Research Report
Statistical Research Unit
Department of Economics
Göteborg University
Sweden

Semiparametric estimation of outbreak regression.

Frisén, M., Andersson, E. & Pettersson, K.

Research Report 2007:13
ISSN 0349-8034

Mailing address:	Fax	Phone	Home Page:
Statistical Research Unit	Nat: 031-786 12 74	Nat: 031-786 00 00	http://www.statistics.gu.se/
P.O. Box 640	Int: +46 31 786 12 74	Int: +46 31 786 00 00	
SE 405 30 Göteborg			
Sweden			

Semiparametric estimation of outbreak regression

MARIANNE FRISÉN*, EVA ANDERSSON and KJELL PETTERSSON
Statistical Research Unit, Department of Economics, University of Gothenburg, Sweden

A regression may be constant for small values of the independent variable (for example time), but then a monotonic increase starts. Such an “outbreak” regression is of interest for example in the study of the outbreak of an epidemic disease. We give the least square estimators for this outbreak regression without assumption of a parametric regression function. It is shown that the least squares estimators are also the maximum likelihood estimators for distributions in the regular exponential family such as the Gaussian or Poisson distribution. The approach is thus semiparametric. The method is applied to Swedish data on influenza, and the properties are demonstrated by a simulation study. The consistency of the estimator is proved.

Keywords: Constant Base-line; Monotonic change, Exponential family Influenza outbreak

* E-mail: marianne.frisen@statistics.gu.se

1. Introduction

Model selection is important and different adaptive and model-free approaches have been suggested (see e.g. [1]). Without including available assumptions on the shape of the regression, the estimates would be unnecessary inefficient. On the other hand, wrong assumptions might cause wrong conclusions from the data. Thus, limited constrains on a regression, focused on the issues that are important for the application, are of interest.

One important aim in public health surveillance is to detect disease outbreaks. An outbreak can be characterised as a change from a constant level to a monotonically increasing incidence. Outbreak detection is an important part of surveillance for bioterrorism as well as of surveillance for the detection of new diseases such as the recent SARS and avian flu. Outbreaks are also important in the study of ordinary influenza. For likelihood-based surveillance methods ([2], [3]) maximum likelihood estimates are needed. Such estimators will be given in this article. However, this article will not deal with the sequential issues of surveillance.

In many applications the “normal” or base-line state can be described by a constant level. At a possibly unknown time, the process changes to a monotonically increasing (or decreasing) regression. In this paper we will treat the case of a monotonically increasing regression following the change point, but the statistical problem is the same for a decreasing regression. This “outbreak” regression is of interest not only at the outbreak of an epidemic disease. We have a similar statistical problem when investigating whether data deviate from a specified econometric model by analysing whether there is a change point after which the residuals are increasing.

Often a parametric regression is used to estimate the expected incidence during the outbreak. In many cases, however, the parameters would vary from case to case. One example of this is the outbreak of influenza, where the parameters describing the outbreak do vary from one year to the next. The character of the outbreak also varies from one period to the next, thus making it difficult to use a parametric model without misspecification. In [4] and [5] it is concluded that parametric models are not suitable when the parameters vary much from year to year, as they do for influenza data. The importance of avoiding the effects of estimation errors is also discussed in [6]. Thus, here we suggest a nonparametric approach (with respect to the regression function) utilising only the characteristics of a constant start followed by a monotonic increase.

There are several related nonparametric regression problems. Unimodal or “J-shaped” regression is treated in e.g. [7], [8] and [9]. Concave regression is treated for example in [10]. A broken-line estimation is suggested in [11], where the parameter, in a distribution belonging to the exponential family, is constant at first, but at an unknown time there is an onset of a positive constant change. The authors point out that also nonlinear regression can be treated by this approach, after a parametric transformation, and they study conditions for consistent estimation of the time of the change-point. They consider the case where the behaviour of the parameter is known after the change, while this paper requires only that the expected value is monotonically changing with time. Smoothing by kernel methods (see e.g. [12]) is often used. In [13] [14] and [15] there are discussions on the use of the extra information by monotonicity restriction in connection with smoothing methods. Smoothing methods are very useful for illustrating the outbreak behaviour, but for some purposes, such as alarm systems and hypothesis testing, maximum likelihood estimates are useful.

The aim of this paper is to derive the least squares and maximum likelihood estimators of the localization parameter for outbreak regression under monotonicity restrictions. We study both the case of a known and an unknown change point. The normal distribution and the Poisson distribution are of special interest but other members of the exponential family are also considered. The estimator is semiparametric in the sense that the regression function is nonparametric while the distributions used for the maximum likelihood estimators are parametric. The result of this paper is used in derivation of sequential likelihood based surveillance in [16] and [17].

In Section 2 the model is specified and notations are given. In Section 3, the least squares estimators are derived. In Section 4 the method is illustrated by an example. Consistency is discussed in Section 5. Maximum likelihood properties are given in Section 6. The properties are demonstrated by a simulation study in Section 7. In Section 8 some concluding remarks are given.

