Folia Oeconomica Stetinensia
DOI: 10.2478/v10031-011-0043-9

Wydział Nauk Ekonomicznych i Zarządzania
Uniwersytetu Szczecińskiego

SIMPLIFIED METHOD OF GED DISTRIBUTION PARAMETERS ESTIMATION

Jan Purczyński, Prof.

Department of Quantitative Methods
Faculty of Management and Economics of Services
University of Szczecin
Cukrowa 8, 71-004 Szczecin
e-mail: jan.purczynski@wzieu.pl

Received 19 September 2011, Accepted 2 January 2012

Abstract

In this paper a simplified method of estimating GED distribution parameters has been proposed. The method uses first, second and 0.5-th order absolute moments. Unlike in maximum likelihood method, which involves solving a set of equations including special mathematical functions, the solution is given in the form of a simple relation. Application of three different approximations of Euler's gamma function value results in three different sets of results for which the χ^2 test is conducted. As a final solution (estimation of distribution parameters) the set is chosen which yields the smallest value of the χ^2 test statistic. The method proposed in this paper yields the χ^2 test statistic value which does not exceed the value of statistic for a distribution with parameters obtained with the maximum likelihood method.

Keywords: estimation of parameters of GED distribution.

JEL classification: C02, C13, C46.

Introduction

While determining the risk of investment in stocks a random variable, stock rate of return, is analyzed. When we determine a series of parameters describing this variable the risk of investment in stocks is measured, which is described in detail in the paper[1]. It is usually assumed that the rate of return (ROR) has a normal distribution. Although this assumption is not always legitimate. In the aforementioned paper, weekly rates of return of the WIG index for the period April 1991 – December 2000 were analyzed only to obtain a negative result for the Gaussian distribution[2].

Therefore it is necessary to model empirical rates of return by means of other distributions, for example distributions with the so called ”fat tails”, which are applied to modeling time-varying conditional variance, where – among others[3] – for estimation of the GARCH model as a conditional distribution, the GED distribution was applied. Another example can be paper[4] in which the modeling of distribution of logarithmic daily rate of return of the WIG index for the period April 1991 – July 2002 was done. Positive results of goodness-of-fit tests were obtained for the GED distribution as well as for the Laplace distribution, the parameters of which were estimated by means of the maximum likelihood method. For the normal distribution the negative results in both maximum likelihood and chi-square tests were obtained.

In this paper a simplified method of estimation of the GED distribution (Generalized Error Distribution) parameters was proposed. For the GED probability density function the following symbols were used:

$$f(x)=\frac{\lambda\cdot s}{2\cdot\Gamma\left(\frac{1}{s}\right)}\cdot\exp(-\lambda^{s}\,|x-\mu|^{s}) \tag{1}$$

where $\Gamma(z)$ – Euler's Gamma function.
For $s = 1$, the GED distribution turns into the Laplace distribution (bi-exponential):

$$f(x)=\frac{\lambda}{2}\cdot\exp(-\lambda\,|x-\mu|) \tag{2}$$

In the case of $s = 2$, the normal distribution is obtained:

$$f(x)=\frac{\lambda}{\sqrt{\pi}}\cdot\exp(-\lambda^{2}(x-\mu)^{2}) \tag{3}$$

In order to simplify the considerations, it is assumed that on the basis of a sample the estimation of the parameter μ was determined:

$$\widehat{\mu} = \frac{1}{N} \cdot \sum_{k=1}^{N} x_k \tag{4}$$

and consequently, the series of values x_k was centered by subtracting the value $\widehat{\mu}$. Hence the density function of the form is considered:

$$f(x) = \frac{\lambda \cdot s}{2 \cdot \Gamma\left(\frac{1}{s}\right)} \cdot \exp(-|\lambda \cdot x|^s) \tag{5}$$

In paper[5] the method of estimation of the GED distribution parameters based on absolute moments was proposed. An absolute moment of order m is given by the formula:

$$E_m = \int_{-\infty}^{\infty} |x|^m \cdot f(x) \cdot dx \tag{6}$$

From equations (5) and (8), we obtain:

$$E_m = \frac{\Gamma\left(\frac{m+1}{s}\right)}{\lambda^m \cdot \Gamma\left(\frac{1}{s}\right)} \tag{7}$$

The moment estimator E_m has the form:

$$\widehat{E}_m = \frac{1}{N} \cdot \sum_{k=1}^{N} |x_k|^m \tag{8}$$

Assuming two different values of m_1 and m_2 in (7) and solving the set of equations we obtain[6]:

$$\frac{\Gamma\left(\frac{m_1+1}{s}\right)^{\frac{1}{m_1}}}{\Gamma\left(\frac{m_2+1}{s}\right)^{\frac{1}{m_2}}} \cdot \Gamma\left(\frac{1}{s}\right)^{\frac{1}{m_2} - \frac{1}{m_1}} = \frac{(E_{m_1})^{\frac{1}{m_1}}}{(E_{m_2})^{\frac{1}{m_2}}} \tag{9}$$

$$\lambda = \left[\frac{\Gamma\left(\frac{m+1}{s}\right)}{E_m \cdot \Gamma\left(\frac{1}{s}\right)}\right]^{\frac{1}{m}} \tag{10}$$

where $m = m_1$ or $m = m_2$.

