Physica A 396 (2014) 114–133

Contents lists available at ScienceDirect

Physica A

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

Modeling record-breaking stock prices

Gregor Wergen *

Institut für Theoretische Physik, Universität zu Köln, 50937 Köln, Germany

HIGHLIGHTS

- Record statistics of correlated stochastic processes such as the AR(1) and the GARCH(1, 1) model.
- Record-breaking daily stock prices and daily returns are modeled.
- Record-breaking daily stock prices can be modeled with a autoregressive GARCH(1, 1) model.
- Analysis of the full distribution of the record number of daily stock prices.
- We report on the weekly distribution of record-breaking stock prices.

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form 3 November 2013
Available online 9 November 2013

Keywords:
Record statistics
Extreme value statistics
Extreme events in financial markets
Random walks
Autoregressive processes
GARCH-models

ABSTRACT

We study the statistics of record-breaking events in daily stock prices of 366 stocks from the Standard and Poor's 500 stock index. Both the record events in the daily stock prices themselves and the records in the daily returns are discussed. In both cases we try to describe the record statistics of the stock data with simple theoretical models. The daily returns are compared to i.i.d. RVs and the stock prices are modeled using a biased random walk, for which the record statistics are known. These models agree partly with the behavior of the stock data, but we also identify several interesting deviations. Most importantly, the number of records in the stocks appears to be systematically decreased in comparison with the random walk model. Considering the autoregressive AR(1) process, we can predict the record statistics of the daily stock prices more accurately. We also compare the stock data with simulations of the record statistics of the more complicated GARCH(1, 1) model, which, in combination with the AR(1) model, gives the best agreement with the observational data. To better understand our findings, we discuss the survival and first-passage times of stock prices on certain intervals and analyze the correlations between the individual record events. After recapitulating some recent results for the record statistics of ensembles of N stocks, we also present some new observations for the weekly distributions of record events.

1. Introduction

Not only because of the recent financial crises, the study of extremes in stock markets is of great importance for scientists and economists [1–4]. Traders are eagerly interested in the statistics of extreme events in stock prices at the world's stock exchanges. In the context of extreme and dramatic developments in finance, people also talk a lot about record-breaking events. A record is an entry in a series of events that exceeds all previous values. At the stock markets, a record stock price is often considered to be an important and remarkable event that attracts more attention and media coverage than others.

In recent years, the theory of records has found many applications in various areas of science. Most extensively studied was the statistics of record temperatures and their connection with global warming [5–10]. In 2010, Wergen and Krug [7]

* Tel.: +49 2214707396.
E-mail address: gw@thp.uni-koeln.de.

0378-4371/$ – see front matter
http://dx.doi.org/10.1016/j.physa.2013.11.001

presented a simple analytical model that predicts the effect of climatic change on the occurrence of daily and monthly temperature records to a good accuracy (see also [10,9]). Furthermore, the statistics of records found applications in evolutionary biology [11], physics [12–14], hydrology [15] and of course also in sports [16,17]. Additionally, a lot of progress was made from the mathematical point of view. Often motivated by the multitude of applications, the theory of records from time-dependent and correlated random variables was developed further (see for instance [18]). The interesting problem of the record statistics of independent random variables with a linear drift is well understood now [18–20]. Also the ramifications involved with discrete distributions and ties due to rounding were studied [18,21–23]. But most important for this work was the full characterization of the universal record statistics of symmetric random walks by Majumdar and Ziff [24] in 2008.

Recently, we started analyzing the statistics of record-breaking stock prices. Preliminary results were published in 2011 in the context of a study of the record statistics of biased random walks [25] (see also [26]) and, in 2012, in an analysis of ensembles of independent random walkers [27]. The purpose of this article is to present a more thorough and comprehensive discussion of record-breaking events in stock prices and returns. In particular, we will discuss and explain the occurrence of new upper and lower records in stock data by comparing them to various stochastic models.