2. Models and specifications

We observe the process X and at time t we have m(t) observations $x_1(t), x_2(t), ..., x_{m(t)}(t)$, t= 0, 1, … s. Let τ be the time when the monotonic increase starts. Thus τ is the first time for which the regression function is not constant. The change point τ may be known or unknown. The expected value of $X_i(t)$ is denoted by $\mu^\tau(t)$. The superscript is suppressed when obvious. At time τ the expected value μ changes from a constant level to an increasing regression:

$$\mu(0)=...=\mu(\tau-1) < \mu(\tau) \le ... \le \mu(s). \qquad (1)$$

The monotonicity restriction contains two parts

$$\mu(0)=...=\mu(\tau-1) \qquad (1a)$$

and

$$\mu(\tau-1) < \mu(\tau) \le ... \le \mu(s) \qquad (1b)$$

We will pay special interest to the situation when $X_i(t)$ is normally distributed and the situation when $X_i(t)$ follows a Poisson distribution, but some results are relevant to all members of the exponential family.

3. Least squares estimation of an outbreak regression

Least squares estimation with monotonicity restrictions was described for example in [18] and [19]. We need optimisation under two restrictions, (1a) and (1b). We will prove that if we first optimise under (1a) and then optimise the resulting series under (1b), we will get estimators with the desired properties. In a situation with more that 1 observation at a specific time (i.e. m(t)>1), the mean is calculated. The mean is the least square estimator of μ. The same vector $\hat{\mu}$ that minimizes the sum of square around the observations will also minimize the sum of square around the means. For a specific value τ the suggested estimator is constructed by first considering condition (1a), which is the base for the computation of a provisional series y(t) where

$$Y^\tau(t) = \begin{cases} \sum_{j=0}^{\tau-1}\sum_{i=1}^{m(t)}\left(X_i(j)\right) / \sum_{t=0}^{\tau-1} m(t), & t < \tau \\ \sum_{i=1}^{m(t)}\left(X_i(t)\right)/m(t), & t \ge \tau \end{cases} \qquad (2)$$

The next step is to consider condition (1b):

$$\hat{\mu}^\tau(t) = g(t \mid Y^\tau(0), Y^\tau(1), ..., Y^\tau(s)), \qquad (3)$$

where the function g(t) is the least squares estimator of the provisional series $Y^\tau(t)$ under the monotonicity restriction (1b).

The order in which the two conditions (1a and 1b) are used will matter and only this ordering will result in estimators which satisfy the least squares and maximum likelihood conditions under the combined restrictions. The estimator can also be seen as a pool-adjacent-violators algorithm (PAVA) [19] as will be demonstrated below.

Theorem 1 For a fixed number of observations s and a fixed time point τ from which $\mu(t)$ is increasing, the least squares estimator under the order restriction (1) is given by $\hat{\mu}^\tau(t)$, given in (3).

Proof Since the ordering of the observations before τ is irrelevant, we can formulate the problem as having τ-1 observations at time τ-1 and 1 observation at each time τ, τ+1,...,s, and the restriction for this new problem is:

$$\mu(\tau-1) < \mu(\tau) \leq \ldots \leq \mu(s)$$

which is on the border of

$$\mu(\tau-1) \leq \mu(\tau) \leq \ldots \leq \mu(s).$$

This problem is an ordinary monotonic regression and the LS estimator is given by PAVA. See for example Section 2.4.1 of [20].■

The estimator $\hat{\mu}^{\tau}(t)$ is weighted by the number of observations. It could also be weighted by using special weights, for example $w(t) = 1/\sigma(t)$ where $\sigma^2(t)$ is the variance of each of these observations.

Theorem 2 *When the change point is unknown, the least squares estimator of $\mu(t)$ is*

$$\hat{\mu}(t) = \hat{\mu}^{1}(t) \qquad (4)$$

Proof All other restrictions are included in the monotonic restriction of τ=1. Thus, no other joint estimators could have a smaller value of $\sum_{t=0}^{s}\sum_{i=1}^{m(t)}(x_i(t) - \hat{\mu}^{j}(t))^2$ = Q(j) than Q(1).■

One conclusion from Theorem 2 is that it is not possible to estimate the value of τ without further restrictions as discussed in Section 8.

4. Calculations of influenza incidences

In order to illustrate the computation of the estimator, we give the details for an example with a few observations. This is the number of laboratory-identified cases of influenza in Sweden during the first weeks of the winter 2003/2004.

