Journal of the Korean Statistical Society 45 (2016) 409–421

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Modeling discrete stock price changes using a mixture of Poisson distributions

Rasitha R. Jayasekare [a,*], Ryan Gill [b], Kiseop Lee [b]

[a] Department of Mathematics and Actuarial Science, Butler University, Indianapolis, IN 46208, United States

[b] Department of Mathematics, University of Louisville, Louisville, KY 40292, United States

ARTICLE INFO

Article history:
Received 12 February 2015
Accepted 6 January 2016
Available online 1 February 2016

AMS 2000 subject classifications:
62F12
62P20

Keywords:
Poisson mixture model
Asymptotic confidence estimation
Liquidity risk
Market microstructure

ABSTRACT

We study discrete price changes due to the size of a trade in the market microstructure model. We use a mixture of Poisson distributions to model the discrete changes in stock price. The parameters are estimated using the Expectation–Maximization (EM) algorithm with mixing probabilities which depend on order size. Consistency and asymptotic normality of a sequence of estimators are proved, and asymptotic confidence intervals for functions of the parameters are derived. We test the method with simulated and real data.

Contents

* Corresponding author.
E-mail addresses: rjayasek@butler.edu (R.R. Jayasekare), ryan.gill@louisville.edu (R. Gill), kiseop.lee@louisville.edu (K. Lee).

http://dx.doi.org/10.1016/j.jkss.2016.01.002
1226-3192/

1. Introduction

Most traditional price models in hedging, pricing and optimization in financial mathematics such as the Black–Scholes model are macro-movement models in the sense that these models are focused on the improvement of prices in a relatively long time scale such as weekly and monthly closing prices. As the market becomes more computerized and trading frequencies get much higher, modeling of micro-movement data becomes more important. Many traders nowadays are day-traders, who make several trading decisions in a few hours, differently from traditional hedgers and option traders.

Micro-movements of the stock price are recorded hourly or even tick-by-tick. Those micro-movements are quite different from the macro-movements, since we can no longer ignore microstructure noises. Discussion on market microstructure noises can be found in Zeng (2003).

A tick-by-tick, or trade-by-trade stock price moves discretely, and there are many sources of noise affecting the movements such as clustering, roundoff, and news arrivals. In this study, we focus on the effect of the order size. Traditional asset pricing models ignored market frictions such as transaction cost and liquidity cost. This means that one can buy or sell infinite amount of shares without changing the price. Certainly this is not the case we observe in a real market. Intuitively, it is expected that the price goes up as the size of a buy order goes up, and it goes down as the size of the sale order goes up, although that phenomenon is noised by other factors related to the market.

The effect of the order size on the price change has been studied in Bank and Baum (2004), Çetin, Jarrow, and Protter (2004), Çetin and Rogers (2007) and Gill, Lee, and Song (2007). Recently, related studies using the limit order book dynamics have been done by Biais and Weill (2009), Cont, Stoikov, and Talreja (2010) and Rosu (2009). In this study, we suggest a new method to explain the tick by tick price change using a Poisson mixture model.

It is important to consider separately increases and decreases in the stock price when analyzing changes of the stock price. Therefore, we investigate how the stock price changes with respect to the order size as a mixture of 'stock price increments' and 'stock price decrements'. In this situation where there are heterogeneous subpopulations within the population, separate models of the subpopulations as well as a model for the relative frequency with which observations are observed from each subpopulation are needed to appropriately model the entire population. The idea of using mixture models in such cases goes back to a classic paper by Karl Pearson in 1894 which used a mixture of two normal distributions (McLachlan & Peel, 2000). Since then, mixture models have been applied to many different disciplines including astronomy, biology, genetics, medicine, economics, engineering, marketing, and many more. Some interesting applications of mixture models are provided by Blekas, Likas, Galatsanos, and Lagaris (2005), Brijs et al. (2004), Caudill, Gropper, and Hartarska (2009), Feng et al. (2008), Gerdtham and Trivedi (2001), Greenspan, Goldberger, and Eshet (2001), Hernandez and Phillips (2006), Park and Lord (2009), Rocha, Soldevilla, and Burtenshaw (2007), Su et al. (2011) and Zhai, Velivelli, and Yu (2004).

With the stock market regulations, a 'tick-size' is maintained to provide a minimum size that a stock price can change, and 'tick-size' has been mandated by electronic exchanges as the 'smallest currency unit'. The 'tick-size' governs an indirect discreteness to the changes in stock price where any given time a stock price can be changed only as a multiple of the established 'tick-size'. This allows the stock prices to cluster among a smaller set of values. This discreteness and clustering add an advantage towards lowering the cost of negotiation and limiting the information exchange between buyers and sellers (Harris, 1991).

The proposed model provides a new way to investigate the stock price movements as a mixture model of count data. For a set of counts, the most natural distribution to use is the Poisson distribution. In the regression setting, classical linear models can be extended to generalized linear models (GLMs) which allow the response variable to follow a distribution from an exponential family (such as the Poisson distribution). A link function is used to link the mean of the response variable to a linear function of the predictors. In the GLM setting, parameters are often estimated by the method of maximum likelihood. For a GLM with a canonical link function, the likelihood function is convex so the maximum likelihood estimate (MLE) is unique when it exists (for example, see Madsen & Thyregod, 2011 and the references therein), and the Newton–Raphson is an effective method for obtaining the MLE.

However, in models with missing data or hidden parameters, direct calculation of MLEs can be very difficult. In these situations, the Expectation–Maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977) is a widely used method to compute MLEs. Advantages of the EM algorithm in mixture models which accommodate the likelihood approach are discussed in Berlinet and Roland (2012), Xu and Jordan (1996) and Yao (2013). Parameter estimation of Poisson mixtures is faster than in Gaussian mixtures due to the reduced number of parameters as stated by Su et al. (2011). Stock movements based on tick-size result in a set of integers as stock prices move up and down. This results in the use of different parametric forms of Poisson distributions in the mixture model. Nevertheless, according to Xu and Jordan (1996), the EM algorithm works faster when the mixtures are clearly separated.

The rest of the paper is organized as follows. The mixture model is proposed in Section 2, and the EM algorithm is used to compute maximum likelihood estimates for the model parameters described. Results proving the consistency and asymptotic normality of the estimators are presented in Section 3, and confidence intervals with the correct asymptotic size are derived based on the asymptotic normality. Proofs of the theoretical results are given in the appendices. In Section 4, simulation studies are presented to illustrate properties of the estimation procedures, and a real example with Federal Express (FDX) tick data is also presented.