An Empirical Likelihood Goodness-of-Fit Test for Time Series

Song Xi Chen

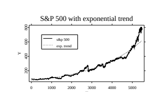
Department of Statistics and Applied Probability

National University of Singapore

Wolfgang Härdle and Torsten Kleinow

Institut für Statistik und Ökonometrie

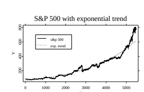
Humboldt-Universität zu Berlin



1. Introduction

Tests of time series models are standard tasks in data analysis.

For nested parametric models there exist a large toolbox.
Goodness of fit tests are designed for this situation.
Comparison with smooth alternatives is natural especially for financial processes, where many observations are available.
The Empirical Likelihood technique is a versatile tool for such situations.



Assume that $\{(X_i, Y_i)\}_{i=1}^n$ is a strictly stationary time series with $X_i \in \mathbb{R}^d$ and $Y_i \in \mathbb{R}$.

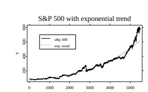
X may be the lagged d-dimensional past

$$Z_t = m(Z_{t-1}, \dots, Z_{t-d}) + \epsilon_t \tag{1}$$

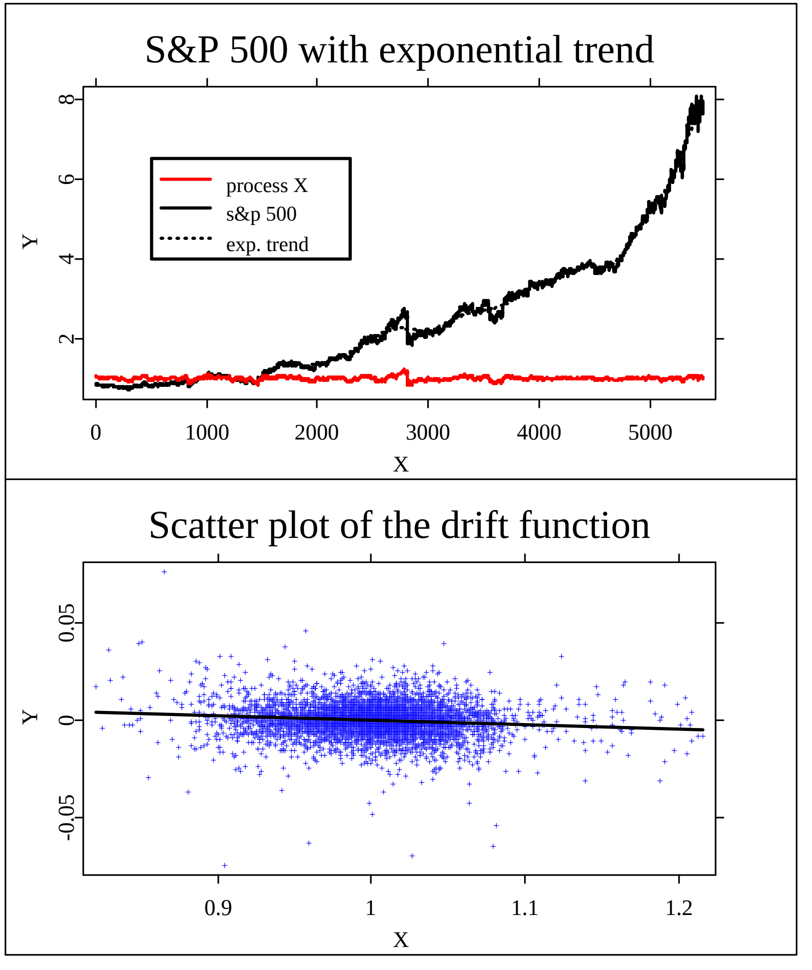
ARCH type processes

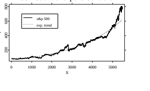
$$\begin{aligned} Z_t &= \sigma_t \xi_t \\ \sigma_t^2 &= \omega + \alpha Z_{t-1}^2 \\ Y_t &= Z_t^2 \\ X_t &= Z_{t-1} \end{aligned}$$

no symmetry of the news impact function $E(Y|x) = m(x)$ is well known (Engle and Gonzalez-Rivera (1991)).



Diffusion models are widely applied in finance.





