X(k)-b\cdot i>g-b\cdot i\right)=$$

$$\prod_{k=1}^{i-1}\left(\Phi(g-b\cdot k)\right)\cdot\left(1-\Phi(g-b\cdot i)\right)$$

This means that:

$$ARL^1=\sum_{i=1}^{\infty}i\cdot P(RL=i)=\sum_{i=1}^{\infty}i\cdot\prod_{k=1}^{i-1}\left(\Phi(g-b\cdot k)\right)\cdot\left(1-\Phi(g-b\cdot i)\right)$$

In order to compare these two situations above I have chosen $\mu^1=1$ and g=1.64 and obtained for ARL^0=20 and ARL^1=3.8 in case of a sudden shift,. In order to get a comparable situation for the linear case let g=1.64, ARL^0=20 and ARL^1=3.8 also in the case of critical event is a linear increase. This will give us the value b=0.33 which will be used in the calculations.

7.2. The False Alarm Probability

The probability of a false alarm and the cumulative false alarm probability is the same for the two situations mentioned in Section 7.1.

The false alarm probabilities are the same for all t. The limit g=1.64 gives:

$$\alpha(t)=0.05$$

As we could see in Section 6.2 the cumulative false alarm probability could be simplified to:

$$\alpha_t=1-P\left(A^c(1)\middle|D(1),A_0^c\right)\cdot\ldots\cdot P\left(A^c(t)\middle|D(t),A_{t-1}^c\right)$$

The variables $X(i)$ are normally distributed and independent as specified in Section 5. For the Shewhart method that implies that all events A(t) are also independent. Thus:

$$\alpha_t=1-(\Phi(g))^t$$

where $\Phi(g)$ is the standardized normal probability distribution function.

In Figure 7.2.1 we could see the probability of a false alarm for g=1.64 and t=1,...,100.

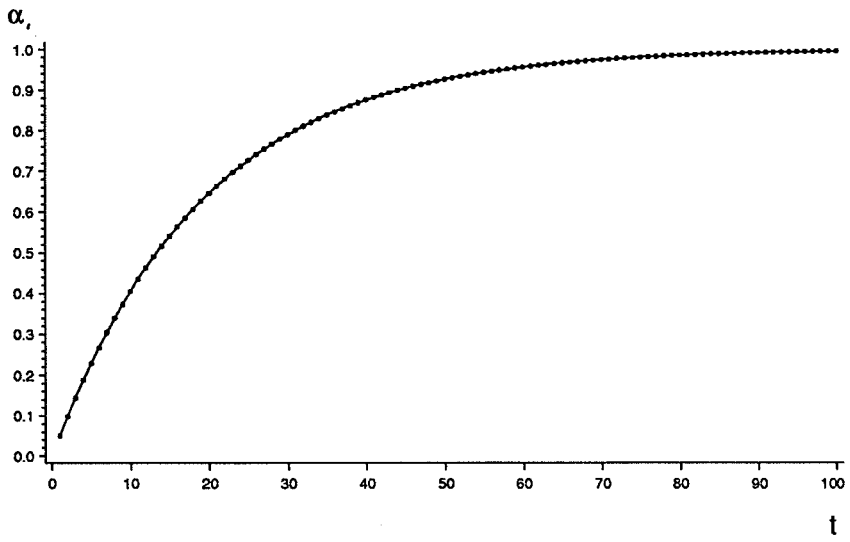

Figure 7.2.1: The cumulative probability of false alarm where g=1.64 and t=1,..,100

7.3. The Probability of Successful Detection

For the Shewhart method and when we study the case of the critical event being a sudden shift the probability of successful detection is:

$$PSD(d,t,\mu(t)) = 1-\left(P\left(A^c(t)\middle|\mu' = \mu^1\right)\cdot P\left(A^c(t+1)\middle|\mu' = \mu^1\right)\cdot\ldots\cdot P\left(A^c(t+d-1)\middle|\mu' = \mu^1\right)\right) =$$

$$= 1-\left(\Phi\left(g-\mu^1\right)\right)^d$$

If we look at the case when the change is linear we have:

$$PSD(d,t,\mu(t)) = 1-\left(P\left(A^c(t)\middle|\mu' = b\right)\cdot P\left(A^c(t+1)\middle|\mu' = 2b\right)\cdot\ldots\cdot P\left(A^c(t+d-1)\middle|\mu' = b(t+d)\right)\right) =$$

$$= 1-\left(\Phi(g-b)\cdot\Phi(g-2b)\cdot\ldots\cdot\Phi(g-db)\right)$$

For g=1.64 and b=0.33, μ^1=1 in Figure 7.3.1 we could see a comparison of these two cases above.

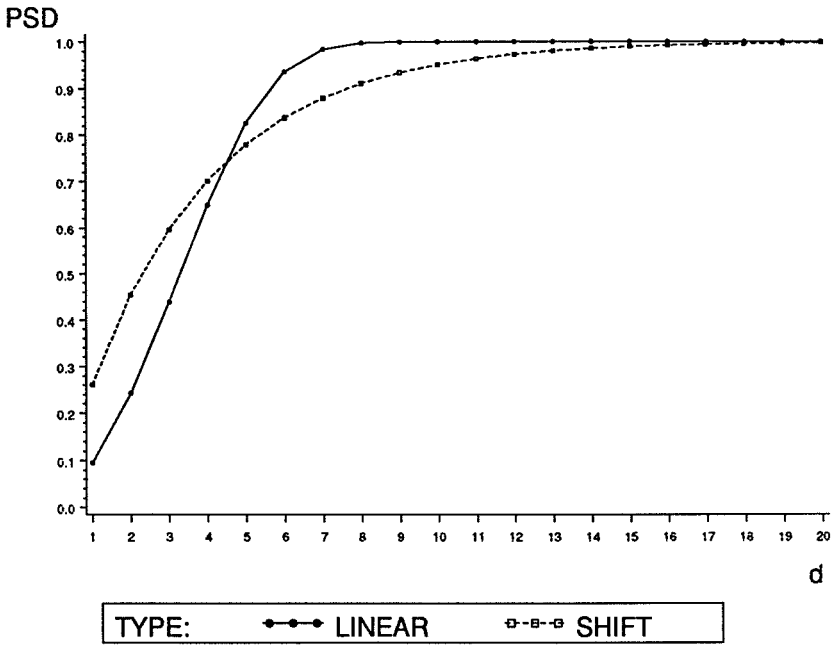

Figure 7.3.1: The probability for successful detection for g=1.64 in case of a sudden shift to a level μ^1=1 and a linear increase with b=0.33.

