STATISTICS IN TRANSITION new series, Autumn 2013
Vol. 14, No. 3, pp. 495–506

A MODIFICATION OF THE PROBABILITY WEIGHTED METHOD OF MOMENTS AND ITS APPLICATION TO ESTIMATE THE FINANCIAL RETURN DISTRIBUTION TAIL

Marta Małecka[1], Dorota Pekasiewicz[2]

ABSTRACT

The issue of fitting the tail of the random variable with an unknown distribution plays a pivotal role in finance statistics since it paves the ground for estimation of high quantiles and subsequently offers risk measures. The parametric estimation of fat tails is based on the convergence to the generalized Pareto distribution (GPD). The paper explored the probability weighted method of moments (PWMM) applied to estimation of the GPD parameters. The focus of the study was on the tail index, commonly used to characterize the degree of tail fatness. The PWMM algorithm requires specification of the cdf estimate of the so-called excess variable and depends on the choice of the order of the probability weighted moments. We suggested modification of the PWMM method through the application of the level crossing empirical distribution function. Through the simulation study, the paper investigated statistical properties of the GPD shape parameter estimates with reference to the PWMM algorithm specification. The simulation experiment was designed with the use of fat-tailed distributions with parameters assessed on the basis of the empirical daily data for DJIA index. The results showed that, in comparison to the commonly used cdf formula, the choice of the level crossing empirical distribution function improved the statistical properties of the PWMM estimates. As a complementary analysis, the PWMM tail estimate of DJIA log returns distribution was presented.

Key words: PWMM, generalized Pareto distribution, tail estimate, distribution function estimate.

1. Introduction

In finance the estimates of the loss distribution tails are used for risk valuation. Using the extreme value theory, risk is assessed on the basis the extreme quantile of the loss distribution or the expected overshoot of a specified

[1] University of Lodz, Department of Statistical Methods. E-mail: marta.malecka@uni.lodz.pl.
[2] University of Lodz, Department of Statistical Methods. E-mail: pekasiewicz@uni.lodz.pl.

threshold. The estimation of the fat tails of the loss distribution is based on the empirical distribution of the auxiliary variable which is defined as the excess over the threshold and converges asymptotically to the generalized Pareto distribution (GPD). Since the shape parameter of the GPD distribution is used to characterize the degree of tail fatness, its estimation receives particular attention.

The paper explored the PWMM applied to estimation of the generalized Pareto distribution parameters (McNeil, Frey and Embrachts, 2005, Rasmussen, 2001). The PWMM requires specification of the cdf estimate of the excess variable and depends on the choice of the order of the probability weighted moments used for parameters estimation. In the paper we suggested the modification of the PWMM through the application of the level crossing empirical distribution function. In the simulation study we compared the GPD shape parameter estimate properties obtained with the use of the level crossing distribution function and the other empirical cdf proposed in the literature. The study involved the bias and variance of the estimates. The simulation experiment was designed with the use of fat-tailed distributions: t-Student, Pareto, log gamma and Burr with parameters assessed on the empirical daily data for DJIA index. As a complementary analysis, the PWMM tail estimate of DJIA log returns distribution was presented.

The paper is organized as follows. The next section sets the notation and introduces the probability weighted method of moments. Section 3 outlines the estimation procedure of the generalized Pareto distribution parameters and the tail of the underlying loss variable with the use of the PWMM. That section is followed by a comparative study of PWMM estimates of GPD parameters according to the chosen empirical distribution function and order of moments. At the end of section 4 we turn to empirical study for the daily DJIA data. The final section summarizes and concludes the article.

2. Probability weighted method of moments

The probability weighted method of moments is the procedure of estimation of the distribution parameters $\theta_1,...,\theta_s$ of the random variable X with cdf F that utilizes the idea of probability weighted moments of the distribution and their estimates. The probability weighted moment of the X distribution is defined as (if the relevant expectation exists):

$$M_{1,j,0} = E\left(XF^{j}(x)\right), \tag{1}$$

$$M_{1,0,k} = E\left(X\left(1-F(x)\right)^{k}\right), \tag{2}$$

for $j,k = 0,1, ...$ (Rasmussen, 2001).