There are observations x(0), x(1), … , x(7) at time points t=0, 1,..., 7 (in this example m(t)≡1). We calculate the estimates for the cases when τ=3 and when τ=6. For τ=3, it is assumed that μ(0)=μ(1)=μ(2) and μ(2)<μ(3)≤μ(4)≤μ(5)≤μ(6)≤μ(7), and for τ=6 it is assumed that μ(0)=μ(1)=μ(2)=μ(3)=μ(4)=μ(5) and μ(5)<μ(6)≤μ(7) respectively. The data (x), the provisional series (y) and the least squares estimators ($\hat{\mu}$) are given in Table 1. The sum of squares is smaller for τ=3, compared to τ=6, thus the earlier change point fits the data best. The observations and the least squares estimates are seen in Figure 1

Table 1. The computations for the restrictions τ = 3 and τ = 6

t	x(t)	$y^3(t)$	$\hat{\mu}^3(t)$	$y^6(t)$	$\hat{\mu}^6(t)$
0	0	0	0	1	1
1	0	0	0	1	1
2	0	0	0	1	1
3	2	2	1	1	1
4	0	0	1	1	1
5	4	4	4	1	1
6	23	23	23	23	23
7	38	38	38	38	38

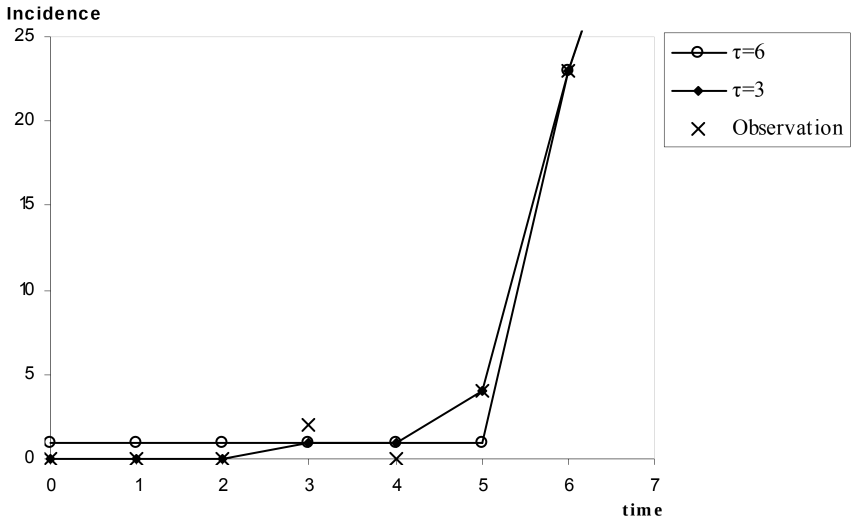

Figure 1: The observed data x and the estimates conditional on the monotonicity restriction τ=3 and τ=6, respectively.

5. Consistency

When the number of observations at each time t, m(t), increases we get consistent estimators for the μ vector.

Theorem 3 If the distribution belongs to the exponential family, then $\hat{\mu}^{\tau}(t)$ will give a consistent estimate of μ(t) which fulfils condition (1).

Proof Let m= $\min_t \{m(t)\}$. The estimator will use the averages $\overline{X}(t) = \sum_{i=1}^{m(t)} X_i(t)/m(t)$, as commented in Section 3. Since $\overline{X}(t)$ is a strongly consistent estimator of the expected value in the exponential family, so is $\hat{\mu}^{\tau}(t)$, since only averaging and PAVA are used in the transformations of $\overline{x}(t)$. It follows that, with probability one,

$$\lim_{m\to\infty} \max_t |Y^{\tau}(t) - \mu(t)| = 0 .$$

Thus, with probability one $\hat{\mu}(t)$ satisfies the condition (1) as m goes to infinity. ■

Unfortunately this consistency does not carry over to the case where there is only one observation for each time but the number of time points increases. For the case when we have a pre-grouping of the time points into classes, the consistency property carries over to the expected values in these time-classes if the number of observations in each time-class increases.

6. Maximum likelihood estimation

For certain distributions the least squares estimators given above are also maximum likelihood estimators. We will consider the regular exponential family with the conditions of the derivatives of the parameters as specified on page 34 of [19] and give special cases of this family.

Theorem 4 The least squares solutions of Theorem 1 and 2 are the maximum likelihood solutions if the values of the dispersion parameter in the exponential distribution are equal for all times (but possibly unknown).

Proof This follows from properties of ordinary isotonic regression since the current problem can be expressed in these terms, as demonstrated in the proof of Theorem 1. See for example Section 2.4.2 of [20].■

Theorem 5 The weighted least squares estimator is the maximum likelihood solution for known but possibly different values of the dispersion parameter.

Proof This follows from properties of ordinary isotonic regression. See for example Section 2.4.2 of [20].■

Corollary 1 The least squares solutions $\hat{\mu}^{\tau}(t)$ and $\hat{\mu}(t)$ in (3) and (4) are the maximum likelihood solutions when the observations at each time follow a normal distribution with equal variances.

Corollary 2 The weighted least squares solutions $\hat{\mu}^{\tau}(t)$ and $\hat{\mu}(t)$ in (3) and (4) (with the weights $w(t) = 1/\sigma(t)$) are the maximum likelihood solutions when the observations at each time follow a normal distribution with unequal variances, where the variances are known (or their relation to each other is known).