7.4. The Predictive Value

For the Shewhart method and when we study the case of the critical event being a sudden shift the probability for false alarm is:

$$PFA(t) = (1-\gamma)^{t} \cdot (\Phi(g))^{t-1} \cdot (1-\Phi(g))$$

and the probability for motivated alarm

$$PMA(t) = \sum_{j=1}^{t} (1-\gamma)^{j-1} \cdot \gamma \cdot (\Phi(g))^{j-1} \cdot (\Phi(g-\mu^1))^{t-j} \cdot (1-\Phi(g-\mu^1))$$

The probability of false alarm for the Shewhart method when a linear increase is of interest we have the same probability of a false alarm as in the case of a sudden shift.
The probability of motivated alarm in this case is:

$$PMA(t) = \sum_{j=1}^{t} (1-\gamma)^{j-1} \cdot \gamma \cdot (\Phi(g))^{j-1} \cdot \prod_{i=j}^{t-1} (\Phi(g-b\cdot(i-j+1))) \cdot (\Phi(-g+b\cdot(t-j+1)))$$

In Figure 7.4.1 with g=1.64, b=0.33, μ^1=1 and γ=0.01 a comparison of these two cases is made for different values of t. In Figure 7.4.2 a comparison of the two cases above is made for different values of t for the same values of the parameters as above, except for the incidence γ=0.1.

As we can see the choice of the incidence γ is very crucial for the predicted value. A medical knowledge about the drug under surveillance is important.

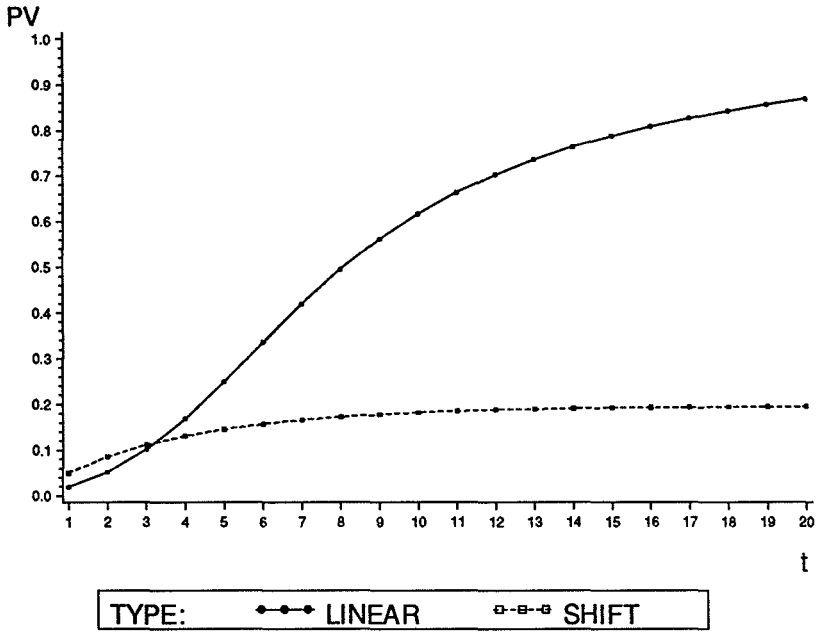

Figure 7.4.1: The predictive value for g=1.64 and the incidence γ=0.01 in case of a sudden shift to a level μ^1=1 and a linear increase with b=0.33.

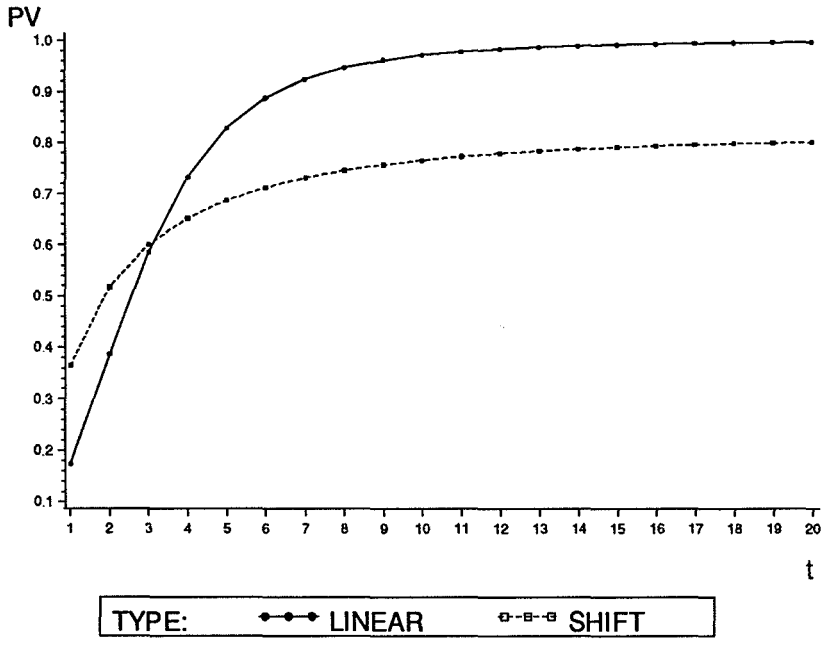

Figure 7.4.2: The predictive value for g=1.64 and the incidence γ=0.1 in case of a sudden shift to a level μ^1=1 and a linear increase with b=0.33.

8. THE LIKELIHOOD RATIO METHOD

The Likelihood Ratio method which is optimal according to conditions a) and b) as described in Section 6, to detect a change, is constructed by Frisén and de Maré (1991). To meet optimality condition a) in Section 6 we will have an alarm if the alarm function $p(x_t) > G_t$ where $p(x_t)$ is the likelihood ratio.

