

Necessity of Use and Cognition Sense of Moments of Higher Rank (Asymmetry and Excess Coefficients)

Assoc. Prof. Todor Kaloianov, Ph.D.

*Department of Statistics and Econometrics,
UNWE*

Summary: The main part of the statistical methods for investigation of factor influences is based on comparison of conditional distributions. Common practice is the analyses to be limited only to measurement and interpretation of differences between the amounts of the arithmetic mean and the mean square deviation, due to the assumption that the distributions are normal. The arithmetic mean and the standard deviation are absolutely sufficient for definition of normal distribution. Very often in practice statistical sets can be found, whose units do not have normal distribution. In these cases additional information about the form of distribution is needed, which the arithmetic mean and the standard deviation could not provide.

The article explains the necessity to use moments of higher rank – third and forth. The cognitive sense of parameters of the statistical distribution, constructed on their basis – coefficients of asymmetry and excess, as well as those which are based on the positional means is discussed.

Key words: moments, distributions, asymmetry, excess.

JEL: C10, C16.

Introduction

Most statistical methods studying factor influences are based on the comparison of conditional distributions. This comparison and the assessment of differences can be achieved in another way – using tests of correspondence with a given theoretical distribution, comparing the importance of the respective differences between distributions, when data are obtained from representative samples. The situation is somewhat different when it is necessary to study the dynamics of a given phenomenon. In this case, the task consists in comparing unconditional distributions.

The opportunities offered by dynamic analysis of values of statistical distribution parameters – arithmetical mean, mean quadratic deviation, skewness and kurtosis coefficients – they all deserve our attention. The meaning of this approach is also determined by the necessity of performing such comparisons, when the methods of the theory of statistical conclusion and inference are not applicable, i.e. the data are obtained from exhaustive studies.

It is common practice to limit the analyzes only to finding and interpreting the arithmetic mean and the standard deviation, which is due to the assumption that the distributions are normal.

The arithmetic mean and standard deviation are indeed sufficient to define the normal distribution.

However, statistical populations whose units are not normally distributed are quite frequent in practice. In such cases, additional information is needed concerning the shape of the distribution, and it cannot be provided by the arithmetic mean and standard deviation. Because of the character of their construction, these parameters cannot show the direction and strength of those system and non-system influences, which affect only some of the units. Indeed, the measures of variation, standard deviation in particular, are also influenced by factors affecting part of the units, but they cannot separate and measure their effect. This problem is solved by the skewness coefficient. Usually, this parameter is described only as a measure of the degree of deviation of the empirical distribution from the respective normal one with respect to symmetry [Venezkii, Venezkaia, 1979; Gatev, 1986, 210]. This explanation can be considered insufficient in the sense that the question remains open concerning part of the influences, which the units have been subject to in the process of their genesis and formation of their values according to a certain characteristic. On this issue, it is important to point out that the following can be established using the skewness coefficient:

- The presence and strength of influences affecting only part of the units of the considered population.
- The cumulative direction of these influences resulting in left (negative) or right (positive) skewness.

Kurtosis coefficient is another important parameter of the statistical distribution. On this coefficient as well, the statistical literature usually only mentions that it measures the deviation of the empirical distribution from the respective normal distribution regarding the narrowness of its peak [Gatev, 1986; Mansfield, 1987]. Indeed, the different values of this parameter are also an indicator of the different influences on the units of a given population, but in another sense [Stefanov,

Totev, 1960]. If the distribution is symmetric, a high value of the kurtosis coefficient can mean a negative correlation between the different realizations of the individual influencing factors. In this case, the units are concentrated around the centre of the distribution to a higher degree than in the normal distribution, while "the skirts" of the empirical distribution go beyond the limits of normal distribution. In case of positive series correlations, the symmetric kurtosis turns out to be blunt, negative (other conditions unchanged). Using the kurtosis coefficient, again thanks to its construction, the answer is given to the question whether there is "overconcentration" of the units in the symmetric distribution, or the opposite takes place – "overdispersal" under the influence of certain factors.

The skewness and kurtosis coefficients can provide valuable information about the influences affecting statistical population units and assist decision-making [Kaloyanov, 1998, pp. 59-66]. It is not by coincidence that tens of years ago, these two coefficients were used for control of the production process [Borodachev, 1946, 1950]. The skewness coefficient (also known as a normalized third central moment) is also applied as one of the methods for comparison of different families of curves to reveal the difference between them, being used as a basis for the empirical analysis [Tsonev, 1971; Cox, Oakes, 1988]. The two measures are at the basis of the Jarque-Berra test of the coincidence with normal distribution.

