TR 05-245

05-245

De Brabanter J., Pelckmans K., Suykens J.A.K., De Moor B., ``generalized likelihood ratio statistics based on bootstrap techniques for autoregression models", in *Proc. of the 14th IFAC Symposium on System Identification (SYSID 2006)*, Newcastle, Australia, Mar. 2006, 6 p., Lirias number: 181229.

GENERALIZED LIKELIHOOD RATIO STATISTICS BASED ON BOOTSTRAP TECHNIQUES FOR AUTOREGRESSIVE MODELS

De Brabanter, J.*, Pelckmans, K.****
Suykens, J.A.K. De Moor B.****

* *Hogeschool KaHo Sint-Lieven (Associatie KULeuven), Departement Industrieel Ingenieur, B-9000 Gent, Belgium*

** *Katholieke Universiteit Leuven, ESAT-SCD/SISTA, Kasteelpark Arenberg 10, B-3001 Leuven, Belgium*

Abstract: Parametric models provide explanatory power and can give a parsimonious description for dynamical systems modelling. But, there is a risk that misspecification of an underlying stochastic model can lead to misunderstanding of the system, wrong conclusions, and erroneous predicting. It is common practice to check whether a parametric model fits a given data set reasonably well. To achieve this, in the Neyman-Pearson framework, one needs to specify a class of alternative (nonparametric models). In this paper, we use the idea of the generalized likelihood ratio test based on bootstrap techniques. This technique can be applied for each parametric model structure and its nonparametric counterpart. We demonstrate the use of these methods for testing the significance of NARX versus ARX models.

Keywords: Generalized Likelihood Ratio Statistics, Bootstrap, Autoregressive models

1. INTRODUCTION

The appropriateness of parametric modelling of regression data in stochastic dynamical systems may be judged by comparison with a nonparametric smoothing estimator. In practical studies the importance of comparing parametric with nonparametric modelling has been pointed out in (Gasser *et al.*, 1985; Hildenbrand and Hildenbrand, 1986). An intuitive approach to handling the nonparametric testing problems is based on discrepancy measures (such as L_2 and L_∞ distances) between the estimators under the null and alternative hypothesis (Hardle and Mammen, 1993). Such a method is not as fundamental as likelihood ratio based tests (Fan *et al.*, 2001).

The maximum likelihood ratio test forms a useful principle that is generally applicable to most parametric hypothesis testing problems. An important fundamental property of the maximum likelihood ratio test is that their asymptotic null distributions are independent of nuisance parameters (Wilks phenomenon). Nuisance parameters are parameters of a model that must be determined but are not of interest to the investigator, e.g. error variance. In an effort to extend the scope of the likelihood approach Owen (1988, 1990, 2001) introduced empirical likelihood. The empirical likelihood has traditionally been used for providing confidence regions for multivariate means and, more generally, for parameters in estimating equations, under various standard as-

sumptions: the number of estimating equations is fixed, they do not involve nuisance parameters, and the parameters of interest can be estimated at rate $\sqrt{n}$, where n is the sample size. Under such assumptions and with i.i.d. observations (or even dependent observations, see Owen, 2001), empirical likelihood based confidence regions can be calibrated using a nonparametric version of Wilks' theorem involving a chi-squared limiting distribution. The method of profile likelihood (Owen, 2001) for dealing with nuisance parameters in estimating equations is often not applicable for infinite number of nuisance parameters, and even when it is applicable, its implementation can be computationally difficult. For applications of empirical likelihood see (Owen, 1990; Hall and Owen, 1993; Chen and Qin, 1993). A further extension of the empirical likelihood, called the "random-sieve likelihood," can be found in Shen *et al.*, (1999). Nevertheless, it cannot be directly applied to a nonparametric function estimation setting. Zhang and Gijbels (1999) incorporated the idea of local modeling into the framework of empirical likelihood and proposed an approximate empirical likelihood, called the "sieve empirical likelihood". The sieve empirical likelihood can efficiently handle the estimation of nonparametric functions. Generalized likelihood ratio statistics have been proposed in Fan *et al.* (2001) as a generally applicable method for testing nonparametric hypotheses about nonparametric models.

