9th Workshop on
Stochastic Models
and Their Applications
March 3 - 6, 2009

Proceedings of the 9th Workshop on Stochastic Models and Their Applications

March 3-6, 2009

RWTH Aachen University

Edited by
A. Steland

RWTH AACHEN UNIVERSITY

Institut für Statistik & Wirtschaftsmathematik RWTH AACHEN

Preface

This volume presents the extended abstracts of the talks contributed to the 9^{th} Workshop on Stochastic Models and Their Applications (March 3-6, 2009). The scientific committee thanks all authors for their valuable contribution. This workshop was organized by the Institute of Statistics at RWTH Aachen University. As in previous years, it is devoted to reliability, sequential change-point analysis, lifetime and survival analysis, time series analysis, and applications in engineering and computer science.

However, various stochastic models and methods encountered in these fields are also employed and applied to solve statistical problems and analyse data in finance, and vice versa. This particularly applies to models and methods from modern time series and regression analysis. Thus, we organized an integrated Spring School on Time Series in Finance and Engineering consisting of three additional lectures on these issues. The lectures highlight nonparametric, semiparametric, and parametric approaches and discuss their role in contemporary research. Although they particularly address graduate students, Ph.D. students and PostDocs, we hope that all participants interested in these issues will benefit.

We hope you enjoy your stay and the talks.

Aachen, March 1, 2009

Ansgar Steland.

Acknowledgements

The Spring School gratefully acknowledges financial support by RWTH's fund "Undergraduates into Research", a measure financed by the German Excellence Initiative.

Logo design by André Weimar (http://www.andreweimar.de), organized by A. Thrun.

We thank H. Satvat and H. Schottmüller for technical support.

Scientific Committee

Uwe Jensen (University of Hohenheim, Germany)

Mirosław Pawlak (University of Manitoba, Canada)

Ansgar Steland (RWTH Aachen University, Germany)

Local Organizing Committee

M. Abujarad (Institute of Statistics)

E. Cramer (Institute of Statistics)

O. Holtemöller (Faculty of Business and Economics)

U. Kamps (Institute of Statistics)

R. Mathar (Institute for Theoretical Information Technology and UMIC)

H. Ney (Chair of Computer Science 6)

A. Steland (Institute of Statistics)

Secretary and Address

Ms. Simone Gerwert
Institute of Statistics
RWTH Aachen University
D-52056 Aachen, Germany

Fon: +49-241-80-94571
Fax: +49-241-80-92130

e-mail: gerwert@stochastik.rwth-aachen.de
URL: http://sma2009.rwth-aachen.de/home.html

Contents

Nonlinear Locally Weighted Kriging Prediction for Spatio-Temporal Environmental Processes

Olha Bodnar *Department of Statistics, European University Viadrina, Germany*
Wolfgang Schmid *Department of Statistics, European University Viadrina, Germany*

Abstract: In the paper a non-linear interpolation procedure for the spatial prediction of an environmental process is proposed. The suggested interpolation is based on the locally weighted scatterplot smoothing method (c.f. [1]). This approach is applied to a non-linear spatio-temporal model. In an empirical study the PM10 concentration in the Berlin-Brandenburg region of Germany is considered. It is shown that the local approach permits a more structured interpolation of the air pollution.

1 Introduction

The prediction of the values of a spatio-temporal environmental process at sites, where no station of the monitoring network is available, is an important problem for environmental statistics today. Usually, the linear kriging predictor is used, which is obtained by minimizing the mean squared error (see, e.g. [3], [5]). The problem of an optimal network design for spatial interpolation using linear kriging predictor is discussed in [5]. The author dealt with the problem, when the parameters of the spatio-temporal process are unknown and have to be estimated before the linear predictor is constructed. The problem of separable approximations of space-time covariance matrices is considered in [3]. This procedure is used to reduce the dimensionality of the inverse covariance matrix arising in the equation of the linear predictor.

In this paper we propose a new approach for interpolating spatial data. The suggested approach is based on the idea of locally weighted regression and smoothing scatterplots. It allows us to derive a non-linear spatial interpolation of environmental data. The main advantages of the approach are: (1) The LOESS predictor does not require any specification of the prediction function. Thus, it is fully non-parametric in this context. (2) It is more flexible than the linear predictor, usually, used in the environmetrics literature. (3) Because not all monitoring stations are used for calculating the prediction in a given site the dimension of the problem is reduced and this leads to a simpler determination of the inverse covariance matrix.

2 The method

In the present paper, we propose a non-linear predictor for spatial prediction which is based on the idea of the locally weighted scatterplot smoothing (LOESS) regression

model (see, e.g. [1]). A subset of stations of the monitoring network is used which are located nearest to the site, where the prediction has to be done. This approach allows us to define a non-linear predictor without specifying an exact expression of its functional form and to apply it to the spatial prediction of non-stationary spatio-temporal processes. Because only a subset of the stations of the monitoring network is taken into account for each site, the suggested approach reduces the dimensionality of the inverse covariance matrix used in the predictor's expression. In the paper, we also extend the general spatio-temporal process of [2] to a non-linear one and apply the LOESS predictor to the new process. In order to estimate the unknown parameters of the suggested process, the maximum likelihood method is applied which is performed using the iterative procedures of the generalized EM algorithm (see, e.g., [4]).

3 Example

The suggested approach is used to construct a PM10 mapping for the Berlin-Brandenburg region in Germany. In the empirical study, we compare the LOESS predictor with the linear predictor. We interpolate the concentration of the PM10 in the Berlin-Brandenburg region based on both predictors. Daily data from 2007 for the Berlin-Brandenburg monitoring network is used. We observe that the LOESS predictor is much more flexible than the linear predictor and it is more suitable to detect local changes. Consequently this approach seems to be of great interest in practice.

References

[1] Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74: 829-836.

[2] Fassò, A. and Cameletti, M. (2007). A general spatio-temporal model for environmental data. *Technical report, Graspa - The Italian Group of Environmental Statistics - www.graspa.org*, 27.

[3] Genton, M.G. (2007). Separable approximations of space-time covariance matrices. *Environmetrics*, **18**, 681-695.

[4] McLachlan, G. J., and Krishnan, T. (1997). *The EM Algorithm and Extensions*, Wiley: New York.

[5] Zimmerman, D. L. (2006). Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction, *Environmetrics*, 17: 635-652.

Surveillance of the Covariance Matrix Based on the Properties of the Singular Wishart Distribution

Olha Bodnar *Department of Statistics, European University Viadrina, Germany*
Taras Bodnar *Department of Statistics, European University Viadrina, Germany*
Yarema Okhrin *Department of Economics, University of Bern, Switzerland*

Abstract: In this paper we develop a methodology which allows applying the standard monitoring techniques for the mean behavior of Gaussian processes in the detection of shifts in the covariance matrix. Moreover, the proposed methodology not only allows the use of an estimator of the covariance matrix based on a single observation, but it also outperforms existing ones according to an extensive simulation study.

1 Introduction

Statistical process control methods play an important role in quality improvement. They have been widely applied in Engineering for a long time. The aim is to detect a structural change in the process of interest as soon as possible after its occurrence. However, this problem setting is also important in other fields and recently several papers discussed the applications of sequential procedures to Economics, Medicine, Chemistry, and Finance. The main tools of statistical process control are control charts. A control chart is characterized by a control statistic which is updated using current information at each time point.

Note that usually control charts are developed to detect shifts in the mean behavior of the process. Application to other areas makes the monitoring of the variance or the covariance matrix increasingly important. In general we can apply the techniques of the mean charts to different volatility measures. Usually squared observations, their logarithms or other transformations are used. This leads, however, to non-standard distributions of the control statistics and substantially complicates the monitoring process. This fact is even more critical in the multivariate case due to a large number of the components in the covariance matrix. In this paper we introduce a new technique, which allows us to apply the standard control charts for the mean directly to monitor the variance of a Gaussian vector. Moreover, the method can be applied to the estimators of the covariance matrix based on a single recent observation.

2 The method

As an estimator of the covariance matrix we use the point estimator based on a single observation, i.e. at time point t the covariance matrix is estimated by $\mathbf{V}_t = \mathbf{X}_t\mathbf{X}_t'$.

The matrix $\mathbf{V}_t$ follows a singular Wishart distributed $\mathbf{V}_t \sim W_p(1, \mathbf{\Sigma})$ (see, e.g. [1]), where $\mathbf{\Sigma} = \mathbf{\Sigma}_0$ if the process is in control and $\mathbf{\Sigma} = \mathbf{\Sigma}_\Delta$ if the process is out of control. Its rank is equal to one with probability one in both cases. It is not new to exploit the unbiased point estimator $\mathbf{V}_t$ for monitoring purposes. In [3] it is used $\mathbf{V}_t$ to update the matrix variate EWMA recursion.

However, the standard control schemes cannot be directly applied to the point estimator $\mathbf{V}_t$. For the EWMA chart, the variance of the control statistic has to be computed and this is a nontrivial task. Moreover, because the distribution of the control statistic is not symmetric, the two-sided EWMA chart for the variance depends on two critical values (see, e.g. [2]). This generates additional computational difficulties. In this paper we use the properties of the singular Wishart distribution to transform $\mathbf{V}_t$ to a set of Gaussian vectors. Then the mean charts can be immediately applied to monitoring the shifts in the variance.

3 Example

Using a Monte Carlo simulation study, we compare the derived control charts with the benchmark charts used to detect changes in the covariance matrix of the multivariate normal distribution. As the performance measures the average run length and the maximum expected delay are used. The simulation study shows that the proposed technique has a good performance. However, it is not possible to provide a unique ranking of the control procedures. While for positive changes in the variances the best results are reached the MC2 and the MEWMA control charts, for negative shifts the M1A2 and M1z2 are the best charts. For positive changes in the correlation the best results are obtained by the MC1 and the MCUSUM schemes.

References

[1] Bodnar, T. and Okrin, Y. (2008). Properties of the partitioned singular, inverse and generalized inverse Wishart distribution, *Journal of Multivariate Analysis*, **99**, 2389-2405.

[2] Chan, L.K. and Zhang, J. (2001). Cumulative sum control charts for the covariance matrix, *Statistica Sinica*, **11**, 767-790.

[3] Yeh, A.B., Huwang, L. and Wu, C.-W. (2005). A multivariate EWMA control chart for monitoring process variability with individual observations, *IIE Transactions*, **37**, 1023-1035.

Progressively Type-II Censored Lifetime Experiments

Marco Burkschat *Institute of Statistics, RWTH Aachen University, Germany*

Abstract: In a lifetime experiment, a progressively Type-II censored sample is obtained by withdrawing a number of surviving units after each failure. Apart from describing accidental losses of experimental units, progressive Type-II censoring may also be applied for the purpose of planning a life test in the sense of experimental design. The flexibility gained by considering different censoring schemes may be exploited in order to improve the outcome of the experiment with respect to a given optimality criterion. If it has been decided in advance to carry out a progressively Type-II censored experiment, the question arises which censoring scheme is optimal for a particular criterion. In the talk optimal censoring schemes for several time-based criteria are presented. Moreover, some recent results related to parameter estimation are discussed.

Estimators and Tests based on Likelihood Depth for Copulas and the Weibull-Distribution

Liesa Denecke *Department of Mathematics, University of Kassel, Germany*

Abstract: New estimators based on the Likelihood Depth for the examples of 2-dimensional, one-parametric Gumbel-Copula (and 2-dimensional Gauss-distribution), and in a second step for the Weibull-distribution, are presented. The Likelihood-Depth is a concept to generalize the one-dimensional median to multidimensional data or to data with more than one parameter. The parameters with maximum depth are the generalized medians. However, in some cases these generalized medians are biased estimators. This is in particular the case for the Gumbel-Copula, the Gauss- and for the Weibull-distribution. But the bias can be corrected leading to new robust estimators in all considered cases. Also tests for the parameters were developed. The power-functions of these tests were simulated exemplary for the cases above. The theory can be used also to find robust estimators for other distributions.

1 Introduction

The copula model has a variety of applications because it models dependence structures, e.g. in finance, in the analysis of credit risks. Copulas can also be used in the simulation of technical production processes to model the occurrence of coupled failures. For an introduction see [9]. Different estimation procedures for copulas were introduced, parametric, semi-parametric and nonparametric methods are proposed, see e.g. [1],[2] or [3].

The Weibull-distribution is often used in Survival Analysis, see for example [6]. It can model constant, de- and increasing Hazard-functions. Because of this and due to the fact, that the survival-function has a rather simple form, it is used in many applications. The distribution function is one-dimensional and depends on two parameters. Most times the Maximum-Likelihood-Estimator is used for parametric estimation, it can be found e.g. in [6].

We derive estimators and tests for one-parametric two-dimensional copulas and for the Weibull-distribution via likelihood depth and simplicial likelihood depth. These are rather general notions of data depth, see [7] and [8]. They extend the half space depth of Tukey ([10]) and the simplicial depth of Liu ([4],[5]) which lead to outlier robust generalizations of the median for multivariate data. They belong to a broad class of depth notions introduced and studied in the last 20 years. Although likelihood depth bases on a parametric approach, it can lead to distribution-free estimators and tests as [7] demonstrated for location-scale estimation and [8] for regression. [8] also showed that simplicial likelihood depth is in particular appropriate for testing since it

is an U-statistic. Thereby rather general hypotheses can be tested and the resulting tests are outlier robust.

Copulas are often given by distributional assumptions on the form of the copula. This distributional assumptions for the copula will be used here to define likelihood depth and simplicial likelihood depth for copulas. The approach is demonstrated for the Gaussian copula and the Gumbel copula for two dimensions which are based on one parameter only. In a next step, the Likelihood depth shall be used to find estimators and tests for the two parameters of the Weibull distribution. We will show that the two parameters can be estimated and corrected step by step, so the same methods as in the case of one parameter can be used.

References

[1] Genest, C.; Ghoudi, K.; Rivest, L.-P. (1995). A semiparametric estimation procedure in multivariate families of distribution, Biometrika 82, 543-552

[2] Hoff, P.D. (2007). Extending the rank likelihood for semiparametric copula estimation, *The Annals of Applied Statistics*, Vol.1, No. 1, June 2007

[3] Kim, G.; Silvapulle, M.J.; Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas, *Computational Statistics & Data Analysis* 51, 2836-2850

[4] Liu, R.Y. (1988). On a notion of simplicial depth, *Proc. Nat. Acad. Sci. USA* 85, 1732-1734

[5] Liu, R.Y. (1990). On a notion of data depth based on random simplices, *Ann. Statist.* 18, 405-414

[6] Lee, E. T., Wang, J. W. (2003). *Statistical Methods for survival data analysis*, Wiley 2003, New Jersey, Third Ed.

[7] Mizera, I.; Müller, Ch.H. (2004). Location-Scale Depth, *J.Am.Stat.Assoc.* 99, No. 468, 949-989

[8] Müller, Ch.H. (2005). Depth estimators and tests based on the likelihood principle with applicationsto regression, *J.Multivariate Anal.* 95, No.1, 153-181

[9] Nelsen, R.B. (2006). *An Introduction to Copulas*, Springer Series in Statistics, Springer 2006, Second Ed.

[10] Tukey, J.W. (1975). Mathematics and the picturing of data, *Proc. International Congress of MAthematicians 2*, Canad.Math.Congress, Montreal

Multiple Change-Point Estimation with U-Statistics

Maik Doering *Institute for Applied Mathematics and Statistics, University of Hohenheim, Germany*

Abstract: We consider a multiple change-point problem: In a sequences of independent random elements there occurs several changes of distributions at unknown places. Here we don't assume that the underlying distributions are known.

