

Bayesian Modelling of Integer Data Using the Generalised Poisson Difference Distribution

Golnaz Shahtahmassebi¹ & Rana Moyeed^{1,2}

¹ School of Science and Technology, Nottingham Trent University, Nottingham, UK

² School of Computing and Mathematics, Plymouth University, Plymouth, UK

Correspondence: Golnaz Shahtahmassebi, School of Science and Technology, Nottingham Trent University, Nottingham, UK. Tel: 44-115-848-8333. E-mail: golnaz.shahtahmassebi@ntu.ac.uk

Received: September 11, 2013 Accepted: November 12, 2013 Online Published: January 6, 2014

doi:10.5539/ijsp.v3n1p35 URL: <http://dx.doi.org/10.5539/ijsp.v3n1p35>

Abstract

Integer-valued random variables arising from the difference of two discrete variables can be seen frequently in various applications. In this paper, we obtain the distribution and derive the properties of the difference of two generalised Poisson variables with unequal parameters. This distribution is adopted to model a set of ultra high frequency (UHF) data relating to FTSE100 index futures using covariates. The unique characteristics of UHF data have introduced new theoretical and computational challenges to both statistical and financial studies. Such data consist of discrete-valued observations and unequally spaced time intervals. We also extend the model to its zero inflated version in order to capture the excess of zeros in the given data set. The analysis is carried out in a Bayesian framework using Markov Chain Monte Carlo methods. Various model diagnostics and model comparisons were undertaken which showed that index changes were explained well by the fitted model.

Keywords: generalised Poisson difference distribution, probability generating function, cumulant generating function, zero inflated, ultra high frequency data

1. Introduction

Integer-valued variables (with negative and positive values) arising from the difference of two discrete variables can be seen frequently in various applications, among others, in medicine (Karlis & Ntzoufras, 2006), sport (Karlis & Ntzoufras, 2009), image analysis (Li, Emmerich, Eggermont, Bovenkamp, Bäck, Dijkstra, & Reiber, 2009), finance and risk analysis (Consul, 1986; Shahtahmassebi, 2011). While, in the literature, there are many techniques available for dealing with the difference of two continuous or binary variables, methods for modelling integer-valued variables are rare. Available methods are mainly based on normal approximations of discrete distributions. However, in most cases normal approximations are not valid, since such data may take on a small range of integers Karlis and Ntzoufras (2006). Thus, techniques that are based on the distributions defined over both negative and positive integer values may improve our inference of such variables.

A possible choice for such a distribution may be the Poisson difference (PD) distribution by Irwin (1937) and Skellam (1946). This distribution has been used for the difference of any two Poisson variables, even when the two variables are not independent (Karlis & Ntzoufras, 2009). However, there are cases where the PD distribution has a tendency to underestimate values in the tails (Shahtahmassebi, 2011).

To address the problem of underestimation of the tails, there are several possible alternatives to the PD distribution, for example the distribution of the difference of two geometric distributions random variables (Kozubowski & Inusah, 2006) or the distribution of the difference of two generalised Poisson (GP) random variables. In this paper, we investigate the latter because the distribution of the difference of two GP variables may have a wider application as it could have longer or shorter tails compared to the PD distribution and also can be considered as a generalisation of the PD distribution. The distribution of the difference of two GP variables with unequal parameters is based on the idea of the GP distribution put forward by Consul (1986). The generalised Poisson difference (GPD) distribution has two additional parameters. We obtain the probability generating function (pgf) and the cumulant generating function (cgf) of the GPD distribution as well as a recurrence relation for all the cumulants.

We show that the GPD distribution with three parameters (Consul, 1986) is a special case of our GPD distribution. We introduce a zero inflated version of the GPD (ZGPD) distribution and use it for modelling the FTSE100 index changes from March 25 2008. In order to validate the fitted model, we obtain one-step ahead predictions of the index change for the next trading day (March 26, 2008). We assess the performance of the posterior predictive distribution with the probability integral transform (PIT) modified for the case of integer-valued variables (Liesenfeld, Nolte, & Pohlmeier, 2006). Finally, model comparisons were performed using the deviance information criterion (DIC) and the Bayes factor (BF). We carried out all our analyses in R. The R functions for evaluating and generating random samples from the GPD and ZGPD densities were written by the authors.

The remainder of the paper is organised as follows. In Section 2, we introduce the GPD distribution and obtain its properties. Section 3 provides a graphical presentation of the GPD distribution. Section 4 describes the data set available for this study and introduces the ZGPD model. The fitting of the ZGPD model using the Bayesian framework is illustrated in Section 5. Results are presented in Section 6. Finally, we conclude and provide an overview of possible extensions in Section 7.

2. Generalised Poisson Difference Distribution

Consider the GP distribution (Chandra, Roy, & Ghosh, 2013; Consul & Jain, 1973; Hubert, Lauretto, & Stern, 2009; Lerner, Lone, & Rao, 1997; Srivastava & Chen, 2010; Tuenter, 2006) (following the notation of Consul & Jain, 1973) of a random variable X with the following probability function

$$f_{GP}(x|\lambda, \theta) = \begin{cases} \lambda(\lambda + \theta x)^{x-1} e^{-\lambda - \theta x} / x!, & x = 0, 1, 2, \dots, \\ 0, & \text{for } x > m \text{ if } \theta < 0, \end{cases} \quad (1)$$

where $\lambda > 0$, $\max(-1, -\lambda/m) \leq \theta \leq 1$, and $m(\geq 4)$ is the largest positive integer for which $\lambda + m\theta > 0$ when $\theta < 0$. This condition is considered to ensure that there are at least five classes with non-zero probability when θ is negative. It can be seen that the GP distribution introduces over or underdispersion whether θ is positive or negative. Also, it reduces to the Poisson distribution when $\theta = 0$; see Consul and Famoye (2006, Chapter 9) and Nikoloulopoulos and Karlis (2008) for further details.

Now, let us assume $X \sim GP_X(\lambda_1, \theta_1)$ and $Y \sim GP_Y(\lambda_2, \theta_2)$ are independently distributed. The joint distribution of X and Y is given by (following the notation of Consul, 1986)

$$f_{X,Y}(x, y|\lambda_1, \lambda_2, \theta_1, \theta_2) = f_X(x|\lambda_1, \theta_1) f_Y(y|\lambda_2, \theta_2).$$

Therefore, the distribution of the difference $Z = X - Y$ is

$$\begin{aligned} f_{GPD}(Z = X - Y = z) &= \sum_{y=0}^{\infty} f_{X,Y}(z+y, y) \\ &= \sum_{y=0}^{\infty} \frac{[\lambda_1(\lambda_1 + \theta_1(z+y))^{z+y-1}]}{(z+y)!} e^{-\lambda_1 - \theta_1(z+y)} \times \frac{[\lambda_2(\lambda_2 + \theta_2 y)^{y-1}]}{y!} e^{-\lambda_2 - \theta_2 y} \\ &= e^{-\lambda_1 - \lambda_2 - \theta_1 z} \sum_{y=0}^{\infty} (\lambda_1, \theta_1)_{z+y} (\lambda_2, \theta_2)_y e^{-(\theta_1 + \theta_2)y}, \end{aligned} \quad (2)$$

for any value of $z \in \mathbb{Z}$, where

$$(\lambda, \theta)_x = \frac{\lambda(\lambda + x\theta)^{x-1}}{x!}.$$

We set lower limits for θ_1 and θ_2 to ensure that there are at least five classes of non-zero probabilities at both tails when $\theta_1 < 0$ or $\theta_2 < 0$. Thus, we set

$$\max(-1, -\lambda_1/m_1) < \theta_1 < 1,$$

$$\max(-1, -\lambda_2/m_2) < \theta_2 < 1,$$

where $m_1, m_2 \geq 4$ are the largest positive integers in which $\lambda_1 + m_1\theta_1 > 0$ and $\lambda_2 + m_2\theta_2 > 0$. Therefore, for any $z > m_1$ when $\theta_1 < 0$, or $z < -m_2$ when $\theta_2 < 0$, we have

$$f_{GPD}(z|\lambda_1, \lambda_2, \theta_1, \theta_2) = 0.$$

Furthermore, parameters λ_1 and θ_1 refer to the positive half and parameters λ_2 and θ_2 to the negative half of the GPD distribution. It is difficult to simplify the probability function defined in (2) in a compact form. We generally denote this distribution by $\text{GPD}(\lambda_1, \lambda_2, \theta_1, \theta_2)$ in which it can have the following particular forms when some of the parameters are zero.