In the analysis of empirical distributions, it should be taken into account that the assumption of a great part of distributions being normal or close to normal has resulted in the preponderant use of the arithmetic mean and variance measures in many studies. This assumption has reflected upon the development of different statistical methods for the study of links and relationships. Higher moments – third and fourth – are not used by these methods.

Necessary use and cognitive sense of higher moments (skewness and kurtosis coefficients)

Finance, and portfolio selection in particular, is one of the fields where statistics and the assumption of normal distribution are widely applied. In 1952, two studies were published, starting a new epoch in modern portfolio analysis. The first study was by Roy, A. D., "Safety first and the holding of assets", and the second – by Harry M. Markowitz, entitled "Portfolio Selection". According to financial specialists, Markowitz study was a revolution in the field of modern investment theory and practice. The essence of the author's idea consists in taking into account simultaneously the return and the variance of the return of the portfolio as a whole when making investment choices – Figure1.

This means to minimize variance at a given return or to ensure maximum return at a given risk level (variance). Variance is assumed not constant in this case; like in a number of statistical methods, it is taken to be variable and an optimal ratio is pursued between the values of the mean and variance. In finance, the return is measured by

mathematical expectation (the arithmetic mean), and variance by the standard deviation or by variance. The mathematical expectation (the expected rate of return) is calculated according to the formula:

$$E(r) = \sum_{s=1}^n \text{Pr}(s)r_s \tag{1}$$

where:
 $s=1, 2, \dots, n$ are the possible portfolio outcomes;
 r_s – the rate of return for scenario s ;
 $\text{Pr}(s)$ – the probability of outcome s taking place.

The variance measuring the risk is calculated according to the following formula, where the symbols are familiar:

$$\sigma^2 = \sum_{s=1}^n \text{Pr}(s)[r_s - E(r)]^2 \tag{2}$$

Financial specialists have accepted to work with the terms first, second, third, and fourth moments, which are more convenient for them because of the more frequent use of probabilities

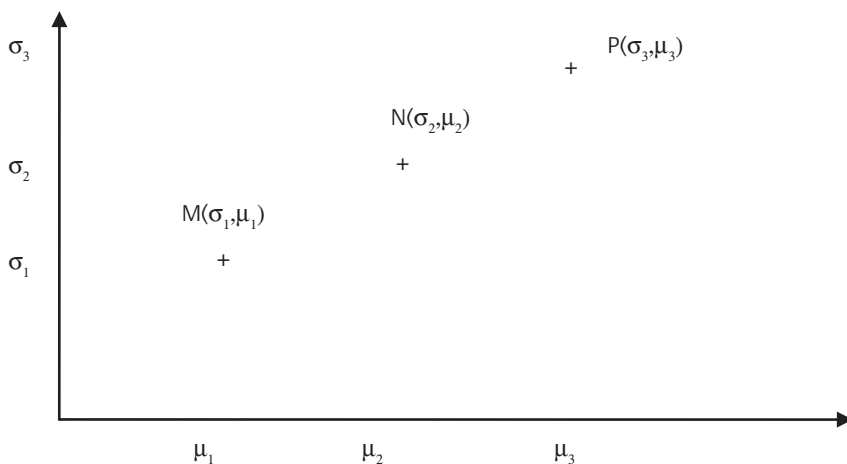


Figure 1. Alternative portfolios with different returns and risk levels

as weights, instead of frequencies. With variance, regardless of whether the average quadratic (standard) deviation or variance (dispersion) is used, they measure the risk of the respective investment. Variance is calculated as a difference between the expected and actual return. Obviously, the smaller is the difference between the expected and actual return, the lower is the risk and vice versa.

In reality, the assessment of and comparison between different investment portfolio allocations is done on the basis of two characteristics – expected value and standard deviation of portfolio returns. The mean value-standard deviation criterion (expectation-dispersion, mean-variance, or M-V) is built on the basis of these two characteristics. On p. 131, Bodie, Z, Kane, A, and Marcus, A (2000) define the criterion in the following way: “A is better than B, if

$$E(r_A) \geq E(r_B) \text{ and } \sigma_A \leq \sigma_B$$

and at least one of the equations is strict (i.e. no equality)”.