We propose a class of estimators of the unknown multiple change-point, that are maximizers of weighted multisample U-statistics. The aim of this work is the analysis of the asymptotic behavior of our estimators.

1 Introduction

We assume that the number of changes $q \in \mathbb{N}$ is known. So let $X_{1,n}, \ldots, X_{n,n}$, $n \in \mathbb{N}$ be a triangular array of rowwise independent random variables defined on a common probability space $(\Omega, \mathfrak{A}, P)$ with values in a measurable space $(\mathfrak{X}, \mathfrak{F})$. We assume that there exists a vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_q) \in \mathbb{R}^q$ with

$$0 = \theta_0 < \theta_1 < \cdots < \theta_q < \theta_{q+1} = 1$$

and distributions ν_i for $0 \leq i \leq q$, such that

$$P \circ X_{j,n}^{-1} = \nu_i \quad \text{for} \quad [n\theta_i] < j \leq [n\theta_{i+1}].$$

The goal is to estimate the unknown multiple change-point $\boldsymbol{\theta}$. Here nothing is known about the underlying distributions ν_i except that $\nu_{i-1} \neq \nu_i$ for $1 \leq i \leq q$.

U-statistics in change-point analysis has been introduced by Csörgö and Horvath, see for instance [1]. Our estimator is based on weighted $(q+1)$-sample U-statistics with a chosen kernel h of degree $\mathbf{m} = (m_0, \ldots, m_q)$, where the i-th sample consists of the random variables $X_{[nt_i]+1,n}, \ldots, X_{[nt_{i+1}],n}$. Thereby let $m_i \in \mathbb{N}_0$, $m := \sum_{i=0}^{q} m_i > 0$ and the kernel $h : \mathfrak{X}^m \to \mathbb{R}$ be a $\mathfrak{F}^m$-$\mathfrak{B}(\mathbb{R})$-measurable and suitable integrable function, which is symmetrical in each of the m_i coordinates. We define for $n \in \mathbb{N}$ and $\mathbf{t} \in H_{\mathbf{m},n}$, where

$$H_{\mathbf{m},n} := \left\{ \mathbf{t} \in \mathbb{R}^q : m_i \leq [nt_{i+1}] - [nt_i] \text{ for } 0 \leq i \leq q \text{ with } t_0 = 0,\ t_{q+1} = 1 \right\},$$

a $(q+1)$-sample U-statistic by

$$U_{n,\mathbf{t}}(h) := \prod_{i=0}^{q} \binom{[nt_{i+1}] - [nt_i]}{m_i}^{-1} \sum_{1 \leq j_1^0 < \cdots < j_{m_0}^0 \leq [nt_1]} \cdots \sum_{[nt_i] \leq j_1^i < \cdots < j_{m_i}^i \leq [nt_{i+1}]}$$
$$\cdots \sum_{[nt_q]+1 \leq j_1^q < \cdots < j_{m_q}^q \leq n} h\left(X_{j_1^0,n}, \ldots, X_{j_{m_0}^0,n}, \ldots, X_{j_1^q,n}, \ldots, X_{j_{m_q}^q,n} \right).$$

We will see that in the context of change-point estimation it is also worthwhile to work with weight functions $w : \mathbb{R}^q \to \mathbb{R}$. They are used to overcome boundary effects, which typically occur when any of the distances $\theta_{i+1} - \theta_i$ for $0 \leq i \leq q$ is close to zero. We define a sequence of stochastic processes $(\rho_n)_{n\in\mathbb{N}}$ with $\rho_n = \{\rho_n(\mathbf{t}) : \mathbf{t} \in \mathbb{R}^q\}$ by

$$\rho_n(\mathbf{t}) := \begin{cases} w\left(\frac{[nt_1]}{n}, \ldots, \frac{[nt_q]}{n}\right) U_{n,\mathbf{t}}(h) & \mathbf{t} \in H_{\mathbf{m},n} \\ 0 & \text{otherwise.} \end{cases}$$

We define a class of estimators for the change-point $\boldsymbol{\theta}$ by

$$\hat{\boldsymbol{\theta}}_n := argmax\left(|\rho_n(\mathbf{t})|,\ \mathbf{t} \in \left\{\left(\frac{k_1}{n}, \ldots, \frac{k_q}{n}\right) \,\middle|\, k_i \in \mathbb{N}_0, 1 \leq i \leq q\right\}\right).$$

This construction goes back to the work of Ferger [2, 3]. He considered the case $q = 1$ and $\mathbf{m} = (1, 1)$.

The aim of this talk is to analyze the asymptotic behavior of our estimators. Dependent on the order of integration of the chosen kernel function of the U-statistic we obtain consistency of our estimator. Furthermore it turns out that under suitable conditions $[n\boldsymbol{\theta}] - n\hat{\boldsymbol{\theta}}_n$ converge in distribution to a maximizer of a random walk with drift.

References

[1] Csörgö, M., Horvath, L. (1997). *Limit Theorems in Change-Point Analysis*, John Wiley, Chichester.

[2] Ferger, D.(1995). *Change-point estimators based on weighted empirical processes with application to the two sample problem in general measurable spaces*, Habilitationsschrift, University of Gießen (in German)

[3] Ferger, D. (2001). Exponential and polynomial tailbounds for change-point estimators, *J. Statist. Planning Inference* 92, 73-109.

Resampling Approach to the Estimation of Stochastic System Models

Maxim Fioshin *Department of Mathematical Methods of Transport Systems Control, Riga Technical University, Latvia*

Abstract: We review some basic results concerning the resampling method properties and its applications to the analysis of stochastic systems. This is one of non-parametrical intensive statistical methods. Historically, this method has been studied for reliability problems. However, it may be applied for a wide range of statistical problems, including regression and stochastic processes.

1 Introduction

Some statistical methods, called intensive statistical methods, require a big amount of computations instead of complex inference. This class includes such methods as Monte Carlo, Randomization, including Cross Validation and Jackknife, Bootstrap. We consider the Resampling, which is one of the intensive statistical methods.

The Resampling has a number of advantages:

- It does not have limitations to the system structure complexity;
- It does not require a prior information about the distributions of initial data;
- It works efficiently in the case of small initial samples;
- In the case of hierarchical structures it allows parallel computations.

These advantages make the resampling a good alternative to classical methods.

2 The method

Many problems in the estimation of stochastic systems may be reduced to the estimation of the expectation of some known function $\psi(X_1, X_2, \ldots, X_n)$ of n random arguments. In the case of insufficient information about arguments and complex function structure the resampling approach allows to obtain good estimators for $E\ \psi$. The method was originally developed by Ivnitsky [1], the present modification proposed by Andronov [2]. The method has some common properties with bootstrap, see [3].

Suppose $X_1, X_2, \ldots, X_n$ are independent r.v.s with unknown distributions $F_n(x)$. Only the small samples H_i are available for each X_i. The problem is to estimate $\mu = E\ \psi(X_1, X_2, \ldots, X_n)$, where $\psi(\cdot)$ is some known function.

To perform an estimation, we extract values from samples H_i, $i = 1, 2, \ldots, n$, forming the resamples H_i, $i = n+1, n+2, \ldots$. If the function $\psi(\cdot)$ has the complex structure, this procedure may be continued forming new resamples from existing resamples, until we obtain a realization of a value of the function ψ.

This procedure is reiterated r times, forming the resampling estimators μ^{*q}, $q = 1, 2, \ldots, r$ of μ. The resampling estimator μ^* is an average of all realizations

$$\mu^* \quad = \quad r^{-1} \sum\nolimits_{q=1}^{r} \mu^{*q}. \tag{1}$$

It maybe proved that under some general conditions

$$\mu^* \xrightarrow{d} \mu,$$

as $r \to \infty$. For details, we refer to [2].

The measure of the estimator efficiency is given by $Var\ \mu^*$. It depends on the moments $E\ \mu^*$ and $E\ [\mu^*]^2$, defined by the properties of $\psi(\cdot)$ and X as well as the mixed moment $E\ \mu^{*p}\mu^{*q}$, $p\widetilde{\chi}_q^0$, defined by the extraction procedure. Usually $E\ \mu^{*p}\mu^{*q}$ is calculated by conditioning it by different factors, for example, by a number of the same elements extracted for μ^{*p} and μ^{*q} from the initial samples.

3 Applications

There are different applications of the method [3], for example, estimation in the case of partially-known distributions, the confidence intervals construction, renewal processes analysis, regression models analysis, analysis of reliability systems, analysis of shot-noise processes etc.

References

[1] Ivnitsky V. (1967). Approach to Statistical Estimation of System Reliability (in Russian), *Engineering Cybernetics* No. 4.

[2] Andronov A. and Merkuryev Yu. (2000). Optimization of Statistical Sample Sizes in Simulation *Journal of Statistical Planning and Inference*, 85: 93-102.

[3] Davison A.C. and Hinkley D.V. (1997). *Bootstrap Methods and their Application*, Cambridge university Press: UK.

[4] Andronov A., Fioshina H. and Fioshin M. (2009). Statistical Estimation for a Failure Model with Damage Accumulation in a Case of Small Samples. *Journal of Statistical Planning and Inference*, in press, doi:10.1016/j.jspi.2008.05.026.

The Resampling Approach to the Estimation of a Reliability Model Based on Shot-Noise Processes

Helen Fioshina *Department of Mathematical Methods of Transport Systems Control, Riga Technical University, Latvia*

Abstract: A failure model based on shot-noise processes is considered. Two sample populations are available: a sample of intervals between shocks and a sample of magnitude values. The purpose is to estimate the expectation of the system stress level at time t. We consider the plug-in and resampling estimators of the above mentioned characteristics.

1 Introduction

We consider a reliability system, subjected to shocks [1]. The shocks occur in accordance to a homogeneous Poisson process with rate λ. The i.i.d. r.v. X_i, independent on this process, represent shocks magnitude. The magnitudes of shocks are additive and decrease over time at a deterministic function $h(t) = exp(-\alpha t)$.

We are interested in the total shock value or stress level at time t

$$X(t) = \sum_{i=1}^{N(t)} X_i exp(-\alpha(t - S_i)). \tag{1}$$

Let $E[X(t)]$ be an expectation of $X(t)$. It is known [2] that it is

$$E[X(t)] = \lambda E[X](1 - exp(-\alpha t))/\alpha. \tag{2}$$

The rate λ and the cdf $F(x)$ are unknown, but the sample $\mathbf{A}$ of the intervals between shocks and the sample $\mathbf{B}$ of shock values are given. We need to estimate $E[X(t)]$.

2 The method

We consider two estimation methods: the *plug-in* and the *resampling* estimators [3]. The plug-in approach uses the estimates $\hat{\lambda}$ and $\hat{E}[X]$ obtained from samples $\mathbf{A}$ and $\mathbf{B}$. Using them instead of true values in (2), one obtains the estimator

$$\hat{\mu}_{X_t} = \hat{\lambda}\hat{E}[X]\frac{(1 - exp(-\alpha t))}{\alpha}. \tag{3}$$

In the resampling approach the initial variables are randomly extracted from the samples $\mathbf{A}$ and $\mathbf{B}$. In the q-th realization we extract elements from $\mathbf{A}$ and $\mathbf{B}$ without

replacement and calculate the shock appearance times $\{S_i^{*q}\}$ and shock magnitudes $\{B_i^{*q}\}$. Then we construct trajectory of the process and calculate the stress level

$$X^{*q}(t) = \sum_{i:S_i \le t} B_i^{*q} exp(-\alpha(t - S_i^{*q})). \tag{4}$$

The procedure is reiterated r times. The average of $X^{*q}(t)$ is the resampling estimator of μ_{X_t}. Our purpose is to obtain the expressions for properies of the plug-in and resampling estimators, such as the expectation, variance and mean squared error.

3 Example

Let us consider a Poisson process of shocks with rate λ=0.66, exponential distribution of shock magnitudes with parameter $\lambda_B = 1$ and the rate $\alpha = 0.1$. The sizes of the samples **A** and **B** are supposed to be equal: $n = n_A = n_B$. The number of resamples $r = 3000$. The comparison of the plug-in and resampling estimators was performed, see Fig. 1

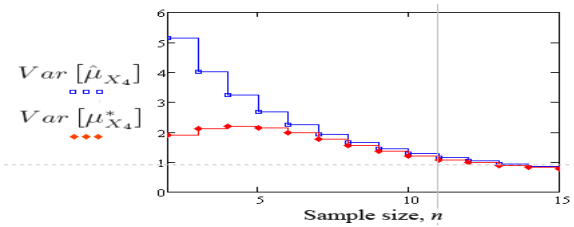

Figure 1: Variance of the plug-in and resampling estimators of μ_4.

We can see that in this case the resampling estimator has smaller variance, espetially in the case of small sample sizes.

References

[1] Cox D.R. and Isham V. (1980). *Point Processes*. Chapman and Hall, London - New York.

[2] Ross Sheldon M. (1996). *Stochastic processes*. John Wiley & Sons, New York.

[3] Andronov A., Fioshina H. and Fioshin M. (2009). Statistical Estimation for a Failure Model with Damage Accumulation in a Case of Small Samples. *Journal of Statistical Planning and Inference*, in press, doi:10.1016/j.jspi.2008.05.026.

A Prediction Problem in Interval-Censored Nonhomogeneous Poisson Processes

Jürgen Franz *TU Dresden, Germany*

Let us consider a failure-repair model where minimal repair actions and neglected repair times are assumed. The model is described by counting processes, especially by nonhomogeneous Poisson processes. Moreover it is assumed that failures can only be observed in an interval $[\sigma, \tau]$. As an additional information the number of failure events before σ is known.

1 Description of the model

Let $(T_i)_{i\geq 1}$ be the sequence of failure time points on a probability space $[\Omega, \mathfrak{F}, \mathrm{P}]$. Then, $\{N(t), t \geq 0\}$ is the corresponding counting process with

$$N(t) = \sum_{i=1}^{\infty} \mathbf{1}\{T_i \leq t\}$$

and is assumed to be a nonhomogeneous Poisson process (NHPP) with intensity function $\lambda(t, \theta)$, cumulative intensity $\Lambda(t, \theta) = \int_0^t \lambda(u, \theta) du < \infty$ and

$$P(N(t) - N(s) = k) = \frac{[\Lambda(t) - \Lambda(s)]^k}{k!} \exp\{-[\Lambda(t) - \Lambda(s)]\}$$

(k=0,1,2,...)

Example: Weibull process: $\lambda(t, \theta) = \alpha \cdot \beta t^{\beta-1} \quad (t \geq 0,\ \theta = (\alpha, \beta) \in \mathbb{R}_+^2)$

For statistical investigations we need the likelihood function. For the considered NHPP we propose that the intensity structure has the factorization $\lambda(t, \theta) = \alpha \cdot \mu(t, \beta))$ $(t \geq 0;\ \alpha > 0, \beta > 0)$. Then

$$L(\theta) \quad = \quad \frac{M(\sigma, \beta)^m}{m!} \alpha^n \prod_{i=m+1}^{n} \mu(t_i, \beta) \exp[-\alpha M(\tau, \beta)]$$

where t_i are the observed failure time points, $N(\sigma) = m$, $N(\tau) = n$ and $M(t, \beta) = \int_0^t \mu(u, \beta) du$.

(Weibull-process case: $L(\theta) = \frac{\alpha^m \sigma^{m\beta}}{m!} (\alpha\beta)^{n-m} \prod_{i=m+1}^{n} t_i^{\beta-1} \exp[-\alpha\tau^\beta]$.)