As in Ref. [25], we study record events in daily stock data from the Standard and Poor's 500 (S&P 500) stock index [28]. Our data set contains 5000 consecutive trading days from 366 stocks that stayed in the S&P 500 for the entire time-span from January 1, 1990 to March 31, 2009. We are interested both in the record events in the time series of the stocks themselves and in record-breaking daily returns. In a series of stock prices $S_0, S_1, \ldots, S_n$, we have a record at the nth day if

$$S_n > \max\{S_0, S_1, \ldots, S_{n-1}\}. \tag{1}$$

Analogously, we have a record breaking return $\Delta_n := S_n - S_{n-1}$, if

$$\Delta_n > \max\{\Delta_0, \Delta_1 \ldots, \Delta_{n-1}\}, \tag{2}$$

where we set $\Delta_0 = 0$. When we consider the time series of stock prices S_i or returns Δ_i, the most important quantity for us is the probability that a certain entry in such a series is a record. This probability P_n for a stock price S_n is defined as

$$P_n := \text{Prob}\left[S_n > \max\{S_0, S_1 \ldots, S_{n-1}\}\right]. \tag{3}$$

For the returns Δ_i we define the probability p_n in the same manner:

$$p_n := \text{Prob}\left[\Delta_n > \max\{\Delta_0, \Delta_1, \ldots, \Delta_{n-1}\}\right]. \tag{4}$$

In the following, we will also refer to these quantities as the *record rates*. Of similar importance are the closely related record numbers of the stock prices R_n and the returns r_n, the numbers of records that occur in a time series up to step n. One obtains the important mean record numbers $\langle R_n\rangle$ and $\langle r_n\rangle$ of the stocks and the returns by summing over the respective record rates:

$$\langle R_n\rangle := \sum_{k=0}^{n} P_k \quad \text{and} \quad \langle r_n\rangle := \sum_{k=0}^{n} p_k. \tag{5}$$

This article discusses the record rates and record numbers of daily stock prices and returns in the S&P 500 and compares them with several simple stochastic models such as simple i.i.d. RVs or biased random walks. The aim of this work is to better understand the occurrence of record-breaking events in the stock markets and to find useful and accurate models that reproduce and predict them correctly.

Since this work summarizes multiple observations and results, we will now give a short outline of the rest of this article. We start by briefly recapitulating some important classical results from the theory of records in time series of independent and identically distributed (i.i.d.) random variables (RVs) in Section 2. There, we also present the findings for the symmetric random walk derived by Majumdar and Ziff [24]. Thereafter we discuss the record statistics of biased random walks with a Gaussian jump distribution following the results derived in Wergen et al. [25] and Majumdar et al. [29]. Subsequently, in Section 3, we introduce the more complicated, so-called autoregressive AR(1) (see [30,31]) process, which might be able to describe the statistics of record-breaking stock prices more accurately. We analyze and discuss its record statistics numerically.

As a second alternative to the simple random walk model we will also introduce the generalized autoregressive conditional heteroscedasticity (GARCH) model [32], which is particularly often used to model financial time series [33,34]. In this more complicated model a second stochastic process is employed that describes the time dependence of the variance of the increments. We will discuss the record statistics of a specific GARCH process employing numerical simulations.

Section 4 is about the record statistics of individual stocks from the S&P 500 index. After introducing the data and analyzing some of the basic statistical properties of the time series of stock prices, we first have a look at the record statistics of the daily returns. We compute the record rate and the mean record number of the daily returns and discuss the correlations between individual return records. Then the record statistics of the stock prices themselves are analyzed. Here, we first compare the stocks to the simple Geometric Random Walk model of stock prices, before we discuss the stock records in the context of the more complicated AR(1) model and the GARCH model. Following this, we also consider the full distribution of the record number of stock prices and compare this distribution with theoretical predictions. We will briefly discuss the first-passage statistics and survival probabilities of the stocks, since they are closely related to the statistics of records in random walks.