$$\begin{aligned}\text{GPD}(\lambda_1, \lambda_2 = 0, \theta_1, \theta_2 = 0) &= \text{GP}(\lambda_1, \theta_1) = \frac{\lambda_1(\lambda_1 + \theta_1 z)^{z-1}}{z!} e^{-\lambda_1 - \theta_1 z}, \\ \text{GPD}(\lambda_1 = 0, \lambda_2, \theta_1 = 0, \theta_2) &= \frac{\lambda_2(\lambda_2 - \theta_2 z)^{-z-1}}{(-z)!} e^{-\lambda_2 + \theta_2 z}, \\ \text{GPD}(\lambda_1, \lambda_2, \theta_1 = 0, \theta_2 = 0) &= \text{PD}(\lambda_1, \lambda_2) = e^{-\lambda_1 - \lambda_2} \left(\frac{\lambda_1}{\lambda_2}\right)^{z/2} I_{|z|}(2\sqrt{\lambda_1 \lambda_2}).\end{aligned}$$

It can be seen that the third case where $\theta_1 = \theta_2 = 0$ is the Poisson difference distribution (Skellam, 1946).

We obtain the following recurrence relation for the cumulants of the probability distribution of the random variable Z (Appendix A)

$$\begin{aligned}(1 - \theta_1)(1 - \theta_2)L_{k+1} &= (2 - \theta_2)\lambda_1\theta_1 \frac{\partial^2 L_k}{\partial \theta_1 \partial \lambda_1} + (2 - \theta_2)\lambda_1 \frac{\partial L_k}{\partial \lambda_1} + \theta_1\theta_2\lambda_2 \frac{\partial^2 L_k}{\partial \theta_2 \partial \lambda_2} + \theta_1\lambda_2 \frac{\partial L_k}{\partial \lambda_2} \\ &\quad - \theta_1 \frac{\partial L_k}{\partial \theta_1} - \theta_2 \frac{\partial L_k}{\partial \theta_2} - L_k,\end{aligned}\quad (3)$$

where $L_1 = [\lambda_1/(1 - \theta_1)] - [\lambda_2/(1 - \theta_2)]$ (mean of the GPD distribution $\mu = L_1$). The expression for the other cumulants of the GPD distribution can be obtained by using (3) recursively for $k = 1, 2, 3, \dots$. For example L_k , for $k = 2, 3, 4$, is as follows

$$\begin{aligned}L_2 &= \frac{\lambda_1}{(1 - \theta_1)^3} + \frac{\lambda_2}{(1 - \theta_2)^3}, \\ L_3 &= \frac{\lambda_1(2\theta_1 + 1)}{(1 - \theta_1)^5} - \frac{\lambda_2(2\theta_2 + 1)}{(1 - \theta_2)^5}, \\ L_4 &= \frac{\lambda_1(1 + 8\theta_1 + 6\theta_1^2)}{(1 - \theta_1)^7} + \frac{\lambda_2(1 + 8\theta_2 + 6\theta_2^2)}{(1 - \theta_2)^7}.\end{aligned}$$

where $L_2 = \sigma^2$, the variance of the GPD distribution. Furthermore, the coefficients of skewness and kurtosis for the random variable Z are obtained as follows

$$\begin{aligned}\beta_1 &= \frac{[\lambda_1(1 - \theta_2)^5(1 + 2\theta_1)] - [\lambda_2(1 - \theta_1)^5(1 + 2\theta_2)]}{[\lambda_1(1 - \theta_2)^3 + \lambda_2(1 - \theta_1)^3]^{3/2} [(1 - \theta_1)(1 - \theta_2)]^{1/2}}, \\ \beta_2 &= \frac{\lambda_1(1 + 8\theta_1 + 6\theta_1^2)(1 - \theta_2)^7 + \lambda_2(1 + 8\theta_2 + 6\theta_2^2)(1 - \theta_1)^7}{[\lambda_1(1 - \theta_2)^3 + \lambda_2(1 - \theta_1)^3]^2 (1 - \theta_1)(1 - \theta_2)}.\end{aligned}$$

By having $\theta_1 = \theta_2$, it can be seen that the GPD distribution of Consul (1986) is a special case of the GPD distribution we introduced here.

It can be seen that the GPD distribution may arise as the difference of any two discrete variables which do not necessarily follow the GP distribution. To see this consider X_1 and X_2 follow GP distribution and X_0 follow any discrete distribution. Let $Y_1 = X_1 + X_0$ and $Y_2 = X_2 + X_0$. If we now take the difference $Y_1 - Y_2 = X_1 - X_2$, we can clearly see that this is equivalent to the difference of two GP variables (Karlis & Ntzoufras, 2009).

3. Graphical Presentation of GPD Distribution

The behaviour of a random variable Z following the GPD distribution with different values of λ_1 , λ_2 , θ_1 and θ_2 is illustrated in this section. To study only the effect of θ_1 and θ_2 on the GPD distribution, first we assume $\lambda_1 = \lambda_2 = \lambda$ and change the values of θ_1 and θ_2 (Figure 1). In general, the distribution has a wider span over Z as either θ_1 or θ_2 increases. The distribution also becomes short tailed as either of the parameters decrease. For a fixed λ , the sign of $[\lambda/(1 - \theta_1)] - [\lambda/(1 - \theta_2)]$ determines the skewness of the GPD distribution.

Now, let us assume $\theta_1 = \theta_2 = \theta$. It can be seen from Figure 2 that when $\lambda_1 < \lambda_2$, the GPD distribution becomes a skewed distribution to the left. For any given value of θ , when λ_1 decreases, the GPD tends to have a shorter tail.

Since the following type of symmetry holds for the GPD distribution

$$f_{\text{GPD}}(Z = z | \lambda_1, \lambda_2, \theta_1, \theta_2) = f_{\text{GPD}}(Z = -z | \lambda_2, \lambda_1, \theta_2, \theta_1),$$

therefore for a given value of θ the GPD distribution tends to be skewed to the right when $\lambda_1 > \lambda_2$ (not illustrated). Also, for any given values of λ_1 and λ_2 , the form of the GPD is substantially altered by varying θ . Figure 3 suggests that an increase in the value of θ increases the variance of the distribution and both the tails of the GPD distribution become long and thin. For any value of λ , when θ becomes smaller most of the mass becomes concentrated at $Z = 0$.

Figures 1(a) and 3(b) show the GPD (with $\theta_1 > \theta_2$) and PD ($\theta_1 = \theta_2$) distributions with identical values of $\lambda_1 = \lambda_2 = \lambda$. It can be seen that the GPD distribution has a shorter left tail and a longer right tail in comparison with the PD distribution.