2 Parameter estimation

Based on the likelihood function the ML-parameter estimators are obtained, for instance in the Weibull-process case:

$$\widehat{\alpha}_{ML}(\tau) = \frac{N(\tau)}{\tau^{\widehat{\beta}_{ML}(\tau)}}, \qquad \widehat{\beta}_{ML}(\tau) = \frac{N(\tau) - m}{N(\tau)\ln\tau - m\ln\sigma - \sum_{i=m+1}^{N(\tau)} \ln T_i}$$

Using the quadratic loss function $V(\theta, \widehat{\theta}(\tau)) = (\theta - \widehat{\theta}(\tau))^2$ and the (semi-conjugate) prior density

$$q(\theta) = \frac{a^b}{\Gamma(b)}\alpha^{b-1}e^{-a\alpha} \cdot p(\beta),$$

the posterior density (Weibull-process case)

$$\widetilde{q}_\tau(\theta) = \frac{(a+\tau^\beta)^{b+n}}{\Gamma(b+n)}\alpha^{b+n-1}e^{-(a+\tau^\beta)\alpha} \cdot \widetilde{p}_\tau(\beta)$$

with $\widetilde{p}_\tau(\beta) = \frac{w(\beta)}{\int w(\beta)d\beta}$ and $w(\beta) = \frac{p(\beta)\sigma^{m\beta}\beta^{n-m}\prod_{i=m+1}^{n} t_i^{\beta-1}}{(a+\tau^\beta)^{b+n}}$,

and the Bayes point estimators $\widehat{\theta}(\tau) = \int_\Theta \theta\widetilde{q}_\tau(\theta)d\theta$ follows, especially:

$$\widehat{\alpha}_B(\tau) = \int_0^\infty \frac{N(\tau)+b}{\left(a+\tau^{\widehat{\beta}_B(\tau)}\right)}\widetilde{p}_\tau(\beta)d\beta \qquad \widehat{\beta}_B(\tau) = \int_0^\infty \beta\widetilde{p}_\tau(\beta)d\beta$$

3 Bayes prediction

The process is observed up to the time τ with the state $N(\tau) = n$. Now we are interested in approximate values of future time points T_{n+k} (k=1,2,...).
For any k the Bayes prediction density of T_{n+k} is:

$$\begin{aligned} g_{T_{n+k}}(t) &= \frac{\Gamma(n+b+k)}{\Gamma(n+b)\Gamma(k)}\int_0^\infty \mu(t,\beta)\widetilde{p}_\tau(\beta)Q(t,s,\beta)d\beta \\ Q(t,s,\beta) &= \frac{[M(t,\beta)-M(s,\beta)]^{k-1}[M(\tau,\beta)+a]^{n+b}}{[M(t,\beta)-M(s,\beta)+M(\tau,\beta)+a]^{n+b+k}} \end{aligned}$$

As Bayes point estimator for T_{n+k} could used

$$\widehat{T_{n+k}} = \int_{T_n}^\infty t g_{T_{n+k}}(t)dt.$$

Bayes prediction intervals I_B for T_{n+k} given the cover probability $1-\varepsilon$ are obtained by

$$\int_s^{t_*} g_{T_{n+k}}(t)(t)dt = \varepsilon_1, \quad \int_s^{t^*} g_{T_{n+k}}(t)(t)dt = 1-\varepsilon_2, \quad \varepsilon_1+\varepsilon_2 = \varepsilon$$

and $I_B = [t_*, t^*]$.
Numerical results are given.

Risk-Adjusted Monitoring of Time to Event

Axel Gandy *Department of Mathematics, Imperial College London, U.K.*

Abstract: Statistical process control has been a common tool in industry for some time. Recently there has been interest in risk-adjusted cumulative sum charts (CUSUM) to monitor the performance of e.g. hospitals, taking into account the heterogeneity of the patients. Even though many outcomes involve time, only conventional regression models are being commonly used. In this talk we investigate how time to event/survival analysis models can be used for monitoring purposes and in what situations the new methods lead to shorter alarm times. We consider monitoring using CUSUMs based on the partial likelihood ratio between an out-of-control state and an in-control state. We consider both proportional and non-proportional alternatives, as well as a head start. Against proportional alternatives, we present an analytic method of computing average run lengths. We present two examples, one from the retail financial industry and one medical application.

Nonstationary-Volatility Robust Panel Unit Root Tests and the Great Moderation

Christoph Hanck *Department of Quantitative Economics, Universiteit Maastricht, NL*

Abstract: This paper proposes a new testing approach for panel unit roots that is, unlike previously suggested tests, robust to nonstationarity in the volatility process of the innovations of the time series in the panel. Nonstationarity volatility arises for instance when there are structural breaks in the innovation variances. A prominent example is the reduction in GDP growth variances enjoyed by many industrialized countries, known as the 'Great Moderation.' The panel test is based on the classical multiple test of [1], which combines evidence from time series unit root tests of the series in the panel. As time series unit root tests, we employ recently proposed tests of [2]. The panel test is robust to general patterns of cross-sectional dependence and yet straightforward to implement, only requiring valid p-values of time series unit root tests, and no resampling. Monte Carlo experiments show that other panel unit root tests suffer from sometimes severe size distortions in the presence of nonstationary volatility, and that this defect can be remedied using the test proposed here. The new test is applied to test for a unit root in an OECD panel of gross domestic products, yielding inference robust to the 'Great Moderation.' We find little evidence of trend stationarity.

References

[1] Simes, R.J. (1986). An improved Bonferroni procedure for multiple tests of significance, *Biometrika* 73(3):751–754.

[2] Cavaliere, G. and Taylor, A. M. R. (2008). Time-transformed unit root tests for models with non-stationary volatility, *Journal of Time Series Analysis* 29(2):300–330.

Identifying Sources of Business Cycle Fluctuations in Germany 1975-1998

Oliver Holtemöller *Faculty of Business and Economics, RWTH Aachen University, Germany*
Torsten Schmidt *RWI Essen, Germany*

Abstract: In this paper, we estimate a small New Keynesian dynamic stochastic general equilibrium (DSGE) model for Germany for the period from 1975 to 1998 and use it to identify the structural shocks, which have driven the business cycle. For this purpose we apply indirect inference methods, that is we specify the parameters of the theoretical model such that simulated data mimics observed data as closely as possible. In addition to the identification of structural shocks, we uncover the unobservable output gap, which is a prominent indicator in business cycle analysis. Furthermore, we show to which extent each identified shock has contributed to the business cycle fluctuations.

Two-Sample Survival Tests with Estimated Scores

Arnold Janssen *Heinrich-Heine Universität Düsseldorf, Germany*

Abstract: In this talk we consider two-sample testing problems for randomly censored survival data. There exist efficient two-sample tests for all kind of semiparametric hazard models for a given direction of hazard alternatives. The tests are weighted log-rank tests, see Fleming and Harrington. For instance they can be used to separate proportional hazards or early (late, central) hazard differences but the statistician should know the shape of the alternative. In this work an adaptive testing procedure is proposed with good power for whole cones of alternatives. The advantage of the tests is explored by two examples given by real data.

References

[1] Andersen, P.K.; Borgan, O.; Gill, R.D.; Keiding, N. (1993). *Statistical models based on counting processes.* Springer, New York.

[2] Behnen, K.; Neuhaus, G. (1989). *Rank tests with estimated scores and their application.* Teubner-Skripten zur Mathematischen Stochastik. Stuttgart.

[3] Fleming, T.R.; Harrington, D.P. (1991). *Counting processes and survival analysis.* Wiley, New York.

[4] Janssen, A. (1991). Conditional rank tests for randomly censored data. Ann. Stat. **19**, 1434-1456.

[5] Janssen, A.; Mayer, C.-D. (2001) Conditional Studentized survival tests for randomly censored models. Scand. J. Stat. **28**, 283-293.

[6] Janssen, A.; Werft, W. (2004) *A survey about the efficiency of two-sample survival tests for randomly censored data.* Mitteilungen aus dem Mathematischen Seminar Giessen **254**, 1-47.

[7] Klein, J.P.; Moeschberger, M.L. (2003) *Survival analysis. Techniques for censored and truncated data. 2nd ed.* Statistics for Biology and Health. New York.

[8] Mayer, C.-D. (1996). *Projektionstests für das Zweistichprobenproblem mit zensierten Daten.* Dissertation, Heinrich-Heine Universität Düsseldorf.

[9] Neuhaus, G. (1993). Conditional rank tests for the two-sample problem under random censorship. Ann. Stat. **21**, 1760-1779.

On Change-Point Models in Survival Analysis with Applications in Reliability

Uwe Jensen *Department of Applied Mathematics and Statistics, University of Hohenheim, Stuttgart, Germany*

Abstract: A Cox-type regression model is considered which admits change-points in the covariates. A change-point specifies the unknown threshold at which the influence of a covariate shifts smoothly, i.e. the regression parameter may change over the range of a covariate and the underlying regression function is continuous but not differentiable. The model can be used to describe change-points in different covariates but also to model more than one change-point in a single covariate. Estimates of the change-points and of the regression parameters are derived and their properties are investigated. It is shown that not only the estimates of the regression parameters are $\sqrt{n}$-consistent but also the estimates of the change-points in contrast to the conjecture of other authors. Asymptotic normality is shown by using results developed for M-estimators. Finally the model is applied to data sets which stem from engineering experiments.

1 Introduction

Consider a multivariate counting process $\mathbf{N}(t) = (N_1(t), \ldots, N_n(t))$, where $N_i(t)$ counts observed events in the life of the ith individual, $i = 1, \ldots, n$, over the time interval $[0, \tau]$. The sample paths of $\mathbf{N}(t)$ are step functions, zero at time zero with jumps of size one only and no two components jump at the same time. The counting process $\mathbf{N}(t)$ admits an intensity $\boldsymbol{\lambda}(t) = (\lambda_1(t), \ldots, \lambda_n(t))$ such that the processes $M_i(t) = N_i(t) - \int_0^t \lambda_i(u)du$, $i = 1, \ldots, n$, and $t \in [0, \tau]$ are martingales. Different models are determined by their intensities. The intensity of the basic Cox model with baseline hazard $\lambda_0(t)$ and covariate vector $\mathbf{Z}(t)$ is given by $\lambda(t) = \lambda_0(t) \exp\{\boldsymbol{\beta}_0^T \mathbf{Z}(t)\}$. In this model it is assumed that the influence of a covariate is constant in time and over the range of the covariate. By analyzing different datasets we found out that some covariates exhibit deviations from this assumption. Therefore, we proposed a new variant of the Cox model with a smooth change at an unknown threshold ξ (see [1]).

In the literature several extensions of the Cox model have been investigated. Among these extensions is a model introduced by Pons [2]

$$\lambda(t) = \lambda_0(t) \exp\{\boldsymbol{\alpha}^T \boldsymbol{Z}_1(t) + \boldsymbol{\beta}^T \boldsymbol{Z}_2(t) I_{\{Z_3 \leq \zeta\}} + \boldsymbol{\gamma}^T \boldsymbol{Z}_2(t) I_{\{Z_3 > \zeta\}}\},$$

where the influence of a covariate jumps at a certain threshold ζ. The estimate of the jump change-point parameter was shown to be n-consistent in this case.

2 The model

In our model we allow for more than only one smooth change-point, the covariates may be time-dependent and the counting process may jump more than once. The model involving m change-points and p ordinary covariates (without change-points) is given as follows:

$$\lambda_i(t, \boldsymbol{\theta}) = \lambda_0(t) R_i(t) \exp\left\{ {\boldsymbol{\beta}_1}^T \mathbf{Z}_{1i}(t) + \boldsymbol{\beta}_2^T \mathbf{Z}_{2i}(t) + \boldsymbol{\beta}_3^T (\mathbf{Z}_{2i}(t) - \boldsymbol{\xi})^+ \right\},$$

where $\boldsymbol{\theta} = (\boldsymbol{\xi}^T, \boldsymbol{\beta}^T)^T$ with $\boldsymbol{\xi} \in \boldsymbol{\Xi} \subset \mathbb{R}^m$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T, \boldsymbol{\beta}_3^T)^T \in \mathcal{B} \subset \mathbb{R}^{p+2m}$. Here $\boldsymbol{\xi}$ and $\boldsymbol{\beta}$ are the vectors of change-points and regression parameters respectively, $\lambda_0(t)$ is the baseline intensity and $R_i(t)$ is a process taking only values 1 or 0 to indicate whether a subject is at risk or not.

The parameter vector $\boldsymbol{\theta}_0$ is estimated by the value $\hat{\boldsymbol{\theta}}_n$ that maximizes the logarithm of the partial likelihood. In particular, we show that the estimates of the change-points are only $\sqrt{n}$-consistent and not n-consistent as one might have guessed. As usual, the cumulative hazard function $\Lambda_0(t) = \int_0^t \lambda_0(u)du$ is estimated by the Breslow estimator $\hat{\Lambda}_n(t)$.

Besides consistency we can show asymptotic normality of the estimators by means of results developed for M-estimators (see [3]).

3 Applications

We applied our model to some datasets. Among these are an actuarial dataset and the well known PBC dataset. It can be shown by means of a goodness-of-fit test that we get a better fit of the data using our change-point model instead of using the classical models. In addition we analysed datasets consisting of survival times and covariates recorded for electric motors and transmissions.

References

[1] Jensen U., Lütkebohmert, C. (2008). A Cox-type regression model with change-points in the covariates, *Lifetime Data Analysis* 14: 267 - 285.

[2] Pons, O. (2003). Estimation in a Cox regression model with a change-point according to a threshold in a covariate, *Ann. Stat.* 31: 442 - 463.

[3] Van der Vaart, A. (1998). *Asymptotic Statistics*, Cambridge University Press, New York.

Bayes Sequential Estimation for a Time-Transformed Exponential Model

Alicja Jokiel-Rokita *Institute of Mathematics and Computer Science, Wroclaw University of Technology, Poland*

Abstract: A problem of Bayesian sequential estimating an unknown parameter of a time transformed exponential model is considered. It is assumed that the loss associated with the error of estimation is a weighted squared or precautionary loss, and the cost of observing the process is a linear function of time and the number of observations. Bayes sequential procedures for estimating the unknown parameter are presented.

1 Introduction

The paper deals with the problem of sequentially estimating the parameter ϑ of the distribution defined by the density

$$f(x;\vartheta) = s'(x)\vartheta \exp[-\vartheta s(x)]\mathbf{1}_{(\mu,\infty)}(x), \tag{1}$$

where $s(x)$ is a strictly increasing and differentiable function with $\lim_{x\to\infty} s(x) = \infty$ and $s(\mu) = 0$. The parameter ϑ is unknown and μ is known. This subclass of exponential family is called the time-transformed exponential model (see [2]). It covers many distributions serving as lifetime distributions in reliability models. It is easy to show that the family of gamma distributions $\mathcal{G}(\alpha, \beta)$ with density

$$\Gamma(\alpha)^{-1}\beta^{\alpha}\vartheta^{\alpha-1}\exp(-\beta\vartheta)\mathbf{1}_{(0,\infty)}(\vartheta),$$

where $\alpha, \beta > 0$ are known parameters, is the conjugate one to the family of distributions given by (1).

We want to estimate the parameter ϑ on the basis of at most n independent observations $X_1, \ldots, X_n$ from the distribution given by (1). If the observation process is stopped at time t, the values of $X_1, \ldots, X_n$ not exceeding t are exactly known, whereas the other X_i are only known to be larger than t. The observations available in this way are sometimes called in literature as longitudinal observational data (see [1]).