This method is also exemplified for the normal distribution for the shift case. That is when we assume unit variance and $\mu^0 = 0$ as described in Section 5:

$$p(x_t) = \frac{\exp\left(-(t+1)\cdot(\mu^1)^2/2\right)}{P(\tau \le t)} \sum_{k=1}^{t} P(\tau = k)\cdot \exp\left((\mu^1)^2 \cdot \frac{1}{2}\cdot k\right)\cdot \exp\left(\mu^1 \cdot \sum_{u=k}^{t} X(u)\right)$$

The method derived below is the likelihood ratio method but it is optimal to detect a linear increase (see Section 5).

8.1. The Likelihood Ratio method for gradual changes

Let us consider the situation where all changes before time t are considered and τ is a generalized random variable with known density P(τ=k) as described in Section 5. That is:

$C(t) = \{\tau \le t\}$

and the alternative

$D(t) = \{\tau > t\}$

We obtain the alarm function

$$p(x_t) = \sum_{k=1}^{t} \frac{P(\tau = k)}{P(\tau \le t)} \prod_{u=k}^{t} \frac{f\{X(u)|\mu(u) = (u-k+1)b\}}{f\{X(u)|\mu(u) = \mu^0\}}$$

If we assume a normal distribution with unit variance and $\mu^0 = 0$ as described in Section 5 we have

$$p(x_t) = \sum_{k=1}^{t} \frac{P(\tau = k)}{P(\tau \le t)} \prod_{u=k}^{t} \frac{\exp(-\frac{1}{2}(X(u) - b(u-k+1))^2}{\exp(-\frac{1}{2}(X(u))^2} =$$

$$\sum_{k=1}^{t} \frac{P(\tau = k)}{P(\tau \le t)} \prod_{u=k}^{t} \exp\left(-\frac{1}{2}(-2b(u-k+1)X(u) + b^2(u-k+1)^2 \right) =$$

$$= \frac{1}{P(\tau \le t)} \sum_{k=1}^{t} P(\tau = k) \exp\left(\sum_{u=k}^{t} b(u-k+1)X(u) - \frac{1}{2} b^2 \sum_{u=k}^{t} (u-k+1)^2 \right) =$$

$$\frac{1}{P(\tau \le t)} \sum_{k=1}^{t} P(\tau = k) \exp\left(\sum_{u=k}^{t} b(u-k+1)X(u) - \frac{b^2}{12} \cdot \left((t-k+1) \cdot (t-k+2) \cdot (2 \cdot (t-k)+3) \right) \right)$$

Let us factorise $p(x_t)$ in:

$$p(x_t) = g(t) p_t(x_t)$$

where

$$g(t) = \frac{1}{P(\tau \le t)}$$

and

$$p_t(x_t) = \sum_{k=1}^{t} P(\tau = k) \exp\left(\sum_{u=k}^{t} b(u-k+1)X(u) - \frac{b^2}{12} \cdot \left((t-k+1) \cdot (t-k+2) \cdot (2 \cdot (t-k)+3) \right) \right)$$

Let us assume that the incidence γ is constant. Then τ has a geometric distribution. That is

$$P(\tau = k) = (1-\gamma)^{k-1} \cdot \gamma$$

Then

$$p_t(x_t) = \sum_{k=1}^{t} (1-\gamma)^{k-1} \cdot \gamma \cdot \exp\left(\sum_{u=k}^{t} b(u-k+1)X(u) - \frac{b^2}{12} \cdot \left((t-k+1) \cdot (t-k+2) \cdot (2 \cdot (t-k)+3) \right) \right)$$

To achieve the optimal error probabilities according to condition a) in Section 6, an alarm should be given as soon as $p_t(x_t) > G_t$.

If we also wish to achieve maximization of the utilities as described in Section 6 condition b), it is also required that an alarm is given as soon as $p(x_t) > G$. In this case we must also consider the function g(t).

8.2. Linear approximation of the Likelihood Ratio statistic

To obtain a method which is easier to use, a linear approximation of $p_t(x_t)$ is of interest. The exponential functions will be approximated by linear functions. In order to get a good approximation of the limit of an alarm the Taylor expressions around values that might cause an alarm are important. If tight limits of an alarm are used then a Taylor expression around 0 is appropriate.

The Taylor expression around 0 is:

$$\exp\left(b\sum_{u=k}^{t}(u-k+1)X(u)\right)\approx 1+b\sum_{u=k}^{t}(u-k+1)X(u)$$

If we put this approximation in $p_t(x_t)$ we get:

$$p_t'(x_t)=\sum_{k=1}^{t}(1-\gamma)^{k-1}\cdot\gamma\cdot\exp\left(-\frac{b^2}{12}\cdot\left((t-k+1)\cdot(t-k+2)\cdot(2\cdot(t-k)+3)\right)\right)\left(1+b\sum_{u=k}^{t}(u-k+1)X(u)\right)=$$

$$\sum_{k=1}^{t}(1-\gamma)^{k-1}\cdot\gamma\cdot\exp\left(-\frac{b^2}{12}\cdot\left((t-k+1)\cdot(t-k+2)\cdot(2\cdot(t-k)+3)\right)\right)+$$

$$\sum_{k=1}^{t}(1-\gamma)^{k-1}\cdot\gamma\cdot\exp\left(-\frac{b^2}{12}\cdot\left((t-k+1)\cdot(t-k+2)\cdot(2\cdot(t-k)+3)\right)\right)\cdot b\cdot\sum_{u=k}^{t}(u-k+1)X(u)$$

The first term in the expression is not depending on X. We rewrite the second term as:

$$p_t''(x_t)=b\cdot\sum_{i=1}^{t}X(i)\sum_{j=1}^{i}(i-j+1)\cdot(1-\gamma)^{j-1}\cdot\gamma\cdot\exp\left(-\frac{b^2}{12}\cdot\left((t-j+1)\cdot(t-j+2)\cdot(2\cdot(t-j)+3)\right)\right)=$$

$b\cdot\gamma\cdot\sum_{i=1}^{t}X(i)m(i,t)$, where

$$m(i,t)=\sum_{j=1}^{i}(i-j+1)\cdot(1-\gamma)^{j-1}\cdot\exp\left(-\frac{b^2}{12}\cdot\left((t-j+1)\cdot(t-j+2)\cdot(2\cdot(t-j)+3)\right)\right)$$

If we now look at the quotient between m(2,t) and m(1,t) we get:

$$\frac{m(2,t)}{m(1,t)}=2+(1-\gamma)\cdot\exp\left(\frac{1}{2}b^2\cdot t^2\right)$$

This indicates that the larger t the larger weight we put on time point two than on time point one.