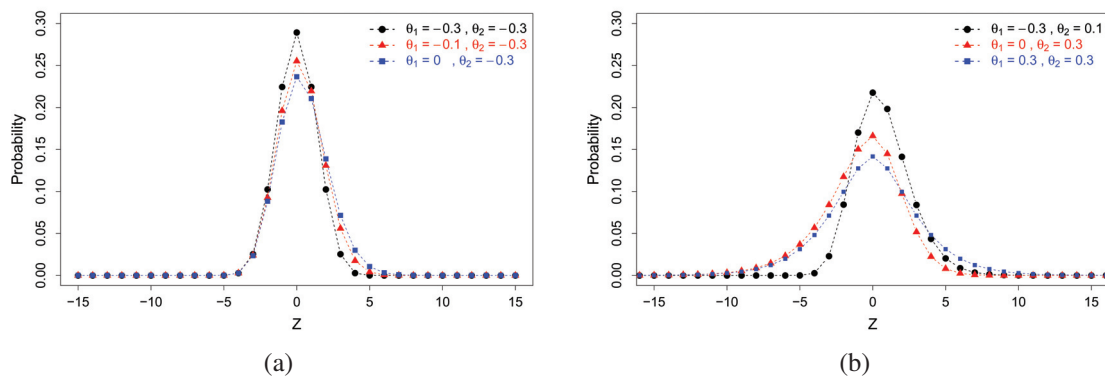


Figure 1. The GPD distribution for different values of θ_1 and θ_2 with $\lambda_1 = \lambda_2 = 2$

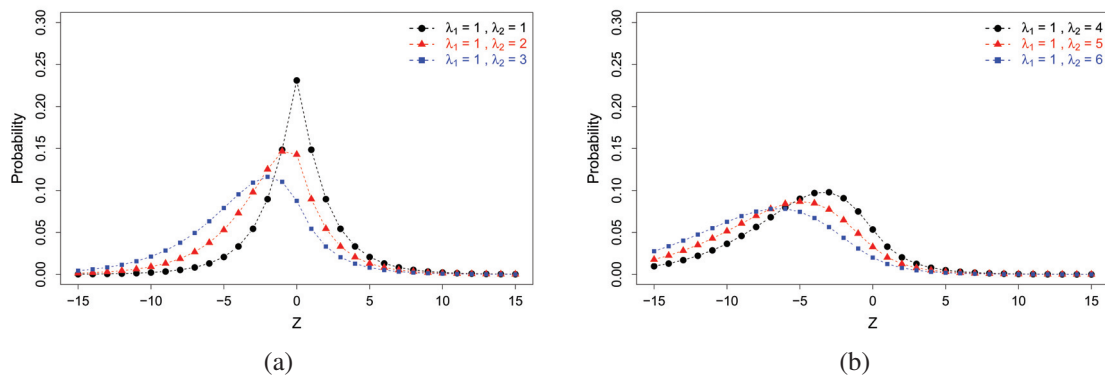


Figure 2. The GPD distribution for different values of λ_1 and λ_2 with $\theta_1 = \theta_2 = 0.4$

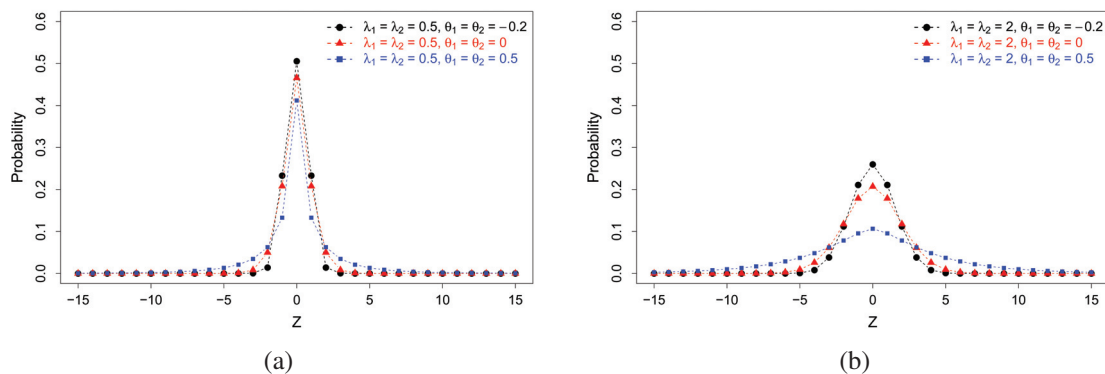


Figure 3. The GPD distribution for different values of $\lambda_1 = \lambda_2 = \lambda$ and $\theta_1 = \theta_2 = \theta$

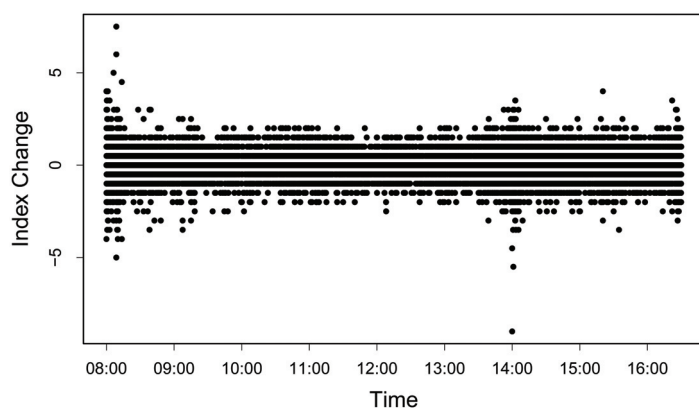
4. Application

The availability of trade-by-trade (commonly known as ultra high-frequency (UHF)) data on transactions, in the recent decade, has revolutionised data processing and statistical modelling techniques in finance (McCulloch & Tsay, 2001; Tsay, 2005; Liesenfeld, Nolte, & Pohlmeier, 2006; Pacurar, 2008). The unique characteristics of UHF data has introduced new theoretical and computational challenges to both statistical and financial studies (Liesenfeld, Nolte, & Pohlmeier, 2006; Pacurar, 2008; Tsay, 2005). Such data consist of (i) discrete-valued observations, as the price change in consecutive transactions is in a multiple of *tick* size, where one tick is defined as the minimum amount by which the price of the market can change and (ii) unequally spaced time intervals; see Tsay (2005, Chapter 5) and Liesenfeld, Nolte, and Pohlmeier (2006) and references therein for further details on analysing UHF data sets.

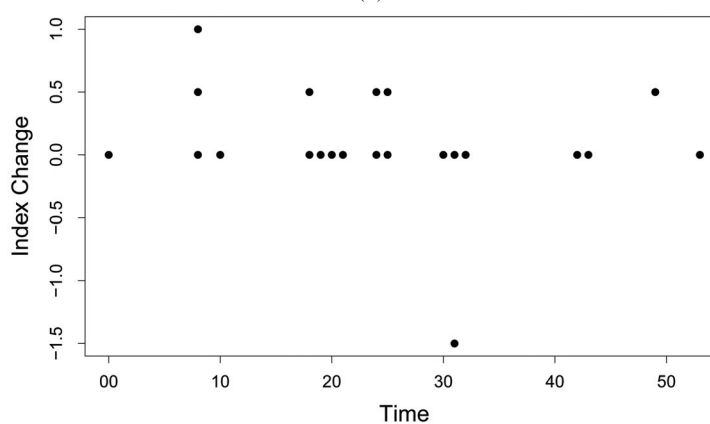
4.1 Data: FTSE100 Index Change

Table 1. Relative frequencies of index change in multiples of tick size for 8:00-16:30, March 25, 2008

Number(tick)	≤ -4	-3	-2	-1	0	1	2	3	≥ 4
Percentage	0.5	1.0	3.6	10.3	69.3	10.0	3.7	1.1	0.5



(a)



(b)

Figure 4. (a) FTSE100 index change futures: the discrete structure of index change process can be observed for March 25, 2008. (b) One minute of FTSE100 index change presents the main features of UHF data, e.g. price discreteness and unequally spaced time intervals

Here, we analyse a set of ultra high frequency index changes from FTSE100 index futures (Figure 4(a)) recorded on March 25, 2008, traded on the *London International Financial Futures and Option Exchange* (LIFFE). The normal trading hours of the FTSE100 are from 08:00 until 16:30. Therefore, for simplicity, any transaction beyond these hours are discarded from the analysis. Thus, the number of transactions on March 25, eliminating 913 transactions, reduces from 46,180 to 45,267 transactions (Figure 4(a)). Furthermore, one tick size is considered to

be 0.5 FTSE100 index. Each contract is valued as £10 per index point, therefore the value of one tick is £5.