Denote $\mathcal{F}_t = \sigma\{(X_1 \wedge t, \mathbf{1}_{(0,t]}(X_1)), \ldots, (X_n \wedge t, \mathbf{1}_{(0,t]}(X_n))\}$, for $t \geq 0$. If observation is stopped at time t, the loss incurred is defined by

$$L_t(\vartheta, d(t)) = L(\vartheta, d(t)) + c_A K(t) + c(t),$$

where $L(\vartheta, d(t))$ denotes the loss associated with the error of estimation, when ϑ is the true value of the parameter and $d(t)$ is the chosen $\mathcal{F}_t$-measurable estimate; $K(t) := \sum_{i=1}^{n} \mathbf{1}_{(0,t]}(X_i)$, c_A is a known nonnegative constant, $c(t)$ is a known nondecreasing function with $c(0) = 0$. We assume the following loss function

$$L(\vartheta, d(t)) = \frac{(\vartheta - d(t))^2}{d^k(t)\vartheta^r}, \tag{2}$$

where $k \in \{0, 1, 2\}$, $r \in \{0, 1, 2\}$. For $k = 0$ we have the usual weighted squared loss function. For $k \neq 0$ we obtain the precautionary loss function introduced in [4]. The condition $k \leq 2$ ensures that the cost increases as the difference $\vartheta - d(t)$ grows. This loss function approaches infinity near the origin to prevent underestimation and thus gives a conservative estimator, especially when low failure rates are being estimated. This estimator is useful when underestimation may lead to serious consequences. A precautionary index k indicates how downside damaging the loss function is.

The special models with $s(x) = x$ and the estimation loss function $\vartheta^{-r}(\vartheta - d(t))^2$ for some $r \geq 0$ have been treated in the papers [3] and [5].

In the paper a class of Bayes sequential procedures for estimating the unknown parameter ϑ is derived, under the loss given by (2) with $k \in \{0, 1, 2\}$ and $r \in \{0, 1, 2\}$.

References

[1] Arjas, E. (1989): Survival models and martingale dynamics. *Scand. J. Statist.*, *16*, 177–225.

[2] Barlow, R.E. and Proschan, F. (1988): Life distribution models and incoplete data. In: P.R. Krishnaiah and C.R. Rao (Eds.): *Handbook of Statistics*. Elsevier Science Publishiers B.V., 225–249.

[3] Chen, E.E. and Wardrop, R.L. (1980): Bayes sequential estimation in a life test and asymptotic properties. *Comm. Statist. Theor. Meth.*, *A9*, 659–672.

[4] Norstrøm, J.G. (1996): The use of precautionary loss functions in risk analysis. *IEEE Trans. Reliability*, *45*, 400–403.

[5] Stadje, W. (1990): A sequential estimation procedure for the parameter of an exponential distribution. *Statistics 21*, 239–250.

Optimal General Maintenance for Discrete Lifetime Distributions

Waltraud Kahle *Institute of Mathematical Statistics, Otto-von-Guericke-University Magdeburg, Germany*

Abstract: In many applications the clock time is not the best scale in which to describe lifetimes. As example, if a unit has a sequence of tasks to perform, then its lifetime is measured by the number of tasks performed before its failure. We consider discrete lifetime distributions. When the item fails, it is minimally repaired. Additionally, from time to time maintenance actions are carried out. We assume, that these maintenance actions do not renew the item, but reset the age of the item to some level between zero and the real age. An optimal maintenance policy will be found for some special cases.

1 The model

We consider an discrete lifetime distribution. As an example, let T be Poisson distributed:

$$\mathrm{P}(T=t) = \frac{\lambda^t}{t!} e^{-\lambda} , \quad t = 0, 1, \dots . \tag{1}$$

The failure rate for discrete distributions (and especially for the Poisson distribution) is given by

$$h(t) = \frac{\mathrm{P}(T=t)}{\mathrm{P}(T \geq t)} = \frac{\lambda^t/t!}{\sum_{k=t}^{\infty} \lambda^k/k!}, \quad t = 0, 1, \dots .$$

It can be shown, that the failure rate of the Poisson distribution is increasing.

2 Optimal maintenance interval

Now, we assume that at failure the item is repaired minimally, that is, after repair the age of the item is the same as just before failure. The cost of a failure is c_F. Further, we have preventive maintenance actions at time point τ, 2τ, ... Let be c_M the costs of an maintenance action. We do not make the assumption that after maintenance the item is as good as new. Every maintenance action leads to an age v, $v = 1, \dots, \tau$. Then the costs per time unit are given by

$$C(\tau) = \frac{c_F \sum_{k=v}^{\tau-1+v} h(k) + c_M}{\tau}, \quad c_F > c_M , \tag{2}$$

where v is the age of the item after maintenance. It can be proved that $C(\tau)$ has an unique minimum.

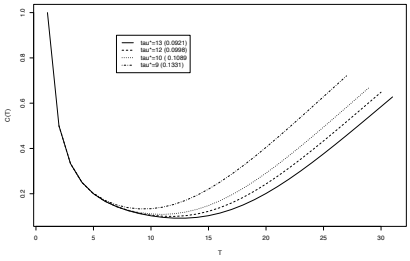

Figure 1: Cost functions for different impact of repairs

In figure 1 are shown the costs for $\lambda = 20$, different impact of repairs $v = 0, 1, 2, 4$ and $c_F/c_M = 5$. It can be seen that with increasing v the optimal time between preventive maintenances decreases and the costs are increasing.

3 Optimal degree of repair

Now, let us consider an other problem. We assume now that the time interval τ between preventive maintenance actions is fixed. Our aim is to find the optimal degree of repair, that is, the optimal age after pm. For this, we must describe the costs of an pm in dependence on the degree of repair. Instead of c_M in (2) we assume the costs of a pm to be

$$c_M \cdot \left(\frac{\tau - v}{\tau}\right)^{\alpha}.$$

For $v = 0$ we get the full costs c_M and for $v = \tau$ the costs are 0. For $\alpha = 1$ the costs are linear in v, otherwise we have an convex ($\alpha > 1$) or a concave ($\alpha < 1$) function.

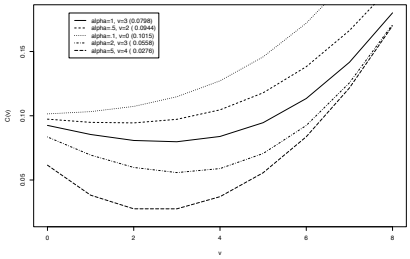

Figure 2: Cost functions for different *alpha*

In figure 2 are shown the costs for $\lambda = 20$, $\tau = 10$, $c_F/c_M = 5$, and different α. For very small α we get that the optimal maintenance is an renewal and the optimal optimal age after pm increases with α.

Two-Dimensional Diffusion Processes as Models in Lifetime Studies

Mario Lefebvre *Department of Mathematics and Industrial Engineering, École Polytechnique de Montréal, Canada*

Abstract: Let $X(t)$ denote the remaining useful lifetime of a machine, and let $Y(t)$ be a standard Brownian motion. Assume that the derivative $\rho[X(t), Y(t)]$ of $X(t)$ is a deterministic function of (at least) $Y(t)$. We consider the two-dimensional degenerate diffusion process $(X(t), Y(t))$. We obtain explicit expressions for the expected value of the random variable $T(x, y)$ denoting the first time the machine must be replaced or repaired for various functions $\rho[X(t), Y(t)]$.

1 Introduction

Assume that a given product possesses a certain quality characteristic, whose value at time t is denoted by $D(t)$, that is closely correlated with its lifetime L, so that L can be defined as follows:

$$L = \inf\{t > 0 : D(t) \leq c\},$$

where c is the critical level for $D(t)$.

Tseng and Peng [3] assumed that $D(t)$ is a decreasing function of t, and that the continuous time (one-dimensional) stochastic process $\{D(t), t \geq 0\}$ satisfies the following equation:

$$D(t) = M(t) + \int_0^t s(u)\, dB(u), \tag{1}$$

in which $M(t)$ is the mean value of $D(t)$, $\{B(t), t \geq 0\}$ is a standard Brownian motion, and $s(u)$ could be a constant. However, a stochastic process that satisfies (1) can be *both* increasing and decreasing on any interval.

In a related problem, Rishel [2] proposed to model the wear $X(t)$ of a machine at time t by the system of two stochastic differential equations

$$dX(t) = \rho[X(t), Y(t)]\, dt, \tag{2}$$

$$dY(t) = f[X(t), Y(t)]\, dt + \sigma[X(t), Y(t)]\, dB(t), \tag{3}$$

where ρ, f and σ are differentiable functions, and ρ and σ are non-negative in the domain of interest. In this model, $Y(t)$ is a variable that directly influences the wear, for instance the temperature or the operating speed of the machine.

In [1], the author computed the expected value of

$$\tau(x,y) := \inf\{t > 0 : X(t) = 0 \mid X(0) = x, Y(0) = y\}$$

when $\rho[X(t), Y(t)] = X^j(t)\, Y^k(t)$, with $j \in \{0, -2, -4, \ldots\}$ and $k \neq 0$, and the process $\{Y(t), t \geq 0\}$ is a geometric Brownian motion.

Let

$$m(x,y) := E\left[T(x,y)\right],$$

where $T(x,y)$ denotes the first time that the machine must be replaced or repaired. In the current paper, we will consider the case when $\{Y(t), t \geq 0\}$ is a standard Brownian motion and we will solve the Kolmogorov backward equation satisfied by the function $m(x,y)$, namely

$$\frac{1}{2}\, m_{yy}(x,y) + \rho(x,y)\, m_x(x,y) = -1$$

(subject to the appropriate boundary conditions) for various functions $\rho(x,y)$.

References

[1] Lefebvre, M. (2009). Mean first-passage time to zero for wear processes. (Submitted for publication)

[2] Rishel, R. (1991). Controlled wear processes: modeling optimal control, *IEEE Trans. Automat. Contr.* 36: 1100-1102.

[3] Tseng, S. T. and Peng, C. Y. (2007). Stochastic diffusion modeling of degradation data, *J. Data Sci.* 5: 315-333.

On Estimation of Parameters of Nonhomogeneous Poisson Process Models for Software Reliability

Alicja Jokiel-Rokita *Institute of Mathematics and Computer Science, Wroclaw University of Technology, Poland*
Ryszard Magiera *Institute of Mathematics and Computer Science, Wroclaw University of Technology, Poland*

Abstract: We consider a subclass of nonhomogeneous Poisson processes (NHPP's) which besides of its theoretically interesting structure it can be used to model software reliability. In contrast to the maximum likelihood (ML) method, we propose two alternative methods which yield satisfactory estimates of unknown parameters of the process models and can be applied when the ML method fails.

1 Introduction

A NHPP is considered with the intensity function $\lambda(t)$ and the mean value function $\Lambda(t) = \int_0^t \lambda(u)du$ having the following parametric form

$$\Lambda(t;\alpha,\beta) = \alpha F(t/\beta). \tag{1}$$

The parameter α is called the expected number of faults to be eventually detected (see [3]), β is a scale parameter, and $F(t)$ is a known increasing function of t. It is assumed that $F(0) = 0$ and $F(\infty) = 1$.

The model defined by (1) has bounded mean value function. Because a software system contains only a finite number of faults, this model is preferable in comparison with NHPP software reliability models with unbounded mean value function.

As a special case of the model (1) we consider the NHPP model defined by

$$\Lambda(t;\alpha,\beta) = \alpha\left[1 - \exp(-t/\beta)\sum_{i=0}^{k}\frac{(t/\beta)^i}{i!}\right], \quad \alpha,\beta > 0, \tag{2}$$

or equivalently, by

$$\lambda(t;\alpha,\beta) = \frac{\alpha(t/\beta)^k}{\beta k!}\exp(-t/\beta).$$

The model (2) was first mentioned in the paper [1] and is called the k-stage Erlangian NHPP software reliability model.

2 The LS and CLS methods as alternatives to the ML method

The ML estimators of the parameters α and β of the model (1) do not always exist. In particular, it follows from Theorem 2.1 of [4] that for the model (2) the ML estimators do not exist with the probability $P\Big(\frac{1}{N(T)}\sum_{i=1}^{N(T)} t_i \geq \frac{k+1}{k+2}T\Big)$, where $N(T)$ is the number of arrives up to time T and $t_1, \ldots, t_{N(T)}$ are the observed arrival times.

We propose the following two alternative methods for evaluating estimators of the unknown parameters α and β: 1) the least squares (LS) method consisting in determining the values of α and β that minimize the quantity

$$\sum_{i=1}^{N(T)} [\Lambda(t_i; \alpha, \beta) - \Lambda(t_{i-1}; \alpha, \beta) - 1]^2; \tag{3}$$

2) the constrained least squares (CLS) method consisting in determining the values of α and β that minimize the quantity (3) subject to the constraint $\frac{1}{N(T)}\sum_{i=1}^{N(T)}[\Lambda(t_i; \alpha, \beta) - \Lambda(t_{i-1}; \alpha, \beta)] = 1$.

The LS and/or the CLS method can be used when the ML method fails. The LS and CLS estimates are compared with the ML estimates and their accuracy is illustrated numerically for an Erlangian NHPP software reliability model.

Based on the papers [2] and [4] we discuss the problem of inconsistency of estimators of some parameter functions in the NHPP models considered.

References

[1] Khoshgoftaar, T. (1988). Nonhomogeneous Poisson processes for software reliability growth. In *COMSTAT'88*, pages 13–14, Copenhagen, Denmark.

[2] Nayak, T. K., Bose, S., and Kundu, S. (2008). On inconsistency of estimators of parameters of non-homogeneous Poisson process models for software reliability. *Statistics & Probability Letter*, 78:2217–2221.

[3] Yamada, S. and Osaki, S. (1985). Software reliability growth modelling: models and applications. *IEEE Trans. Software Engng*, 11:1431–1437.

[4] Zhao, M. and Xie, M. (1996). On maximum likelihood estimation for a general non-homogeneous Poisson process. *Scandinavian Journal of Statistics*, 23(4):597–607.

Stochastic Modelling for Speech Recognition and Language Translation

Hermann Ney *Computer Science Department, RWTH Aachen University, Aachen*

The last two decades have seen a dramatic progress in the area of automatic processing of speech and language. This talk gives an overview of the stochastic approach to both automatic speech recognition and language translation. It is remarkable that, in both tasks, the stochastic approach makes use of the same four principles:

- Bayes decision rule for minimum error rate,
- probabilistic models ((like Hidden Markov models) for handling strings of observations (like acoustic vectors for speech recognition and written words for language translation),
- training criteria and algorithms for estimating the free model parameters from large amounts of data,
- the generation or search process that generates the recognition or translation result.

The work on speech recognition and language translation we report has been and is being performed in various large-scale projects:

- TC-Star (supported by EU): recognition of speeches given in the European parliament, translation of Spanish and English in both directions;
- GALE (funded by US DARPA): recognition of broadcast news and conversations in Arabic and Chinese, translation of both speech and text from Arabic and Chinese into English;
- QUAERO (funded by French government): recognition and translation of French, German and other European languages for broadcast news, parliamentary speeches, audio archives, podcasts etc.

Speech Recognition

The progress in automatic speech recognition is based on improvements along various dimensions:

- improved acoustic features;

- speaker adaptive methods like MLLR (maximum likelihood linear regression) and VTLN (vocal tract length normalization) and SAT (speaker adaptive training);
- discriminative training, which unlike conventional maximum likelihood better focuses on recognition errors;
- use of more training data (order of 1000 hours) and also unsupervised training, which is able to make use of speech data without manual transcription;
- improved search algorithms;
- system combination.