Suppose that we have a sequence of i.i.d. observations $X_1, X_2, ..., X_n$ from an unknown distribution *F*. The moments $M_{1,j,0}$ and $M_{1,0,k}$ estimates are sample probability weighted moments given by:

$$m_{1,j,0} = \frac{1}{n}\sum_{i=1}^{n} X_{(i)}^{(n)}\left[F_n\left(x_{(i)}^{(n)}\right)\right]^j, \tag{3}$$

$$m_{1,0,k} = \frac{1}{n}\sum_{i=1}^{n} X_{(i)}^{(n)}\left[1 - F_n\left(x_{(i)}^{(n)}\right)\right]^k, \tag{4}$$

where $X_{(i)}^{(n)}$ are position statistics, F_n is the empirical cdf *F* and $j, k = 0, 1, ...$

Solving the system of *s* equations of the form:

$$\begin{cases} m_{1,j,0} = M_{1,j,0}, \\ m_{1,0,k} = M_{1,0,k}, \end{cases} \tag{5}$$

where $j = 0, 1, ..., s_1$, $k = 1, ..., s_2$ and $s_1 + s_2 = s$, gives the estimates of $\theta_1, ..., \theta_s$.

Statistical properties of the estimates depend on the form of empirical distribution function. Various cdf estimates are presented in Hosting and Wallis (1987), Landwehr, Matalas and Wallis (1979), Seckin et al., (2010). We considered the commonly used cdf estimate:

$$F_n(x) = \frac{1}{n-1}\sum_{i=1}^{n}\left(I_{\left(-\infty,x\right\rangle}\left(x_{(i)}^{(n)}\right) - 1\right), \tag{6}$$

and proposed the level crossing empirical distribution function (Huang and Brill, 1999):

$$F_n(x) = \sum_{i=1}^{n} w_{n,i} I_{\left(-\infty,x\right\rangle}\left(x_{(i)}^{(n)}\right), \tag{7}$$

where

$$w_{n,i} = \begin{cases} \frac{1}{2}\left[1 - \frac{n-2}{\sqrt{n(n-1)}}\right] & \text{dla } i = 1, n, \\ \frac{1}{\sqrt{n(n-1)}} & \text{dla } i = 2, 3, ..., n-1. \end{cases} \tag{8}$$

and *I* is the indicator taking the value of one for $x_i \le x$ and 0 in the opposite case.

In the initial study also the popular formula $F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\left(-\infty,x\right\rangle}\left(x_{(i)}^{(n)}\right)$ for the cdf was used. The comparative analysis including this cdf formula can be found in (Małecka and Pekasiewicz, 2013).

3. Estimation of the generalized Pareto distribution parameters and the distribution tail

Fitting the tail of the variable X distribution with the cdf F gives grounds for estimation of high quantiles and is therefore used for risk valuation. Of particular importance are distributions with fat tails which demonstrate financial data features. In such cases the estimate of the distribution function F for $x > u$ is given by (McNeil, Frey and Embrechts, 2005):

$$\hat{F}(x) = 1 - \frac{n_u}{n}\left(1 + \hat{\xi}\,\frac{x-u}{\hat{\beta}}\right)^{-\frac{1}{\hat{\xi}}}, \tag{9}$$

where u is the fixed threshold, n_u is the number of observations larger than u and $\hat{\xi}, \hat{\beta}$ are the estimates of the ξ, β parameters of the generalized Pareto distribution with the distribution function:

$$F_{\beta,\xi}(y) = 1 - \left(1 + \xi\,\frac{y}{\beta}\right)^{-\frac{1}{\xi}} \quad \text{for } \xi > 0 \text{ and } y \ge 0, \tag{10}$$

where β is the scale parameter and ξ is the shape parameter, whose reciprocal is commonly referred to as a tail index.