By solving equation (9), using e.g. the secant method, the estimation of the shape parameter $\hat{s}$ is determined. And by applying formula (10), the estimation of the parameter $\hat{\lambda}$ is calculated.

1. Proposed method

In paper[7] the method was proposed which does not require solving equation (11). An approximate dependency was given for estimating the parameter s for four variants of moment orders m_1 and m_2: ($m_1 = 0.1$; $m_2 = 0.5$), ($m_1 = 0.5$; $m_2 = 1$), ($m_1 = 1$; $m_2 = 2$), ($m_1 = 2$; $m_2 = 3$). The variant for which the estimated parameter s fulfils inequality $m_1 \le \hat{s} \le m_2$ should be chosen. On the basis of the author's experience (hundreds of cases of the estimated value of the shape parameter) it can be claimed that in the case of modeling the ROR of stock market indexes and companies the following inequality holds:

$$0.667 \le \hat{s} \le 2 \tag{11}$$

Hence the following expressions can be provided:

$$\hat{s}_1 = \left[\frac{0.2217}{\ln\left(\frac{\sqrt{\hat{E}_2}}{\hat{E}_1}\right) - 0.125} \right]^{0,879} \quad \text{where} \quad \hat{E}_1 = \frac{1}{N} \cdot \sum_{k=1}^{N} | x_k |; \;\; \hat{E}_2 = \frac{1}{N} \cdot \sum_{k=1}^{N} | x_k |^2 ,$$

$$\hat{s}_2 = \left[\frac{0.0747}{\ln\left(\frac{\sqrt{\hat{E}_1}}{\hat{E}_{0,5}}\right) - 0.0461} \right]^{0,935} \quad \text{where} \quad \hat{E}_{0,5} = \frac{1}{N} \cdot \sum_{k=1}^{N} | x_k |^{0,5} \tag{12}$$

However, the estimation of the parameter λ can be determined based on the dependency derived from equation (12) for $m = 1$:

$$\hat{\lambda} = \frac{\Gamma\left(\frac{2}{\hat{s}}\right)}{\hat{E}_1 \cdot \Gamma\left(\frac{1}{\hat{s}}\right)} \tag{13}$$

On the basis of equation (14) two values of the parameter s estimation are obtained. In this paper the following estimation dependency of the parameter s is proposed, based on estimations $\hat{s}_1$ and $\hat{s}_2$:

$$sA = \begin{cases} \hat{s}_1 & for \quad \hat{s}_1 > 1 \\ \hat{s}_2 & for \quad \hat{s}_1 < 1 \end{cases} \quad \text{for} \quad (\hat{s}_1 - 1) \cdot (\hat{s}_2 - 1) > 0 \,,$$

$$sA = \frac{\hat{s}_1 + \hat{s}_2}{2} \quad \text{for} \quad (\hat{s}_1 - 1) \cdot (\hat{s}_2 - 1) < 0 \tag{14}$$

Furthermore, yet another form of the parameter s estimation was considered, which was obtained as a result of modification of (14):

$$sP = \left[\frac{0.23}{\ln\left(\frac{\sqrt{\hat{E}_2}}{\hat{E}_1} \right) - 0.14} \right]^{0.88} \tag{15}$$

where $\hat{E}_1$ and $\hat{E}_2$ are defined by (14).

For the Gamma Euler's function the following formula holds[8]:

$$\frac{\Gamma\left(\frac{2}{s}\right)}{\Gamma\left(\frac{1}{s}\right)} = \frac{4^{\frac{1}{s}}}{2 \cdot \sqrt{\pi}} \Gamma\left(\frac{1}{s} + \frac{1}{2} \right) \tag{16}$$

From equations (13) and (16) we obtain:

$$\hat{\lambda} = \frac{4^{\frac{1}{\hat{s}}} \cdot \Gamma\left(\frac{1}{\hat{s}} + \frac{1}{2} \right)}{2 \cdot \sqrt{\pi} \cdot \hat{E}_1} \tag{17}$$

Taking into account inequality (11), it can be observed that the argument of the gamma function in equation (17) fulfils the condition:

$$1 \le \frac{1}{\hat{s}} + \frac{1}{2} \le 2 \qquad (18),$$

i.e. is within the range (1,2).