If we also calculate the quotient between m(t,t) and m(t-1,t) we get:

$$\frac{m(t,t)}{m(t-1,t)} = 1 + \frac{(1-\gamma)^{t-1} \cdot \exp\left(\frac{1}{2}b^2\right)}{\sum_{j=1}^{t-1}(t-j)\cdot(1-\gamma)^{j-1} \cdot \exp\left(-\frac{b^2}{12}\cdot(t-j+1)\cdot(t-j+2)\cdot(2\cdot(t-j)+3)\right)}$$

For b=0.33, γ=0.01 and t=2 this quotient is 3.23. For the same values for the parameters, but for t=5 this quotient is 1.97.

Let us compare it with the approximation of the Likelihood Ratio method optimal to detect a sudden shift to a constant level, suggested by Frisén (1994).
This approximation is:

$$p_t^{''}(x_t) = \gamma \cdot \sum_{i=1}^{t} x(i) \cdot m(i)$$

where

$$m(i) = \sum_{j=1}^{i}\left(\exp\left(\frac{(\mu^1)^2}{2}\right)\right)^j \cdot (1-\gamma)$$

If we, for this approximation, look at the quotient between m(2) and m(1) we get:

$$\frac{m(2)}{m(1)} = 1 + (1-\gamma)\cdot\exp\left(\frac{(\mu^1)^2}{2}\right)$$

This indicates that we put larger weight on time point two than on time point one but the difference is not so great as for the approximation of the Likelihood Ratio method optimal to detect a linear change. As we can see these weights above are independent of t.

The calculation for the quotient between m(t) and m(t-1) for this approximation is:

$$\frac{m(t)}{m(t-1)} = \frac{d^t - 1}{d^{t-1} - 1} \text{ where } d = (1-\gamma)\cdot\left(\exp\left(\frac{(\mu^1)^2}{2}\right)\right)$$

For μ^1=1, γ=0.01 and t=2 this quotient is 2.63. For the same values for the parameters, but for t=5 this quotient is 1.74.

The Shewhart method applies the weight zero for all points up to the last one where we put the weight one. Thus we could see for the examples that this method approximates the optimal one better at linear changes than at sudden shifts.

9. EVALUATION OF THE FDA RECOMMENDATIONS

9.1. The expected values for the statistics of the FDA recommendations

The transformation of the data suggested by FDA to survey adverse reactions as described in Section 3.2 are as follows:

For a drug marketed for less than three years:

$$Z_A(t) = \frac{Y(t)}{f(t)} - \frac{\sum_{i=1}^{t-1} Y(i)}{\sum_{i=1}^{t-1} f(i)}$$, where every i denotes one quarter

The variance for $Z_A(s)$ is:

$$Var[Z_A(t)] = \sigma^2 \left(\frac{1}{f^2(t)} + \frac{t-1}{\left(\sum_{i=1}^{t-1} f(i)\right)^2} \right)$$

In order to get a unit variance we have to divide Z(s) with its standard deviation. That is:

$$X_A(t) = \frac{Z_A(t)}{\sqrt{Var(Z_A(t))}}.$$

This transformation will, in the following, be called transformation A.

For a drug marketed for three years or longer:

$$Z_B(t) = \frac{Y(t)}{f(t)} - \frac{Y(t-1)}{f(t-1)}$$

where t and t-1 denotes the most recent year and the year preceding respectively.

The variance for $Z_B(t)$is:

$$Var[Z_B(t)] = \sigma^2 \cdot \left(\frac{1}{f^2(t)} + \frac{1}{f^2(t-1)} \right)$$

For the same reason as above we have to divide $Z_B(t)$ with its standard deviation. That is:

$$X_B(t) = \frac{Z_B(t)}{\sqrt{Var(Z_B(t))}}$$

This transformation will, in the following, be called transformation B.

It is of interest to evaluate these two transformations when the change occurs at time τ for the variable under surveillance Y(t) for the two situations:

I. $\{E[Y(t)] = (t - \tau + 1) \cdot b\}$ if $t \geq \tau, b > 0$

$\{E[Y(t)] = \mu^0\}$ if $t < \tau$

II. $\{E[Y(t)] = \mu^1\}$ if $t \geq \tau$

$\{E[Y(t)] = \mu^0\}$ if $t < \tau$

What happens with the transformation A when we want to detect a linear increase at time τ for E[Y(t)]? Could we also detect a linear increase with the transformation E[X(t)]? The same question arises for the situation II. If the purpose is to detect a sudden shift with E[Y(t)] would we, with the transformations suggested, detect a sudden shift or not.

We now assume, for simplicity, that the sales volumes are constant in time. Thus $f(t) \equiv f$. In all calculations below t>2.