The GPD parameters are estimated from the sample $Y_1, Y_2, ..., Y_n$, where $Y_i = X_i - u$ and $Y_i > 0$ for $i = 1, 2, ..., n_u$.

If $\xi < 0.5$ parameters ξ and β of the GPD may be estimated by the method of moments. Then $\hat{\xi} = \frac{1}{2} - \frac{\overline{Y}^2}{2S^2}$ and $\hat{\beta} = \frac{\overline{Y}\left(\overline{Y}^2/S^2 + 1\right)}{2}$, where $\overline{Y}$, S^2 are subsequently the mean and the variance of n_u - element sample $Y_1, Y_2, ..., Y_n$.

Since for $0.5 \le \xi < 1$ the variance of the GPD does not exist, the parameter estimates may only be obtained with the use of the expectation and the probability weighted moments. In such case, solving the system of equations (5) for the expected value ($M_{1,0,0}$) and the moment $M_{1,0,1}$ yields the estimates of the form:

$$\hat{\xi}^{mMWP} = 2 - \frac{\overline{Y}}{\overline{Y} - 2\alpha}, \tag{11}$$

$$\hat{\beta}^{mMWP} = \frac{2\alpha\overline{Y}}{\overline{Y} - 2\alpha}, \tag{12}$$

where α is the estimate of the moment $M_{1,0,1}$. Depending on the chosen form of the empirical distribution function we have subsequently:

- for the empirical distribution (6):

$$\alpha = \frac{1}{n_u}\sum_{i=1}^{n_u}\frac{n_u - i}{n_u - 1}Y_{(i)}^{(n_u)} \tag{13}$$

- for the proposed level crossing empirical distribution function (7):

$$\alpha = \frac{1}{n_u}\left(\frac{1}{2}+\frac{n_u - 2}{2\sqrt{n_u(n_u - 1)}}\right)Y_{(1)}^{(n_u)} + \frac{1}{n_u}\sum_{i=2}^{n_u - 1}\left(\frac{1}{2}+\frac{n_u - 2i}{2\sqrt{n_u(n_u - 1)}}\right)Y_{(i)}^{(n_u)}, \tag{14}$$

where $Y_{(i)}^{(n)}$ is the *i*-th position statistic

For $\xi \geq 1$ the expectation and the variance of the GPD do not exist. The estimates of its parameters are based on moments $M_{1,0,k}$ and $M_{1,0,k+1}$ (if they exist) and, for the empirical distribution function (6), it holds that (Seckin et al., 2010)

$$\hat{\xi} = \frac{(k+1)^2\alpha_1 - (k+2)^2\alpha_2}{(k+1)\alpha_1 - (k+2)\alpha_2}, \tag{15}$$

$$\hat{\beta} = \frac{(k+2)(k+1)\alpha_1\alpha_2}{(k+1)\alpha_1 - (k+2)\alpha_2}, \tag{16}$$

where α_1 and α_2 take the following forms:

$$\alpha_1 = \frac{1}{n_u}\sum_{i=1}^{n_u}\frac{(n_u - i)(n_u - i - 1)\ldots(n_u - i - k + 1)}{(n_u - 1)(n_u - 2)\ldots(n_u - k)}Y_{(i)}^{(n_u)}, \tag{17}$$

$$\alpha_2 = \frac{1}{n_u}\sum_{i=1}^{n_u}\frac{(n_u - i)(n_u - i - 1)\ldots(n_u - i - k)}{(n_u - 1)(n_u - 2)\ldots(n_u - k - 1)}Y_{(i)}^{(n_u)}. \tag{18}$$