It is assumed that by applying equation (17) the gamma function is derived in an approximate manner. In order to do that, the formula provided in paper[9] can be applied, which ensures the relative error smaller than $5 \cdot 10^{-7}$. Considering computational complexity of this formula, the following approximations of function $\Gamma(x)$ were derived, correct when $1 \le x \le 2$:

$$f1(x) = \exp(0.953 - 1.442x + 0.484x^2) \tag{19a}$$

$$f2(x) = \exp(1.4087 - 2.41x + 1.1487x^2 + 0.148x^3) \tag{19b}$$

$$f3(x) = \exp(1.701375 - 3.23755x + 2.00873x^2 + 0.53763x^3 + 0.064995x^4) \tag{19c}$$

For individual formulas (19a)–(19c) the maximum error B fulfils the inequality: $|Ba| \leq 0.5$; $|Bb| \leq 0.06$; $|Bc| \leq 0.008$. The proposed algorithm is as follows:

1. In accordance with equation (12) moments $\widehat{E}_1$, $\widehat{E}_2$, $\widehat{E}_{0.5}$ are determined – as well as approximation values $\hat{s}_1$ and $\hat{s}_2$;
2. On the basis of equations (14) and (15), approximations sA and sP are calculated;
3. Substituting $\hat{s} = sA$ into equation (17) and applying equations (19a), (19b) and (19c), values $\lambda A1$, $\lambda A2$ and $\lambda A3$ are determined;
4. Substituting $\hat{s} = sP$ into equation (17) and applying equations (19a), (19b) and (19c), values $\lambda P1$, $\lambda P2$ and $\lambda P3$ are determined.

Consequently, six solution sets are obtained: (sA, $\lambda A1$, $\lambda A2$, $\lambda A3$) and (sP, $\lambda P1$, $\lambda P2$, $\lambda P3$), for which distribution goodness-of-fit tests are conducted. As a solution we choose the set that yields the smallest value of the test statistic.

2. Computational example

As an example the modeling of ROR distribution of WIG20 closing stock prices for the period 1994–2010 (daily data) was considered. Following the proposed algorithm estimations of parameters (sA, $\lambda A1$, $\lambda A2$, $\lambda A3$) were determined, for which the chi-square test was conducted. Subsequently, normalized statistics $hPA1$, $hPA2$, $hPA3$ were determined, being the ratio of test χ^2 statistic value and the critical value. For individual years – out of the three statistic values – the one with the smallest value was selected. The set of these values was labeled *hpa* and presented in Figure 1 with a dotted line with '+'. Another curve *hPA* represents the normalized statistic of the chi-square test obtained for estimations sA (equation (14)) and λA (equation (17)).

Similarly, the estimations of parameters (sP, $\lambda P1$, $\lambda P2$, $\lambda P3$) and normalized statistics $hP1$, $hP2$, $hP3$ were determined – the smallest values were labeled *hp* and presented in Figure 2 using a dotted line with '+'. Another curve *hP* represents the normalized statistic of the chi-square test obtained for estimations sP (equation (15)) and λA (equation (17)).

Figures 4 and 5 present comparison of the results of the proposed method and the results of the maximum likelihood method. The solid line with circles *hW* represents the values of the normalized statistic of the chi-square test for the maximum likelihood method. The dotted line with '+' represents the already defined values *hp* (Figure 4) and *hpa* (Figure 5).

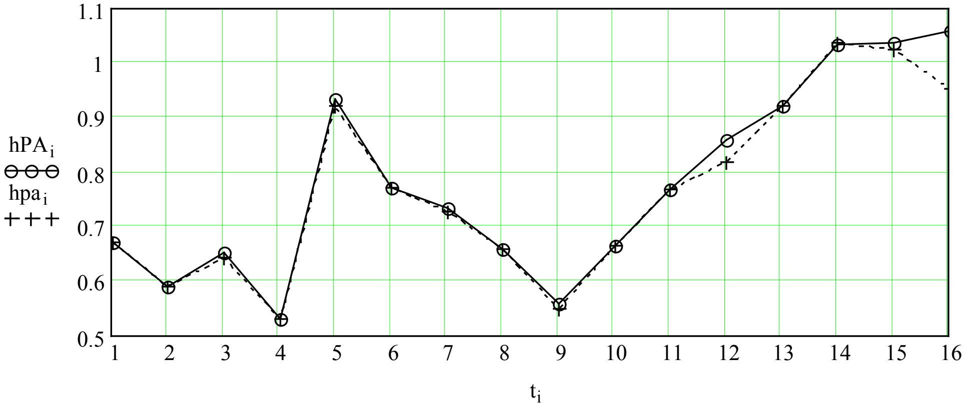

Fig. 1. Values of the normalized statistic of the χ^2 test for the ROR of the WIG20 index (daily data) for the period 1994–2010 for estimation *sA* (equation (14))

Source: Author's own study.