Let us calculate the expected values for transformations A and B for the situations I and II. For transformation A the expected value for the case of no change in the process, that is when t<τ, is:

$$E[X_A(t)] = \frac{f}{\sigma \cdot \sqrt{\frac{t}{t-1}}} \cdot \left(\frac{\mu^0}{f} - \frac{(t-1) \cdot \mu^0}{(t-1) \cdot f} \right) = 0$$

The expected value for the situation I, when $t \geq \tau$, for transformation A is:

$$E[X_A(t)] = \frac{f}{\sigma \cdot \sqrt{\frac{t}{t-1}}} \cdot \left(\frac{b \cdot (t-\tau+1)}{f} - \frac{b}{(t-1) \cdot f} \cdot \sum_{i=1}^{t-\tau} i \right) = \frac{b}{2 \cdot \sigma} \cdot \left(\frac{t^2 - t - \tau^2 + \tau - 2}{\sqrt{t \cdot (t-1)}} \right) =$$

$$\frac{b}{2 \cdot \sigma} \cdot \left(\sqrt{t \cdot (t-1)} + \frac{\tau - \tau^2 - 2}{\sqrt{t \cdot (t-1)}} \right) \qquad \text{(IA)}$$

When $t \rightarrow \infty$ and τ is small we will get the approximation: $E[X_A(t)] = \frac{b}{2 \cdot \sigma} \cdot t$

If we assume unit variance this means that the regression coefficient is $E[X_A(t)] = b/2$.

In Figure 9.1.1 and 9.1.2 we could see E[Y(t)] and the transformation $E[X_A(t)]$ as in the equation IA for b=0.33, σ=1 and τ=1,5,20 .

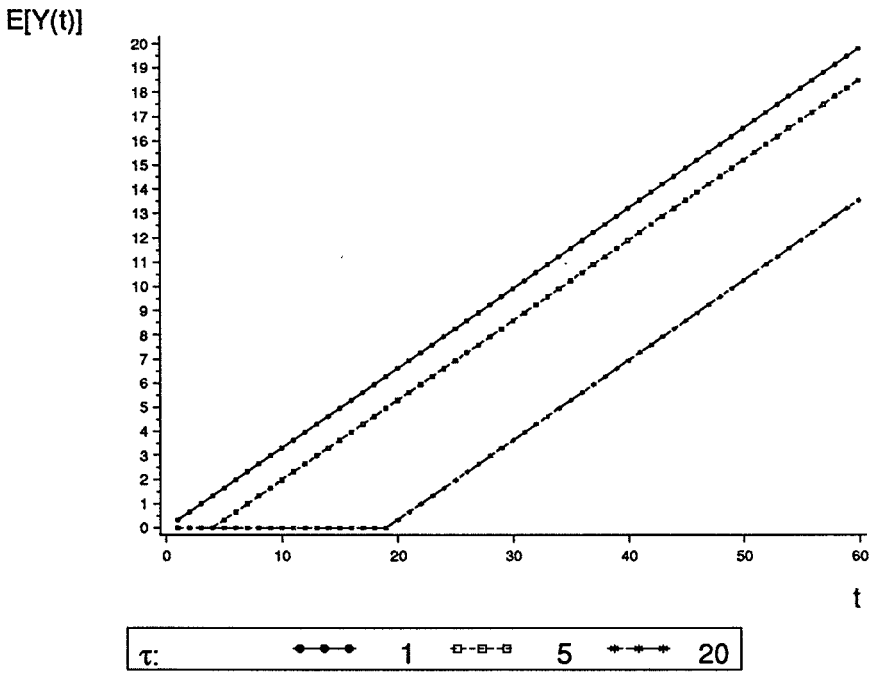

Figure 9.1.1: The expected value for the variable Y(t) when the critical event is a linear increase (b*(t-τ+1)) with b=0.33. The time τ denotes the time of the change.

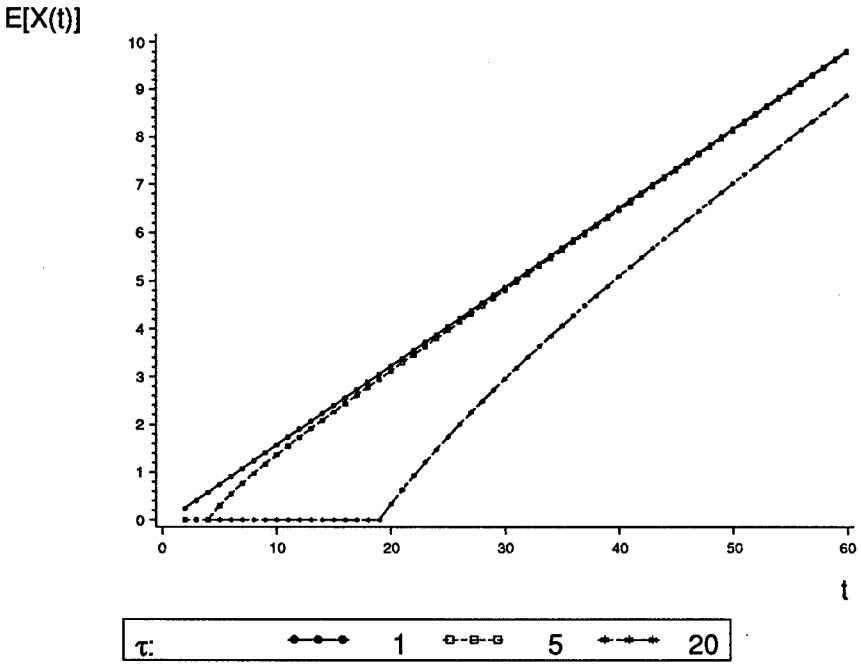

Figure 9.1.2: The expected value for the transformed variable $X_A(t)$ by transformation A when the critical event is a linear increase (b*(t-τ+1)) with b=0.33. The time τ denotes the time of the change.

The expected value for situation II, when $t \geq \tau$, for transformation A is:

$$E[X_A(t)] = \frac{f}{\sigma \cdot \sqrt{\frac{t}{t-1}}} \cdot \left(\frac{\mu^1}{f} - \frac{(t-\tau)\cdot\mu^1}{(t-1)\cdot f} \right) = \frac{\mu^1 \cdot (\tau - 1)}{\sigma \cdot \sqrt{(t-1)\cdot t}} \qquad \text{(IIA)}$$

When $t \to \infty$ and τ is fixed we will get the approximation: $E[X_A(t)] = \frac{\mu^1 \cdot (\tau - 1)}{t \cdot \sigma}$

In Figure 9.1.3 and 9.1.4 we could see the result for μ^1=1, σ=1 and τ=5,20 for E[Y(t)] and the transformation as in the equation IIA.