Daily closing value $S(t)$ of the S&P 500 share index from 31. Dec 1977 to 31. Dec 1997 ($n = 5479$).

Residual series $S(t)/\bar{S}(t)$ is modeled as

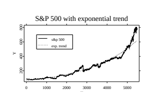
CIR model

$$dZ(t) = \beta\{1 - Z(t)\}dt + \gamma\sqrt{Z(t)}dW(t).$$

OU model

$$dZ(t) = \beta\{1 - Z(t)\}dt + \gamma dW(t).$$

Goal: Test the parametric form of drift and diffusion functions.



Discretising the series leads to (X_i, Y_i) with

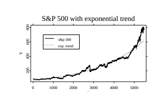
$$\begin{aligned} X_i &= X_{i\Delta} \\ Y_i &= X_{(i+1)\Delta} - X_{i\Delta} \end{aligned}$$

This series is α-mixing and the form of $m(x) = \alpha(1 - x)$ may then be tested using the empirical likelihood method.

Formula framework:

$$\begin{aligned} m(x) &= E(Y|X = x) \\ \sigma^2(x) &= Var(Y|X = x) \end{aligned}$$

$\{m_\theta | \theta \in \Theta\}$ a parametric model.



We are interested in testing

$$H_0 : m(x) = m_\theta(x) \quad \text{for all}\, x \in S$$

$$H_1 : m(x) = m_\theta(x) + c_n \Delta_n(x),$$

where $c_n \to 0$, $\Delta_n(x)$ are bounded functions.

Semiparametric testing problem:

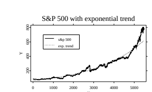
Horowitz (1997)

Bosq (1998)

Empirical likelihood:

Owen (1988)

Hall and La Scala (1990)



√ 1. Introduction

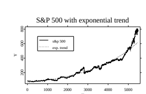
2. Kernel Estimator and Empirical Likelihood

3. Empirical Likelihood Goodness-of-fit Statistic

4. Goodness-of-fit Test

5. Simulation and Application

6. Conclusion



2. Kernel Estimator and Empirical Likelihood

Let K be a d-dimensional standard kernel.

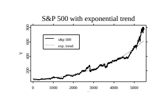
$$K_h(u) = h^{-d} K(h^{-1} u)$$

NW estimator

$$\hat{m}(x) = \frac{\sum_{i=1}^{n} Y_i K_h(x - X_i)}{\sum_{i=1}^{n} K_h(x - X_i)}. \tag{2}$$

smoothed parametric model

$$\tilde{m}_{\hat{\theta}}(x) = \frac{\sum K_h(x - X_i) m_{\hat{\theta}}(X_i)}{\sum_{i=1}^{n} K_h(x - X_i)}$$



Empirical Likelihood (EL)

$$L\{\mu(x)\} = \max \prod_{i=1}^{n} p_i(x)$$

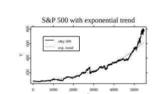
subject to

$$\sum_{i=1}^{n} p_i(x) = 1$$

and

$$\sum_{i=1}^{n} p_i(x) K\left(\frac{x - X_i}{h}\right)\{Y_i - \mu(x)\} = 0$$

Compute the EL $L\{\mu\}$ for $\mu(x) = \hat{m}(x)$ and $\mu(x) = \tilde{m}_{\hat{\theta}}(x)$.

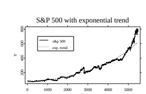


Introduce Lagrange multipliers to obtain

$$p_i(x) = \frac{1}{n}\left[1 + \lambda(x)K\left(\frac{x - X_i}{h}\right)\{Y_i - \mu(x)\}\right]^{-1} \tag{3}$$

where

$$\sum_{i=1}^{n} \frac{K\left(\frac{x-X_i}{h}\right)\{Y_i - \mu(x)\}}{1 + \lambda(x)K\left(\frac{x-X_i}{h}\right)\{Y_i - \mu(x)\}} = 0. \tag{4}$$



The maximum EL is

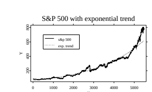
$$p_i(x) = \frac{1}{n} \qquad L\{\mu(x)\} = n^{-n}$$

which is equivalent to

$$\mu(x) = \hat{m}(x).$$

The log-EL ratio test statistic is

$$\ell\{\tilde{m}_{\hat{\theta}}(x)\} = -2\log\frac{L\{\tilde{m}_{\hat{\theta}}(x)\}}{L(\hat{m}(x))} = -2\log[L\{\tilde{m}_{\hat{\theta}}(x)\}n^n].$$