Corollary 3 The (unweighted) least squares solutions $\hat{\mu}^{\tau}(t)$ and $\hat{\mu}(t)$ in (3) and (4) are the maximum likelihood solutions for a Poisson distribution.

This follows from the fact that there is no additional dispersion parameter for the Poisson distribution. One might have expected that weights should be used since the parameter of the Poisson distribution also reflects the variance. However, the only places where the regression differs from the observations are where the estimates by the PAVA are constants. A weighted regression should thus have constant weight.■

The estimated curve (and the corresponding likelihood) may be used for inference such as hypothesis testing or surveillance concerning the start of the influenza season, but such inference will not be treated here.

7. Simulation study of properties

We performed a simulation study in order to illustrate bias, variance and the influence of the value of τ (the time when the increase starts). We generated data similar to those that can be expected at an influenza outbreak according to [4]. In Sweden, the monitoring of influenza starts at week 40 each year but the time of the onset varies considerably between years and thus also the length of the constant phase. Here is a situation where there is need of a surveillance system in order to get early warnings regarding the start of the influenza. We investigated several possible scenarios. The reported results are based on at least one million replicates. In some cases, such as for Figures 3 and 7, we used 50 million replicates in order to avoid influence of random fluctuations in the reported simulation results. We report results for Poisson and normally distributed variables.

To illustrate the case of a Poisson distribution we generated weekly numbers of laboratory-diagnosed influenza cases (LDI) according to their similarity with the influenza season 2003/2004, which was a "typical" season. The observed process X follows a Poisson distribution with the parameter μ(t), where

$$\mu(t) = \begin{cases} \mu_0, & t < \tau \\ \exp(\beta_0 + \beta_1 \cdot (t - \tau + 1)), & t \geq \tau \end{cases}$$

where $\mu_0=1$, $\beta_0=-0.26$, $\beta_1=0.826$.

In Figure 2 the mean and standard deviation (by 2 SD bars) of the estimates of 1 000 000 replicates are given. The cases are generated for different values of τ (τ=4 and τ=8). The estimates were produced with knowledge of the true value. The variation of the estimates is smaller than without the restriction, thus Var($\hat{\mu}^{\tau}(t)$)<Var(X(t)). The effect of the restriction of a constant phase has a major influence on Var($\hat{\mu}^{\tau}(t)$) during this phase, and this variance is smaller than the variance for the mean of all the observations during the constant phase. The monotonicity restriction has a small variance-reducing effect when the slope is large in comparison with the variance.

There is a bias, but this is too small to be seen in the scale of Figure 2. Thus the two series (the mean of the estimate and the expected value of the generated data) coincide in Figure 2. The bias ($E[\hat{\mu}] - \mu$) is illustrated in a larger scale in Figure 3.

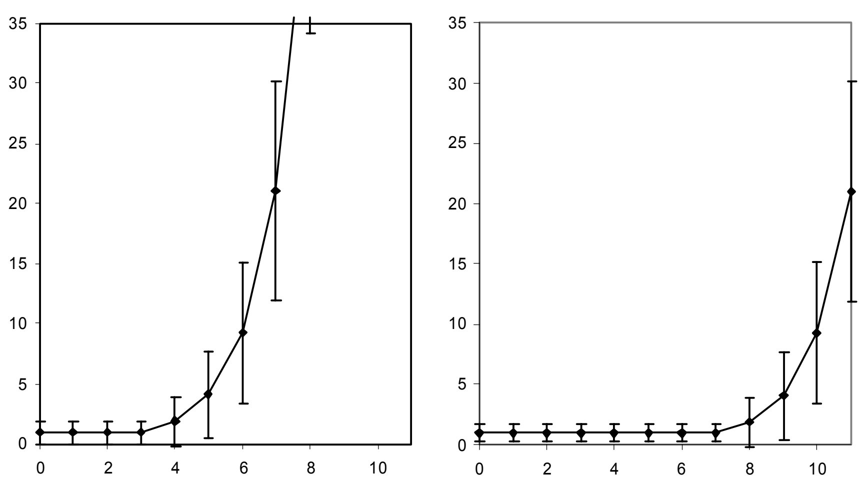

Figure 2. The mean of the estimated values at each time point (dot) and the variation of the estimated values, illustrated by ± 2SD (bars). The true expected value, μ(t), cannot be distinguished from the mean of the estimates, E[$\hat{\mu}$ (t)], in the scale of the figure. The left figure is estimated under the true restriction of τ=4, and the right figure is estimated under the true restriction that τ=8.

