Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

Q1 Regime switching model for financial data: Empirical risk analysis

Q2 Khaled Salhi [a,b,c,d], Madalina Deaconu [a,b,c], Antoine Lejay [a,b,c,*], Nicolas Champagnat [a,b,c], Nicolas Navet [d,e]

[a] *Université de Lorraine, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France*
[b] *CNRS, Institut Elie Cartan de Lorraine, UMR 7502, Vandoeuvre-lès-Nancy, F-54506, France*
[c] *Inria, Villers-lès-Nancy, F-54600, France*
[d] *Alphability, F-54500 Vandoeuvre-lś-Nancy, France*
[e] *FSTC/CSC Research Unit/Lassy lab., Université du Luxembourg, L-1359, Luxembourg*

HIGHLIGHTS

- This paper introduces a regime switching model for Value-at-Risk estimation.
- Hidden Markov models and extreme value theory are combined into a hybrid model.
- The regime switching model is applied to real data NYSE Euronext stocks.
- Classifying data in two states permits a fast detection of regime switching.
- This new model increases predictive performance of VaR forecasting.

ARTICLE INFO

Article history:
Received 3 October 2013
Received in revised form 11 February 2015
Available online xxxx

Keywords:
Value-at-Risk
Power tail distribution
Hidden Markov Model
Regime switching

ABSTRACT

This paper constructs a regime switching model for the univariate Value-at-Risk estimation. Extreme value theory (EVT) and hidden Markov models (HMM) are combined to estimate a hybrid model that takes volatility clustering into account. In the first stage, HMM is used to classify data in crisis and steady periods, while in the second stage, EVT is applied to the previously classified data to rub out the delay between regime switching and their detection. This new model is applied to prices of numerous stocks exchanged on NYSE Euronext Paris over the period 2001–2011. We focus on daily returns for which calibration has to be done on a small dataset. The relative performance of the regime switching model is benchmarked against other well-known modeling techniques, such as stable, power laws and GARCH models. The empirical results show that the regime switching model increases predictive performance of financial forecasting according to the number of violations and tail-loss tests. This suggests that the regime switching model is a robust forecasting variant of power laws model while remaining practical to implement the VaR measurement.

1. Introduction

The Value-at-Risk (VaR) is one of the main risk indicators for management of financial portfolios [1]. It is the threshold above which a loss over a chosen time horizon occurs with at most a given level of confidence.

* Corresponding author at: Inria, Villers-lès-Nancy, F-54600, France.
E-mail addresses: Khaled.Salhi@inria.fr (K. Salhi), Madalina.Deaconu@inria.fr (M. Deaconu), Antoine.Lejay@inria.fr (A. Lejay), nicolas.navet@uni.lu (N. Champagnat), Nicolas.Champagnat@inria.fr (N. Navet).

http://dx.doi.org/10.1016/j.physa.2016.05.002
0378-4371/

The VaR may be estimated either by parametric or non-parametric techniques. The non-parametric ones use only the empirical distributions (historical, resampling) without fitting a model. Due to the small amount of available data, it does not provide an accurate estimate for the probability of extreme events. The parametric approach overrides partly the problem induced by the lack of data by fitting the parameters of a model on historical data and computing afterwards the VaR, either by analytical or numerical methods.

The RiskMetric methodology [2] is widely used to estimate the risk associated to a portfolio by establishing quantitative relations between variations of the risk indicator with respect to risk factors (stocks and prices of derivatives for example). Nowadays this methodology incorporates heavy tail distributions, but it was initially developed in the Gaussian context, which is still prevalent in risk management and enforced by Basel Accord [3].

However, actual regulations and standard procedures for computing the VaR, based mainly on the Gaussian world, have been invalidated by many studies (see *e.g.* Refs. [4,5]) as they strongly underestimate the extreme events observed in the market.

The successive financial crises since 1987 led to a greater attention in modeling tail behavior of the induced returns distributions, and to the use of Extreme Value Theory (EVT) as a central concept in Risk Management [6].