The application of the level crossing empirical distribution function leads to the sample moments:

$$m_{1,j,0} = \left(\frac{1}{2}-\frac{n-2}{2\sqrt{n(n-1)}}\right)^j\frac{X_{(1)}^{(n)}}{n} + \sum_{i=2}^{n-1}\frac{X_{(i)}^{(n)}}{n}\left(\frac{1}{2}-\frac{n-2i}{2\sqrt{n(n-1)}}\right)^j + \frac{1}{n}X_{(i)}^{(n)}, \tag{19}$$

$$m_{1,0,k} = \left(\frac{1}{2}+\frac{n-2}{2\sqrt{n(n-1)}}\right)^k\frac{X_{(1)}^{(n)}}{n} + \sum_{i=2}^{n-1}\frac{X_{(i)}^{(n)}}{n}\left(\frac{1}{2}+\frac{n-2i}{2\sqrt{n(n-1)}}\right)^k. \tag{20}$$

and

$$\hat{\xi} = \frac{(k+1)^2 m_{1,0,k} - (k+2)^2 m_{1,0,k+1}}{(k+1)m_{1,0,k} - (k+2)m_{1,0,k+1}}, \tag{21}$$

$$\hat{\beta} = \frac{(k+2)(k+1)m_{1,0,k} m_{1,0,k+1}}{(k+1)m_{1,0,k} - (k+2)m_{1,0,k+1}}. \tag{22}$$

4. Statistical properties of GPD parameter estimates for chosen return distributions

For its practical importance, in the simulation study we focused on parameter ξ estimation. Statistical properties of the estimates were examined in a simulation study. We considered four fat-tailed distributions: t-Student, Pareto, log gamma and Burr. The distribution parameters assumed for the purpose of the simulation experiment were assessed on the basis of the historical distribution of the daily returns from the stock market index. In order to reflect typical features of real financial series the DJIA index was chosen, which has a long tradition and is frequently used as a benchmark index in empirical analysis.

As the initial sample estimates showed that the ξ values usually fell into the interval (0,1) therefore we considered ξ from that range and presented the results for $\xi = 0.2, 0.4, 0.6, 0.8$. We set the distribution parameters on relevant levels to obtain such values. The threshold u was fixed at the level of the quantile of order 0.95 and the sample size was set to 1000, which guaranteed the number of observations over u large enough for the use of the asymptotic theorems.

In cases where values of ξ were below 0.5 we considered the use of the method of moments and PWMM estimates for weighted moments of orders 0 to 3. In cases with ξ between 0.5 and 1 it was only possible to use the expected value and further probability weighted moments of orders higher than 0. We used moments of orders 0 to 3. In the first step we investigated how the moment choice affects the estimate properties. The initial study was conducted with the use of the usual cdf formula (6)[1]. The results, obtained over 20000 replications[2], are presented in Tables 1 - 4. The simulation results showed that, while the estimate bias did not always exhibit a regular behavior, the variance influence resulted in the mean square error (MSE) increase for higher weighted moments. On the other hand, the ordinary methods of moments in most cases gave larger MSE than PWMM. Hence the results support the choice of the PWMM with expectation and the weighted moment of order 1, for ξ estimation, according to formulas (11) and (12).

[1] The simulations were afterwards repeated for the modified method with the crossing level distribution function, which gave the same conclusions about the choice of the order of weighted moments. The results are available from the authors upon request.

[2] In most cases 10000 is a sufficient number of repetitions in simulations (Białek, 2014), however, in our study we noticed differences in the second decimal point when diminishing the number of repetitions from 20000 to 10000.

Table 1. Statistical properties of parameter $\xi = 0.2$ estimates for different weighted moments

Distribution	Moments	$BIAS(\hat{\xi})$	$D^2(\hat{\xi})$	$MSE(\hat{\xi})$
t-Student (5)	EX , D^2X	-0.139	0.041	0.060
	EX , $M_{1,0,1}$	-0.099	0.041	0.051
	$M_{1,0,1}$, $M_{1,0,2}$	-0.118	0.116	0.130
	$M_{1,0,2}$, $M_{1,0,3}$	-0.134	0.292	0.310
Pareto (1, 5) $\xi = 0.2$	EX , D^2X	-0.088	0.027	0.035
	EX , $M_{1,0,1}$	-0.027	0.029	0.030
	$M_{1,0,1}$, $M_{1,0,2}$	-0.020	0.093	0.094
	$M_{1,0,2}$, $M_{1,0,3}$	-0.023	0.257	0.254
Log gamma (2, 0.2, 0)	EX , D^2X	-0.084	0.026	0.033
	EX , $M_{1,0,1}$	-0.023	0.030	0.031
	$M_{1,0,1}$, $M_{1,0,2}$	-0.016	0.093	0.093
	$M_{1,0,2}$, $M_{1,0,3}$	-0.021	0.254	0.254
Burr (0.5, 10)	EX , D^2X ,	0.015	0.015	0.015
	EX , $M_{1,0,1}$	0.130	0.043	0.060
	$M_{1,0,1}$, $M_{1,0,2}$	0.188	0.113	0.149
	$M_{1,0,2}$, $M_{1,0,3}$	0.208	0.253	0.297