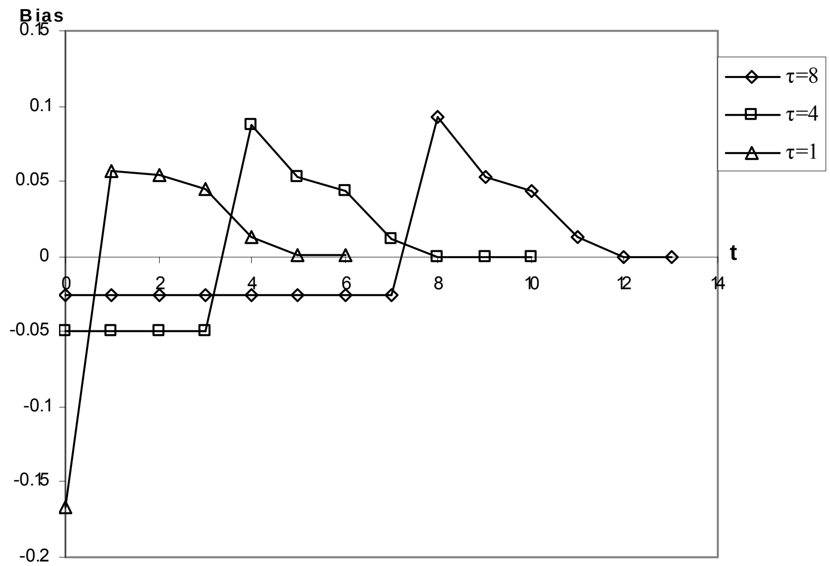

Figure 3. The bias $E[\hat{\mu}]-\mu$ for the situations when data are generated for τ=1 τ=4 and τ=8 and the true value of τ is known in the estimation.

As could be expected the bias in the constant phase is small since the first step of forming the mean (provisional series) produces an unbiased estimate. In the next step the isotonic regression will produce a too low estimate of the constant phase. The weight of the unbiased estimate is (τ-1)/s, thus the bias will be small for a large value of τ. For the next part of the regression the bias is as expected for an isotonic curve; namely, there is a negative bias for early time points and a positive bias for late ones. This is illustrated in Figure 4 where a constant value is generated and the estimation is made under the restriction of τ=1.

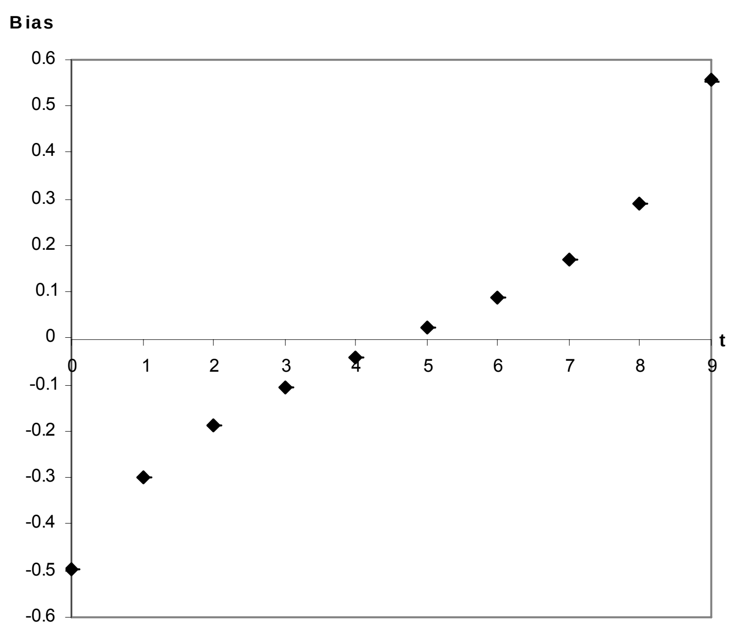

Figure 4. Estimation under the restriction of τ=1 for data generated by a Poisson distribution with a constant mean (i.e. generated under the condition that τ=∞).

The pattern in Figure 3 will not completely agree with the one in Figure 4 even at the isotonic phase, since we have an exponential increase as soon as the influenza has started. Thus, we will very soon have very little influence of the isotonic regression. The later points will almost always be estimated by the observed values, and the bias will thus decrease to zero.

The effect of misspecification of τ is illustrated in Figure 5. Both curves (τ=4 and τ=8) from which data are generated are the same as those in Figure 2.