A one-minute index change (10:00:00-10:01:00) is illustrated in Figure 4(b) which shows unequally spaced time intervals and price discreteness. Table 1 provides the relative frequencies of index change for March 25, 2008. It shows that the positive and the negative index changes are almost symmetrically distributed around zero. It can be seen that about 69% of the indices have not changed at all and about 20% had a change of 1 tick size. In addition, changes of 4 or more ticks are rare (235 and 225 in left and right tails respectively, around 1% in total), in which the minimum and maximum changes occurred with the size of -18 and 15 ticks, respectively (see Appendix B).

4.2 Model for Index Change

Assume P_{t_i} is the index of the most recent transaction at time t_i . Here t_i is a continuous time, but indices are only updated when a transaction actually occurs. We may consider the index, P_{t_i} , as

$$P_{t_i} = P_{t_0} + \sum_{i=1}^{N_{t_i}} Z_i, \quad (4)$$

where N_{t_i} is the number of transactions recorded in the interval from time 0 up to time t_i and Z_i is the index change associated with the i th transaction. In practice many of the Z_i 's are zero. Instead of modelling the index process itself, we model the index change process as a non-stationary and non-linear process (Rydberg & Shephard, 2003). Here, we assume only the difference of two consecutive indices (an integer-valued variable) is observed and follows the GPD distribution with the probability mass function given in Equation (2).

In real life applications, count (or integer-valued) data sets, such as dental (Karlis & Ntzoufras, 2006; Mwalili, 2008), spatial (Agarwal, Gelfand, & Citron-Pousty, 2002) and sports (Karlis & Ntzoufras, 2009), may contain an excess of zero values, i.e. more than what the model would predict. We have already mentioned that in our data set about 69% of index changes were zero, a figure which may be underestimated by the GPD model. By introducing an extra probability parameter, a zero-inflated distribution enables us to capture the possible excess of zero values.

4.3 Zero-Inflated Generalised Poisson Difference Model

Zero-inflated distributions are used when there is an excess of zeros, relative to the expected frequency of zeros. In the current data set about 70% of index changes are zero. Thus, we extend the GPD distribution to its zero inflated version in order to capture the possible excess of zeros in a given data set. Let Z be a ZGPD random variable with the following probability function

$$f_{\text{ZGPD}}(z|\lambda_1, \lambda_2, \theta_1, \theta_2, p) = \begin{cases} p + (1-p)f_{\text{GPD}}(z|\lambda_1, \lambda_2, \theta_1, \theta_2) & \text{if } z = 0, \\ (1-p)f_{\text{GPD}}(z|\lambda_1, \lambda_2, \theta_1, \theta_2) & \text{if } z \neq 0, \end{cases} \quad (5)$$

for $z \in \mathbb{Z}$, where $f_{\text{GPD}}(z|\lambda_1, \theta_1, \lambda_2, \theta_2)$ is the GPD probability function given by (2) and $p \in (0, 1)$ is the proportion of extra zeros. According to (5), the first part of the model represents the probability of all zero values, while the probability of $z \neq 0$ is given by the second part. Thus, the mean and the variance of a ZGPD random variable Z are

$$E(Z) = (1-p) E(Z_{\text{GPD}}), \quad (6)$$

$$\text{Var}(Z) = (1-p) \left(\text{Var}(Z_{\text{GPD}}) + p [E(Z_{\text{GPD}})]^2 \right) \quad (7)$$

Further details on the zero inflated distributions can be found in (Johnson, Kotz, & Kemp, 2005) and references therein. Now let us assume Z , an integer-valued variable, is a response variable which follows the $\text{ZGPD}(\lambda_1, \lambda_2, \theta_1, \theta_2, p)$ distribution. We adopt a set up similar to a ZPD model which has been used widely in statistics literature (see Karlis and Ntzoufras (2009) and references therein). Suppose there are n observations and k explanatory variables. We define the parameters of the model as follows

$$\log(\lambda_1) = X\alpha_1, \quad (8)$$

$$\log(\lambda_2) = X\alpha_2, \quad (9)$$

where $\lambda_1 = (\lambda_{1,1}, \dots, \lambda_{1,n})^T$ and $\lambda_2 = (\lambda_{2,1}, \dots, \lambda_{2,n})^T$, X is a matrix of covariates of size $n \times k$ and $\alpha_1 = (\alpha_{1,0}, \dots, \alpha_{1,k})^T$ and $\alpha_2 = (\alpha_{2,0}, \dots, \alpha_{2,k})^T$ are vectors of parameters of the ZGPD model. For simplicity, we assume θ_1, θ_2 and p are fixed over the whole trading day. Furthermore, we set $\log(1-\theta_1) = \beta_1$ and $\log(1-\theta_2) = \beta_2$ in the model which allows us to have parameters over the real line. The advantage of modelling λ_1 and λ_2 in this

manner is that it allows us to identify whether each covariate has a similar or different affect on the negative and the positive values within the same model.

We assess the effect of the the previous index change (z_{i-1}), the volume of the previous transaction (v_{i-1}) and the time duration between two consecutive transactions ($\Delta t_i = t_i - t_{i-1}$), on the current index change. Let us point out that having Δt allows us to consider unequally spaced time intervals in the model. Thus, the ZGPD model in Equations (8)-(9), in terms of the above covariates, is

$$\begin{aligned}\log(\lambda_{1,i}) &= \alpha_{1,0} + \alpha_{1,1}z_{i-1} + \alpha_{1,2}\log(v_{i-1}) + \alpha_{1,3}\log(\Delta t_i), \\ \log(\lambda_{2,i}) &= \alpha_{2,0} + \alpha_{2,1}z_{i-1} + \alpha_{2,2}\log(v_{i-1}) + \alpha_{2,3}\log(\Delta t_i).\end{aligned}\quad (10)$$

5. Bayesian Analysis

5.1 Prior Distribution

To fully specify a Bayesian model, we need to specify the prior distribution. When no information is available, we propose to use normal prior distributions for the parameters of the ZGPD model with mean equal to zero and a large variance (e.g. 10^4) to express prior ignorance. For the proportion of the excess of zeros p , we propose a uniform distribution defined in the $(0, 1)$ interval.

Nevertheless, the Bayesian approach enables us to incorporate external information to our inference via our prior distribution. Also, the Bayesian approach can be used sequentially by using the previous posterior distribution as the current prior distribution which speeds up the process of updating our model. In other words, the information from the previous day's transactions can be used to specify our prior distribution (Karlis & Ntzoufras, 2009).