Table 2. Statistical properties of parameter $\xi = 0.4$ estimates for different weighted moments

Distribution	Moments	$BIAS(\hat{\xi})$	$D^2(\hat{\xi})$	$MSE(\hat{\xi})$
t-Student (2.5)	EX , D^2X	-0.185	0.0522	0.086
	EX , $M_{1,0,1}$	-0.081	0.0398	0.046
	$M_{1,0,1}$, $M_{1,0,2}$	-0.076	0.092	0.099
	$M_{1,0,2}$, $M_{1,0,3}$	-0.086	0.236	0.243
Pareto (1, 2.5)	EX , D^2X	-0.168	0.044	0.072
	EX , $M_{1,0,1}$	-0.050	0.033	0.036
	$M_{1,0,1}$, $M_{1,0,2}$	-0.022	0.082	0.082
	$M_{1,0,2}$, $M_{1,0,3}$	-0.024	0.221	0.222
Log gamma (2, 0.4, 0)	EX , D^2X	-0.151	0.038	0.061
	EX , $M_{1,0,1}$	-0.022	0.031	0.031
	$M_{1,0,1}$, $M_{1,0,2}$	-0.011	0.081	0.081
	$M_{1,0,2}$, $M_{1,0,3}$	-0.009	0.2218	0.222

Table 2. Statistical properties of parameter $\xi = 0.4$ estimates for different weighted moments (cont.)

Distribution	Moments	$BIAS(\hat{\xi})$	$D^2(\hat{\xi})$	$MSE(\hat{\xi})$
Burr (0.5, 5)	EX, D^2X	-0.111	0.025	0.037
	EX, $M_{1,0,1}$	0.055	0.031	0.034
	$M_{1,0,1}$, $M_{1,0,2}$	0.140	0.094	0.114
	$M_{1,0,2}$, $M_{1,0,3}$	0.166	0.227	0.255

Table 3. Statistical properties of parameter $\xi = 0.6$ estimates for different weighted moments

Distribution	Moments	$BIAS(\hat{\xi})$	$D^2(\hat{\xi})$	$MSE(\hat{\xi})$
t-Student (5/3)	$M_{1,0,0}$, $M_{1,0,1}$	-0.111	0.046	0.058
	$M_{1,0,1}$, $M_{1,0,2}$	-0.057	0.080	0.083
	$M_{1,0,2}$, $M_{1,0,3}$	-0.059	0.205	0.208
Pareto (1, 5/3) $\xi = 0.2$	$M_{1,0,0}$, $M_{1,0,1}$	-0.099	0.042	0.052
	$M_{1,0,1}$, $M_{1,0,2}$	-0.034	0.077	0.078
	$M_{1,0,2}$, $M_{1,0,3}$	-0.033	0.197	0.198
Log gamma (2, 0.6, 0)	$M_{1,0,0}$, $M_{1,0,1}$	-0.056	0.034	0.037
	$M_{1,0,1}$, $M_{1,0,2}$	-0.032	0.074	0.075
	$M_{1,0,2}$, $M_{1,0,3}$	-0.039	0.189	0.191
Burr (0.5, 3.33)	$M_{1,0,0}$, $M_{1,0,1}$	-0.032	0.029	0.030
	$M_{1,0,1}$, $M_{1,0,2}$	0.086	0.075	0.082
	$M_{1,0,2}$, $M_{1,0,3}$	0.114	0.190	0.203