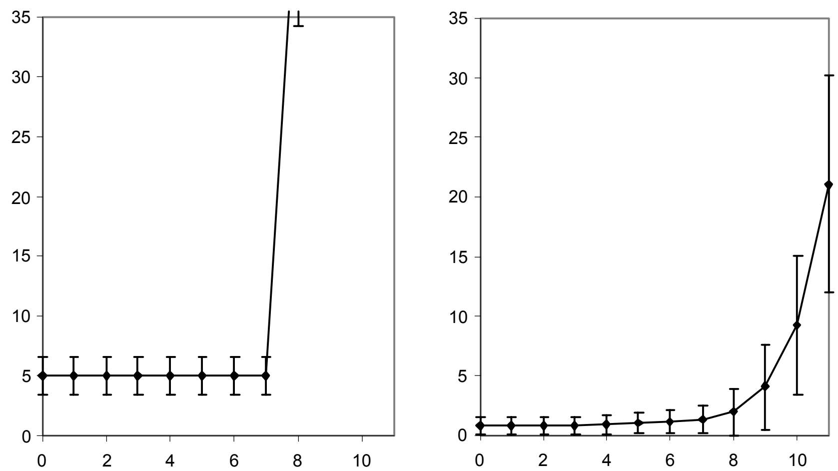

Figure 5. The mean of the estimated values at each time point (dot) and the variation of the estimated values, illustrated by ± 2SD (bars). The effect of error in the assumption of τ is illustrated: in the left figure, the true τ equals 4 but the restriction τ=8 is imposed in the estimation, and in the right figure, the true τ equals 8 but the restriction τ=4 is imposed in the estimation.

In Figure 5 we can see that a restriction of a later change than the true one will give a constant phase at a too high level. In Figure 5 (right) we can see that a restriction of an earlier change than the true one has very little impact. In Figure 6 we illustrate the bias and the standard deviation when no assumption of the value of τ is made but the general maximum likelihood estimator $\hat{\mu}(t)$ is used.

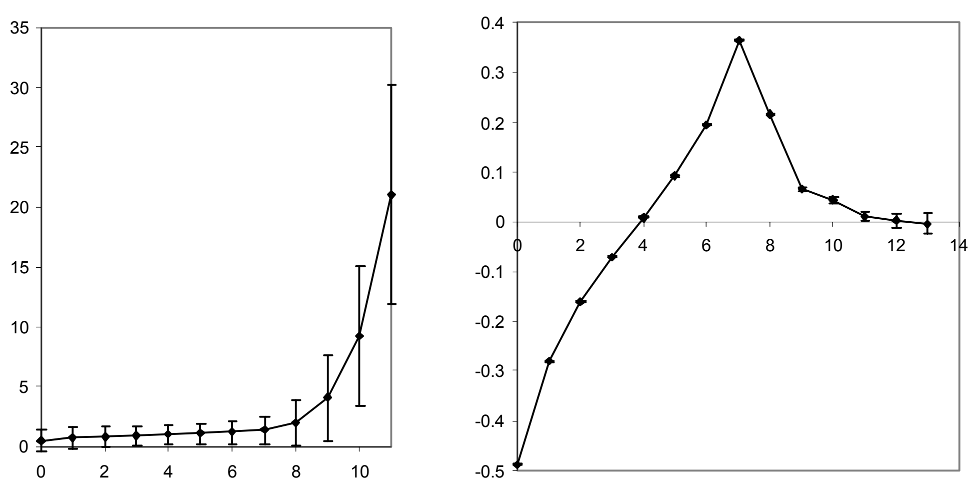

Figure 6. The maximum likelihood estimator, without any information about τ (true τ equals 8). Left: The mean of the estimated values at each time point (dot) and the variation of the estimated values, illustrated by ± 2SD (bars). The true expected value, $\mu(t)$, cannot be distinguished from the mean of the estimates, E[$\hat{\mu}$(t)], in the scale of the figure. Right: The bias at each time point (dot).

In Figure 6 (left) we see that the mean of the estimated curve, even without information on τ, is very close to the real curve in the current scale. Thus, even without knowledge of τ, the estimator produces a reasonable estimate. However, by comparing the bias in the right panel of Figure 6 with the one in Figure 3 (for τ=8), we can conclude that the knowledge of the value of τ decreases the bias – especially for the constant phase. By comparing the variation of the estimates (±2SD) in Figure 6 (left) with that of Figure 2, we can see that the correct restriction (knowledge about τ) decreases the variation – especially during the constant phase.

For the Poisson distribution, the variance and the expected value have the same value. Therefore, normally distributed data are used to examine the effect of the variance. To illustrate the properties for normal distributions with different variances we generated data with means similar to the number of influenza-like cases (ILI) during the winter 2003/2004 and a constant variance. The following model was used for the observed process X:

$$X(t) \sim N(\mu(t); \sigma^2),$$

where

$$\mu(t) = \begin{cases} \mu_0, & t < \tau \\ \exp(\beta_0 + \beta_1 \cdot (t - \tau + 1)), & t \geq \tau \end{cases}$$

and μ_0=20, β_0=2.67, β_1=0.68 and different values of the variance σ^2 are used. A normal distribution is a reasonable approximation here since the incidences are rather high. Different scenarios were considered regarding the length of the constant phase.