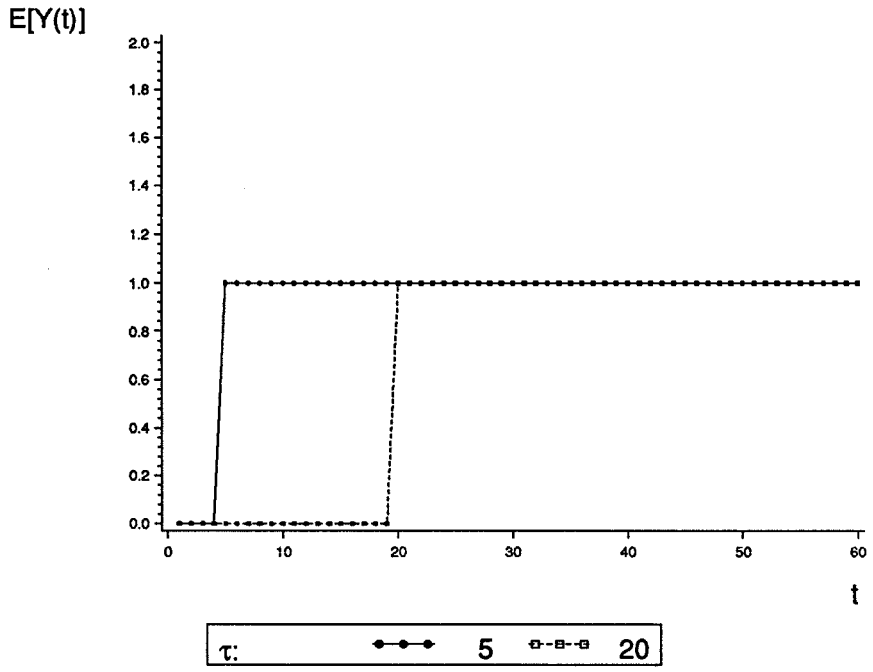

Figure 9.1.3: The expected value for the variable Y(t) when the critical event is a sudden shift to the level μ^1=1. The time τ denotes the time of the change.

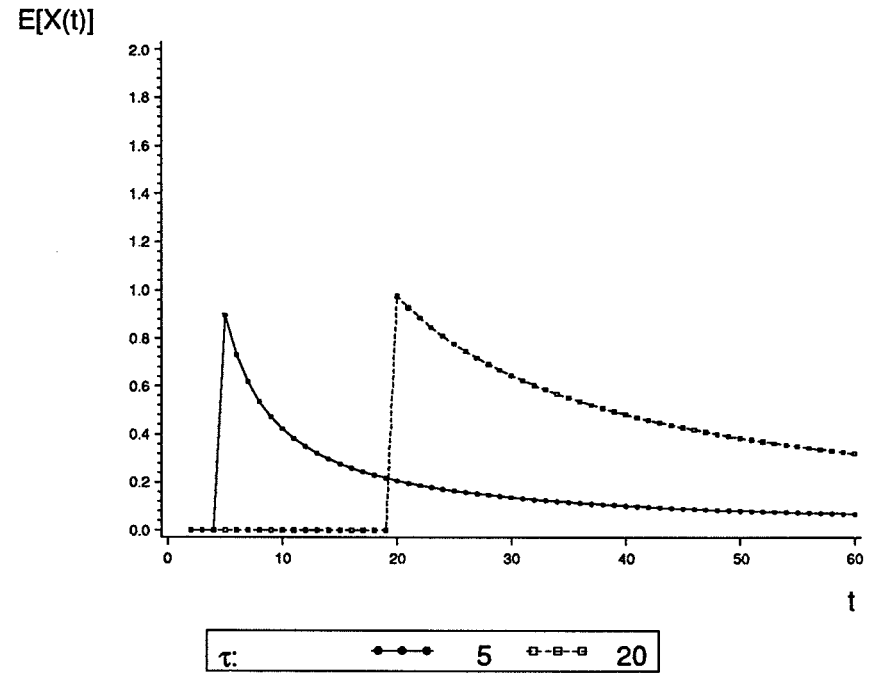

Figure 9.1.4: The expected value for the transformed variable $X_A(t)$ by transformation A when E[Y(t)] has a sudden shift to the level μ^1=1 . The time τ denotes the time of the change.

The expected value for the case of no change in the process, that is when $t < \tau$ for transformation B is:

$$E[X_B(t)] = \frac{f}{\sigma \cdot \sqrt{2}} \cdot \left(\frac{\mu^0}{f} - \frac{\mu^0}{f} \right) = 0$$

The expected value for the situation at I, when $t \geq \tau$ for transformation B is:

$$E[X_B(t)] = \frac{f}{\sigma \cdot \sqrt{2}} \cdot \left(\frac{b \cdot (t - \tau + 1)}{f} - \frac{b \cdot (t - \tau)}{f} \right) = \frac{b}{\sigma \cdot \sqrt{2}} \qquad \text{(IB)}$$

In Figure 9.1.5 we could see the result for b=0.33,σ=1 and τ=5,20 for E[Y(t)] and the transformation as in the equation IB.

