Computational Statistics and Data Analysis 56 (2012) 3080–3090

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

On the online estimation of local constant volatilities

Roland Fried *

Department of Statistics, TU Dortmund University, 44221 Dortmund, Germany

ARTICLE INFO

Article history:
Available online 27 February 2011

Keywords:
Heteroscedasticity
Structural breaks
Heavy tails
Outliers
Tests for equality of variances

ABSTRACT

Time varying volatilities in financial time series are commonly modeled by GARCH or by stochastic volatility models. Models with piecewise constant volatilities have been proposed recently as nonparametric alternatives. Following the latter approach, a procedure for online approximation of the current volatility is constructed by combining one-sided localized estimation of the variability with sequential testing for a change in it. A robust nonparametric framework is assumed since many financial time series show tails heavier than the Gaussian. A two-sample test for a change in variability is proposed, which works well even in case of skewed distributions.

1. Introduction

The returns $z(t) = \log(p(t)/p(t-1))$ of risky assets in a period $t \in \mathbb{Z}$, with $p(t)$ being the price at the end of period t, are commonly modeled as

$$Z(t) = \sigma(t)E(t), \tag{1}$$

where $(\sigma(t) : t \in \mathbb{Z})$ is a sequence of time-varying volatilities and $(E(t) : t \in \mathbb{Z})$ is a white noise process with unit variance, which is independent from the volatilities $(\sigma(t) : t \in \mathbb{Z})$. The GARCH(1, 1) model is popular for describing the time-varying behavior of the latter process,

$$\sigma^2(t) = \alpha_0 + \alpha_1 Z^2(t-1) + \beta_1 \sigma^2(t-1). \tag{2}$$

As an alternative, Mercurio and Spokoiny (2004) and Granger and Stărică (2005) point out that the volatility of many financial time series can be represented adequately by piecewise constant models. A piecewise constant behavior of the volatility provides an explanation of the long memory effects found in many financial time series, since these can be artificially generated by structural breaks (Mikosch and Stărică, 2000).

For fitting piecewise constant volatilities one needs to determine time intervals within which the volatility can be approximated by a constant and detect changes between subsequent intervals. Davies et al. (2012) make use of the fact that under the additional assumptions that $(E(t) : t \in \mathbb{Z})$ is Gaussian and $(\sigma(t) : t \in \mathbb{Z})$ a sequence of constants we have in model (1)

$$\sum_{t \in I} \frac{Z^2(t)}{\sigma^2(t)} \sim \chi^2_{|I|}, \tag{3}$$

where χ^2_m denotes the χ^2-distribution with m degrees of freedom and $I \subset \{1, 2, \ldots, N\}$ is a nonempty time interval of width $|I|$. These authors then search a piecewise constant volatility function which minimizes the number of intervals on

* Tel.: +49 231 755 3119; fax: +49 231 755 3454.
E-mail address: fried@statistik.tu-dortmund.de.

0167-9473/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2011.02.012

which $(\sigma(t) : t \in \mathbb{Z})$ is constant under the restriction that it is locally adequate, meaning that

$$\chi^2_{|I|,(1-\alpha_N)/2} \leq \sum_{t\in I} \frac{Z^2(t)}{\sigma^2(t)} \leq \chi^2_{|I|,(1+\alpha_N)/2}$$

for all intervals $I \subset \{1, 2, \ldots, N\}$ within which $\sigma^2(t)$ is constant. Here $1 - \alpha_N \in (0, 1)$ is a confidence level depending on the number N of observation times.

Spokoiny (2009) puts forward another local approach based on a local multiscale change point analysis and likelihood ratio tests. The volatility is considered not to be constant within a time interval $\{t_1, \ldots, t_2\}$ whenever any of the likelihood ratio tests against the alternative of two constant pieces rejects the null hypothesis. Under the Gaussian assumption, the likelihood ratio test statistic for a change after time point $\tau \in \{t_1, \ldots, t_2 - 1\}$ equals the weighted sum of the Kullback Leibler information KL between the normal distributions corresponding to the maximum likelihood estimator $\hat{\sigma}^2$ derived from the full interval on the one side, and $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ derived from the two subintervals $\{t_1, \ldots, \tau\}$ and $\{\tau + 1, \ldots, t_2\}$ on the other side,

$$\begin{aligned}(\tau - t_1 + 1)\mathrm{KL}(\hat{\sigma}_1^2, \hat{\sigma}^2) + (t_2 - \tau)\mathrm{KL}(\hat{\sigma}_2^2, \hat{\sigma}^2) &= -(\tau - t_1 + 1)\frac{\log(\hat{\sigma}_1^2/\hat{\sigma}^2) + 1 - \hat{\sigma}_1^2/\hat{\sigma}^2}{2} \\ &\quad - (t_2 - \tau)\frac{\log(\hat{\sigma}_2^2/\hat{\sigma}^2) + 1 - \hat{\sigma}_2^2/\hat{\sigma}^2}{2}. \end{aligned} \tag{4}$$