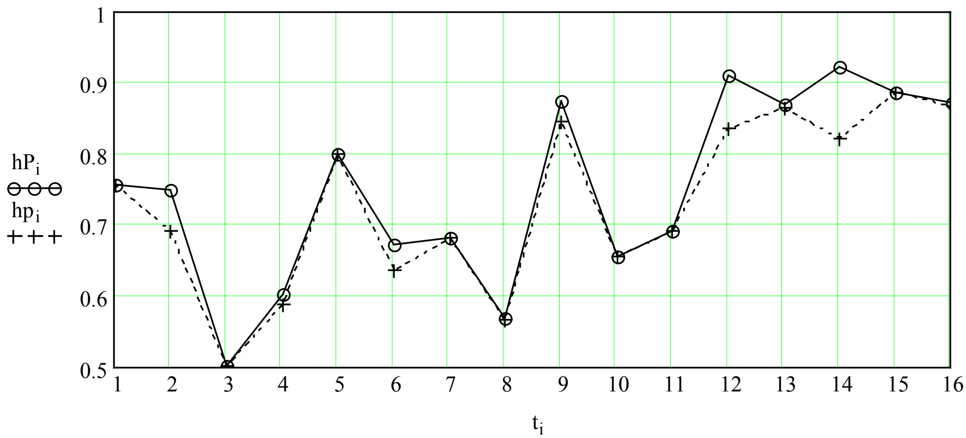

Fig. 2. Values of the normalized statistic of the χ^2 test for the ROR of the WIG20 index (daily data) for the period 1994–2010 for estimation *sP* (equation (15))

Source: Author's own study.

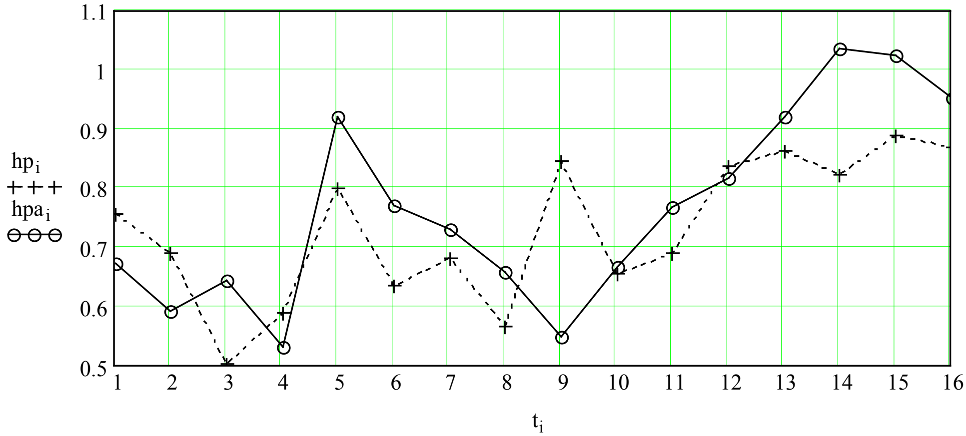

Fig. 3. Comparison of normalized statistic values of the χ^2 test for the ROR of the WIG20 index (daily data) for the period 1994–2010 for estimation *sA* (solid line with circles *hpa*) and estimation *sP* (dotted line with '+' *hp*)

Source: Author's own study.

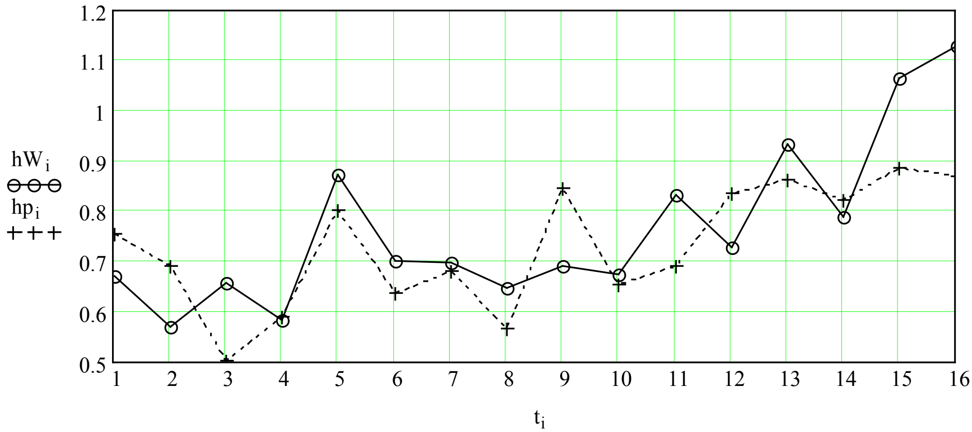

Fig. 4. Comparison of normalized statistic values of the χ^2 test for the ROR of the WIG20 index (daily data) for the period 1994–2010. The solid line with circles *hW* corresponds to estimations obtained through the maximum likelihood method, line *hp* as in Figure 2

Source: Author's own study.