5.2 Posterior Distribution

We obtain our inference based on the posterior distribution of the parameter vector $\psi = (\alpha_1, \alpha_2, \beta_1, \beta_2, p)$. The posterior distribution of ψ can be obtained as

$$f_{\text{post}}(\psi|z) \propto f_{\text{like}}(z|\psi)f_{\text{prior}}(\psi),$$

where the likelihood is given by

$$\begin{aligned}L(\lambda_1, \lambda_2, \theta_1, \theta_2, p) &= \prod_{i=1}^n (p I_{\{0\}} + (1-p)f_{\text{GPD}}(z_i|\lambda_{1,i}, \lambda_{2,i}, \theta_{1,i}, \theta_{2,i})) \\ &= \prod_{i=1}^n [p I_{\{0\}} + (1-p)f_{\text{GPD}}(z_i|\exp(X_i\alpha_1), \exp(X_i\alpha_2), 1 - \exp(\beta_1), 1 - \exp(\beta_2))],\end{aligned}\quad (11)$$

where $I_{\{0\}}$ is an indicator function and $f_{\text{prior}}(\psi)$ is the prior distribution given by

$$f_{\text{prior}}(\psi) = f_r(p)f_{\alpha_1}(\alpha_1)f_{\alpha_2}(\alpha_2)f_{\beta_1}(\beta_1)f_{\beta_2}(\beta_2).$$

It can be seen that the posterior distribution is not analytically tractable. Thus, in order to generate samples from the posterior distribution we use MCMC methods, more specifically, the random walk Metropolis-Hastings (M-H) algorithm.

5.3 Metropolis Hastings Algorithm

In the M-H algorithm, we choose the normal distribution to be our proposal distribution, such that in the k th iteration, for $k = 1, \dots, m$, where m is the number of iterations, the normal distribution is centred at the values from the previous iteration, i.e. $\psi^{\text{cand}} \sim N(\psi^{(k-1)}, \sigma^2)$, for some value of σ^2 . In addition,

(i) An important point to consider in the case of random walk chains is the choice of the value of the dispersion parameter of the proposal distribution. A large value for the variance allows a greater variation from the previous value, but will lead to a very small acceptance rate. On the other hand, a small value of the variance results in draws which are close to the previous value with a high acceptance rate (Gamerman & Lopes, 2006). The optimal choice for the variance of the normal proposal is $\sigma^2 = c^2\Sigma$, where $c \approx 2.4/\sqrt{d}$ (d is the dimension of the parameter vector) and Σ is the variance-covariance matrix based on the curvature of the posterior at the mode (Tanner, 1998).

(ii) We use the logit transform for p , i.e. we assume $y = \log(p/(1-p))$ and draw $y^{\text{cand}} \sim N(y^{(k-1)}, \sigma_y^2)$.

(iii) Parameters are updated one at the time.

5.4 Simulating Posterior Predictive Distribution

Let $Z^* = (Z_1^*, \dots, Z_s^*)^T$ be the s values which we wish to predict. The posterior predictive distribution is defined as

$$f_{\text{pred}}(Z^*|z) = \int f_{\text{like}}(Z^*|\psi) f_{\text{post}}(\psi|z) d\psi, \quad (12)$$

where $f_{\text{like}}(Z^*|\psi)$ may be given as

$$f_{\text{like}}(Z^*|\psi) = \prod_{i=1}^s f(Z_i^*|\psi). \quad (13)$$

Therefore, by obtaining the predictive distribution for each component of index change we may be able to evaluate the predictive distribution of index change. For this purpose, we add the following steps to the MCMC algorithm. At the j th iteration, $j = 1, \dots, m$, where m is the number MCMC iterations, repeat the following steps, for $i = 1, \dots, s$:

(i) Obtain

$$\begin{aligned} \lambda_{1,i}^{*(j)} &= \exp(Y_i \alpha_1^{(j)}) \text{ and } \lambda_{2,i}^{*(j)} = \exp(Y_i \alpha_2^{(j)}), \\ \theta_1^{*(j)} &= 1 - \exp(\beta_1^{(j)}) \text{ and } \theta_2^{*(j)} = 1 - \exp(\beta_2^{(j)}), \end{aligned}$$

where $\lambda_{1,i}^{*(j)}$, $\lambda_{2,i}^{*(j)}$, $\theta_1^{*(j)}$ and $\theta_2^{*(j)}$ are the ZGPD model parameters which are obtained by substituting the values of $\alpha_1^{(j)}$, $\alpha_2^{(j)}$, $\beta_1^{(j)}$ and $\beta_2^{(j)}$ at the j th iteration in (8)-(9), and Y_i is the i th row of a matrix of covariates, Y , of size $s \times k$.

(ii) Draw $b_i^{(j)} \sim \text{Bernoulli}(p^{(j)})$:

- If $b_i^{(j)} = 1$, then $Z_i^{*(j)} = 0$, otherwise draw

$$Z_i^{*(j)} \sim \text{GPD}(\lambda_{1,i}^{*(j)}, \lambda_{2,i}^{*(j)}, \theta_1^{*(j)}, \theta_2^{*(j)}),$$

where a random sample from the GPD distribution is obtained as the difference of the two generalised Poisson random variables.

5.5 Model Checking

The posterior predictive distribution can be used to probabilistically quantify the response variable which enables us to assess the goodness of fit and overall performance of the model. Therefore, if the predictive distribution, in general, is in agreement with the observed data, this implies a good fit of the model (Karlis & Ntzoufras, 2009). However, considering the size of the data and the fact that most of the values (70%) of index change is zero (no index change), it may be difficult to find suitable diagnostics measures. Here, a randomised version of the PIT is implemented to measure how well the predictive distribution, especially in the tails, is able to explain the density of index change for the next trading day (Liesenfeld, Nolte, & Pohlmeier, 2006). In order to use the PIT, we have to construct intervals based on the cumulative predictive distribution of the i th observed index change and the i th observed index change minus 1, as follows

$$u_i^u = \hat{p}(Z_i \leq z_i) = \sum_{j=-\infty}^{z_i} \hat{p}(Z = j), \quad (14)$$

$$u_i^l = \hat{p}(Z_i \leq z_i - 1) = \sum_{j=-\infty}^{z_i-1} \hat{p}(Z = j), \quad (15)$$

where z_i is the i th index change in the data set from the next trading day and $\hat{p}(\cdot)$ represents the estimated counterpart of the conditional probability given in Equation (5). It can be seen that Equations (14) and (15) form intervals with upper and lower limits of u_i^u and u_i^l , respectively. If the model is correctly specified, random draws from such intervals will have a standard uniform distribution. To test the idea of the uniformity of the constructed values, one can use the Kolmogorov-Smirnov (K-S) test or transform the values to the standard normal distribution and plot a quantile-quantile (Q-Q) plot of the standard normal distribution against the transformed values of random draws. Here, we used both the K-S test and the Q-Q plot to judge the efficacy of the predicted values.

Furthermore, in order to decide how much complexity is necessary to fit the data, model comparison is also undertaken. That is, whether the full model fits the data properly or a model with fewer parameters has a similar performance. For this purpose, we use the deviance information criterion (DIC):

$$DIC = p_D + \hat{D}_{avg}(z),$$

where $p_D = \hat{D}_{avg}(z) - D_{\hat{\theta}}(z)$ is the effective number of parameters and $\hat{D}_{avg}(z) = \frac{1}{L} \sum_{l=1}^L D(z, \theta^l)$, where L is the number of MCMC iterations after burn-in, is a measure of goodness of fit, in which $D_{\hat{\theta}}(z) = D(z, \hat{\theta}(z))$ where $D(z, \theta) = -2 \log f(z|\theta)$; see Gelman, Carlin, Stern, and Donald (2003, Chapter 6) for further details. It should be noted that when comparing the Bayesian models, a smaller value of DIC suggests a better fit.