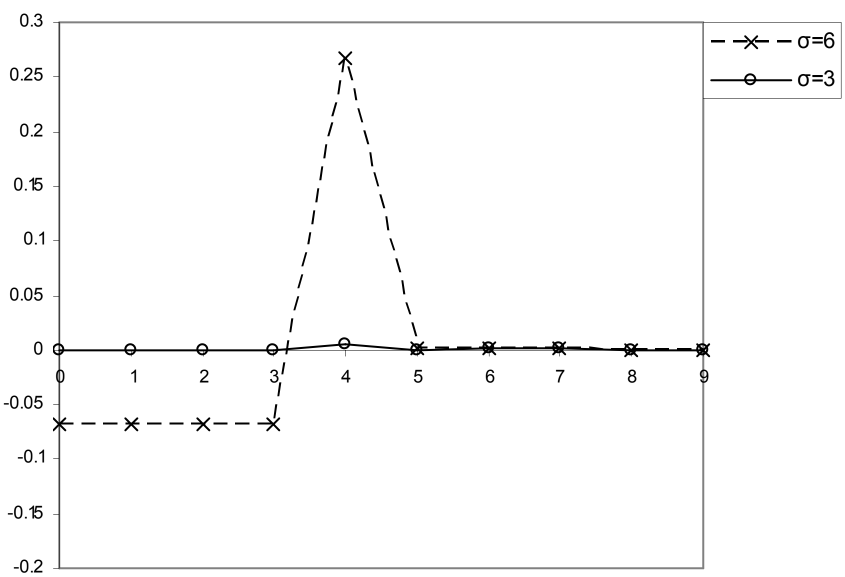

Figure 7. Average bias of the ML-estimator for normally distributed data with standard deviation σ=3 and σ=6 respectively. The data are generated for τ=4 and this is used as a restriction in the estimation.

In Figure 7 we can see that the bias is small for a small variance. This illustrates that the estimator is consistent.

8. Concluding remarks

The results presented here on outbreak regression are of importance not only in pure estimation contexts but also for on-line surveillance based on maximum likelihood estimators. The outbreak of a disease can often be characterised as a change from a constant level to a monotonically increasing incidence. Surveillance systems for detecting outbreaks are crucial in surveillance for bioterrorism as well as in surveillance for the detection of new diseases, see

[21]. Outbreak detection is also important in the study of ordinary influenza [22]. Surveillance systems based on likelihood ratios have important optimality properties [3]. For nonparametric surveillance as in [5] (nonparametric with respect to the shape of the curve), maximum likelihood estimates are useful as a basis for maximum likelihood ratios. The estimators presented here can be used for likelihood-based surveillance to detect the onset of an increasing incidence. Smoothing methods are useful for the description of the outbreak behaviour but will not give the required maximum likelihood estimators.

Sometimes it is reasonable to believe that the regression is continuous and has continuous derivatives. However, this condition can always be satisfied by some definition of estimates between the discretely observed times. Thus this is no restriction to the estimates. When a smooth curve is needed for illustration, it is possible to fit a smooth curve (such as a spline [23]) to the maximum likelihood estimates.

One may be interested in estimating the time, τ, of the onset of the increasing phase and also the level of the constant phase, $\mu(0)$. The maximum likelihood estimation of the curve by the proposed method will also give maximum likelihood estimates of these parameters. Generally, however, there will not be one unique maximum likelihood estimator of τ. No other value of τ can give a larger value of the likelihood than $\tau=1$, since $\mu(0)= \ldots =\mu(i-1)<\mu(i)\leq \ldots\leq\mu(s)$ is a special case of, or on the limit of, $\mu(0)<\mu(1)\leq\ldots\mu(i)\ldots\leq\mu(s)$. The maximum likelihood estimator of $\mu(0)$ will be unique but biased since the maximum likelihood estimators of τ and $\mu(0)$ are closely related. This problem of bias in the endpoints is shared with other maximum likelihood estimators of ordered statistics such as the usual monotonic regression. In order to get unbiased estimators of τ and $\mu(0)$, more (parametric) structure could be used, for example a certain size of the change. In [24] a penalization of the distance between $\mu(0)$- $\mu(s)$ was used to avoid bias in connection with a test of whether the regression is constant or monotonic for the whole period.

When the maximum likelihood statistic derived here is used for test or surveillance purposes, the bias is not a problem. In such cases there are natural false alarm requirements which give the user the opportunity to state that only important deviations should be detected. This corresponds to the above-mentioned parametric size condition for the estimator but is expressed by probability and does not require any parametric assumption for the curve.

The estimator is consistent (for a large number of observations at each time) but not unbiased. The direction of the bias is that the estimates are too low. However, the bias is very small for the constant phase. For the increasing phase the bias is smaller to start with due to the stabilisation by the constant phase. A long constant phase exaggerates this tendency.

Acknowledgements

Linus Schiöler has provided expert technical and computational help. The data were made available to us by the Swedish Institute for Infectious Disease Control. The research was supported by the Swedish Emergency Management Agency (grant 0314/206).

References

[1] O. Bunke, et al., Model Selection, Transformations and Variance Estimation in Nonlinear Regression, in Statistics, 1999, pp. 197 - 240.

[2] E. Andersson, Monitoring cyclical processes - a nonparametric approach, in Journal of Applied Statistics, 2002, pp. 973-990.