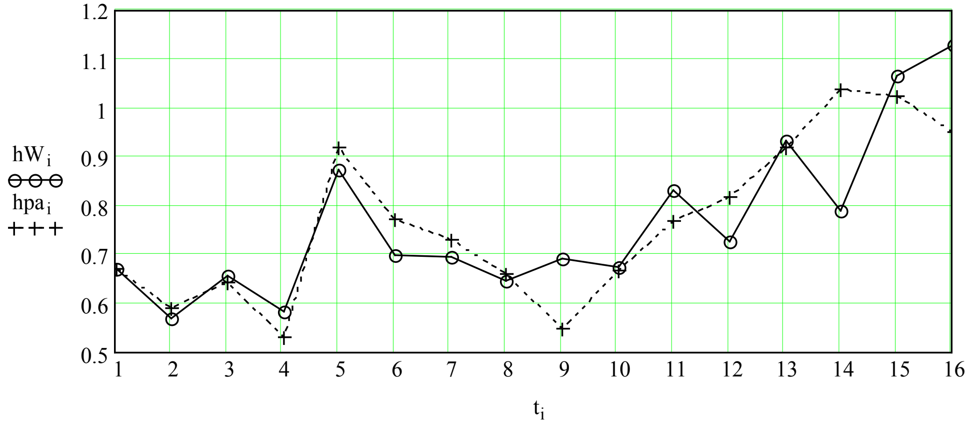

Fig. 5. Comparison of normalized statistic values of the χ^2 test for the ROR of the WIG20 index(daily data) for the period 1994–2010. The solid line with circles *hW* corresponds to estimations obtained through the maximum likelihood method, line *hpa* as in Figure 1

Source: Author's own study.

Figures 3, 4 and 5 clearly show which method is the best – yields the smallest value of the normalized statistic. Therefore, the mean value for 16 years of observations was calculated obtaining: *hpas* = 0.765; *hps* = 0.730; *hWs* = 0.764; *hps* is the smallest value, i.e. the proposed method based on equation (15). However the maximum likelihood method and the method based on equation (14) yield similar vales of the χ^2 test.

Figure 6 proves that the maximum likelihood method (solid line with rectangles *sW*) and the proposed method with equation (14) (dotted line with '+' *sPA*) yield similar values of estimations of the shape parameter *s.* Yet the method based on equation (15) estimates inflated values of the parameter *s* estimation in comparison with the maximum likelihood method. On the other hand, the method based on equation (15) yields smaller values of the statistic in the χ^2 test than the maximum likelihood method.

As another example, the issue of modeling of ROR distribution of SP500 closing stock prices for the period 1996–2010 (daily data) was considered. Following the proposed algorithm, the same calculations as in the case of the WIG20 index were conducted.

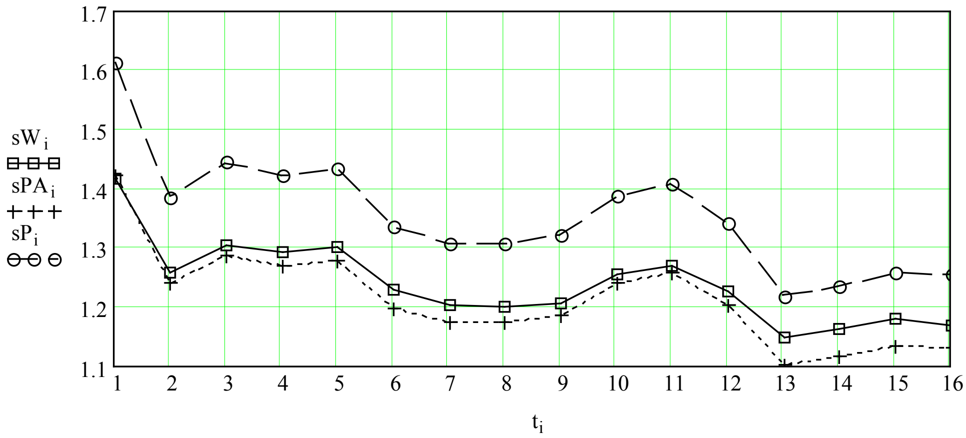

Fig. 6. Values of estimation of the shape parameter *s* obtained using the particular methods – WIG20 index. Solid line with rectangles *sW* – the maximum likelihood method; dotted line with '+' *sPA* – equation (14); dashed line with circles *sP* – equation (15)

Source: Author's own study.