6. Results

The Gelman and Rubin's convergence diagnostic with the statistic value of 1.01 suggests that we can consider the convergence of the MCMC chains after 15000 iterations (Gelman & Rubin, 1992). A further 5000 samples were collected after burn-in. Table 2 confirms that the model with three covariates and 11 parameters has the lowest value of DIC.

Table 2. DIC, point-estimate, average deviance and estimated number of parameters for each of five models fitted to the FTSE100 index change

Model	DIC	$D_{\hat{\theta}}$	\hat{D}_{ave}
Intercept	101402.0	101392.7	101397.3
z_{i-1}	101021.7	101008.9	101015.3
$z_{i-1} + \log(v_{i-1})$	101021.6	101000.0	101010.8
$z_{i-1} + \log(\Delta t_{i-1})$	99949.5	99930.49	99940.0
$z_{i-1} + \log(v_{i-1}) + \log(\Delta t_i)$	99944.7	99922.4	99933.6

Table 3. Posterior means, standard deviations and 95% credible intervals of parameters of the ZGPD model based on the FTSE100 data

Coefficients	Mean	SD	95% credible intervals	
$\alpha_{1,0}$	-1.407	0.044	-1.492	-1.327
$\alpha_{1,1}$	-0.137	0.013	-0.164	-0.110
$\alpha_{1,2}$	-0.026	0.018	-0.060	0.009
$\alpha_{1,3}$	0.599	0.020	0.558	0.637
θ_1	0.228	0.007	0.214	0.241
p	0.171	0.024	0.124	0.215
Coefficients	Mean	SD	95% credible intervals	
$\alpha_{2,0}$	-1.387	0.044	-1.477	-1.301
$\alpha_{2,1}$	0.153	0.013	0.129	0.180
$\alpha_{2,2}$	0.032	0.017	-0.002	0.066
$\alpha_{2,3}$	0.577	0.020	0.537	0.616
θ_2	0.217	0.007	0.203	0.231

The posterior mean, standard deviation and 95% credible intervals of the posterior distribution of model parameters are provided in Table 3. For none of the model parameters the 95% credible intervals contained zero, indicating that all three covariates have a significant effect on the index change. It can be seen that the previous index change has a larger effect on λ_2 compared to λ_1 . This implies that in the next transaction a switch from a positive to a negative index change is more likely than a switch from a negative to a positive one. We can also see that the effect of volume on the expected value of the index change tends to be small. Since the posterior means of $\alpha_{1,3}$ and $\alpha_{2,3}$ are almost the same, the mean index change at the next transaction, irrespective of when it occurs, is expected to be zero, assuming all the other variables (previous volume and index change) were fixed. The posterior means of θ_1 and θ_2 have values of 0.228 and 0.217 suggesting that the fitted model has longer tails than the corresponding ZPD model. Finally, the 95% credible interval of p is (0.124, 0.215) which clearly indicates the presence of excess of zeros and, therefore, the necessity of fitting a zero-inflated model.

6.1 ZGPD and ZPD Models

In order to show that the ZGPD model with $\theta_1 \neq \theta_2$ (ZGPD $_{\theta_1 \neq \theta_2}$) outperforms a ZGPD with $\theta_1 = \theta_2$ (ZGPD $_{\theta_1 = \theta_2}$) model and a ZPD model, we use the Bayes factor (BF), which is defined as the ratio of marginal likelihood under one model to the marginal likelihood under the other model (Gelman, Carlin, Stern, & Donald, 2003, July; Kass & Raftery, 1995):

$$BF(M_1, M_2) = \frac{f_{like, M_1}(z|M_1)}{f_{like, M_2}(z|M_2)} = \frac{\int f_{prior, M_1}(\psi|M_1) f_{like, M_1}(z|\psi, M_1) d\psi}{\int f_{prior, M_2}(\psi'|M_2) f_{like, M_2}(z|\psi', M_2) d\psi'}, \quad (16)$$

where ψ and ψ' are vectors of parameters under the M_1 and M_2 distributions, respectively and $f_{like, M_1}(z|M_1)$ and $f_{like, M_2}(z|M_2)$ are marginal likelihoods under M_1 and M_2 models. A Bayes factor greater than one (or $\log(BF) > 0$) supports M_1 model over M_2 model. We obtain the BF using the “MCMCpack” package in R.

The logarithm of $BF(\text{ZGPD}_{\theta_1 \neq \theta_2}, \text{ZGPD}_{\theta_1 = \theta_2})$ and $BF(\text{ZGPD}_{\theta_1 \neq \theta_2}, \text{ZPD})$ are respectively 210.7 and 5538.7 suggesting that the ZGPD $_{\theta_1 \neq \theta_2}$ model performs considerably better than a ZGPD $_{\theta_1 = \theta_2}$ and a ZPD model. Finally, Figure 5(a) provides the Q-Q plot for the values of PIT which suggests that the ZGPD model explains the behaviour of the data well, especially for the values in the tails. Comparing Figures 5(a) and 5(b), we see that the ZGPD model outperforms the ZPD model.

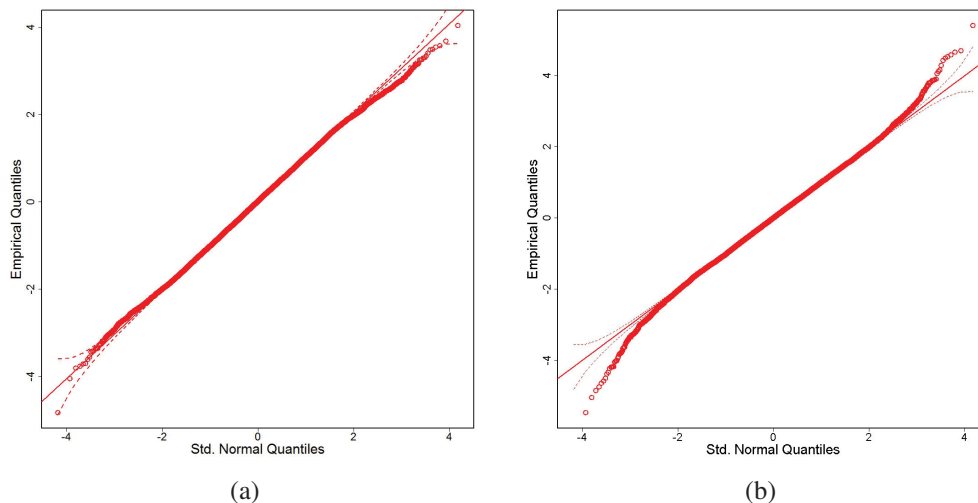


Figure 5. A normal Q-Q plot of random samples drawn from intervals based on the cumulative predictive distribution (a) ZGPD and (b) ZPD models

7. Conclusion

In this paper we introduced the GPD distribution with four parameters. The GPD distribution can be considered as an alternative to the PD distribution when the data has longer or shorter tails compared to the PD distribution. We obtained the pgf and the cgf for the GPD distribution and explored the theoretical characteristics of the distribution. We established a recurrence relation for all the cumulants of a GPD random variable and derived the first four cumulants. In addition, we showed that the GPD distribution suggested by Consul (1986) is a special case of our GPD distribution.