[3] M. Frisén, Statistical surveillance. Optimality and methods., in International Statistical Review, 2003, pp. 403-434.

[4] E. Andersson, et al., Modeling influenza incidence for the purpose of on-line monitoring, in Statistical Methods in Medical Research, 2007, pp. doi:10.1177/0962280206078986.

[5] D. Bock, et al., Statistical surveillance of epidemics: Peak detection of influenza in Sweden, in Biometrical Journal, 2008, pp. 71-85.

[6] R.A. Ion, and C.A.J. Klaassen, Non-parametric Shewhart control charts, in Journal of Nonparametric Statistics, 2005, pp. 971-988.

[7] M. Frisén, *Unimodal regression*, in *The Statistician*, 1986, pp. 479-485.

[8] L. Andersson, and M. Frisén, Verifications of turning points, in Journal of Nonparametric Statistics, 2002, pp. 623-645.

[9] E. Andersson, et al., Statistical surveillance of cyclical processes. Detection of turning points in business cycles, in Journal of Forecasting, 2005, pp. 465-490.

[10] C. Hildreth, Point estimation of ordinates of concave functions, in Journal of the American Statistical Association, 1954, pp. 599-619.

[11] R. Gill, and M. Baron, Consistent estimation in generalized broken-line regression, in Journal of Statistical Planning and Inference, 2004, pp. 460.

[12] W. Härdle, Smoothing techniques. With implementations in S., Springer 1991.

[13] E. Mammen, Nonparametric regression under qualitative smoothness assumptions, in The Annals of Statistics, 1991, pp. 741-759.

[14] S. Efromovich, Density estimation under random censorship and order restrictions: From asymptotic to small samples, in Journal of the American Statistical Association, 2001, pp. 667-684.

[15] M. Frisén, Discussion on "Sequential design and estimation in heteroscedastic nonparametric regression" by Sam Efromovich, in Sequential Analysis, 2007, pp. 45-47.

[16] M. Frisén, and E. Andersson, *Semiparametric surveillance of outbreaks*, in *Research Report*, Statistical Research Unit, Department of Economics, Göteborg University, Sweden : 2007:11, 2007.

[17] M. Frisén, et al., *Robust outbreak surveillance of epidemics in Sweden*, in *Research Report*, Statistical Research Unit, Department of Economics, Göteborg University, Sweden : 2007:12, 2007.

[18] R.E. Barlow, et al., *Statistical inference under order restrictions*, Wiley, London, 1972.

[19] T. Robertson, et al., *Order restricted statistical inference*, John Wiley & Sons Ltd, Chichester, 1988.

[20] M. Silvapulle, and P.K. Sen, Constrained statistical inference. Inequality, order and shape restriction, Wiley, 2005.

[21] C. Sonesson, and D. Bock, A review and discussion of prospective statistical surveillance in public health, in Journal of the Royal Statistical Society A, 2003, pp. 5-21.

[22] E. Andersson, et al., *Predictions by early indicators of the time and height of yearly influenza outbreaks in Sweden*, Research Report, Statistical Research Unit, Göteborg University 2007.

[23] B.W. Silverman, Some aspects of the spline smoothing approach to non-parametric regression curve fitting, in Journal of the Royal Statistical Society B, Royal Statistical Society, 1985, pp. 1-52.

[24] W. Wu, et al., Isotonic Regression: another look at the changepoint problem, in Biometrika, 2001, pp. 793-804.

Research Report

2007:1	Andersson, E.:	Effect of dependency in systems for multivariate surveillance.
2007:2	Frisén, M.:	Optimal Sequential Surveillance for Finance, Public Health and other areas.
2007:3	Bock, D.:	Consequences of using the probability of a false alarm as the false alarm measure.
2007:4	Frisén, M.:	Principles for Multivariate Surveillance.
2007:5	Andersson, E., Bock, D. & Frisén, M.:	Modeling influenza incidence for the purpose of on-line monitoring.
2007:6	Bock, D., Andersson, E. & Frisén, M.:	Statistical Surveillance of Epidemics: Peak Detection of Influenza in Sweden.
2007:7	Andersson, E., Kühlmann-Berenzon, S., Linde, A., Schiöler, L., Rubinova, S. & Frisén, M.:	Predictions by early indicators of the time and height of yearly influenza outbreaks in Sweden.
2007:8	Bock, D., Andersson, E. & Frisén, M.:	Similarities and differences between statistical surveillance and certain decision rules in finance.
2007:9	Bock, D.:	Evaluations of likelihood based surveillance of volatility.
2007:10	Bock, D. & Pettersson, K.	Explorative analysis of spatial aspects on the Swedish influenza data.
2007:11	Frisén, M. & Andersson, E.	Semiparametric surveillance of outbreaks.
2007:12	Frisén, M., Andersson, E. & Schiöler, L.	Robust outbreak surveillance of epidemics in Sweden.