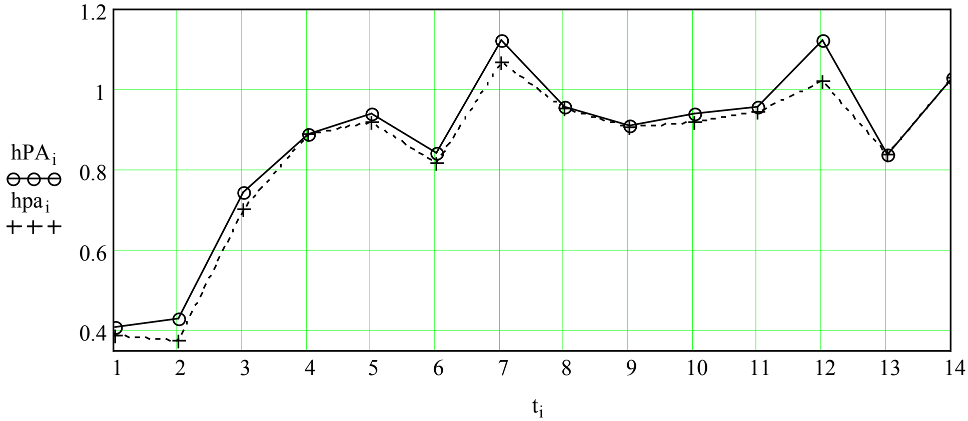

Fig. 7. Values of the normalized statistic of the χ^2 test for the ROR of the SP500 index (daily data) for the period 1996–2010 for estimation *sA* (equation (14))

Source: Author's own study.

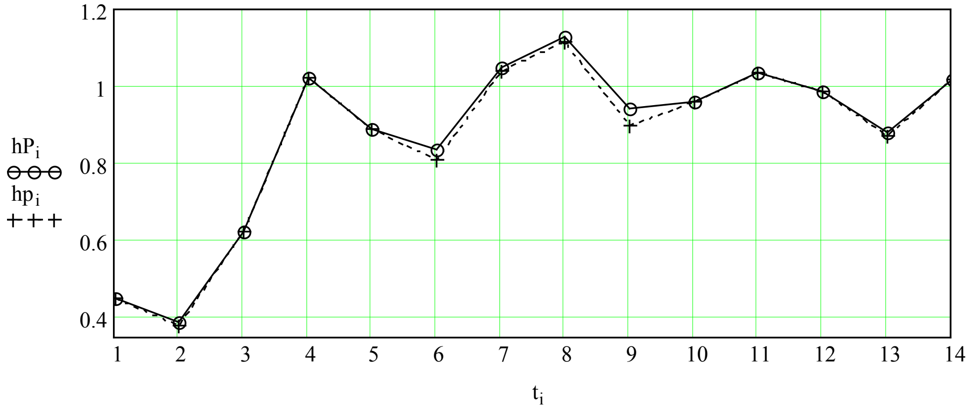

Fig. 8. Values of the normalized statistic of the χ^2 test for the ROR of the SP500 index (daily data) for the period 1996–2010 for estimation *sP* (equation (15))

Source: Author's own study.

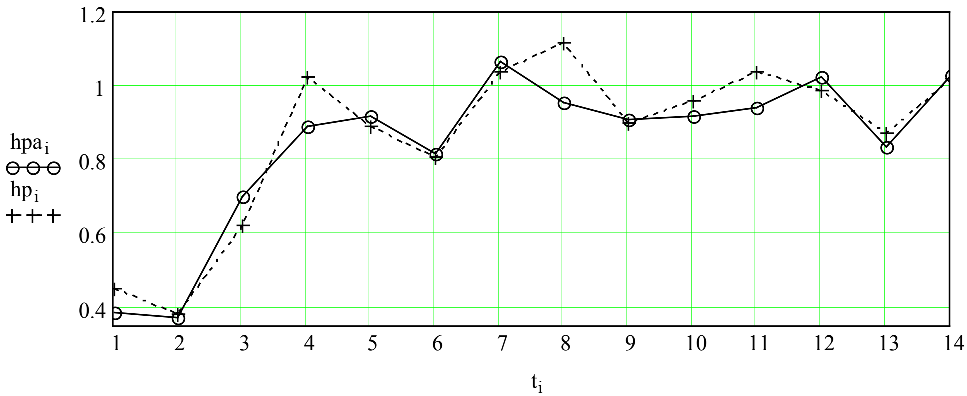

Fig. 9. Comparison of normalized statistic values of the χ^2 test for the ROR of the SP500 index (daily data) for the period 1996–2010 for estimation *sA* (solid line with circles *hpa*) and estimation *sP* (dotted line with '+' *hp*)

Source: Author's own study.

Figures 9, 10 and 11 do not clearly show which method is the best – yields the smallest value of the normalized statistic. Therefore, the mean value for 14 years of observations was calculated obtaining: *hpas* = 0.839; *hps* = 0.871; *hWs* = 0.863.

Hpas is the smallest value, i.e. the proposed method based on equation (14). However the maximum likelihood method and the method based on equation (15) yield similar values of the χ^2 test.