Language Translation

Unlike speech recognition, the stochastic approach to language translation does not have a long history yet. It began in 1988 when IBM research started its pioneering work on statistical translation. At that time, the use of statistics for translation was very controversial. Now twenty years later, the best systems for language translations are based on the stochastic approach. Today, a typical state-of-the-art statistical system for language translation has the following four components:

- **Training:** For each sentence pair of the training data, an alignment matrix is computed, typically by using the set of so-called IBM-1 to IBM-5 alignment models and a Hidden Markov model.
- **Phrase extraction:** From the alignment matrices of all training sentence pairs, source-target fragments are excised and used to define the so-called phrase tables.
- **Log-linear model combination:** For each source-target phrase pair in the phrase table, so-called scoring functions are defined. Based on the training data, these scoring functions compute a probabilistic score of the hypothesis that the source fragment and the target fragment under consideration are translations of each other. These scoring functions are complemented with a word and/or phrase re-ordering model. All these scoring functions are combined in a so-called log-linear model. The weight of each scoring function is tuned for optimal translation quality or a related criterion.
- **Generation or search:** For the given source sentence, the goal is to select the target sentence with the highest probabilistic score in the log-linear model. To this purpose, the search algorithm has to generate and score hypotheses over various unknowns: unknown segmentation of the source sentence, unknown target phrases and unknown order of these phrases in the target sentence.

Modeling and Forecasting Realized Volatility via State-Space Representation

Vasyl Golosnoy *Institute of Statistics and Econometrics, Kiel University, Germany*
Iryna Okhrin *Department of Statistics, European University Viadrina, Frankfurt (Oder), Germany*
Wolfgang Schmid *Department of Statistics, European University Viadrina, Frankfurt (Oder), Germany*

Abstract: Availability of ultra high frequency data allows to construct precise estimators for daily realized volatility. Econometric modeling and forecasting of realized volatilities of asset returns is of interest for an investor. This paper models the realized volatility process using a state-space representation. The validity of the state-space model is checked at every new point in time in a sequential way. Our methodology is applied to empirical data.

1 Modeling realized volatilities

This paper aims to check the validity of time series models for the realized volatility. Since the process of volatility is unknown in practice, our model relies on the state-space representation. It could also be represented as a type of measurement error equations.

It is assumed the log-price of a stock follows an Itô process. Then the log-returns for the day t are normally distributed, $R_t|\mu_t, \sigma_t^2 \sim \mathcal{N}(\mu_t, \sigma_t^2), t = 1, 2, \ldots$. The variance σ_t^2 is called actual volatility and is given by $\sigma_t^2 = \int_{(t-1)\Delta}^{t\Delta} \sigma^2(u)\mathrm{d}u, \quad t \in \mathbb{N}$. The process σ_t^2 is unobservable and can be estimated by a sum of squared intraday log-returns, $s_t^2 = \sum_{i=1}^{m} R_{t,i}^2$. It is called realized volatility ([2], [1]).

The volatility dynamics is assessed via the state-space model. In particular, we assume that (unobservable) log-actual volatility follows an $AR(1)$ process. Then the state equation is given by

$$log(\sigma_{t+1}^2) - log(\sigma^2) \;\; = \;\; \phi(log(\sigma_t^2) - log(\sigma^2)) + v_{t+1}.$$

Since the paper [1] shows that log-realized volatility is asymptotically normally distributed, we write the observation equation as

$$log(s_t^2) \;\; = \;\; log(\sigma_t^2) + w_t.$$

The model of these two equations can be estimated by Kalman filter. In the classical setup the residuals v_t and w_t assumed to be normally distributed and not correlated. Since the residuals are not homoscedastic in our case we adopt the estimation procedure for the time varying residual variances.

2 Control charts and empirical application

We suggest to check the model validity in a sequential way by monitoring the residual process. For this purpose we normalize the difference between the observed value s_t^2 and its forecast based on information set $\mathcal{I}_{t-1}$ available at $t-1$:

$$\eta_t = \frac{s_t^2 - \hat{s}_{t|t-1}^2}{var(s_t^2 - \hat{s}_{t|t-1}^2)^{1/2}}, \qquad t \geq T. \tag{1}$$

These residuals form the process of interest $\{\eta_t\}$. If the model is correct, the residuals should be normally distributed with mean zero and variance one and not autocorrelated. The classical Shewhart chart gives a signal if $|\eta_t| > c$, where the control limit is defined as a quantile of the standard normal distribution $c = z_{1-p/2}$, $p = (ARL_0)^{-1}$ ([5]).

The four stocks from NYSE are selected for the empirical analysis: General Motors, Hewlett-Packard, Coca-Cola, and United Technologies. The ultra high frequency data is available for the period from April 15, 2005 to October 27, 2005. The estimation of realized volatility should be done taking into account the market microstructure noise. We follow the common proposition to take two scale estimator based on the 5 minutes frequent data ([4],[3]).

In all four cases the normality assumption for the model residuals is not rejected. Moreover, the estimated residuals show no autocorrelation. The Shewhart control charts provide signals during the observed period, which indicate on the model instability and should be investigated for their causes.

References

[1] Barndorff-Nielsen O. E. and Shephard N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society, B* 64: 253-280.

[2] Andersen T. G. and Bollerslev T. (1998). Answering the skeptics: standard volatility models do provide accurate forecasts, *International Economic Review* 39: 885-905.

[3] Barndorff-Nielsen O. E. and Shephard N. (2006). Impact of Jumps on Returns and Realized Variances: Econometric Analysis of Time-Deformed Lévy Processes, *Journal of Econometrics* 131: 217-252.

[4] Zhang L., Mykland P.A. and Aït-Sahalia Y. (2005). A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data, *Journal of the American Statistical Association* 100: 1394-1411.

[5] Montgomery, D. C. (2005). *Introduction to Statistical Quality Control*, 5th ed., Wiley: New York.

Semiparametric Inference with Applications to Image Symmetry and Nonlinear Time Series

Mirosław Pawlak, *Electrical and Computer Engineering, University of Manitoba, Canada*

Abstract: In this paper, we examine the problem of the semiparametric statistical inference in the context of the regression analysis. Here one wishes to estimate a finite-dimensional parameter in the presence of a nonparametric class of functions that define an infinitely dimensional nuisance parameter. We identify two distinct cases, in the first one we can obtain an estimate of the parameter without using a nonparametric smoothing technique. On the other hand, in the second case the parameter estimator requires a pilot nonparametric smoothing algorithm. In both cases we are interested in establishing conditions for the $\sqrt{n}-$consistency of the proposed parametric estimators. The first scenario of the developed theory is illustrated by the problem of estimating the axis of reflectional symmetry of a nonparametric image function. The second semiparametric situation is explained in the context of random dynamical systems being a low-dimensional approximation of general nonlinear time series models.
Keywords: semiparametric inference, $\sqrt{n}-$consistency, nonparametric estimation, image analysis, image symmetry, nonlinear time series.

1 Summary

There are two traditional ways concerning the statistical inference for estimating regression function, i.e., parametric and nonparametric approaches. In the former approach, one specifies the regression function up to the unknown finite-dimensional parameter. On the contrary, in the latter case no specific form of the regression function is assumed. There are numerous parameter estimation techniques that enjoy the optimal $\sqrt{n}-$consistency, provided that parametric regression model is correctly specified. In the nonparametric setting there is no risk of model misspecification. Nevertheless, corresponding nonparametric estimators exhibit low convergence rates which additionally deteriorates with increasing dimensionality of input variables.

In practice, we can often accept a semiparametric (intermediate) model which lies between parametric and fully nonparametric cases. The model is characterized by a finite-dimensional parameter embedded in unknown infinite-dimensional parameters that run typically through a nonparametric class of low-dimensional functions. Hence, the semiparametric model is defined by the pair $(\theta, g(\bullet))$, where θ a finite-dimensional parameter and $g(\bullet)$ is from a class of nonparametric functions. The goal is to estimate θ treating $g(\bullet)$ as an infinite-dimensional nuisance parameter. The regression analysis is defined by the vector (X, Y) distributed according to the probability law $P(\bullet, \bullet)$. In the parametric case $P(\bullet, \bullet)$ is known up to the unknown finite-dimensional parameter θ, i.e., $P(\bullet, \bullet) = P_\theta(\bullet, \bullet)$. On the contrary, in the fully nonparametric situation $P(\bullet, \bullet)$ is completely unknown. In the semiparametric setting

we have a natural parametrization $P_{\theta,g(\bullet)}(\bullet,\bullet)$. The process of recovering θ from the training set $\{(X_1,Y_1),\ldots,(X_n,Y_n)\}$ is usually based on the formation of the contrast function

$$Q_n(\theta,g(\bullet))=n^{-1}\sum_{i=1}^{n}\Psi_{\theta,g(\bullet)}(X_i,Y_i), \tag{1}$$

where $\Psi_{\theta,g(\bullet)}(\bullet,\bullet)$ is the known score function parametrized by $(\theta,g(\bullet))$. The basic idea of the semiparametric inference (in finding an estimate $\hat{\theta}$ of θ) is to eliminate the dependence of $Q_n(\theta,g(\bullet))$ on the nonparametric function $g(\bullet)$. Hence, we wish to analyze the parametric part of the semiparametric model as if the nonparametric part were given. This usually is done by forming a pilot nonparametric estimate of $g(\bullet)$ for a given θ. This leads to a nonparametric estimate $\hat{g}(\bullet;\theta)$ that itself depends on θ. As a result, we can replace $g(\bullet)$ in (1) by $\hat{g}(\bullet;\theta)$ and consequently obtain the following criterion depending solely on θ

$$\hat{Q}_n(\theta)=n^{-1}\sum_{i=1}^{n}\Psi_{\theta,\hat{g}(\bullet;\theta)}(X_i,Y_i). \tag{2}$$

It is now natural to define an estimate $\hat{\theta}$ of the true θ as the minimizer of $\hat{Q}_n(\theta)$. In this paper we examine sufficient conditions yielding the $\sqrt{n}-$consistency of $\hat{\theta}$. The dependence of $\hat{g}(\bullet;\theta)$ on θ plays an important role in the accuracy of the estimate. Nevertheless, we also identify cases where the dependence of θ on $g(\bullet)$ is weak and no pilot nonparametric estimate of $g(\bullet)$ is required. In particular, this phenomenon is explained by an important problem of estimating the axis of mirror symmetry of a reflection symmetric image $g(x_1,x_2)$ observed on the square grid points $\{(x_{1,i},x_{2,j}),1\leq i,j\leq\sqrt{n}\}$. Hence, we observe

$$Y_{i,j}=g(x_{1,i},x_{2,j})+Z_{i,j}, \tag{3}$$

where $Z_{i,j}$ is a noise process, and we wish to estimate the angle θ of the axis of mirror symmetry. Thus, the image function $g(x_1,x_2)$ is completely unknown but it is assumed to be symmetric, i.e., it satisfies

$$g(x_1,x_2)=g(x_1(\theta),x_2(\theta)),$$

where $x_1(\theta)=x_1\cos(2\theta)+x_2\sin(2\theta)$, $x_2(\theta)=x_1\sin(2\theta)-x_2\cos(2\theta)$ defines the mirror symmetry transformation with respect to the axis symmetry line with the unknown angle θ. In this paper we propose an estimation procedure $\hat{\theta}$ for estimating θ and we show that for any image functions $g(x_1,x_2)$ of bounded variation the estimate $\hat{\theta}$ converges at the optimal parametric rate of $\sqrt{n}$. The estimation procedure is based on minimizing over θ the L_2 distance calculated in the projection domain spanned by a class of orthogonal radial polynomials defined on the image domain. This approach does not need an explicit nonparametric smoothing algorithm of the unknown image function $g(x_1,x_2)$.

In our second case study we examine the nonlinear stochastic dynamical model of the following form

$$Y_t=g(\sum_{i=0}^{p}\theta_iX_{t-i})+Z_t. \tag{4}$$

Here we wish estimate the vector $\theta = (\theta_0, \ldots, \theta_p)$ regardless of the shape of the nonlinear function $g(\bullet)$. We can achieve this goal by using the least squares method with a data splitting strategy where the training set is divided into two nonoverlapping parts. The first part plays the role of a testing sequence that defines the least squares fitting function for estimating θ, whereas the other part is used as a training sequence to form preliminary nonparametric regression estimates for $g(\bullet)$. We establish sufficient conditions for the convergence of our estimation algorithm and we proof its $\sqrt{n}-$consistency. In addition, we present an alternative approach for estimating θ which does not need any optimization procedures. This direct strategy relies on the concept of the average derivative estimation of a regression function. This may be an appealing method in many applications because it is simple and noniterative. However, it is important to note that the method requires smooth density functions of the input signal $\{X_t\}$ and can be applicable to only a limited class of nonlinear dynamical systems.
It is also worth noting that the dynamical model in (4) forms the one-dimensional approximation of the fully nonparametric nonlinear system

$$Y_t = m(X_t, X_{t-1}, \ldots, X_{t-p}) + Z_t.$$

A direct nonparametric estimate of the regression function $m(\bullet, \ldots, \bullet)$ can be obtained yielding a slow rate of convergence of order $O_P(n^{-2/(5+p)})$. The rate of estimating the model in (4) is $O_P(n^{-2/5})$ for any p.

References

[1] P. J. Bickel, C. A. J. Klaassen, Y. Ritov, J. A. Wellner, *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins University Press, 1993.

[2] N. Bissantz, H. Holzmann, and M. Pawlak, "Testing for image symmetries – with application to confocal microscopy," *IEEE Trans. Inform. Theory*, to appear, 2009.

[3] H. Dette, S. Kusi-Appiah, and N. Neumeyer, "Testing symmetry in nonparametric regression models," *J. Nonparametr. Statist.*, vol. 14, pp. 477–494, 2002.

[4] J. Fan and Q. Yan, *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer, 2003.

[5] W. Greblicki and M. Pawlak, *Nonparametric System Identification*. Cambridge University Press, 2008.

[6] W. Härdle and M. Müller and S. Sperlich and A. Werwatz, *Nonparametric and Semiparametric Models*. Springer, 2004.

[7] F. Leblanc and O. V. Lepski, "Test of symmetry in nonparametric regression," *Theory Probab. Appl.*, vol. 47, pp. 34–52, 2003.

[8] M. Pawlak and S. X. Liao, "On the recovery of a function on a circular domain," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2736–2753, 2002.

[9] A. W. van der Vaart, *Asymptotic Statistics*. Cambridge University Press, 1998.

[10] M. A. G. Viana, *Symmetry Studies: An Introduction to the Analysis Structured Data in Applications*. Cambridge University Press, 2008.

[11] A. Yatchev, *Semiparametric Regression for the Applied Econometrician*. Cambridge University Press, 2003.

Hunting for the Best Control Chart

Mirosłk Pawlak, *Electr. Engineering and Computer Science, University of Manitoba, Canada*
Ewaryst Rafajłowicz, *Institute of Computer Engineering, Control and Robotics, Wrocław University of Technology, Poland*
Ansgar Steland, *Institute of Statistics, RWTH Aachen University, Germany*

1 Introduction

We discuss requirements which are usually imposed on control charts. To fix ideas, we shall briefly present two control charts recently proposed and studied by the authors. For comprehensive reviews of well known control charts we refer to [1], [2], [3], and also [15], [4], and [5]. Control charts for detection of changes in the mean characteristic of produced items is still the most frequently used tool in quality control. On the other hand, they are more and more frequently used in engineering diagnostics, financial economics, detecting of changes in video sequences and many others. All of these fields impose its own requirements on control charts. Even a model of "change" can be different in the above mentioned areas of applications. In addition to the well known step change model, we should consider less abrupt, slower changes, which do not even have a parametric model, or assume that the mean after the change is unknown.

Taking the above into account, one can not hope to select one "universally good" chart, which is uniformly sensitive to small, moderate and large shifts in the mean and still robust against (i) departures from normality, (ii) simultaneous changes of the mean and variance, (iii) correlation effects, and other factors.