Recently, there have been many collective efforts on hypothesis testing in nonparametric regression problems. Most of them focus on one-dimensional nonparametric regression models. For an overview and references, see the book by Hart (1997). Hypothesis testing for multivariate regression problems is difficult due to the curse of dimensionality. Fan and Huang (2001) proposed various testing techniques based on adaptive Neyman tests for various alternative models in a multiple regression setting. These problems become conceptually simple by using the generalized likelihood method of Fan *et al.* (2001).

The goal of this paper is to construct Monte Carlo tests, based on generalized likelihood ratio statistics and bootstrap techniques, for system identification purposes. For an overview of model structures, see the book by Ljung (1987).

This paper is organized as follows. Section 2 introduces the problem of hypothesis testing. We describe the likelihood ratio test and explain the power function of a test. In the second part of this section we present the basic idea of the generalized likelihood test based on Fan *et al.* (2001). In Section 3 we consider the autoregressive model (ARX and NARX) to illustrate the results and explain the bootstrap technique for estimating the null distribution and computing the p-values in detail. Section 4 reports results to evaluate the performance of the proposed procedure empirically.

2. HYPOTHESIS TESTS

2.1 Likelihood Ratio Test

Let $X = (Y, V)$ be a random variable with probability distribution function F and $T(F)$ a statistical functional. Let the estimator $T(\hat{F}_n)$ of $T(F)$ be the functional of the sample distribution $\hat{F}_n$. Typically, a hypothesis test is specified in terms of a functional $T(\hat{F}_n)$, a function of the sample. The goal of a hypothesis test is to decide, based on a sample, which one of two complementary hypotheses is true. If $T(F)$ denotes a functional, the general format of the null and alternative hypothesis is $H_0 : T(F) \in \Theta_0$ and $H_1 : T(F) \in \Theta_0^c$ respectively where Θ_0 is some subset of the statistical functional space and Θ_0^c is its complement. One of the most useful methods for complicated models is the log-likelihood ratio method of test construction. Likelihood ratio tests are defined as follows.

Definition 1. (Lehmann, 1997) The likelihood ratio test statistic for testing H_0 versus H_1 is

$$S(x) = \frac{\sup_{\Theta_0} l(T(F)|x)}{\sup_{\Theta_0^c} l(T(F)|x)}, \tag{1}$$

where $l(T(F)|x)$ is the log-likelihood function. The test has a rejection region of the form

$$\{x : S(x) \leq c\}, \tag{2}$$

where $c \in \mathbb{R}_+$.

The distribution of $S(x)$ is, however, in general not known. For large n, one can uses the following result, which is valid under very general conditions.

Theorem 1. (Shao, 1999) UnderH_0,

$$-2\log(S(x)) \xrightarrow{d} \chi_r^2 \tag{3}$$

with $r = \dim(\theta_0^c) - \dim(\theta_0)$.

The symbol $\xrightarrow{d}$ denotes equality in distribution and χ_r^2 denotes the chi-squared distribution with r degrees of freedom.

In deciding to accept or reject the null hypothesis H_0, one might be making a mistake. Usually, hypothesis tests are evaluated and compared through their probabilities of making mistakes. A hypothesis test of $H_0 : T(F) \in \Theta_0$ versus $H_1 : T(F) \in \Theta_0^c$ might make one of two types of errors. These two types of errors traditionally have

been given the names, type I error and type II error (Pestman, 1991). The probability of making a type I error and a type II error is defined as

$$P(\text{type I error}) = P(\text{Reject } H_0 | H_0 \text{ is True}), \quad (4)$$

$$P(\text{type II error}) = P(\text{Accept } H_0 | H_0 \text{ is False}). \quad (5)$$

Both contain all the information about the test with rejection H_0. This leads to the following definition.