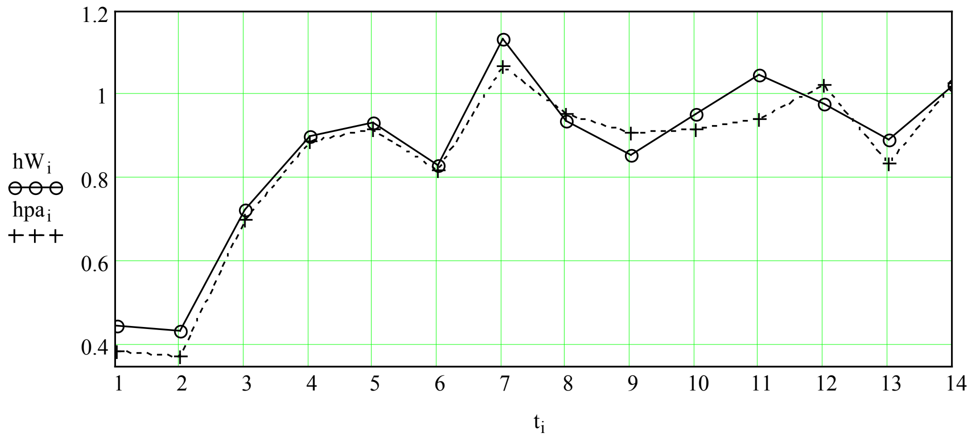

Fig. 10. Comparison of normalized statistic values of the χ^2 test for the ROR of the SP500 index (daily data) for the period 1996–2010. The solid line with circles *hW* corresponds to estimations obtained through the maximum likelihood method, line *hpa* as in Figure 1

Source: Author's own study.

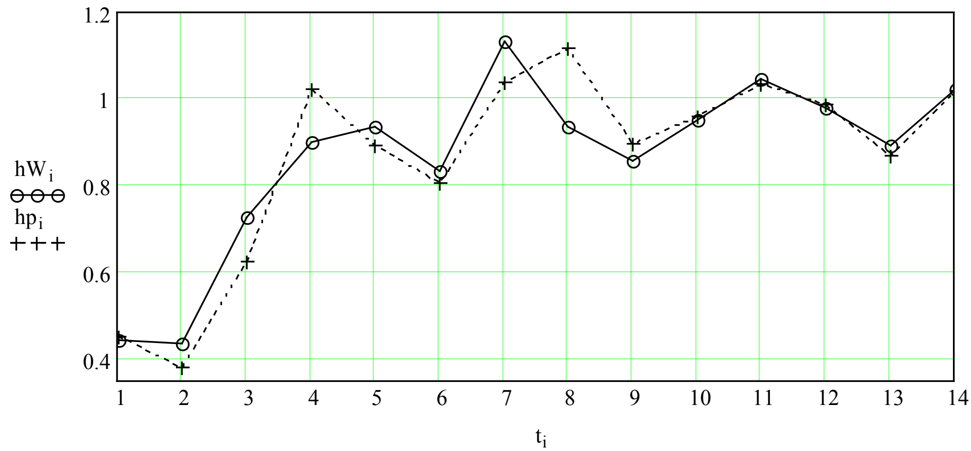

Fig. 11. Comparison of normalized statistic values of the χ^2 test for the ROR of the SP500 index (daily data) for the period 1996–2010. The solid line with circles *hW* corresponds to estimations obtained through the maximum likelihood method, line *hp* as in Figure 2

Source: Author's own study.

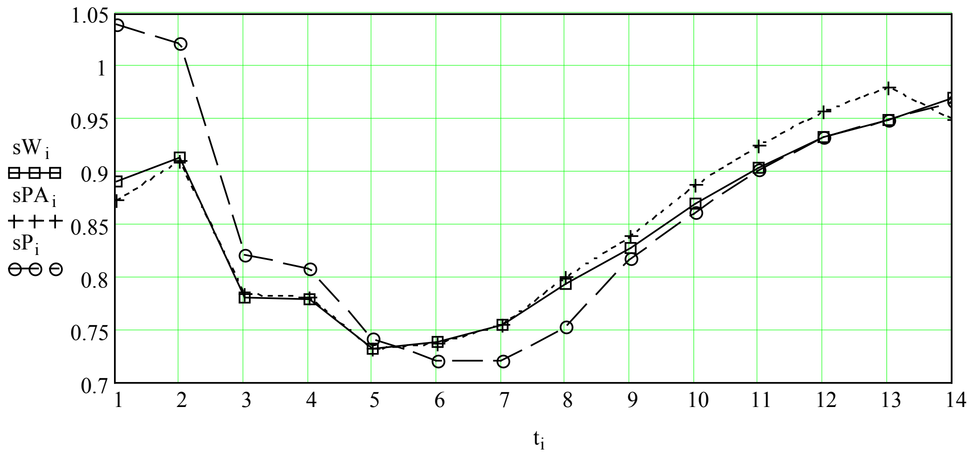

Fig. 12. Values of estimation of the shape parameter *s* obtained using the particular methods – SP500 index. Solid line with rectangles *sW* – the maximum likelihood method; dotted line with '+' *sPA* – equation (14); dashed line with circles *sP* – equation (15)

Source: Author's own study.