Thus, hunting for the best chart is hopeless. What can be done instead ? At least the following four directions of research can be suggested.

- Investigate popular control charts taking into account various quality measures.
- Develop new charts with the hope that they will have the desired properties.
- Explor merits and disadvantages when running several charts in parallel on the same data.
- Identify the most useful features (indices) for a given application.

Feature identification is, indeed, a delicate issue, e.g., when aiming at detecting changes in video sequences. Clearly, the above mentioned research directions are so wide that they can be an outline of a large project rather then the contents of one paper, but we shall try to sketch the main ideas.

2 Criteria for selecting charts

The average run length (ARL) of a chart is the most commonly used criterion for selecting a chart. Usually, the in-control ARL is assumed fixed (typically at 465 or 830) and the chart with smaller out-of-control ARL is considered to be better than a chart with the same in-control ARL but larger out-of-control ARL. However, this criterion does not take into account the variances of ARL, which are known to be large. Fig. 1, which shows the probability density functions of run lengths of two hyphotetical charts, assuming the same shift in the mean and the same in-control ARL. The analysis of Fig. 1 leads to the following question: should we prefer a chart with a smaller out-of-control ARL and large run length variance or a chart with a slightly larger out-of-control ARL and small run length variance ? The answer to this question is even more difficult, if we consider that also out-of-control run lengths can have different variances.

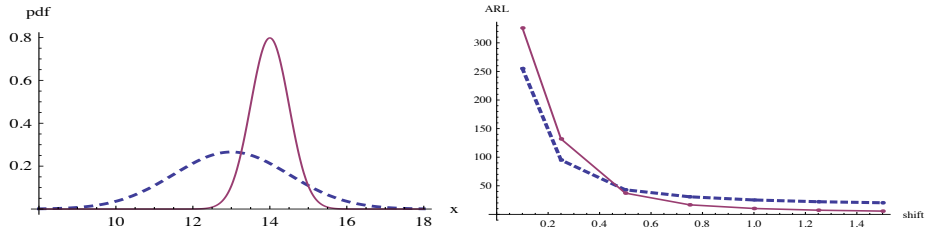

Figure 1: Left panel: Hyphotetical pdf's of out-of-control run lengths of two control charts. Assuming the same shift in the mean and the same in-control ARL, dashed line – pdf of a chart with smaller out-of-control ARL and larger variance, solid line – a chart with larger out-of-control ARL and smaller variance.
Right panel: Dependence of out-of-control ARL on the shift in the mean for two charts: the binary chart (dashed line) and the CUSUM chart (solid line).

Staying within out-of-control ARL paradigm, consider its dependence on changes of the shift, measured in σ (the dispersion of errors) units. Assuming the same out-of-control ARL, namely 435, the in-control ARL's of the CUSUM chart and the binary chart (recently proposed in [11]) are shown in Fig. 1. The answer to the question: which one is better, depends on whether we expect small or larger shifts more frequently.

The next criterion for selecting a good chart can be the probability of detecting a shift immediately after its occurrence (zero-delay detection), assuming that there is a bound on the probability of false alarm or on the in-control ARL. The plots, which are shon in Fig. 2 show the estimated probabilities of zero-delay detection for the following control charts: (i) VBox chart as proposed by the authors in [10], (ii) Shewart1, i.e., the Shewart control chart applied for individual observations, (iii) Shewart5, the Shewart chart applied to the averages of 5 observations, and (iv) the EWMA chart.

The differences between probabilities of zero-delay detection can be noticeable for moderate shifts ($0.5 - 1\,\sigma$), while for large shifts this criterion can not distinguish, which chart is better.

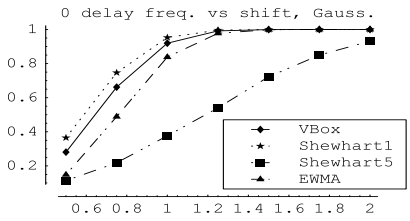

Figure 2: Dependence of the estimated probability of detecting shift just after its occurrence vs the shift magnitude for different charts 9described in the text).

If our aim is to detect moderate to large jumps then charts based on the so called jump-preserving procedures are attractive (see [7], [14], and [8, 9]). Nonparametric kernel control charts have been studied in [12] and [13]. Further, [16] combined a classic Shewhart chart and a conforming run length chart yielding smaller ARLs for shifts larger than 0.8σ, but that method is inferior to the EWMA chart for smaller shifts.

References

[1] Antoch J. and Jarušková M. (2002). On-line statistical process control, in: *Multivariate Total Quality Control, Foundations and Recent Advances*, ed. Lauro C., Antoch J., and Vinzi, V.E., Physica, Heidelberg.

[2] Antoch J., Hušková M., and Jarušková M. (2002). Off-line statistical process control, in: *Multivariate Total Quality Control, Foundations and Recent Advances*, ed. Lauro C., Antoch J., and Vinzi, V.E., Physica, Heidelberg.

[3] Brodsky, B.E., and Darkhovsky, B.S. (2000). *Non-Parametric Statistical Diagnosis: Problems and Methods*. Dordrecht: Kluwer Academic Publishers.

[4] Chakraborti S., van der Laan P., and Bakir S.T. (2001). Nonparametric control charts: An overview and some results, *Journal of Quality Technology*, vol. 33, 304-315.

[5] Montgomery D. C. (2001). *Introduction to Statistical Quality Control*. 4th Edition, New York: Wiley.

[6] Munford A.G. (1980). A control chart based on cumulative scores, *Applied Statistics*, vol. 29, 252-258.

[7] Pawlak, M. and Rafajłowicz, E. (1999). Vertically weighted regression - a tool for non-linear data analysis and constructing control charts. *Journal of the German Statistical Association*, vol. 84, 367-388.

[8] Pawlak, M., Rafajłowicz, E., and Steland A. (2004). On detecting jumps in time series - Nonparametric setting. *Journal of Nonparametric Statistics*, vol. 16, 329-347.

[9] Pawlak, M. and Rafajłowicz, E., and Steland A. (2008). Nonlinear image processing and filtering: A unified approach based on vertically weighted regression. *International Journal of Applied Mathematics and Computer Science*, vol. 18, 1, 49-61.

[10] Rafajłowicz, E., Pawlak, M. and Steland A. (2009) Nonparametric sequential change-point detection by a vertically trimmed box method., *IEEE Trans. Information Theory* (to appear).

[11] Rafajłowicz, E. and Steland A. (2008) A binary control chart to detect small jumps. *Statistics*, in press.

[12] Steland, A. (2004). Sequential control of time series by functionals of kernel-weighted empirical processes under local alternatives. *Metrika*, 60, 229-249.

[13] Steland, A. (2005). Optimal sequential kernel smoothers under local nonparametric alternatives for dependent processes. *Journal of Statistical Planning and Inference*, 132, 131-147.

[14] Steland, A. (2005). On the distribution of the clipping median under a mixture model. *Statistics and Probability Letters*, 70 (1), 1-13.

[15] Woodall W.H. (1997). Control charts based on attribute data: Bibliography and review. *Journal of Quality Technology*, vol. 29, pp 172-183.

[16] Wu Z. and Spedding T.A. A synthetic control chart for detecting small shifts in the process mean, *Journal of Quality Technology*, vol. 32, pp 32-38.

Systems with n Levels of Damage

Christoph Riethmüller *Institut für Mathematische Stochastik, Otto-von-Guericke-Universität Magdeburg, Germany*

Abstract: We investigate the reliability of a system. Changes of the behavior of the system are modelled as random events. We use a Markov chain approach and obtain a lifetime model which may be interpreted easily. Questions arise about repair and inspection policies as well as about statistical inference.

1 Introduction

The main aspect of choosing a model is choosing an appropriate failure rate. We use a Markov-modulated Poisson process as described in e.g. [1]. The developing of the (stochastic) failure rate is given by a homogeneous, absorbing Markov chain with finite state space $\{1, ..., n\}$. The transient states are called *working levels of damage* and the absorbing states are called *failure states*. One motivation for this model is that a user might interprete the levels of damage verbally:

Example: We consider the system *car* with the levels 1-*enough engine oil*, 2-*low engine oil* and 3-*engine damage*. We have two transient and one absorbing state and the corresponding generator might be

$$Q = \begin{pmatrix} -3 & 2 & 1 \\ 0 & -5 & 5 \\ 0 & 0 & 0 \end{pmatrix} .$$

The user (car driver) may continously observe the level of damage, so starting at level 1 two decisions have to be made:

- Should he refill engine oil when he enters level 2?
- Should he refill engine oil (if possible) when he enters level 3?

The first question covers preventive maintenance, the second one covers repair in case of a failure. The latter will be treated here.

2 Optimal repair

It makes sense that the repair action depends on the failure level and the level of the system just before failure, so wlg repair is performed in virtue of a function

$$r : \{1, ..., m\} \times \{m + 1, ..., n\} \to \{1, ..., m\}$$

where m is the number of working levels of damage. Concerning the example we should decide whether $r(2,3) = 1$ or $r(2,3) = 2$ is optimal. If we used mean repair costs as optimality criterion, we need to calculate the expected number of failures of a certain type up to time t:

$$E[N_t^{(i,j)}] = \int_0^t Q_{ij} a' e^{uQ_{(r)}} e_i \, du$$

where a is the initial distribution, e_i is the i-th unit vector and $Q_{(r)}$ is the generator of the Markov chain obtained by repeatedly repairing using repair function r.

We use a technique from [2] to calculate matrix-exponentials but we don't want to survey every possible repair function. We use a heuristic approach to resrict ourselves to a little number of repair functions. We sequentially choose repair functions such that the average costs per time unit are minimized. Here we need the probability p_k^{ij} that starting in level k the next failure is of type (i,j). We introduce bilateral phase type distributions [3] and obtain

$$p_k^{ij} = -W_{ki} Q_{ij} \ ,$$

where W is the inverse of upper-left part of Q that belongs to the working levels of the system.

References

[1] Aven, T., Jensen, U. (1999). *Stochastic Models in Reliability*, Springer: New York.

[2] Dunford, N., Schwartz, J.T. (1988). *Linear Operators Part I: General Theory*, Wiley: New York.

[3] Soohan Ahn, Ramaswami, V. (2005). Bilateral phase type distributions, *Stochastic Models* 21:239-259.

Reliability Centered Design - Method for Design of Products of Individual Reliability

Henrik Schnegas *Hochschule Wismar, Department of Mechanical, Process and Environmental Engineering*

Abstract: All customers in industry have individual requirements for safety and reliability. This publication shows an analytic and numerical method for design of products of individual reliability. Because reliability in Mechanical Engineering is reported as propability of survival we must combine mechanical and stochastic models, for example models of Boolean Algebra and redundancy.

1 Introduction

In DIN EN ISO 12100 we find a new definition for safety of machines. A machine is safe if exist risk is lower then allowable risk. Allowable risk is a function of probability of failure or low-function. So we must use the Reliability Theory. The question in Engineering Design is to integrate a required reliability in construction.

2 Method

We must install the following method in the standard VDI 2221 design process - especially in the period of conception. At first we need the concept of new product with all components.

2.1 Risk analysis by fault tree analysis

If we know the structure of system we can start with Fault Tree Analysis (FTA). The Fault Tree Analysis is most popular Top-Down method. In step one we define failure or low-function as fault. In step two we identify components of high hazard potential. In step three we must find model of boolean algebra to describe structure of reliability.

2.2 Structure of Reliability and numerical modelling

1. Series connection

In Systems of series connections all components have same failure effect. Each component added to the system decreases the probability of system. Reliability according to the formula

$$R_{ges.}(x) = \prod_{i=1}^{n} R_i(x) \tag{1}$$

2. Parallel connection

Systems of parallel connections are systems of redundancy. We know hot, warm or cold redundancy. If one component or system is out of order another component ore system release task of this component. Each component added to the system increases the probability of system. Reliability according to the formula

$$R_{ges.}(x) = 1 - \prod_{i=1}^{n} (1 - R_i(x)) \tag{2}$$

3. Combined connection

In combined connection systems we find parallel and series components. We must combine equations (1) and (2).

On base of Fault Tree Analysis and Boolean Algebra we can find description for reliability for systems of high complexity. It is a kind of risk assessment.

2.3 Reversible allocation of Reliability

If we could not reach required Reliability of system R_{Sys}, we must look for individual required reliability of components R_i. In case of construction of equal reliability we can compute required reliability for components. We can write:

$$\text{for series systems} \qquad R_{erf.i} = \sqrt[n]{R_{sys.erf.}} \tag{3}$$

and

$$\text{for parallel systems} \qquad R_{erf.i} = 1 - \sqrt[n]{1 - R_{sys.erf.}}. \tag{4}$$

At last we must integrate the required reliability in construction especially in geometry

2.4 Reliability Centered Design - Numerical Application

If we know the Reliability for each components, we can start the "Reliability Centered Design". In case of fatigue stress we can use equation 5:

$$B_{zul.} = B_N \cdot \sqrt[a]{\frac{x_N}{x_{Ausl.}}} \cdot \sqrt[a \cdot \beta]{\frac{ln\, R_{Ausl.}}{ln\, R_N}} \tag{5}$$

Allowable stress $B_{zul.}$ is the result of equation 5. Allowable stress is a function of required reliability $R_{Ausl.}$ and required lifetime $x_{Ausl.}$ Now we can use the standard methods for design of components. And in the end we will get an reliability based product.

3 Summary

If we like to design products of individual reliability, we have to use models from Mechanics and also from Stochastics. But there are also any problems in this moment. 1. There are

not so much tables with dates about stochastic parameters of failures. So we need methods to find parameters in test with low samplings 2. Stochastic Models for Reliability Centered Design are not basic Know-ledgement in Mechanical Engineering. In this case we need help of Stochastic People.

References

[1] Schnegas, Henrik: Reliability Centered Design (RCD) - Reliability in Mechanical Engineering, Publication for Distance learning, WINGS, 2008/9

Dimensionality Reduction Using Space-Filling Curves with Applications to Pattern Recognition

Ewa Skubalska-Rafajłowicz *Institute of Computer Engineering, Control and Robotics, Wrocław University of Technology, Poland*

1 Summary

The classical problem of constructing a multidimensional pattern classifier in the bayesian framework is considered. The key idea of the approach is to apply a well defined quasi-inverse of a space-filling curve in order to transform a multidimensional learning sequence to the unit interval. Complexity of such a transformation is linear with respect to d, since we can transform each pattern separately without constructing an approximation of the whole space-filling curve.

Our approach is based on space-filling curves. A space-filling curve is a continuous surjection from unit interval onto d-dimensional unit cube (I_d, $d \leq \infty$), that is, a curve going through each point of I_d. Space-filling curves were at first described by Peano, Hilbert, Sierpiński and [4] as an evidence that a lower dimensional space (for example the unit interval $[0, 1]$) can be mapped continuously onto a space of higher dimension. The space-filling curve cannot be one-to-one because, in general, I_d and I_p are not homeomorphic whenever $d \neq p$ (by invariance of dimension number theorem).

We give theoretical foundation under which consistent classifiers based on data scanned by a space-filling curve can be build. Furthermore, we indicate the properties of the space-filling curves, which ensure the Bayes risk consistency. The classifiers considered here are based on the plug-in principle.

We study the following pattern recognition rules, which are based on space-filling curves:

- k-nearest neighbor algorithms,
- the family of orthogonal series classifiers,
- partitioning rules.

The key result states that the measure preserving space-filling curve based transformation of data does not change the Bayes risk for any distribution with the support in I_d. This leads to a class of classifiers with the following features:

1. asymptotic optimality in the sense that the error probability of the proposed classifier approaches the Bayes risk almost surely.
2. high degree of data compression in the learning sequence
3. fast recognition of new patterns,

4. possibility of graphical interpretation of a decision rule and the learning sequence, after a suitable transform,

5. easy update of the classifier when new observations are added to a learning sequence.