Neither single model nor statistical methodology is acknowledged as standard for dealing with heavy-tails. In this study, we construct a variant of power law tails of the distributions by taking into consideration the regime switching between crisis and steady periods. We compare our model to several known models (stable, Generalized Pareto tails of distribution and GARCH). We consider heavy tail distributions for the losses L with distribution function F, typically, generalized Pareto (power laws):

$$1 - F(x) = \mathbb{P}(L \geqslant x) = \frac{\ell(x)}{x^{\alpha}}, \tag{1}$$

where ℓ is a function of slow variation and α is the *tail index* which summarizes the heaviness of the tail distribution and characterizes also the existence of moments.

In our approach, we aim to benefit from the stability of power law models in the VaR forecasting and the detection of volatility clustering given by conditional models. We assume the existence of two states: crisis and steady, and we classify data in two regimes using hidden Markov models (HMM). Then, the power law tail distribution is estimated from the past crisis and steady periods. The model gives more weight to the current regime, which reduces under-estimations and over-estimations at the beginning of crisis and steady periods.

We focus here only on univariate distributions. The case of several stocks leads to higher complexity, as the notion of VaR itself is not properly defined (see *e.g.* Ref. [7]). The multivariate case will be subject of further studies.

Our models of risk are intensively tested on historical market data from NYSE Euronext Paris. We focus only on daily returns of single stock prices, which is the practical situation of small investors and for which the problem of model calibration is the most difficult because of the small amount of data.

Outline. In Section 2, we give an overview of heavy-tailed models used in finance, the markets on which they have been applied and the estimators for power laws. In Section 3, we introduce our methodology for estimating the parameters and performing the backtesting. Finally, in Section 4, we introduce our dataset and discuss the conclusions we draw from the study over a selection of 56 stocks from NYSE Euronext Paris. Perspectives and conclusions are then highlighted in Section 5.

2. Empirical evidence for heavy tails

2.1. Evidences for the power law in financial markets

It is widely acknowledged that prices and returns of stocks obey to general laws usually called "stylized facts" [8]. Skewness and heavy-tail are the two main properties of observed prices which are not verified by the Black&Scholes models.

Although stable distributions have been proposed since the 60's [9,10] and an alternative model (finite variance subordinated log-normal distributions) has been proposed in 1973 by P. Clark [11], the systematic use of Extreme Value Theory (EVT) is recent, as the crash of 1987 urged for a better understanding of large losses. For early occurrences of the use of EVT focusing only on the tail distributions, let us cite [12–14].

Developed markets have been investigated as well, mainly through market indices: S&P 500, Dow Jones and Nasdaq [15], German DAX Stocks [16], Australian ASX-ALL [17], Nikkei and Eurostoxx 50 [18], etc.

Stable distributions and processes exhibit generalized Pareto tail distributions as in (1). Yet, choosing this model imposes tail index $\alpha < 2$. This parameter is difficult to estimate, especially when close to 2 [19]. Many critical studies (see *e.g.* Refs. [19,20]) show that the tail index is greater than 2, and around 3 for short terms returns.

To estimate the generalized Pareto distribution, let us consider that the return X at a given time satisfies (1) and that n successive returns $(X_1, \ldots, X_n)$ are independent or at least stationary. It is a crucial and complex problem to estimate α and $\ell(x)$ written in a parametric or semi-parametric form (for example, $\ell(x) = C$ or $\ell(x) = C_1 + C_2 x^{-\beta} + o(x^{-\beta})$), as well as the threshold x_0 above which the previous expressions for $\ell(x)$ are valid. For studying the tail of the distribution of X, we use Q3 *order statistics* $(X_{(1)}, \ldots, X_{(n)})$ of $(X_1, \ldots, X_n)$ with $X_{(1)} \leq X_{(2)} \leq \cdots b \leq X_{(n)}$.

Please cite this article in press as: K. Salhi, et al., Regime switching model for financial data: Empirical risk analysis, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2016.05.002