Table 4. Statistical properties of parameter $\xi = 0.8$ estimates for different weighted moments

Distribution	Moments	$BIAS(\hat{\xi})$	$D^2(\hat{\xi})$	$MSE(\hat{\xi})$
t-Student (1.25)	$M_{1,0,0}$, $M_{1,0,1}$	-0.180	0.062	0.095
	$M_{1,0,1}$, $M_{1,0,2}$	-0.059	0.076	0.080
	$M_{1,0,2}$, $M_{1,0,3}$	-0.054	0.180	0.183
Pareto (1; 1.25)	$M_{1,0,0}$, $M_{1,0,1}$	-0.176	0.061	0.093
	$M_{1,0,1}$, $M_{1,0,2}$	-0.051	0.075	0.078
	$M_{1,0,2}$, $M_{1,0,3}$	-0.041	0.181	0.182

Table 4. Statistical properties of parameter $\xi = 0.8$ estimates for different weighted moments (cont.)

Distribution	Moments	$BIAS(\hat{\xi})$	$D^2(\hat{\xi})$	$MSE(\hat{\xi})$
Log gamma (2, 0.8,0)	$M_{1,0,0}$, $M_{1,0,1}$	-0.127	0.044	0.060
	$M_{1,0,1}$, $M_{1,0,2}$	0.041	0.075	0.077
	$M_{1,0,2}$, $M_{1,0,3}$	0.061	0.175	0.179
Burr (0.5, 2.5)	$M_{1,0,0}$, $M_{1,0,1}$	-0.136	0.044	0.063
	$M_{1,0,1}$, $M_{1,0,2}$	0.043	0.070	0.072
	$M_{1,0,2}$, $M_{1,0,3}$	0.075	0.172	0.178

In the next step we evaluated, over 20000 repetitions, the bias and the variance of estimates $\hat{\xi}_1, \hat{\xi}_2$ for empirical distribution functions respectively (6) and (7). According to the initial results we restricted the analysis to the use of PWMM with expectation and the weighted moment of order 1. In cases where $0 < \xi < 0.5$ we also considered the ordinary methods of moments, whose estimates were denoted $\hat{\xi}_0$. The results are presented in Tables 5 and 6 (ξ values below 0.5) and 7 and 8 (ξ values between 0.5 and 1). For the considered range of ξ values, in comparison to the commonly used empirical cdf formula (6), the choice of formulas given by (7) improved both the bias and the variance of parameter ξ estimates. The estimate bias for the formula $\hat{\xi}_1$ was mainly negative and reached values up to -0.18, which showed that this formula produced substantially lower estimates than the real parameter value. For the examined fat-tailed distributions better statistical properties were obtained for estimate $\hat{\xi}_2$, which used the suggested level crossing empirical distribution function defined by formula (7). Nearly in all cases both the bias and variance were lower for the modified method. Moreover, the PWMM estimates for both cdf formulas offered systematically lower bias and variance then the ordinary method of moments.

Table 5. Bias and variance of parameter $\xi = 0.2$ estimates for different cdf estimates

Distribution	$BIAS(\hat{\xi}_0)$	$BIAS(\hat{\xi}_1)$	$BIAS(\hat{\xi}_2)$	$D^2(\hat{\xi}_0)$	$D^2(\hat{\xi}_1)$	$D^2(\hat{\xi}_2)$
t-Student(5)	0.139	0.099	-0.064	0.041	0.041	0.032
Pareto(1, 5)	-0.088	-0.027	0.004	0.027	0.029	0.025
Log gamma(2, 0.2,0)	-0.084	-0.023	0.008	0.030	0.030	0.026
Burr(0.5, 10)	0.015	0.130	0.153	0.015	0.043	0.047

Table 6. Bias and variance of parameter $\xi = 0.4$ estimates for different cdf estimates