The choice of the sample space as the unit cube I_d is not too restrictive, since the Bayes error is invariant under any continuous and strictly monotone transformation (e.g. the logistic function) of the coordinate axis.

We do not claim that a transformation based on a space-filling curve is an easy remedy for the "curse of dimensionality". One-dimensional patterns must be stored with a sufficiently high precision which allows to separate them in $[0, 1]$. This is the price that we pay for working with one-dimensional patterns. We should also note that such a transformation of data is non-linear, and it is not invariant under linear transformations of the feature space.

The presentation is organized as follows: we provide necessary preliminaries, including the problem statement. Then, the basic properties of space-filling curves are discussed. The problem of preserving class separation and the Bayes risk under space-filling curve transformation and its inverse is elucidated. As the next step, the above mentioned families of classifiers are described in some details. Then, their asymptotic optimality is proved in a number of steps. Namely, the upper bounds for the bias and variance provide the basis for the proof of the strong Bayes risk consistency of the classifiers. Attainable convergence rate of the risk is discussed. Finally, we discuss the results of simulation studies and comparisons.

References

[1] A.R. Butz. (1971). Alternative algorithm for Hilbert's space-filling curve. IEEE Trans. Comput, vol. C-20, pp.424–426.

[2] E. Skubalska–Rafajłowicz and A. Krzyżak. (1996). Fast k-NN classification rule using metric on space-filling curves, Proc. of the 13th ICPR, Vienna, Austria, August 25- 29, vol.2, pp.221–225, 1996.

[3] Krzyżak Adam, Skubalska-Rafajłowicz Ewa. (2004). *Combining space-filling curves and radial basis function networks*. Lecture Notes in Computer Science, vol. 3070 Springer, s. 229-234.

[4] H. Sagan. (1994) *Space-filling curves*, Berlin: Springer-Verlag.

[5] Skubalska-Rafajłowicz Ewa.: Pattern recognition algorithms based on space-filling curves and orthogonal expansions. IEEE Transactions on Information Theory. 2001 vol. 47, nr 5, s. 1915-1927.

[6] Skubalska-Rafajłowicz Ewa., (2001). Data compression for pattern recognition based on space-filling curve pseudo-inverse mapping. Nonlinear Analysis. vol. 47, s. 315-326.

[7] Skubalska-Rafajłowicz Ewa., (2005). A new method of estimation of the box-counting dimension of multivariate objects using space-filling curves. Nonlinear Analysis. vol. 63, nr 5-7, s. 1281-1287.

On the Reaction Time of Moving Sum Detectors

Josef G. Steinebach *Mathematical Institute, University of Cologne, Germany*

Abstract: In this talk, we discuss some asymptotics, under the null hypothesis as well as under the alternative, concerning the reaction time of on-line monitoring schemes to detect a change in the mean. The stopping rules are based on "moving sums", that is, they sequentially compare a "training sample" of size m to the average of the $h = h(m)$ most recent observations. Perhaps surprisingly, the limit distributions (as $m \to \infty$) crucially depend on the asymptotic relation of h and m, posing potential problems in applications.

1 The model

In [2] we discuss some "open-end" as well as "closed-end" monitoring procedures for detecting a change in the mean in the following location model:

$$X_i = \mu_i + \varepsilon_i\,, \quad i = 1, 2, \dots\,,$$

where $\{\mu_i\}_{i=1,2,\dots}$ are unknown means and $\{\varepsilon_i\}_{i=1,2,\dots}$ are the unobservable errors. It is assumed that there is no change in the mean of a "training sample" of size m, i.e. that $\mu_i = \mu,\ i = 1, \dots, m$. We are interested in constructing appropriate stopping rules for testing the null hypothesis

$$H_0 :\ \mu_i = \mu, \quad i = m+1, m+2, \dots\,,$$

against the (two-sided) alternative

$$\begin{aligned} H_A :\quad & \text{there is } k^* \geq 1 \text{ such that } \mu_i = \mu,\ m < i < m + k^*, \\ & \text{and } \mu_i = \mu + \Delta,\ i \geq m + k^*, \quad \text{with some } \ \Delta \neq 0. \end{aligned}$$

2 Stopping rules

Our rules for testing H_0 versus H_A are based on "moving sum detectors" (MOSUM's), more precisely, on comparing the quantities

$$\overline{X}_m = \frac{1}{m}\sum_{i=1}^{m} X_i \qquad \text{and} \qquad \overline{X}_{m,k} = \frac{1}{h}\sum_{i=0}^{h-1} X_{m+k-i}, \quad k = 1, 2, \dots\,,$$

where $h = h(m)\ (\leq m)$. For example, we study the (closed-end) stopping rule

$$\tau_{m,T} = \min\Big\{k :\ 1 \leq k \leq mT,\ \frac{1}{\hat{\gamma}}\left|\overline{X}_{m,k} - \overline{X}_m\right| > ch^{-1/2}g(k/m)\Big\}, \tag{1}$$

where $\min \emptyset = +\infty,\ \hat{\gamma}^2$ is a certain variance estimator, and g is a weight function.

In Section 3 it is shown that the constant c in (1) can be chosen such that, under H_0, $\lim_{m\to\infty} P\{\tau_{m,T} < \infty\} = \alpha$, where $0 < \alpha < 1$ is a prescribed level. In addition, some limiting distributions under H_A are discussed in Section 4. Interestingly, it turns out that, under H_0, the asymptotics crucially depend on the relation between h and m, and, under H_A, also on the limits of k^*/h and k^*/m, respectively.

3 Null asymptotics

To obtain the null asymptotics, we assume that the errors $\{\varepsilon_i\}$ satisfy a functional central limit theorem (with asymptotic variance γ^2). Then, for example, if $\lim_{m\to\infty} h/m = b \in (0, 1]$, we have

$$\lim_{m\to\infty} P\{\tau_{m,T} < \infty\} = P\Big\{ \sup_{0\le t\le T/b} \frac{1}{g(t)} \Big| W\Big(\frac{1}{b} + t\Big) - W\Big(\frac{1}{b} + t - 1\Big) - bW\Big(\frac{1}{b}\Big)\Big| > c\Big\}, \quad (2)$$

where $\{W(t), t \ge 0\}$ denotes a standard Wiener process, i.e., the critical value c in (1) can be determined via the weighted Gaussian process from (2). Similar results apply in case of $\lim_{m\to\infty} h/m = 0$, but require a more careful discussion (see [2]).

4 Asymptotics under the alternative

For the limiting results under the alternative, it is assumed that the errors $\{\varepsilon_i\}$ satisfy a Hungarian (KMT) type strong approximation. Various cases and stopping rules can be discussed then, depending on the orders of h/m, k^*/h, and k^*/m, respectively. For example, if $h/m \to b \in (0, 1]$, $k^*/h \to a \ge 0$, and $T > ab$, then, for $y > 0$,

$$\lim_{m\to\infty} P\Big\{\tau_{m,T} > k^* + \frac{\sqrt{h}}{|\Delta|} y \Big| \tau_{m,T} \ge k^*\Big\} = P\Big\{ \sup_{0\le t\le y} \big|Z_1(a) + \frac{t}{\gamma g(a)}\big| \le c \Big| \sup_{0\le t\le a} \big|Z_1(t)\big| \le c\Big\},$$

where $\{Z_1(t), t \ge 0\}$ is the weighted Gaussian process from (2) and c the critical value therein. For other relations between h, k^*, and m, a number of similar asymptotics are available (cf. [1]).

References

[1] Aue, A., Horváth, L., Kühn, M., Steinebach, J. (2009). On the reaction time of moving sum detectors, Preprint, University of California, Davis, University of Utah, Salt Lake City, and University of Cologne.

[2] Horváth, L., Kühn, M., Steinebach, J. (2008). On the performance of the fluctuation test for structural change, *Sequential Analysis* 27: 126-140.

On Local Linear Detectors for Trend Surveillance in Engineering and Finance

Ansgar Steland *Institute of Statistics, RWTH Aachen University, Germany*

Abstract: We study the problem of detecting a change in a trend disturbed by stationary or random walk noise generalizing the related problem dealing with i.i.d. error terms. We study procedures based on sequential local linear estimators and discuss new functional central limit theorems (FCLTs) yielding the asymptotic distribution of the detectors, which is required for the statistical design of the detector.

1 Introduction

Suppose we observe a sequence of random variables $\{Y_n\}$ where Y_n is observed at the nth time instant and observations arrive sequentially. Often the aim is to detect a change in the mean indicating that the process is no longer in a state of statistical control. Methods based on MOSUM, CUSUM, EWMA, and kernel statistics have been studied by many authors. For recent results and references dealing with dependent but stationary error terms we refer to [1, 2] and [4]. In engineering and econometrics methods which can handle the random walk setting are also of interest. Econometric time series such as log exchange rates or log consumer price levels come to mind, whereas in engineering series as (cumulative) damage processes or the workload of a switch in a communication network matter. It worth noting that detectors to distinguish between stationarity and random walk behavior have been recently developed, we refer to [5] and the references given there.

When aiming at consistent nonparametric estimation of the process mean, local linear estimation is nowadays a well established approach. Consistency can be ensured provided that the time instants where observations are available get dense, asymptotically, see [3] for the general methodology. Thus, using sequential versions of these estimators seems to be promising to develop monitoring (surveillance) procedures and, indeed, Monte Carlo experiments have demonstrated the outstanding performance of such procedures in many cases. However, the sequential asymptotic distribution theory has been an open problem.

2 Methods and results

Given a sequence $Y_1, Y_2, \ldots$ of real-valued random variables denote the corresponding (marginal) means by $m(t) = E(Y_t)$. Suppose that locally at the current time instant $t_n = n \in \mathbb{N}$ the approximation

$$m(s) = \beta_{0n}(t_n) + \beta_{1n}(t_n)(s - t_n) + o(1)$$

with unknown local intercept $\beta_{0n} = \beta_{0n}(t_n)$ and slope $\beta_{1n} = \beta_{1n}(t_n)$ holds true. Notice that these local parameters can be very informative in an analysis, since they measure locally

'level' and 'derivative' of the trend. To estimate the local parameters we fit a straight line to the data by weighted least squares, i.e., given the data $Y_1, \ldots, Y_n$ at the current time instant t_n, we minimize the objective

$$\sum_{i=1}^{n} w_{ni}(Y_i - \beta_0 - \beta_1(t_i - t_n))^2$$

with respect to $(\beta_0, \beta_1) \in \mathbb{R}^2$. $\{w_{ni}\}$ are nonnegative weights defined via a kernel function K. Denote the minimizers by $\widehat{\beta}_{0n}$ and $\widehat{\beta}_{1n}$. Notice that these estimates are $\mathcal{F}_n = \sigma(Y_s : s \leq n)$-measureable by construction, thus providing a reasonable basis for the construction of a stopping time, both from an intuitive and a mathematical viewpoint. Specifically, one may consider the stopping rules $L_T^{(i)} = \inf\{k \leq n \leq T : T^{-1/2}\widehat{\beta}_{in} > c_i\}$, $i = 1, 2$, for control limits (critical values) c_1, c_2. Notice that monitoring stops when the time horizon T is reached. Such procedures are also called *closed end* stopping times. We propose to select the control limits to ensure that the type I error does not exceed a pre-specified nominal value $\alpha \in (0, 1)$ as $T \to \infty$.

In [6] a FCLT for the stochastic process associated to the above sequential estimates has been established for a large class of local change-point models in the presence of random walk error terms with possibly dependent increments. The assumptions on the increments are weak and allow for many sets of assumptions encountered in time series models and applied work, and the class of local change-point models covers various settings of practical interest. The relevant limit process is a functional of standard Brownian motion. The asymptotic distribution theory for the case when the error terms are stationary and satisfy certain additional (weak) regularity conditions is subject of current research.

References

[1] Aue, A., Horváth, L., Hušková, M. and Kokoszka, P. (2008a). Testing for changes in polynomial regression. *Bernoulli*, **14**, 3, 637-660.

[2] Aue, A., Horváth, L., Kokoszka, P. and Steinebach J. (2008b). Monitoring shifts in mean: Asymptotic normality of stopping times. *Test*, **17**, 515-530.

[3] Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London.

[4] Gut, A. and Steinebach, J. (2004). EWMA charts for detecting a change-point in the drift of a stochastic process. *Sequential Analysis*, **23**, 2, 195-237.

[5] Steland, A. (2008). Sequentially updated residuals and detection of stationary errors in polynomial regression models. *Sequential Analysis*, **27**, 3, 304-329.

[6] Steland, A. (2009a). A surveillance procedure for random walks based on local linear estimation, *submitted*.

Detection of Stationarity Errors in a Linear Regression Model with an I(1)-Regressor

Sabine Teller *Institute of Statistics, RWTH Aachen University, Germany*

Abstract: This paper discusses the problem to detect sequentially that the error terms in a linear regression model with an $I(1)$-Regressor no longer behave as a random walk but as a stationary process. We provide the asymptotic distribution theory for a monitoring procedure given by a control chart, i.e. a stopping time. The main result is a functional limit theorem for the corresponding stochastic process which implies a central limit theorem for the control chart.

1 Motivation

Random walks have been proposed as reasonable models for discretely observed data in many fields. Particularly, random walks have been proposed as a model for important economic series as the gross domestic product. Therefore, an important problem is to check sequentially whether a time series is compatible with the random walk (in-control) model or follows an alternative (out-of-control) model under which the series is stationary.

Obviously, an erroneous answer to that question can lead to completely wrong statistical conclusions, since even elementary statistics change their convergence rates and limit distributions. Having this in mind, it is of particular interest to study sequential monitoring (surveillance) procedures, which are designed to detect departures from the random walk hypothesis as soon as possible.

2 The model

We aim to investigate a sequential monitoring procedure. This aim was motivated by a preliminary study from Steland [2] where the author considers the problem to test sequentially whether or not the error terms in a polynomial regression model form a random walk.

We assume that we observe sequentially a time series $\{Y_t\}_{t\in\mathbb{N}}$ of real-valued observations satisfying

$$Y_t = \beta_0 + \beta_1 \cdot t + \alpha_1 \cdot z_t + \varepsilon_t, \qquad t \in \mathbb{N}, \tag{1}$$

where $\alpha_1, \beta_0, \beta_1 \in \mathbb{R}$ are unknown regression coefficients, and we assume that the sequence $\{z_t\}_{t\in\mathbb{N}}$ is integrated of order 1, denoted by $\{z_t\}_{t\in\mathbb{N}} \sim I(1)$. We relax the notations of stationarity and random walk and consider a more general concept: We decide in favor of the $I(0)$-property (covering stationarity) or $I(1)$-property (covering random walks).

Under the null hypothesis, we assume that the error terms of the regression model, defined in (1), form a random walk, i.e., are integrated of order 1. We will apply a control chart (stopping time) providing a signal, if there is evidence that the error terms are no longer compatible with the random walk hypothesis.