Definition 2. The power function of a statistical test is the probability function of correctly rejecting a false null hypothesis

$$\begin{aligned} \text{power} &= 1 - P(\text{Accept } H_0 | H_0 \text{ is False}) \\ &= P(\text{Reject } H_0 | H_0 \text{ is False}). \end{aligned} \quad (6)$$

Qualitatively, a good test has a power function near one for most $T(F) \in \Theta_0^c$ and near zero for most $T(F) \in \Theta_0$. Next, we give a simple example that illustrates the power function.

Example 1. Let $X_1, \ldots, X_n$ be a random sample from a normal $\mathcal{N}(0, \sigma^2)$ population with known variance σ^2. A likelihood ratio test of $H_0 : T(F) \leq \theta_0$ versus $T(F) > \theta_0$ is a test that rejects H_0 if $(\overline{X} - \theta_0)/(\sigma/\sqrt{n}) > c$ where $\overline{X} = (1/n)\sum_{i=1}^n X_i$. The power function of this test is

$$\begin{aligned} \text{power} &= P\left(\frac{(\overline{X} - \theta_0)}{(\sigma/\sqrt{n})} > c\right) \\ &= P\left(\frac{(\overline{X} - T(F))}{(\sigma/\sqrt{n})} > c + \frac{(\theta_0 - T(F))}{(\sigma/\sqrt{n})}\right) \\ &= P\left(Z > c + \frac{(\theta_0 - T(F))}{(\sigma/\sqrt{n})}\right) \end{aligned} \quad (7)$$

where Z is a standard normal random variable, since $(\overline{X} - \theta_0)/(\sigma/\sqrt{n}) \sim \mathcal{N}(0, 1)$.

2.2 Generalized Likelihood Ratio Test

The basic idea of the generalized likelihood test can be found in Fan *et al.* (2001). They used an example to illustrate the generalized likelihood test. Let g_θ be a family of parametric functions and let ξ be the nuisance parameters. Suppose that the logarithm of the likelihood of a given set of data is denoted by $l(g_\theta, \xi)$. Let $\hat{g}_n$ be a nonparametric estimator. In general, maximum likelihood estimators (if they exist) under nonparametric regression models are hard to obtain. To attenuate these difficulties, Fan *et al.* (2001) replace the maximum likelihood estimator under the alternative nonparametric model by any nonparametric estimator, leading to the generalized likelihood ratio. The nuisance parameters ξ can be estimated by the profile likelihood by maximizing $l(\hat{g}_n, \xi)$ with respect to ξ, resulting in the profile likelihood estimator $\hat{\xi}$. This gives the profile likelihood $l(\hat{g}_n, \hat{\xi})$, which is not the maximum likelihood since $\hat{g}_n$ is not a maximum likelihood estimator.

Suppose we are testing whether a parametric family g_θ fits a given data set. Formally, the null hypothesis is

$$H_0 : g = g_\theta, \theta \in \Theta_0, \quad (8)$$

and we use the nonparametric model g_n as the alternative model. Let $\hat{\theta}_0$ and $\hat{\xi}_0$ be the maximum likelihood estimator under the null model (8) maximizing the function $l(g_\theta, \xi)$. The generalized likelihood ratio test statistic compares the log-likelihood under the two competing classes of models:

$$S = l(\hat{g}_n, \hat{\xi}) - l(\hat{g}_{\theta_0}, \hat{\xi}_0) \quad (9)$$

where $l(\hat{g}_{\theta_0}, \hat{\xi}_0)$ is the maximum likelihood under the null hypothesis.

Remark 1. Generalized likelihood ratio statistics have been proposed in Fan *et al.* (2001) as a generally applicable method for testing nonparametric hypotheses about nonparametric functions. The likelihood ratio statistics are constructed based on the assumptions that the distribution of stochastic errors belong to a certain parametric family. Fan and Zhang (2004) extended the ratio tests to the case where the error distribution is completely unspecified via newly proposed sieve empirical likelihood ratio tests.