We proposed the ZGPD model and illustrated its application by modelling the FTSE100 index changes in a Bayesian framework implemented via MCMC methods. The Bayesian estimation of the parameters of interest were presented in detail. One of our principle interests concerned the distribution of index change for the whole trading day and we have shown that our approach models the behaviour of index change well. Results from the data revealed that the previous index change and the time duration between two consecutive transactions significantly affect the FTSE100 index change. Finally, various model diagnostics were carried out which showed that index changes were explained well by the fitted ZGPD model. Furthermore, we compared our results with the results from the ZGPD $_{\theta_1 = \theta_2}$ and the ZPD models and showed that the ZGPD $_{\theta_1 \neq \theta_2}$ model fits significantly better than the

both models, especially in the tails. A possible extension to the ZGPD model is a ZGPD duration model so that it captures the dynamics of index change and time duration jointly.

Acknowledgements

We would like to thank Dr Mario Cortina Borja and two reviewers for their helpful and constructive comments.

References

- Adler, D., Gläser, C., Nenadic, O., Oehlschlägel, J., & Zucchini, W. (2010). *ff: memory-efficient storage of large data on disk and fast access functions*. R package version 2.1-2.
- Agarwal, D. K., Gelfand, A. E., & Citron-Pousty, S. (2002). Zero-inflated models with application to spatial count data. *Environmental and Ecological Statistics*, 9, 341-355. <http://dx.doi.org/10.1023/A:1020910605990>
- Chandra, N. K., Roy, D., & Ghosh, T. (2013). A generalized Poisson distribution. *Communications in Statistics - Theory and Methods*, 42(15), 2786-2797. <http://dx.doi.org/10.1080/03610926.2011.620207>
- Consul, P. C. (1986). On the differences of two generalized Poisson variates. *Communications in statistics. Simulation and computation*, 15(3), 761-767. <http://dx.doi.org/10.1080/03610918608812538>
- Consul, P. C., & Famoye, F. (2006). *Lagrangian probability distributions*. Boston: Birkhäuser.
- Consul, P. C., & Jain, G. C. (1973). A generalization of the Poisson distribution. *Technometrics*, 15(4), 791-799. <http://dx.doi.org/10.1080/00401706.1973.10489112>
- Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: stochastic simulation for bayesian inference*. Boca Raton: Chapman & Hall.
- Gelman, A., Carlin, J. B., Stern, H. S., & Donald, R. B. (2003, July). *Bayesian Data Analysis* (2nd ed.). London: Chapman and Hall/CRC.
- Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457-472. <http://dx.doi.org/10.1214/ss/1177011136>
- Hubert, P. C., Lauretto, M. S., & Stern, J. M. (2009). FBST for generalized Poisson distribution. *AIP Conference Proceedings*, 1193(1), 210-217. <http://dx.doi.org/10.1063/1.3275617>
- Irwin, J. O. (1937). The frequency distribution of the difference between two independent variates following the same Poisson distribution. *Journal of the Royal Statistical Society*, 100(3), 415-416. <http://dx.doi.org/10.2307/2980526>
- Johnson, N., Kotz, S., & Kemp, A. (2005). *Univariate Discrete Distributions*. John Wiley. <http://dx.doi.org/10.1002/0471715816>
- Karlis, D., & Ntzoufras, I. (2006). Bayesian analysis of the differences of count data. *Statistics in Medicine*, 25(11), 1885-1905. <http://dx.doi.org/10.1002/sim.2382>
- Karlis, D., & Ntzoufras, I. (2009). Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference. *IMA J Management Math*, 20(2), 133-145. <http://dx.doi.org/10.1093/imaman/dpn026>
- Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430), 773-795. <http://dx.doi.org/10.1080/01621459.1995.10476572>
- Kozubowski, T., & Inusah, S. (2006). A skew Laplace distribution on integers. *Annals of the Institute of Statistical Mathematics*, 58(3), 555-571. <http://dx.doi.org/10.1007/s10463-005-0029-1>
- Lerner, B., Lone, A., & Rao, M. (1997). On generalized Poisson distributions. *Probability and Mathematical Statistics*, 17(2).
- Liesenfeld, R., Nolte, I., & Pohlmeier, W. (2006). Modelling financial transaction price movements: a dynamic integer count data model. *Empirical Economics*, 30(4), 795-825. <http://dx.doi.org/10.1007/s00181-005-0001-1>
- Li, R., Emmerich, M., Eggermont, J., Bovenkamp, E., Bäck, T., Dijkstra, J., & Reiber, J. (2009). Optimizing a medical image analysis system using mixed-integer evolution strategies. In S. Cagnoni (Ed.), *Evolutionary Image Analysis and Signal Processing*, Volume 213 of *Studies in Computational Intelligence*, 91-112. http://dx.doi.org/10.1007/978-3-642-01636-3_6

- McCulloch, R., & Tsay, R. (2001). Nonlinearity in high-frequency financial data and hierarchical models. *Studies in Nonlinear Dynamics and Econometrics*, 5(1). <http://dx.doi.org/10.2202/1558-3708.1067>
- Mwalili, S. M. (2008). The zero-inflated negative Binomial regression model with correction for misclassification: an example in caries research. *Statistical Methods in Medical Research*, 17(2), 123-139. <http://dx.doi.org/10.1177/0962280206071840>
- Nikolouloupoulos, A. K., & Karlis, D. (2008). On modelling count data: a comparison of some well-known discrete distributions. *Journal of Statistical Computation and Simulation*, 78(3), 437-457. <http://dx.doi.org/10.1080/10629360601010760>
- Pacurar, M. (2008). Autoregressive conditional duration models in finance: A survey of the theoretical and empirical literature. *Journal of Economic Surveys*, 22(4), 711-751. <http://dx.doi.org/10.1111/j.1467-6419.2007.00547.x>
- Plummer, M., Best, N., Cowles, K., & Vines, K. (2006). Coda: Convergence diagnosis and output analysis for MCMC. *R News*, 6(1), 7-11.
- Rydberg, T. H., & Shephard, N. (2003). Dynamics of trade-by-trade price movements: Decomposition and models. *Journal of Financial Econometrics*, 1(1), 2-25. <http://dx.doi.org/10.1093/jjfinec/nbg002>
- Shahtahmassebi, G. (2011). *Bayesian modelling of ultra-high frequency financial data*. (Ph. D. thesis, School of Computing and Mathematics, Plymouth University). Retrieved from <http://pearl.plymouth.ac.uk:8080/handle/10026.1/894?show=full>
- Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society*, 109(3), 296. <http://dx.doi.org/10.2307/2981372>
- Srivastava, S., & Chen, L. (2010). A two-parameter generalized Poisson model to improve the analysis of RNA-seq data. *Nucleic Acids Research*, 38(17), e170. <http://dx.doi.org/10.1093/nar/gkq670>
- Tanner, M. A. (1998). *Tools for statistical inference: models for the exploration of posterior distributions and likelihood functions*. New York: Springer.
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Tsay, R. S. (2005). *Analysis of financial time series* (2nd ed.). New Jersey: Wiley-Interscience. <http://dx.doi.org/10.1002/0471746193>
- Tuenter, H. J. H. (2006). *On the generalized Poisson distribution*.