Distribution	$BIAS\left(\hat{\xi}_0\right)$	$BIAS\left(\hat{\xi}_1\right)$	$BIAS\left(\hat{\xi}_2\right)$	$D^2\left(\hat{\xi}_0\right)$	$D^2\left(\hat{\xi}_1\right)$	$D^2\left(\hat{\xi}_2\right)$
t-Student(2.5)	-0.185	-0.081	-0.045	0.052	0.040	0.031
Pareto(1, 2.5)	-0.168	-0.050	-0.015	0.044	0.033	0.027
Log gamma(2, 0.4,0)	-0.151	-0.022	-0.011	0.038	0.031	0.027
Burr(0.5, 5)	-0.111	-0.055	-0.084	0.024	0.031	0.032

Table 7. Bias and variance of parameter $\xi = 0.6$ estimates for different cdf estimates

Distribution	$BIAS\left(\hat{\xi}_1\right)$	$BIAS\left(\hat{\xi}_2\right)$	$D^2\left(\hat{\xi}_1\right)$	$D^2\left(\hat{\xi}_2\right)$
t-Student(5/3)	-0.111	-0.084	0.046	0.037
Pareto(1, 5/3)	-0.099	0.073	0.042	0.034
Log gamma(2, 0.6,0)	-0.056	-0.032	0.034	0.029
Burr(0.5, 3.33)	-0.032	-0.010	0.029	0.025

Table 8. Bias and variance of parameter $\xi = 0.8$ estimates for different cdf estimates

Distribution	$BIAS\left(\hat{\xi}_1\right)$	$BIAS\left(\hat{\xi}_2\right)$	$D^2\left(\hat{\xi}_1\right)$	$D^2\left(\hat{\xi}_2\right)$
t-Student(1.25)	-0.180	-0.169	0.062	0.056
Pareto(1, 1.25)	-0.176	-0.165	0.061	0.055
Log gamma(2, 0.8,0)	-0.127	-0.117	0.044	0.039
Burr(0.5, 2.5)	-0.136	-0.126	0.044	0.040

In finance, the estimates of the generalized Pareto distribution are used for fitting tails of fat-tailed return distributions. The functional form of the tail gives basics for calculating high quantiles of the distribution and in consequence offers risk measures. As an illustration of the simulation results we compared the use of the estimate in the form $\hat{\xi}_2$, characterized by the lowest bias and variance, with the usual $\hat{\xi}_1$ formula for the empirical time series. We applied the ordinary and the modified method to calculate the tail estimate from the sample of 1000 observations of DJIA dating back to March 2009. The estimated distribution function took the form:

$$\hat{F}_2(x) = 1 - 0.05\left(1 + 0.104\frac{x - 0.025}{0.008}\right)^{-\frac{1}{0.104}} \quad \text{for} \quad \hat{\xi}_2 \tag{23}$$

compared to

$$\hat{F}_1(x) = 1 - 0.05\left(1 + 0.052\frac{x - 0.025}{0.008}\right)^{-\frac{1}{0.052}} \quad \text{for} \quad \hat{\xi}_1 \tag{24}$$

In empirical analysis the value of parameter ξ, referred to as an extreme index, is often of particular interest, since it gives information about the degree of fatness of the distribution tail. In Figure 1 we presented the estimated values of this parameter for DJIA index, with the use of the three considered cdf estimates. The results were obtained