To utilize the generalized likelihood ratio test statistic (9), we need to derive the distribution under the null hypothesis H_0. For a number of models and a number of hypotheses, studied by Fan *et al.* (2001), it has been shown that the Wilks phenomenon continues to hold. Like Wilks (1938), Fan *et al.* (2001) are not able to show that the Wilks type of results hold for all problems, but their results indicate that such a phenomenon holds in general. In this paper we use the conditional nonparametric bootstrap as an alternative. The bootstrap principle for estimating the null distribution and computing the p-values is explained in the next section.

The power of the generalized likelihood ratio test has been calculated by Fan *et al.* (2001) based on local linear fits. In this paper we use a Monte Carlo simulation as an alternative. Barnard (1963) suggested the use of Monte Carlo methods to estimate quantiles of a test statistic under the null hypothesis. Marriott (1979) studied the power of the test and found that the power of Monte Carlo tests can be quite good. Monte Carlo simulations for

computing the power of the generalized likelihood ratio test are used in the next section.

3. ARX VERSUS NARX MODEL STRUCTURE

3.1 Generalized likelihood ratio test

Given a training set $\{(Y_t, U_t),\ t = 1, ..., n\}$ sampled from the nonparametric regression model

$$Y_t = g(Z_t) + e_t \tag{10}$$

where $Z_t = (Y_{t-1}, ..., Y_{t-p}, U_{t-1}, ..., U_{t-d})$, $g \in C^\infty(\mathbb{R})$ is an unknown function and $e_t \overset{i.i.d.}{\sim} G$ where G is a parametric error distribution. Consider testing the ARX model structure

$$\begin{aligned} H_0 : g(z) &= -a_1 Y_{t-1} - ... - a_p Y_{t-p} \\ &\quad + b_1 U_{t-1} + ... + b_q U_{t-q} \\ H_1 : g(z) &\neq -a_1 Y_{t-1} - ... - a_p Y_{t-p} \\ &\quad + b_1 U_{t-1} + ... + b_q U_{t-q} \end{aligned} \tag{11}$$

with the nonparametric alternative model (10). Let g_θ denote the ARX model with θ the parameter vector, $\theta = (a_1, ..., a_p, b_1, ..., b_q)^T$. Suppose that the $\{e_t\}_{t=1}^n$ are i.i.d. with density function $\frac{1}{\sigma_\nu} f\left(\frac{e}{\sigma_\nu}\right)$ where σ_ν is a scale parameter. Then the conditional likelihood function is

$$L(g_\theta, \sigma_\nu) = \prod_{t=1}^{n} \left(\frac{1}{\sigma_\nu} f\left(\frac{Y_t - g_\theta(Z_t)}{\sigma_\nu} \right) \right) \tag{12}$$

If for example $e_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$, (12) becomes

$$L(g_\theta, \sigma^2) = C \exp\left(-\frac{1}{2} \sum_{t=1}^{n} \frac{(Y_t - g_\theta(Z_t))^2}{\sigma^2} \right) \tag{13}$$

where $C = (2\pi\sigma^2)^{-\frac{n}{2}}$.

Ignoring constants, we denote the logarithm of (13) by $l(g_\theta, \sigma^2)$ and obtain

$$\begin{aligned} l(g_\theta, \sigma^2) &= -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{n} (Y_t - g_\theta(Z_t))^2 \\ &= -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} RSS_{H_0} \end{aligned} \tag{14}$$

where RSS_{H_0} denotes the residual sum of squares under the null hypothesis.

For a given σ, let $\hat{g}_n(\cdot)$ be the nonparametric estimator based on the data, which is independent of σ. Substituting it into the likelihood, we obtain the profile likelihood

$$\begin{aligned} l(\hat{g}_n, \sigma^2) &= -\frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{n} (Y_t - \hat{g}_n(Z_t))^2 \\ &\quad - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} RSS_{H_1} \end{aligned} \tag{15}$$

where RSS_{H_1} denotes the residual sum of squares under the alternative hypothesis. Substituting (14) and (15) into the definition of the generalized likelihood ratio statistic, we obtain

$$\begin{aligned} S &= l(\hat{g}_n, \sigma^2) - l(g_\theta, \sigma^2) \\ &= \frac{n}{2} \log\left(\frac{RSS_{H_0}}{RSS_{H_1}} \right) \end{aligned} \tag{16}$$

Given a test set $\{(Y_t, U_t),\ t = n+1, ..., n+m\}$, one can in (16) replace the RSS for the training data by an RSS for the test data (as is done in the sequel in all experiments).