Figure 12 proves that the maximum likelihood method (solid line with rectangles *sW*) and the proposed method based on equation (14) (dotted line with '+' *sPA*) yield similar values of the shape parameter *s* estimations. However, the method based on equation (15) estimates values of the parameter *s* estimation different from the maximum likelihood method estimations. In contrast to the WIG20 index (Figure 6), where $sP > sW$ for the whole observation period, in the case of the SP500 index the following holds: $sP > sW$ for $t \in \langle 1,5 \rangle$ and $sP < sW$ for $t \in \langle 6,10 \rangle$.

As far as the WIG20 index is concerned, the best fit of the theoretical curve to empirical data was provided by the method based on equation (15). Yet in the case of the SP500 index an optimum estimation was obtained using equation (14). The reason for that lies in the values of the parameter *s* estimation. For the WIG20 index the following held: $sW > 1$ ($sPA > 1$), and for the SP500 index: $sW < 1$ ($sPA < 1$). It means that if the estimation of the shape parameter $\hat{s} > 1$, then equation (15) should be applied, yet if $\hat{s} < 1$, it is advisable to apply equation (14).

Conclusions

While applying the Maximum Likelihood Method to estimation of the GED distribution parameters given by equation (5), the following set of equations is obtained:

$$\lambda = \left(\frac{N}{s \sum_{k=1}^{N} | x_k |^s} \right)^{\frac{1}{s}} \tag{22}$$

$$s + \Psi\left(\frac{1}{s}\right) + \ln\left(\frac{s}{N} \sum_{k=1}^{N} |x_k|^s\right) - \frac{\sum_{k=1}^{N} |x_k|^s \ln|x_k|}{\sum_{k=1}^{N} |x_k|^s} = 0,$$

where: $\Psi(z) = \frac{d}{dz}\left[\ln\Gamma(z)\right]$.

The set of equations (22) includes the zeta function $\Psi(z)$, which is the derivative of the logarithm of the Euler's Gamma function $\ln(\Gamma(z))$. In order to solve this set of equations professional software is necessary.

In order to simplify calculations, in paper[10] the method of moments was proposed (equations (10), (11) and (12)). Solution of equation (11) is noticeably simpler than solution of equation (22), however the specialist software is still essential.

The method proposed in this paper eliminates these difficulties and makes it possible to perform calculations using a calculator – equations (12), (13), (14) and (15). Furthermore, application of equations (19a), (19b) and (19c) to determine the values of the Euler's Gamma function substantially reduces computational complexity of the algorithm in comparison with the Maximum Likelihood Method and the Moments Method.

In the proposed method the chi-square test is conducted for the six sets of parameter values (sA, $\lambda A1$, $\lambda A2$, $\lambda A3$) and (sP, $\lambda P1$, $\lambda P2$, $\lambda P3$). As a solution we choose the set which yields the smallest value of the test statistic. Such a modus operandi ensures that the proposed simplified method of estimation of the GED distribution parameters is competitive in terms of the solution quality, which is measured with the value of the χ^2 test statistic, compared with the Maximum Likelihood Method (Figure 4 and Figure 10).

Notes

1 Tarczyński, Mojsiewicz (2001), pp. 61–84.

2 Ibidem, pp. 55–58.

3 Nelson (1991), p. 368.

[4] Purczyński (2002), pp. 81–85.

[5] Ibidem, pp. 71–72.

[6] Ibidem, p. 72.

[7] Krupiński, Purczyński (2006), pp. 205–211.

[8] Ryżyk, Gradsztejn (1964), p. 348.

[9] Purczyński (2003), p. 124, eq. (2.1.9).

[10] Purczyński (2002), p. 71–72.

References

Krupiński, R., Purczyński, J. (2006). Approximated fast estimator for the shape parameter of generalized Gaussian distribution, *Signal Processing*, Vol. 86, No. 4.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, Vol. 59, No. 2.

Purczyński, J. (2002). *Estymacja parametrów rozkładu GED*. *Rynek kapitałowy – Skuteczne inwestowanie*, cz. 1, Szczecin: WNUS.

Purczyński, J. (2003). *Wykorzystanie symulacji komputerowych w estymacji wybranych modeli ekonometrycznych i statystycznych*, Rozprawy i Studia T.(DXXV)451, Szczecin: WNUS.

Ryżyk, I.M., Gradsztejn, I.S. (1964). *Tablice całek, sum, szeregów i iloczynów*, Warszawa: PWN.

Tarczyński W., Mojsiewicz M. (2001). *Zarządzanie ryzykiem*, Warszawa: PWE.