Lemma 1. Under appropriate assumptions,

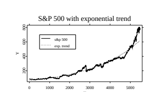
$$\sup_{x\in S}|\lambda(x)| = o_p\{(nh^d)^{-1/2}\log(n)\}.$$

Denote by $\gamma(x) = \tilde{O}_p(\delta_n)$: $\sup_{x\in S}|\gamma(x)| = O_p(\delta_n)$.

Define $\bar{U}_j(x) = (nh^d)^{-1}\sum_{i=1}^n \left[K\left(\frac{x-X_i}{h}\right)\{Y_i - \tilde{m}_{\hat{\theta}}(x)\}\right]^j$

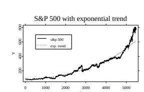
Obtain $\lambda(x) = \bar{U}_2^{-1}(x)\bar{U}_1(x) + \tilde{o}_p\{(nh^d)^{-1}\log^2(n)\}$

$$\begin{aligned}
\ell\{\tilde{m}_{\hat{\theta}}(x)\} &= -2log[L\{\tilde{m}_{\hat{\theta}}(x)\}n^n] \\
&= 2\sum_{i=1}^n \log[1+\lambda(x)K\left(\frac{x-X_i}{h}\right)\{Y_j - \tilde{m}_{\hat{\theta}}(x)\}] \\
&= 2(nh^d)\lambda(x)\bar{U}_1 - (nh^d)\lambda^2(x)\bar{U}_2 + \tilde{o}_p\{(nh^d)^{-1/2}\log^3(n)\} \\
&= (nh)^d\bar{U}_2^{-1}(x)\bar{U}_1^2(x) + \tilde{o}_p\{(nh^d)^{-1/2}\log^3(n)\}.
\end{aligned}$$



The log EL ratio is asymptotically equivalent to a studentized L_2 distance between $\tilde{m}_{\hat{\theta}}$ and $\hat{m}$.

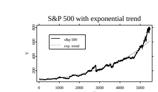
$$
\begin{aligned}
\ell\{\tilde{m}_{\hat{\theta}}(x)\} &= \bar{U}_2^{-1}\bar{U}_1^2 + \tilde{o}_p\{(nh^d)^{-1/2}\log^3(n)\} \\
&= V^{-1}(x;h)\{\hat{m}(x) - \tilde{m}_{\theta}(x)\}^2 + \tilde{O}\{(nh^d)^{-1}h\log^2(n)\}
\end{aligned}
$$



3. Empirical Likelihood Goodness-of-fit Statistic

Choose k_n equally spaced lattice points $t_1, t_2, \cdots, t_{k_n}$ in $[0,1]^d$.
A simple choice: $k_n = (2h)^{-d}$.
Global goodness-of-fit test:

$$\ell_n(\tilde{m}_{\hat{\theta}}) = \sum_{j=1}^{k_n} \ell\{\tilde{m}_{\hat{\theta}}(t_j)\} \tag{5}$$

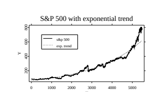


Theorem 1

$$k_n^{-1}\ell_n(\tilde{m}_{\hat{\theta}}) = (nh^d)\int \frac{\{\hat{m}(x)-\tilde{m}_\theta(x)\}^2}{V(x)}dx + O_p\{k_n^{-1}\log^2(n) + h\log^2(n)\}.$$

Härdle and Mammen (1993):

$$T_n = nh^{d/2}\int\{\hat{m}(x)-\tilde{m}_\theta(x)\}^2\pi(x)dx$$



Theorem 2

$$k_n^{-1}\ell_n(\tilde{m}_{\hat{\theta}}) \xrightarrow{\mathcal{L}} \int_S \mathcal{N}^2(s)ds$$

where $\mathcal{N}$ is a normal process on $S = [0,1]^d$ with mean

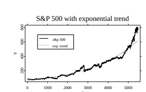
$$E\{\mathcal{N}(s)\} = h^{d/4}\Delta_n(s)/\sqrt{V(s)}$$