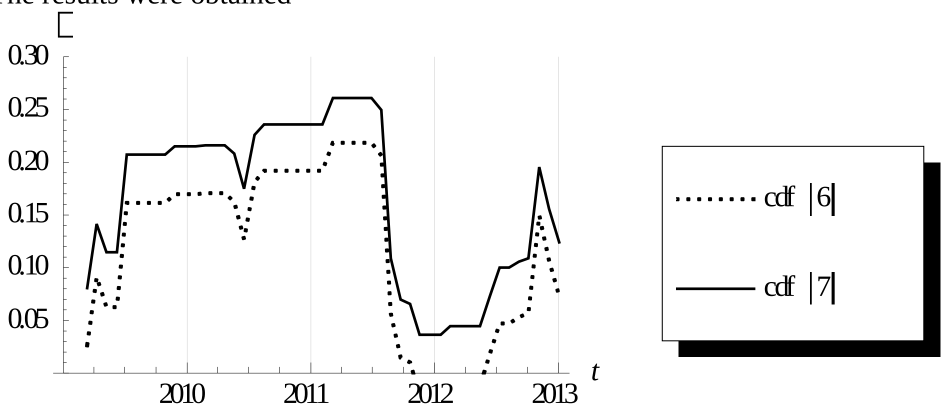

Figure 1. Parameter ξ estimates from the rolling estimation based on daily DJIA observations

from the rolling estimation with the window length of 1000 observations and the step of 20 observations. The dates included in the picture indicate the end point of the estimation window: starting sample ranged from March 2005 to March 2009 and the final sample range was from March 2009 to March 2013. The graphical presentation confirmed that parameter values had tendency to lie below the level of 1. The estimate $\hat{\xi}_1$ obtained with the use of the usual empirical distribution function in the form (6) lied at some distance from the proposed estimate $\hat{\xi}_2$, characterized by lower bias and variance, and were systematically below them.

5. Conclusion

The paper explored the probability weighted method of moments applied to estimation of the parameters of the generalized Pareto distribution. We suggested modification of the PWMM method through the application of the level crossing empirical distribution function. Statistical properties of the estimates of GPD parameter ξ, which characterizes the degree of tail fatness, were examined in a simulation study. We considered four fat-tailed distributions: t-Student, Pareto, log gamma and Burr and the typical parameter ξ values, which were assessed on the basis of the initial study for DJIA daily returns. The simulation exercise showed that suggested modification offered better statistical properties of the ξ estimate for ξ range (0,1) and the examined distributions. In case of other distributions, applicability of our results requires further research.

Since the PWMM method depends on the choice of the order of the probability weighted moments used for parameters estimation, the first step of the presented study concentrated on estimate statistical properties with relation to the chosen moment order. It was shown that there is a general increasing tendency in mean square error with increasing the moment order. This result supported the choice of the expectation and the weighted moment of order 1 for the considered ξ range and distributions.

Following the moment choice analysis we turned to examining statistical properties of parameter ξ estimates with reference to the cdf estimate required for PWMM parameter estimation. The simulation study results indicated that for the considered range of the parameter values it was possible to improve estimate statistical properties by the choice of the empirical distribution function given by formulas different that the commonly used cdf estimate.

REFERENCES

BIAŁEK J., (2014). Simulation Study of an Original Price Index Formula. *Communications in Statistics – Simulation and Computation* 43(2), 285–297.

HOSTING J. R. M., WALLIS J. R., (1987). Parameter and Quantile Estimation for the Generalized Pareto Distribution. *Technometrics* 29, 339–349.

HUANG M. L., BRILL P. H., (1999). *A level crossing quantile estimation method,* Statistics & Probability Letters 45, 111–119.

LANDWEHR J. M., MATALAS N. C., WALLIS J. R., (1979). Probability Weighted Moments Compared with Some Traditional Techniques in Estimating Gumbel Parameters and Quantiles. *Water Resources Research* 15(5), 1055–1064.

MAŁECKA M., PEKASIEWICZ D., (2013). Application of probability weighted method of moments to estimation of financial return distribution tail in: *Proceedings of the 31st International Conference Mathematical Methods in Economics 2013* [ed. Hana Vojáčková], College of Polytechnics Jihlava, 569–574.

MCNEIL A. J., FREY R., EMBRACHTS P., (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press, Princetown.

RASMUSSEN P. F., (2001). Generalized probability weighted moments: Application to the generalized Pareto distribution. *Water Resources Research* 37(6), 1745–1751.

SECKIN N., YURTAL R., HAKTAIR T., DOGAN A., (2010). Comparison of Probability Weighted Moments and Maximum Likelihood Methods Used in Flood Frequency Analysis for Ceyhan River Basin. *The Arabian Journal for Science and Engineering* 35, 49–69.