3.2 Null distribution based on the bootstrap

Here, we give a detailed description of the bootstrap algorithm.

Algorithm 1. (estimating the null distribution and computing the p-values)

(1) Generate bootstrap residuals $\{\hat{e}_t^*\}_{t=1}^n$ of the empirical distribution of the centered residual $\{\hat{e}_t - \tilde{e}_t\}_{t=1}^n$ from the ARX model, where $\hat{e}_t = Y_t - \hat{g}_\theta(Z_t)$ and $\tilde{e}_t$ is the average of $\{\hat{e}_t\}_{t=1}^n$. Construct the bootstrap sample

$$Y_t^* = \hat{g}_\theta(Z_t) + \hat{e}_t^* \tag{17}$$

(2) Calculate the test statistic S^* based on the bootstrap sample $\{(Y_t^*, Z_t)\}_{t=1}^n$. The following is done then on the test data:
   (a) By applying the linear regression technique one obtains $RSS_{H_0}^*$ under the null hypothesis,
   (b) By applying the nonparametric technique one obtains $RSS_{H_1}^*$.
   This gives

$$S^* = \frac{n}{2} \log\left(\frac{RSS_{H_0}^*}{RSS_{H_1}^*} \right).$$

(3) Repeat step 1 and step 2 B times ($B = 1000$) and obtain $S^{*(1)}, ..., S^{*(B)}$.
(4) Use the percentage of $S^{*(i)} > S, \forall i = 1, ..., B$ as an estimate of the p-value of the test.

Remark 2. The bootstrap is a computer-intensive method that provides answers to a large class of statistical inference problems without stringent structural assumptions on the underlying random process generating the data. Since its introduction by Efron (1979), the bootstrap has found its application to a number of statistical problems, including many standard ones, where it has outperformed the existing methodology as well as to many complex problems involving independent data where conventional approaches failed to provide satisfactory answers. Since then, a number of classical statistical texts have been written on the topic (Hall, 1992; Efron and Tibshirani, 1993;

Shao and Tu, 1995; Davison and Hinkley, 1997; Chernick, 1999). The bootstrap does not always work. Examples of bootstrap failure have been reported from the beginning of its development (Bickel and Freedman, 1981; Mammen, 1992; Hu and Hu, 2000). Special care is therefore required when applying the bootstrap in real-life situations (Politis, 1998).

4. ILLUSTRATIVE EXAMPLES

To illustrate the generalized likelihood ratio statistic, we present two examples using the following data set. Consider the nonlinear ARX model defined on

$$Y_t = -0.1Y_{t-1} + 0.2Y_{t-2} + \xi U_{t-1}^2 + 0.3U_{t-2} + e_t, \quad (18)$$

where ξ is a given parameter, e.g. $\xi \in [0,1]$ and $e_t \sim \mathcal{N}(0,\sigma^2)$. We split the data set into a training set ($t = 1,\ldots,100$) and a test set ($t = 101,\ldots,150$).

4.1 Example 1

Let $\xi = 1$ and let the significance level $\alpha = 0.05$. Consider the null hypothesis of linear model

$$H_0 : Y_t = -a_1Y_{t-1} - a_2Y_{t-2} + b_1U_{t-1} + b_2U_{t-2} + e_t, \quad (19)$$

namely, testing whether the generated series was from an ARX(2) model. The LS-SVM regression model (Suykens *et al.*, 2002) with RBF kernel was used to construct the nonparametric estimation. By using 1000 bootstrap samples, we obtained 1000 test statistics S^* (based on the test set). Figure 1 shows the histogram of the S^*. The p-value is estimated as 0.03. Hence the linear ARX model does not fit the data.