Appendix

A. Cumulant Generating Function

Let us assume $\psi(\beta) = \sum_{i=1}^{\infty} \beta^k L_k/k!$, where L_k is the k th cumulant of the random variable Z , is the cgf of the GPD distribution. Using the moment generating function and pgf of the variable Z , we are able to obtain cumulants as follows. Consider the pgf of the random variable $X \sim G_1(u)$, and $Y \sim G_2(u^{-1})$ (Consul, & Famoye, 2006), given by

$$G_1(u) = e^{\lambda_1(t_1-1)} \text{ and } G_2(u^{-1}) = e^{\lambda_2(t_2-1)},$$

where $t_1 = ue^{\theta_1(t_1-1)}$ and $t_2 = u^{-1}e^{\theta_2(t_2-1)}$. Thus, the pgf of a GPD random variable is given by

$$G(u) = G_1(u)G_2(u) = e^{\lambda_1(t_1-1)+\lambda_2(t_2-1)}. \quad (\text{A.1})$$

By replacing t_1 , t_2 and u by e^{T_1} , e^{T_2} and e^β , respectively in (A.1), and taking the logarithm, the cgf, $\psi(\beta)$, of $Z = X - Y$ can be obtained as follows,

$$\psi(\beta) = \lambda_1(e^{T_1} - 1) + \lambda_2(e^{T_2} - 1) \text{ and } = \frac{(T_1 - \beta)\lambda_1}{\theta_1} + \frac{(T_2 + \beta)\lambda_2}{\theta_2}, \quad (\text{A.2})$$

where $T_1 = \beta + \theta_1(e^{T_1} - 1)$ and $T_2 = -\beta + \theta_2(e^{T_2} - 1)$. Using the cgf, we are able to obtain cumulants for the GPD distribution by differentiating w.r.t β , θ_1 and θ_2 .

Let us start by differentiating T_1 and T_2 with respect to β , θ_1 and θ_2 as follows

$$\frac{\partial T_1}{\partial \beta} = (1 - \theta_1 e^{T_1})^{-1} \quad \text{and} \quad \theta_1 \frac{\partial T_1}{\partial \theta_1} = -1 + (1 - \theta_1) \frac{\partial T_1}{\partial \beta}, \quad (\text{A.3})$$

$$\frac{\partial T_2}{\partial \beta} = -(1 - \theta_2 e^{T_2})^{-1} \quad \text{and} \quad \theta_2 \frac{\partial T_2}{\partial \theta_2} = -1 - (1 - \theta_2) \frac{\partial T_2}{\partial \beta}. \quad (\text{A.4})$$

Next, we partially differentiate $\psi(\beta)$ w.r.t β , θ_1 and θ_2 . Furthermore, to simplify $\psi(\beta)/\partial\beta$, we can eliminate $\partial T_1/\partial\beta$ using (A.3)-(A.4) and obtain the following equations

$$\theta_1 \frac{\partial \psi(\beta)}{\partial \theta_1} + \psi(\beta) = \lambda_1 \frac{\partial T_1}{\partial \theta_1} + \frac{(T_2 + \beta)\lambda_2}{\theta_2}, \quad (\text{A.5})$$

$$\theta_2 \frac{\partial \psi(\beta)}{\partial \theta_2} + \psi(\beta) = \frac{(T_1 - \beta)\lambda_1}{\theta_1} + \lambda_2 \frac{\partial T_2}{\partial \theta_2}, \quad (\text{A.6})$$

$$(1 - \theta_1)(1 - \theta_2) \frac{\partial \psi(\beta)}{\partial \beta} - (1 - \theta_2)\lambda_1 + (1 - \theta_1)\lambda_2 = (1 - \theta_2)\lambda_1 \frac{\partial T_1}{\partial \theta_1} - (1 - \theta_1)\lambda_2 \frac{\partial T_2}{\partial \theta_2}, \quad (\text{A.7})$$

then add (A.5), (A.6) and (A.7), which results in

$$(2 - \theta_2)\lambda_1 \frac{\partial T_1}{\partial \theta_1} + \theta_1 \lambda_2 \frac{\partial T_2}{\partial \theta_2} = \theta_1 \frac{\partial \psi(\beta)}{\partial \theta_1} + \theta_2 \frac{\partial \psi(\beta)}{\partial \theta_2} + (1 - \theta_1)(1 - \theta_2) \frac{\partial \psi(\beta)}{\partial \beta} + \psi(\beta) - (1 - \theta_2)\lambda_1 + (1 - \theta_1)\lambda_2. \quad (\text{A.8})$$

By differentiating $\psi(\beta)$ in (A.2) w.r.t λ_1 , then differentiating partially w.r.t θ_1 again, we have

$$\frac{\partial^2 \psi(\beta)}{\partial \theta_1 \partial \lambda_1} = -\frac{1}{\theta_1} \frac{\partial \psi(\beta)}{\partial \lambda_1} + \frac{1}{\theta_1} \frac{\partial T_1}{\partial \theta_1}, \quad (\text{A.9})$$

Similarly, by differentiating $\psi(\beta)$ in (A.2) w.r.t λ_2 , then differentiating partially w.r.t θ_2 again, we get

$$\frac{\partial^2 \psi(\beta)}{\partial \theta_2 \partial \lambda_2} = -\frac{1}{\theta_2} \frac{\partial \psi(\beta)}{\partial \lambda_2} + \frac{1}{\theta_2} \frac{\partial T_2}{\partial \theta_2}. \quad (\text{A.10})$$

By substituting (A.9) and (A.10) into (A.8), we eliminate $\partial T_1/\partial\theta_1$ and $\partial T_2/\partial\theta_2$, and obtain the following relation

$$(2 - \theta_2)\lambda_1 \left[\theta_1 \frac{\partial^2 \psi(\beta)}{\partial \theta_1 \partial \lambda_1} + \frac{\partial \psi(\beta)}{\partial \lambda_1} \right] + \theta_1 \lambda_2 \left[\theta_2 \frac{\partial^2 \psi(\beta)}{\partial \theta_2 \partial \lambda_2} + \frac{\partial \psi(\beta)}{\partial \lambda_2} \right] =$$

$$\theta_1 \frac{\partial \psi(\beta)}{\partial \theta_1} + \theta_2 \frac{\partial \psi(\beta)}{\partial \theta_2} + (1 - \theta_1)(1 - \theta_2) \frac{\partial \psi(\beta)}{\partial \beta} + \psi(\beta) - (1 - \theta_2)\lambda_1 + (1 - \theta_1)\lambda_2. \quad (\text{A.11})$$

by replacing the value of the cgf by $\psi(\beta) = \sum_k \beta^k L_k / k!$ in (A.8), and obtain the following recurrence relation for the cumulants of the probability distribution of the random variable Z

$$\begin{aligned} (1 - \theta_1)(1 - \theta_2)L_{k+1} = & (2 - \theta_2)\lambda_1 \theta_1 \frac{\partial^2 L_k}{\partial \theta_1 \partial \lambda_1} + (2 - \theta_2)\lambda_1 \frac{\partial L_k}{\partial \lambda_1} + \theta_1 \theta_2 \lambda_2 \frac{\partial^2 L_k}{\partial \theta_2 \partial \lambda_2} \\ & + \theta_1 \lambda_2 \frac{\partial L_k}{\partial \lambda_2} - \theta_1 \frac{\partial L_k}{\partial \theta_1} - \theta_2 \frac{\partial L_k}{\partial \theta_2} - L_k, \end{aligned} \quad (\text{A.12})$$

where $L_1 = [\lambda_1 / (1 - \theta_1)] - [\lambda_2 / (1 - \theta_2)]$.

B. Table of Frequency of the Tails

Table B1. Frequency distribution of index change in multiples of tick size with values ≤ -4 (or ≥ 4) for 8:00-16:30, March 25, 2008

Ticks	Freq
-18	1
-11	1
-10	1
-9	1
-8	3
-7	11
-6	18
-5	49
-4	150
4	158
5	35
6	16
7	7
8	5
9	1
10	1
12	1
15	1

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).