Under the null hypothesis of no change, these test statistics are asymptotically χ_1^2-distributed as the length of both subintervals goes to infinity. Spokoiny suggests a sophisticated adaptive testing procedure with critical values determined to achieve good approximations of the true volatility.

Although the approaches of Davies et al. (2012) and Spokoiny (2009) are similar in spirit, a comparison of them is difficult because they aim at different objectives, namely a good approximation with a minimal number of intervals of constancy and an optimal global approximation, respectively. These authors agree that in practice the distribution of $E(t)$ usually has tails heavier than the Gaussian and can be better approximated by a t-distribution with between 5 and 10 degrees of freedom. Nonetheless they work within the Gaussian framework because of its simplicity and its analytical tractability.

We apply a robust nonparametric framework for estimation of the volatilities $(\sigma(t) : t \in \mathbb{Z})$, avoiding stringent assumptions on the innovation process $(E(t) : t \in \mathbb{Z})$. Our approach is designed to work online, albeit some time delays cannot be avoided. We apply sequential testing to detect changes of the variability, comparing the data in a test window of the most recent observations to the data in a reference window of already controlled 'old' observations. This allows application of two-sample tests for a difference in variance. There are many nonparametric tests for a difference of the variance in two samples as alternatives to the Gaussian F-test, but most of them are unreliable if the distributions are skewed (e.g. Shoemaker, 1999).

A further contribution of this paper is the construction of nonparametric two-sample tests for a difference in variance, which are robust against outliers and heavy tails and work under asymmetry. For this construction we transform the problem of detecting a change of the variability from σ_1^2 to σ_2^2 after time τ in model (1) to the problem of detecting a change of location of size $\log(\sigma_2^2) - \log(\sigma_1^2)$ in the log-transformed data

$$X(t) = \log(Z^2(t)) = \log(\sigma^2(t)) + \log(E^2(t)). \tag{5}$$

Since the terms $\log(E^2(t))$, $t \in \mathbb{Z}$, are again i.i.d. random variables if this is true for the $E(t)$, $t \in \mathbb{Z}$, a change of location in the transformed series $(X(t) : t \in \mathbb{Z})$ is equivalent to a jump in $\log(\sigma^2(t))$, $t \in \mathbb{Z}$, i.e. a piecewise constant volatility of $(Z(t) : t \in \mathbb{Z})$ is equivalent to a piecewise constant mean level of $(X(t) : t \in \mathbb{Z})$. To test for a change of location after a given time point τ in the log-transformed series we can apply e.g. the robust nonparametric two-sample test suggested by Fried and Dehling (under revision) to the windows $\{\tau - m + 1, \ldots, \tau\}$ and $\{\tau + 1, \ldots, \tau + n\}$ of widths m and n, respectively.

When the test does not reject the null hypothesis of a locally constant variability, we can estimate $\sigma^2(\tau + n)$ from the combined moving estimation window $Z(\tau - m + 1), \ldots, Z(\tau + n)$, so that the sequence of estimates adapts to slow changes of the volatility. Otherwise the estimates are calculated from the data after time τ, only, and the testing procedure is paused until enough data points have been observed after time τ.

The question of systematic changes of volatility and structural breaks therein has been raised also for other economic time series like gross domestic products or interests, see Cavaliere and Taylor (2009) and the references cited therein. Besides the null hypothesis of constant variance treated here, retrospective tests have been suggested for the more general hypothesis of stationary volatility, under control of a given overall significance level (e.g. Cavaliere and Taylor, 2007). The interest here is online approximation of time varying volatility, which corresponds to the conditional variance of $Z(t)$ given its past in GARCH models, not the marginal variance. The significance level α of the tests for a change in variability at a specific time point merely acts as a smoothing parameter: it determines the width of the time windows within which the volatility is regarded as approximately constant. We do not aim at controlling the overall probability of false detection of a change in variance at any time point. This is also a basic difference to other monitoring procedures for online detection of changes in the variance of heteroscedastic time series, as proposed e.g. by Horváth et al. (2006). We combine the testing procedure