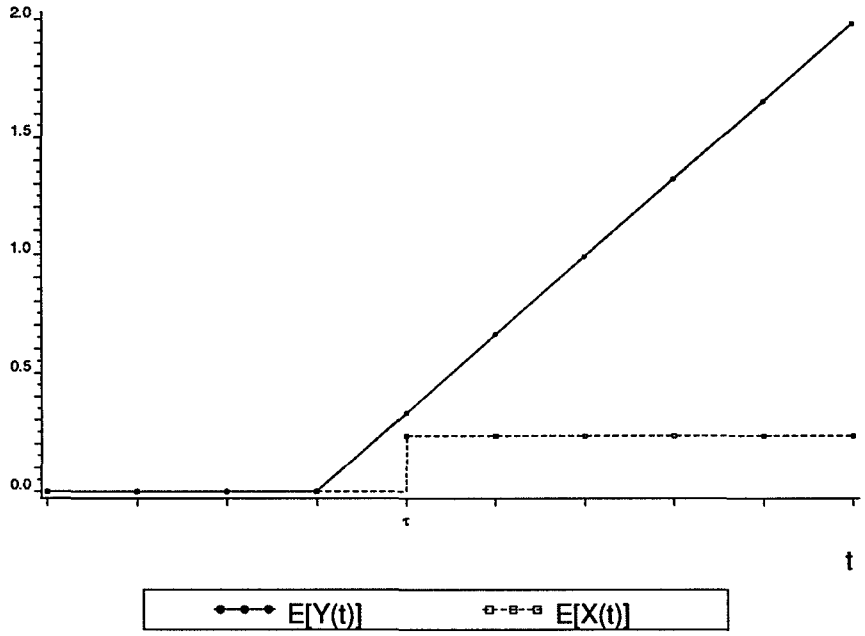

Figure 9.1.5: The expected value of the variable Y(t) when the critical event is a linear increase (b*(s-τ+1)) for b=0.33, compared to the transformed variable by the transformation B. The time τ denotes the time of the change.

For the situation II and transformation B we have to calculate the expected value for $t=\tau$ and $t>\tau$.

For $t=\tau$:

$$E[X_B(t)] = \frac{f}{\sigma \cdot \sqrt{2}} \cdot \frac{\mu^1}{f} = \frac{\mu^1}{\sigma \cdot \sqrt{2}} \qquad \text{(IIB)}$$

For $t>\tau$:

$$E[X_B(t)] = \frac{f}{\sigma \cdot \sqrt{2}} \cdot \left(\frac{\mu^1}{f} - \frac{\mu^1}{f} \right) = 0 \qquad \text{(IIB')}$$

In Figure 9.1.6 and Figure 9.1.7 we could see the result for $\mu^1=1$, $\sigma=1$ and $\tau=5,20$ for E[Y(t)] and the transformation as in the equations IIB and IIB'.

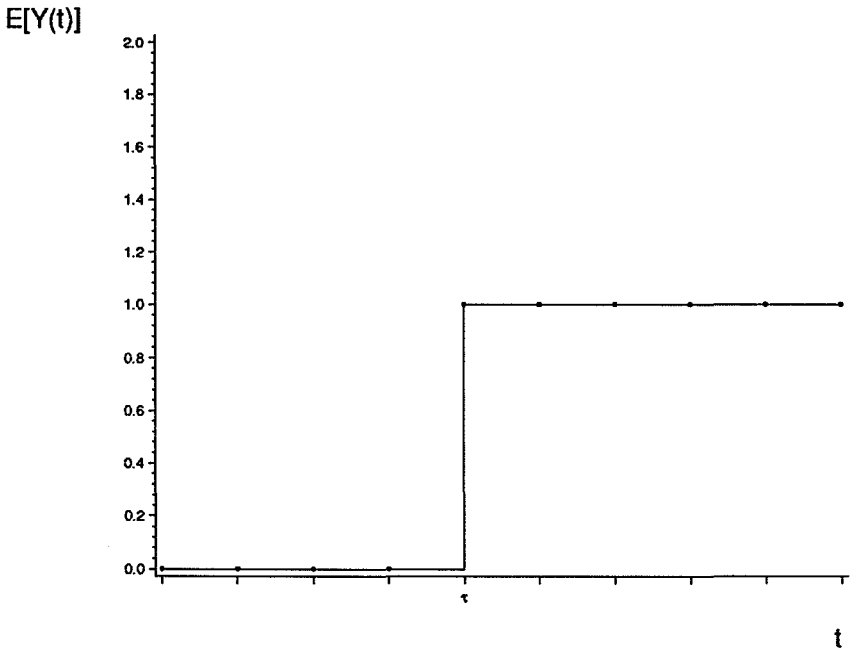

Figure 9.1.6: The expected value of the variable Y(t) when the critical event is a sudden shift to the level $\mu^1=1$. The time τ denotes the time of the change.

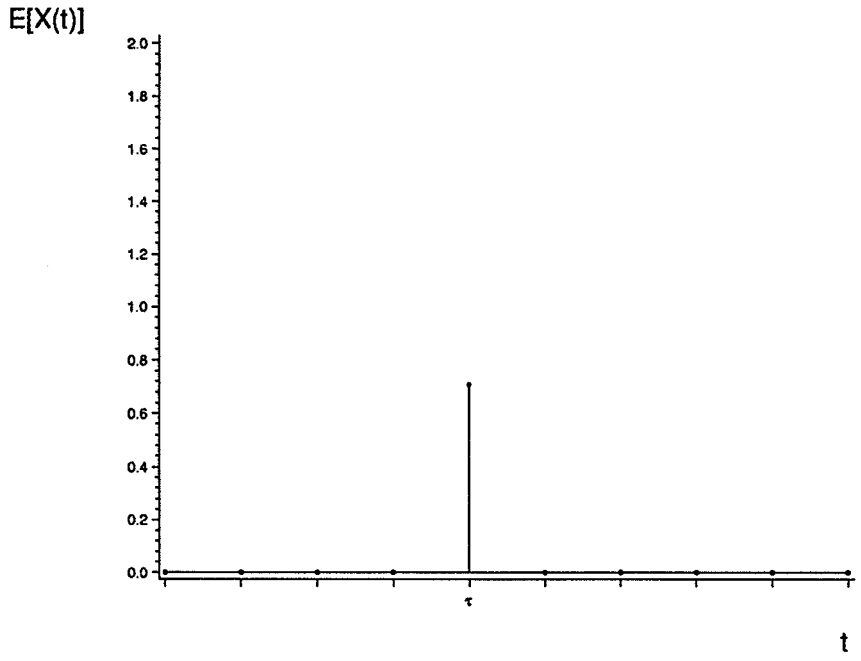

Figure 9.1.7: The expected value of the transformed variable $X_B(t)$ by the transformation B when E[Y(t)] has a sudden shift to the level μ^1=1. The time τ denotes the time of the change.

Now we need the results from Section 7. For example for transformation B we have to detect a sudden shift in E[X(t)] when there is a linear increase in E[Y(t)]. From Section 7 we could see that the predicted value is much lower for the shift situation than the linear case. However, the correlation structure described in the next section will also influence the performance.