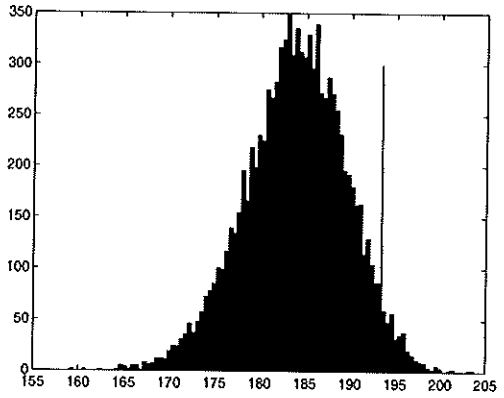

Fig. 1. The distribution by bootstrap of the generalized test statistic based on test data. The line indicates the value of S.

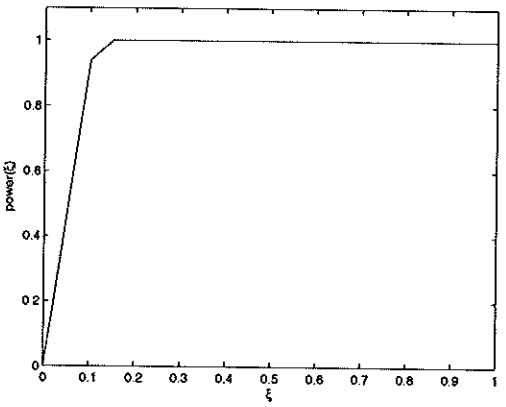

Fig. 2. Power function of the generalized likelihood ratio test based on test data.

4.2 Example 2

Let $\xi \in \{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ and let the significance level $\alpha = 0.05$. Consider the null hypothesis of linear model

$$H_0 : Y_t = -a_1Y_{t-1} - a_2Y_{t-2} + b_1U_{t-1} + b_2U_{t-2} + e_t, \quad (20)$$

The LS-SVM model with RBF kernel was used to construct the nonparametric estimation. The power was computed based on 200 simulations. For each realization, 1000 bootstrap samples were drawn to compute the p-value of the generalized likelihood ratio test. When the p-value is less than 5%, the null hypothesis is rejected. The power is computed as the percentage chance of rejecting the null hypothesis among 200 simulations. The results are summarized in Figure 2. When $\xi = 0$, the power is 5%. This demonstrates that the bootstrap estimate of the null distribution gives a very good approximation of the size of the test. The power increases rapidly to 100% as ξ increases.

5. CONCLUSIONS

We constructed a Monte Carlo test, based on generalized likelihood ratio statistics and bootstrap techniques, for system identification purposes. In the past such tests were only possible for parametric models based on likelihood ratio statistics. The method proposed in this paper is applicable to non-parametric models as well. We illustrated its use for checking the significance of linear ARX models with respect to non-linear ARX models with the linear model as the null hypothesis.

ACKNOWLEDGEMENTS

This research work was carried out at the ESAT laboratory of the K.U. Leuven. Research Council KU Leuven: Concerted Research Action GOA-Mefisto 666, GOA-Ambiorics IDO, several PhD and postdoc & fellow grants; Flemish Government: Fund for Scientific Research Flanders (several PhD/postdoc grants, projects G.0407.02, G.0256.97, G.0115.01, G.0240.99, G.0197.02, G.04-99.04, G.0211.05, G.0080.01, research communities ICCoS, ANMMM), AWI (Bil. Int. Collaboration Hungary/ Poland), IWT (Soft4s, STWW-Genprom, GBOU-McKnow, Eureka-Impact, Eureka-FLiTE, several PhD grants); Belgian Federal Government: DWTC IUAP IV-02 (1996-2001) and IUAP V-10-29 (2002-2006) (2002-2006), Program Sustainable Development PODO-II (CP/40); Direct contract research: Verhaert, Electrabel, Elia, Data4s, IPCOS. JS is an associate professor and BDM is a full professor at K.U.Leuven Belgium, respectively.