and covariance

$$\Omega(s,t) = Cov\{\mathcal{N}(s),\mathcal{N}(t)\} = \sqrt{\frac{f(s)\sigma^2(s)}{f(t)\sigma^2(t)}}\,\frac{W_0^{(2)}(s,t)}{\sqrt{W_0^{(2)}(s,s)W_0^{(2)}(t,t)}}$$

where

$$W_0^{(2)}(s,t) = \int_{y\in S} h^{-d}K\{(s-y)/h\}K\{(t-y)/h\}dy. \qquad (6)$$

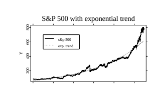


4. Goodness-of-Fit Test

Derive the asymptotic distribution of $k_n^{-1}\ell_n(\tilde{m}_{\hat{\theta}})$ by discretisation of $\int_S \mathcal{N}^2(s)ds$ as $(k_n)^{-1}\sum_{j=1}^{k_n}\mathcal{N}^2(t_j)$.
Choose $k_n = (2h)^{-d}$ with $|t_j - t_k| \geq 2h$, $j \neq k$:

$$\sum_{j=1}^{k_n}\mathcal{N}^2(t_j) \sim \chi^2_{k_n}(\gamma_{k_n}) \tag{7}$$

where $\gamma_{k_n} = h^{d/4}\{\sum_{j=1}^{k_n}\Delta_n^2(t_j)/V(t_j)\}^{1/2}$ is the non centrality parameter.



Asymptotic normality:

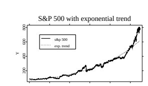
$$k_n^{-1}\ell_n(\tilde{m}_{\hat{\theta}}) \xrightarrow{\mathcal{L}} N\left(1 + h^{1/2}\int \Delta_n^2(s)V^{-1}(s)ds, 2hK^{(4)}(0)\{K^{(2)}(0)\}^{-2}\right) \tag{8}$$

test for H_0:

$$k_n^{-1}\ell_n(\tilde{m}_{\hat{\theta}}) > 1 + z_\alpha\{K^{(2)}(0)\}^{-1}\sqrt{2hK^{(4)}(0)} \tag{9}$$

asymptotic power:

$$1 - \Phi\left\{z_\alpha - \frac{K^{(2)}(0)\int \Delta_n^2(s)V^{-1}(s)ds}{\sqrt{2K^{(4)}(0)}}\right\}. \tag{10}$$

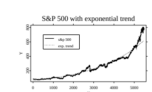


5. Simulation and Application

$$Y_i = 2Y_{i-1}/(1 + Y_{i-1}^2) + c_n Sin(Y_{i-1}) + \sigma(Y_{i-1})\eta_i$$

Here

$$\begin{aligned} X_i &= Y_{i-1} \\ \sigma(x) &= exp(-x^2/4) \\ \eta_i &\sim U[-1, 1] \\ n &= 500, 1000 \\ c_n &= 0, 0.03, 0, 06 \end{aligned}$$



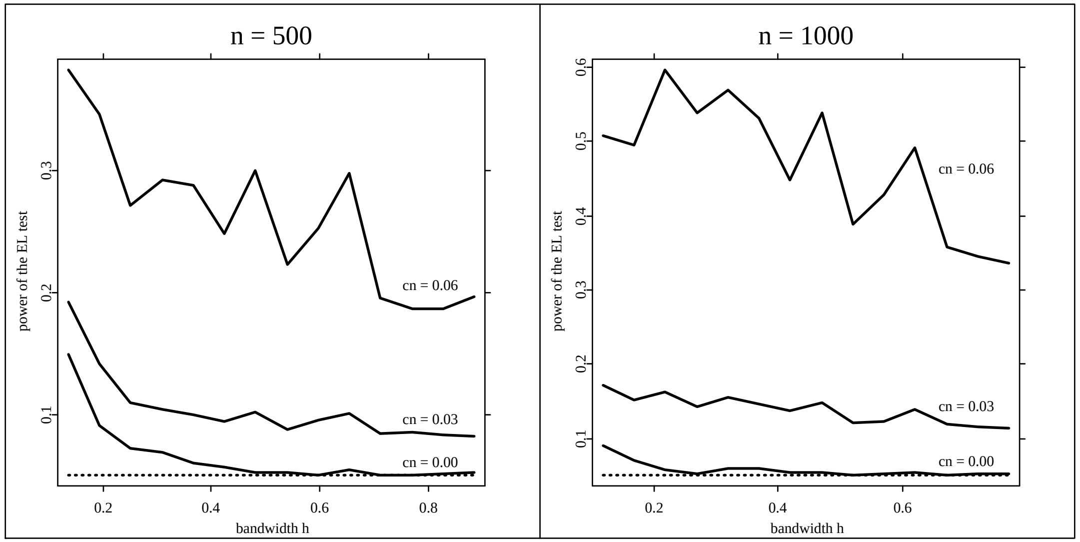
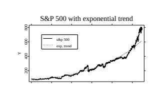