9.2. The correlations for the statistics of the FDA recommendations

As stated in Section 3.2 we assume that Y(t) and Y(t+1) are independent and normally distributed with expected value μ(u) and variance σ^2.

For $Z_A(t)$ in Section 9.1 we will get the following dependence:

$$Cov\left(\frac{Y(t)}{f(t)}-\frac{\sum_{i=1}^{t-1}Y(i)}{\sum_{i=1}^{t-1}f(i)};\frac{Y(t+1)}{f(t+1)}-\frac{\sum_{i=1}^{t}Y(i)}{\sum_{i=1}^{t}f(i)}\right)=\frac{\sigma^2}{\sum_{i=1}^{t}f(i)}\left(-\frac{1}{f(t)}+\frac{(t-1)}{\sum_{i=1}^{t-1}f(i)}\right)=\frac{\sigma^2}{\sum_{i=1}^{t}f(i)}\left(\frac{1}{\overline{f(t-1)}}-\frac{1}{f(t)}\right)$$

where $\overline{f(t-1)}=\frac{1}{t-1}\sum_{i=1}^{t-1}f(i)$.

If all $f(i)\equiv f$, that means that the sales volume is constant for all i, then the covariance will be zero as you can see below:

$$Cov\left(\frac{Y(t)}{f}-\frac{1}{t-1}\cdot\sum_{i=1}^{t-1}\frac{Y(i)}{f};\frac{Y(t+1)}{f}-\frac{1}{t}\cdot\sum_{i=1}^{t}\frac{Y(i)}{f}\right)=-\frac{1}{t}\cdot\frac{\sigma^2}{f^2}+\frac{1}{(t-1)\cdot t}\cdot(t-1)\cdot\frac{\sigma^2}{f^2}=0$$

If $f(i)$ is not constant for every i, in order to get an approximately covariance of zero we must have a small value for

$$\frac{\sigma^2}{\sum_{i=1}^{t}f(i)}\left(\frac{1}{\overline{f(t-1)}}-\frac{1}{f(t)}\right)$$

The data in Norwood and Sampson (1988) suggest that this is usually the case.

For $Z_B(t)$, the covariance of $Z_B(t)$ and $Z_B(t+1)$ is:

$$Cov\left(\frac{Y(t)}{f}-\frac{Y(t-1)}{f};\frac{Y(t+1)}{f}-\frac{Y(t)}{f}\right)=-\frac{\sigma^2}{f^2}$$

The evaluations of methods such as the probability of a false alarm, the probability for a successful detection and the predicted value are dependent of that the correlations are zero. The measures of performance calculated are therefore valid only for transformation A.

10. CONCLUDING REMARKS

The assumption of normality made in this report may not always be appropriate for post marketing surveillance. If the adverse reaction is a rare event, which is often the case, a discrete distribution would be more appropriate to use. However, the assumption of normality is an approximation often used and the main issue in this paper was to study the consequences of gradual changes.

Gradual changes in adverse reactions is a more realistic critical event than the change to a constant level. These gradual changes could be of different arts. Here, I have chosen changes of a linear shape.

The formulas for different measures of performance are illustrated by some examples. The limit for an alarm corresponds to that of the present FDA (Food and Drug Administration) method. The values of the linear increase (b=0.33) and the sudden shift (μ^1=1) are chosen to give the same ARL^0=20 and the same ARL^1=3.8.

The cumulative probability of a false alarm is tending to grow fast. As an example see Figure 7.2.1. At the decision time 4, the false alarm probability is 0.20. When the time lapse between every decision time is a quarter, this corresponds to one year of surveillance. If a more efficient and faster way to gather ADR's is developed ,say the decision time is every month, then the false alarm probability after a year is 0.40. This means that we reach a rather large false alarm probability after just a few decision times. This kind of information is necessary for a realistic view on a surveillance system.

The probability of successful detection (PSD) for the Shewart method is slightly higher for the critical event of a sudden shift to a level μ^1=1 than the linear increase with b=0.33 for d=1,...,4. For d=5,...,14 the PSD is higher for the critical event of a linear increase with b=0.33 than of a sudden shift to a level μ^1=1. This demonstrates differences in the power to detect changes which are not disclosed if only the conventional ARL measures are used. For d>14 the PSD is close to 1 for both cases (see Figure 7.3.1).

ACKNOWLEDGEMENTS

This work has been a part of a project supported by the Swedish Council for Research in the Humanities and Social Sciences. I would like to thank my supervisor Marianne Frisén for her engagement and encouragement through the work on this paper. Her constructive critisism and comments on numerous drafts has been an invaluable help for me in finalizing this paper. I am grateful to my company Astra Hässle AB for giving me the opportunity to do this paper. I also thank all my colleagues at the Department of Statistics, University of Göteborg and the colleagues at Astra Hässle AB for many constructive discussions. Particulary I want to thank my colleague Hrafnhildur Arnkelsdottir for all our enjoyable discussions. I am also grateful to Robert Johnson for valuable comments on the draft manuscript. Finally I would like to thank Margareta Alftenius for all the help she has given me with the manuscript.

1994:1	Frisén, M.	Statistical surveillance of business cycles.
1994:2	Frisén, M.	Characterization of methods for surveillance by optimality.
1994:3	Frisén, M. & Cassel, C.	Visual evaluation of statistical surveillance.
1994:4	Ekman, C.	A comparison of two designs for estimating a second order surface with a known maximum.
1994:5	Palaszewski,B.	Comparing power and multiple significance level for step up and step fown multiple test procedures for correlated estimates.
1994:6	Åkermo, G.	Constant predictive value of an alarm.
1994:7	Åkermo, G.	On performance of methods for statistical surveillance.
1994:8	Palaszewski, B.	An abstract bootstrap based step down test procedure for testing correlated coefficients in linear models.
1994:9	Ekman, C.	Saturated designs for second order models.
1994:10	Ekman, C.	A note on rotatability.
1995:1	Arnkelsdóttir, H	Surveillance of rare events. On evaluation of the sets method.