3 The method

At each time point $t \leq T$ when a new observation is available, we calculate the full set of residuals $\widehat{\varepsilon}_1(t), \ldots, \widehat{\varepsilon}_t(t)$ using all available observations $Y_1, \ldots, Y_t$. We calculate a version of the unit root test statistic U_t, motivated by the KPSS test statistic, see Kwiatkowski et. al. [1], and Steland [3]. The ratio control chart is now defined as

$$R_T := \inf\{k \leq t \leq T \,|\, U_t \leq c_R\}\,, \qquad T \geq k\,, \tag{2}$$

with the convention $\inf \emptyset = \infty$. For our asymptotic results we assume $T \to \infty$, since for applications, approximations of the distribution of R_T for large time horizon T are of interest.
c_R is a control limit (critical value) and since monitoring stops at the latest at time T, we may interpret the stopping time as a hypothesis test with early stopping in favor of the alternative. Thus, one may choose c_R to control asymptotically the type I error rate of a false decision in favor of stationarity of the error terms, i.e.,

$$\lim_{T\to\infty} P_0\left(R_t \leq T\right) = \alpha\,, \tag{3}$$

for some given $\alpha \in (0,1)$. P_0 indicates that the probability is calculated under the null hypothesis that the error terms form a random walk.

References

[1] Kwiatkowski, D.; Phillips, P.C.B.; Schmidt, P.; Shin, Y. (1992), Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, *Journal of Econometrics* 54: 159-178.

[2] Steland, A. (2008), Sequentially Updated Residuals and Detection of Stationarity Errors in Polynomial Regression Models, *Sequential Analysis* 27: 304-329.

[3] Steland, A. (2007), Monitoring Procedures to Detect Unit Roots and Stationarity, *Econometric Theory* 23: 1108-1135.

Monitoring of First-Order Integer-Valued Autoregressive Processes of Poisson Counts

Christian H. Weiß *Dep. of Statistics, Inst. of Math., University of Würzburg, Germany*
Murat Caner Testik *Industrial Engineering Department, Hacettepe University, Turkey*

Abstract: Attributes control charts for counts generally assume that the process being monitored is independent and identically distributed in its in-control state. However, violation of this assumption in practice may significantly degrade a chart's performance and usefulness if the autocorrelation structure is not taken into account. We investigate the cumulative sum (CUSUM) control chart for monitoring autocorrelated processes of counts modeled by a Poisson INAR(1) model. Exact numerical results obtained through a bivariate Markov Chain approach are provided for sustained shifts in any or both of these process parameters. It is shown that the standard CUSUM chart of observations with adjusted control limits to account for the process autocorrelation has a good overall performance in detecting assignable causes in autocorrelated count processes.

1 Introduction

The *INAR(1) model*, originally introduced by [2], is the discrete-valued analogue to the standard AR(1) model and allows to model autocorrelated processes of counts, e. g., having a Poisson marginal distribution. [3] proposed a number of charts designed to control a process, which follows a stationary Poisson INAR(1) model in its in-control state. Based on a simulation study, the ability of these charts to detect several types of out-of-control situations were investigated. It turned out that for some situations neither of the considered charts could be used, e. g., if the autocorrelation is increased while the process mean does not change. Therefore, we shall consider a newly developed CUSUM chart in this article, which proves to be sensitive also to this type of out-of-control situation. In addition, the new CUSUM procedure allows to compute average run lengths (ARL) exactly, while an analysis of the ARL performance of most of the charts considered by [3] must be based on simulations.

2 Poisson INAR(1) CUSUM Chart

Consider the INAR (1) process model defined by the recursion $N_t = \alpha \circ N_{t-1} + \epsilon_t$, where '$\circ$' represents the *binomial thinning operator*. A one-sided *Poisson INAR(1) CUSUM chart* for detecting positive shifts from the in-control parameters λ_0 and/or α_0 of the process N_t is developed: Define the process $\{C_t^+ : t = 0, 1, \ldots\}$ by

$$C_0^+ = 0, \qquad C_t^+ = \max\left(0;\ N_t - k^+ + C_{t-1}^+\right) \quad \text{for } t = 1, 2, \ldots, \tag{1}$$

where $k^+ \in \mathbb{N}$, $k^+ \geq \lambda_0$, is called the *reference value*. In operating an upper Poisson INAR(1) CUSUM, the chart is formed by plotting the quantity C_t^+ against time t. The process is considered as being in control unless an out-of-control signal $C_t^+ \notin \{0, \ldots, h^+ - 1\}$ is triggered, where $h^+ \in \mathbb{N}$ is the control limit. The ARLs of the upper-sided CUSUM chart can be computed exactly with the Markov chain approach of [1], since the bivariate process $(N_t, C_t^+)_{\mathbb{N}}$ proves to be a discrete Markov chain.

3 Performance and design

Performance of the CUSUM chart in detecting changes in both the process mean λ and the dependence parameter α were investigated. We searched for possible integer h^+ and k^+ pairs and designed charts such that the in-control ARL_0 is about 500. For the considered in-control values $\lambda_0 = 2.5$, 5, 7.5, 10 and 15, and $\alpha_0 = 0.25$, 0.50 and 0.75, the performance of the designed charts were evaluated under various shift scenarios in either or both of the λ_0 and α_0 values. It turns out that this CUSUM scheme generally has a superior performance in detecting increases in the parameters than any of the other proposed charts for Poisson INAR(1) models. Based on the ARL results, guidelines were also derived for the chart design.

4 A real-data example

To illustrate design and performance of the Poisson INAR(1) CUSUM chart, we continue the real-data example of [3], who analyzed count data about accesses to the server of the Department of Statistics of the University of Würzburg collected on 29 November 2005. [3] showed that the data can be modeled by an INAR(1) model with parameters $\lambda = 1.28$ and $\alpha = 0.29$. We shall consider the time series collected a week later on 6 December 2005. Based on the above in-control model, we design a CUSUM chart with $(h^+, k^+) = (8, 2)$, leading to an $\text{ARL}_0 = 419.043$, and a CUSUM chart with $(h^+, k^+) = (4, 3)$, having $\text{ARL}_0 = 506.915$.

References

[1] Brook, D., Evans, D.A. (1972). An approach to the probability distribution of cusum run length. *Biometrika* 59(3): 539-549.

[2] McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin* 21(4): 645-650.

[3] Weiß, C.H. (2007). Controlling correlated processes of Poisson counts. *Qual. Reliab. Engng. Int.* 23(6): 741-754.

Surveillance of the Risk Behaviour of a Time Dependent Process

Wolfgang Schmid *Department of Statistics, European University Viadrina, Germany*
Svitlana Zabolotska *Department of Statistics, European University Viadrina, Germany*

Abstract: Two new cumulative sum control charts for detecting changes in the variance of a time series are introduced. In an extensive simulation study the introduced control charts are compared with existing cumulative sum type control schemes for the variance. In order to asses the performance of the schemes both the average run length and the maximum expected delay are used. The target process is assumed to be an autoregressive process of order 1 (AR(1)).

1 Introduction

Several authors showed that control charts for independent variables, like e.g. the Shewhart, the EWMA (exponentially weighted moving average), and the CUSUM (cumulative sum) scheme, cannot be directly applied to time series. It turns out to be necessary to take the structure of the time series into account. In the last years some control charts for time series have been introduced. Most of the papers are focused on monitoring the mean behavior of a time series. In this work we deal with the detection of a changes in the variance. This problem is of great interest in practice since the variance is a frequently applied measure for the risk. This motivates the need of monitoring tools, that can detect a shift in the variance of an observed process as quickly as possible.

Because of the complicated structure of the probability distribution of a stationary Gaussian process variance charts for such processes have not been derived up to now by using the log likelihood ratio approach or the sequential probability ratio test. The starting point of the considerations is usually an independent normal sample. The CUSUM scheme for independent variables is derived and after that the same recursion is applied to stationary processes (cf.[1]). It is obvious that this procedure is unsatisfactory.

2 Main results

In our paper ([2]) we introduce two CUSUM control schemes for Gaussian processes for detecting changes in the variance. We derive these control schemes from the log likelihood ratio approach and the related sequential probability ratio test. As an example we consider the control statistics for autoregressive process of order 1.

Now the obvious question is whether the new control schemes outperform the well-known classical control charts. In an extensive simulation study we compare these control schemes classical control schemes based on i.i.d. assumption. As measures for the performance of

control schemes the average run length (ARL) and the maximum expected delay are taken. All charts are calibrated such that the in-control ARL is the same if no change is present. Their behavior is analyzed with respect to a scale deviation.

It is shown that the new control charts (which was derived over the log likelihood ratio approach) provides better result in many situations. These control charts can be used very effectively for monitoring structural deviations in the volatility of a financial time series.

References

[1] Schipper, S. and Schmid, W. (2001), Sequential methods for detecting changes in the variance of economic time series, *Sequential Analysis* 20(4): 235-262.

[2] Schmid, W. and Zabolotska S. (2008), Surveillance of the Risk Behaviour of a Time Dependent Process, *Viadrina Working Paper Series*.

Specifications for Prediction Procedures

Luis Zitzmann *Volkswirtschaftliches Institut, Universität Würzburg, Germany*

Abstract: The ability to develop stochastic models for making reliable and accurate predictions will probably turn out to be the key technology for society and industry of the 21st century. However, so far, there is no ISO standard for producing prediction procedures safeguarding the usefulness and reproducibility of predictions. As a consequence many different procedures are used yielding divers predictions for the same situation leading almost necessarily to confusion and to wrong decisions. The described circumstances can be observed for instance in the case of predictions of the gross domestic product (GDP), which play an important role in Political Economy. Many institutions, organizations and governmental offices produce and circulate continuously different numbers as predictions that practically never occur.

In my talk I will present specifications for prediction procedures that could be used for a prediction standard similar to standards that exist for almost every product which is sold on the market. If such a standard would exist, a prediction not meeting the specifications could be classified as nonconforming and discarded from any decision-making process.

Participants

Abujarad, Mohammed
RWTH Aachen University, Institut für Statistik und Wirtschaftsmathematik, abujarad@stochastik.rwth-aachen.de

Bodnar, Olha
European University Viadrina, Department of Statistics, obodnar@euv-frankfurt-o.de

Bodnar, Taras
European University Viadrina, Department of Statistics, bodnar@euv.frankfurt-o.de

Buka, Nadzeya
Otto-von-Guericke-University, Institut für Mathematische Stochastik, nadi.buko@gmail.com

Burkschat, Marco
RWTH Aachen University, Institut für Statistik und Wirtschaftsmathematik, burkschat@stochastik.rwth-aachen.de

Christoph, Gerd
OvG-Universität Magdeburg, Institut für Mathematische Stochastik, gerd.christoph@ovgu.de

Denecke, Liesa
University of Kassel, liesad@mathematik.uni-kassel.de

Döring, Maik
Universität Hohenheim, Angewandte Mathematik und Statistik, maik.doering@uni-hohenheim.de

Fioshina, Helen
Riga Technical University, Department of Mathematical Support of Transport Systems Control, jelena.fiosina@rtu.lv

Fioshin, Maxim
Riga Technical University, Department of Mathematical Support of Transport Systems Control, maksims.fiosins@rtu.lv

Franz, Jürgen
TU Dresden, Institut für Mathematische Stochastik, juergen.franz@tu-dresden.de

Gandy, Axel
Imperial College London, Department of Mathematics, a.gandy@imperial.ac.uk

Gupta, Arjun
Bowling Green State University, Mathematics and Statistics, gupta@bgsu.edu

Hanck, Christoph
Universiteit Maastricht, Quantitative Economics, c.hanck@ke.unimaas.nl

Herff, Wolfgang
RWTH Aachen University, Institut für Statistik und Wirtschaftsmathematik, herff@stochastik.rwth-aachen.de

Holtemöller, Oliver
RWTH Aachen University, Fakultät für Wirtschaftswissenschaften, oliver.holtemoeller@rwth-aachen.de

Hörmann, Eva
JLU Gießen, Eva.S.Hoermann@math.uni-giessen.de
Janssen, Arnold
Universität Düsseldorf, Mathematics, janssena@math.uni-duesseldorf.de
Jensen, Uwe
Universtität Hohenheim, Institut für angew. Mathematik und Statistik, jensen@uni-hohenheim.de
Jokiel-Rokita, Alicja
Wroclaw University of Technology, Institute of Mathematics and Computer Science, Alicja.Jokiel-Rokita@pwr.wroc.pl
Kahle,Waltraud
Otto-von-Guericke-University, Institute of Mathematical Stochastics, waltraud.kahle@ovgu.de
Kordnoori, Shaghayegh
Islamic Azad University North Tehran Branch, Department of Statistics, shaghayeghkordnourie@yahoo.com
Lefebvre, Mario
Ecole Polytechnique de Montreal, Mathematics and Industrial Engineering, mlefebvre@polymtl.ca
Lemken, Alexander
RWTH Aachen University, alexander.lemken@rwth-aachen.de
Leucht, Anne
Friedrich Schiller Universtiy of Jena, Institute for Stochastics anne.leucht@minet.uni-jena.de
Lopez, Rodrigo
ITESM, A00888363@itesm.mx
Magiera, Ryszard
Wroclaw University of Technology, Institute of Mathematics and Computer Science, ryszard.magiera@pwr.wroc.pl
Mathar, Rudolf
RWTH Aachen University, Institute for Theoretical Information Technology and UMIC, r.mathar@ti.rwth-aachen.de
Meisen, Sabine
RWTH Aachen University, sabinemeisen@yahoo.de
Mostafaei, Hamidreza
Islamic Azad University North Tehran Branch, statistics, hrmostafaei@yahoo.com
Ney, Hermann
RWTH Aachen University, Computer Science 6, ney@cs.rwth-aachen.de
Okhrin, Iryna
European University Viadrina, Department of Statistics, ryna.okhrin@euv.frankfurt-o.de
Pawlak, Mirosław
University of Manitoba, Electrical and Computer Engineering, pawlak@ee.umanitoba.ca

Rafajłowicz, Ewaryst
Wroclaw University of Technology, Inst. Computer Eng. Control Rob., ewaryst.rafajlowicz@pwr.wroc.pl
Rafajłowicz, Wojciech
Wroclaw University of Technology, Inst. Computer Eng. Control Rob., wojchiech.rafajlowicz@pwr.wroc.pl
Reznikov, Lilia
RWTH Aachen University, liliareznikov@gmail.com
Riethmüller, Christoph
Otto-von-Guericke-University, IMST, christoph.riethmueller@ovgu.de
Schmid, Wolfgang
European University Viadrina, schmid@euv.frankfurt-o.de
Schmiedt, Anja Bettina
RWTH Aachen University, Bettina.Schmiedt@rwth-aachen.de
Schnesgas, Henrik
Hochschule Wismar, Mechanical Engineering, henrik.schnegas@hs-wismar.de
Skubalska-Rafajłowicz, Ewa
Wroclaw University of Technology, Inst. Computer Eng Control Rob ewa.rafajlowicz@pwr.wroc.pl
Stadtmüller, Ulrich
Universität Ulm, Institut für Zahlentheorie und Wahrscheinlichkeitstheorie ulrich.stadtmueller@uni-ulm.de
Steinebach, Josef G.
University of Cologne, Mathematical Institute, jost@math.uni-koeln.de
Steland, Ansgar
RWTH Aachen University, Institut für Statistik und Wirtschaftsmathematik, steland@stochastik.rwth-aachen.de
Teller, Sabine
RWTH Aachen University, Institut für Statistik und Wirtschaftsmathematik, teller@stochastik.rwth-aachen.de
Thrun, André
RWTH Aachen University, Institut für Statistik und Wirtschaftsmathematik, Thrun@stochastik.rwth-aachen.de
Weiß, Christian
Universität Würzburg, Institut für Mathematik, Lehrstuhl für Statistik, christian.weiss@mathematik.uni-wuerzburg.de
Wirtz, Anke
RWTH Aachen University, anke.wirtz@rwth-aachen.de
Zabolotska, Svitlana
European University Viadrina, Department of Statistics, zabolotska@euv-frankfurt-o.de
Zitzmann, Luis
Universität Würzburg, Volkswirtschaftliches Institut, mail@luis-zitzmann.de