Figure 1: Power of the empirical likelihood test. The dotted lines indicate the 5% level.

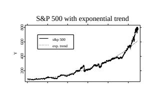
Trend of decreasing power when h increases. This comes from discretisation.



Application to S&P 500 data:

$$H_0 : m(x) = \beta(1 - x)$$

parametric estimate: $\hat{\beta} = 0.00968$.
The estimator is the mean value of $\hat{\beta}_1$ and $\hat{\beta}_2$, where $\hat{\beta}_1$ is based on the marginal distribution of X while $\hat{\beta}_2$ is based on the autocorrelation function of X (Härdle, Kleinow, Korestelev, Logeay, Platen (2001)). Global smoothing bandwidth was determined by cross validation: $h_{cv} = 0.053$.



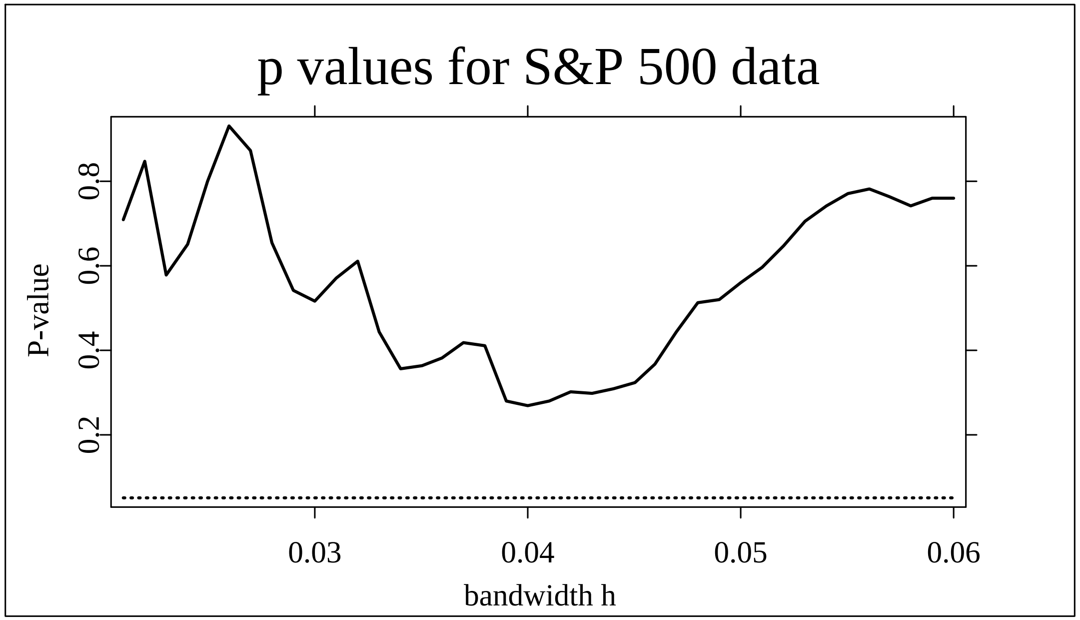
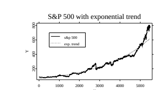


Figure 2: P-values of the empirical likelihood test for the S&P data. The dotted line indicates the 5% level.



6. Conclusion

- The proposed test compares the parametric model with a kernel smoothing estimator.
- The test statistic is based on the asymptotics of the empirical likelihood.
- Its asymptotic distribution is known which avoids bootstrap and secondary plug-in estimation.
- The null hypothesis of a diffusion process with drift $m(x) = a(1 - x)$ is not rejected for the normalized S&P 500 data.