REFERENCES

Barnard, G.A. (1963). Discussion of Bartlett, "The spectral analysis of point processes." *Journal of the Royal Statistical Society, Series B* **25**, 264–296.

Bickel, P.J. and Freedman, D.A. (1981). Some Asymptotic Theory for the Bootstrap. *Ann. Statist.* **9**, 1196–1217.

Chen, J.H. and Qin, J. (1993). Empirical likelihood estimation for finite populations and the effective usage of auxiliary information. *Biometrica* **80**, 107–116.

Chernick, M.R. (1997). *Bootstrap Methods.* John Wiley and Sons.

Davison, A.C. and Hinkley, D.V.(1997). *Bootstrap Methods and their Application.* Cambridge University Press.

Efron, B. (1979). Bootstrap methods: Another look at the jacknife *Ann. Statist.***7**, 1–26.

Efron, B. and Tibshirani, R.J. (1993). *An Introduction to the Bootstrap.* Chapman and Hall, New York.

Fan, J. and Huang, L. (2001). Goodness-of-fit test for parametric regression models. *J. Amer. Statist. Assoc.* **96**, 640–652.

Fan, J., Zhang, C. and Zhang, J. (2001). Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Statist.* **29**, 153–193.

Gasser, T., Kohler, W., Muller, H.G., Lago, R., Molinari, L. and Prader, A. (1985). Human height growth: Correlational and multivariate structure of velocity and acceleration. *Ann. of Human Biology.* **12**, 501–515.

Hall, P. (1992). *The bootstrap and Edgeworth Expansion.* Springer-Verlag, New-York.

Hardle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regrssion fits. *Ann. of Statist.* **21**, 1926–1947.

Hart, J.D. (1997). *Nonparametric Smoothing and Lack-of-fit Tests.* Springer-Verlag, New-York.

Hildenbrand, K. and Hildenbrand, W. (1986). On the mean income effect: A data analysis of the U.K. family expenditure survey. In: *Contribution to Mathematical Economics* (Hildenbrand, W. and Mas-Colell, A. Ed.). North-Holland, Amsterdam.

Hu, F. and Hu, J. (2000). A note on Breakdown Theory for Bootstrap Methods. *Statistics and Probability Letters* **50**, 49–53.

Lehmann, E.L. (1997). *Testing Statistical Hypotheses.* Springer-Verlag, New-York.

Ljung, L. (1987). *System Identification: Theory for the user.* Prentice-Hall, Englewood Cliffs, New Jersey.

Mammen, E. (1992). *When does the Bootstrap work? Asymptotic Results and Simulations.* Lecture notes in Statistics, Volume 77, Springer-Verlag, New-York.

Marriott, F.H. (1979). Barnard's Monte Carlo test: How many simulations? *Applied Statistics* **28**, 75–78.

Owen, A.B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrica* **75**, 237–249.

Owen, A.B. (1990). Empirical likelihood ratio confidence regions. *Ann. of Statist.* **18**, 90–120.

Owen, A.B. (1997). *Empirical Likelihood.* Chapman and Hall.

Pestman, W.R. (1991). *Mathematical Statistics.* Walter de Gruyter, New York.

Politis, D.N. (1998). Computer-Intensive Methods in Statistical Analysis. *IEEE Signal Processing Magazine* **15**, 39–55.

Shao, T. and Tu, D. (1995). *The Jacknife and Bootstrap.* Springer-Verlag, New York.

Shen, X., Shi, J and Wong, W.H. (1999). Random sieve likelihood and general regression models. *J. Amer. Statist. Assoc.* **94**, 835–846.

Suykens, J.A.K., Van Gestel, T., De Brabanter, J., De Moor, B., Vandewalle, J. (2002). *Least Squares Support Vector Machines.* World Scientific, Singapore.

Zhang, J and Gijbels, I. (2003). Sieve empirical likelihood and extensions of generalized least squares. *Scand. J. Statist.* **30**, 1–24.