The criterion is constructed on the basis of only two characteristics of probability distribution – the mathematical expectation and standard deviation. And if it is assumed that the mathematical expectation (the arithmetic mean) is clear enough, then standard deviation has a rather rich content. In this case, it is accepted as a measure of risk and, according to the authors of “*Investments*”, “the idea is to describe the likelihood and magnitudes of “surprises” (deviations from the mean) with as small a set of statistics as is needed for accuracy. The idea is to describe the probabilities and magnitudes of “surprises” (deviations from the mean) using the minimum set of statistics necessary to achieve the needed preciseness” [2000; 140]. Bodie, Z, Kane, A, and Marcus, A emphasize that variance does not provide a full description of risk.

The main issue of concern not only for theoreticians, but also for financial practitioners, is when and under what conditions the mean-variance analysis is applicable. As Samuelson wrote in the very beginning of his article, “The Fundamental Approximation Theorem of Portfolio Analysis in Terms of Means, Variances, and Higher Moments” (1970), “James Tobin (1958, 1965), Harry Markowitz (1952, 1959), and many other writers have made valuable contributions to the problem of optimal risk decisions by emphasizing analyses of means and variances. These writers have realized that the results can be only approximate, but have also realized that approximate and computable results are better than none... But I think, it is important to re-emphasize an aspect of the mean-variance model that seems not to have received sufficient attention in the recent controversy, namely the usefulness of mean and variance in situations involving less and less risks – what I call “compact” probabilities. The present paper states and proves two general theorems involved. In a sense, therefore, it provides a defense of mean-variance analysis – in my judgment the most weighty defense yet given. (In economics, the relevant probability distributions are *not* nearly Gaussian, and the quadratic utility in the large leads to well-known absurdities). But since I improve on the mean-variance method and show its exact limitations – along with those for any *r*-moment model – the paper can also be regarded as a critique of the mean-variance approach.”

Samuelson study provides a theoretical basis for the use of mean-variance analysis.

In their actions, investors suppose that the conditions of applying mean-variance analysis are met and overlook higher moments. Like in many other fields, it is assumed that the distribution to work with in the case of assets’ returns is a normal one. The convenience created by this assumption is well-known. But

at the same time, the problems following from the acceptance of this assumption, when it is not true, are known as well. As Zvi Bodie, Alex Kane, Alan Marcus [2000; 143] note, "There are theoretical objections to the assumption that individual security returns are normally distributed. Since the security prices cannot be negative, the normal distribution cannot be truly representative of the behavior of the return for the holding period as it allows for any outcome, including for the whole range of negative prices."

The main criticism of the mean-variance method, according to Byrne, P. and Lee, S. (1997), consist in:

- Utility as a basis of the approach is subject to serious limitations, resulting in narrow capacity of describing the actual behavior of large groups of investors. The use of this approach for efficient portfolio selection imposes in almost all cases a quadratic utility function with the assertion that it approximates well enough many other functions. But this function presents significant limitations to the practical application of the mean-variance approach.
- The assumption of normality of returns is generally invalid for most securities, including real estate. When the distribution of returns is not normal, it is generally impossible to find the exact optimal solution of distribution problems (portfolio allocation).

Increasingly, researchers look for ways of overcoming the limitations imposed by the assumption of normality of distributions. A direct consequence of this assumption is also the use only of the mean and variance, generally the first and second initial, central, and mixed moments. Despite the convenience related to normal distribution, namely that it is defined unambiguously by two parameters only, it is useful in the solution just of one part of the various practical tasks.

One of the possible methods for achieving more accurate representation of reality and more precise solutions is the use of higher than second order moments. Traditionally, these moments have been applied only in physical and biological science. According to Donoho (2000) „by looking at third, fourth and even higher moments, new philosophical insights may be gained in a wide variety of disciplines, from analysis of genes to random motions of financial data series. Nowadays, including third and higher order moments becomes a must". Harvey and Siddique (2000) arrive to the conclusion that „the market takes into account the presence of skewness in assets evaluation and investors require compensation for maintaining assets with negative skewness". Ang and Bekaert (2001) arrive to the analogical conclusion; according to them „the market evaluates the degree of asymmetric relationship, generating an asymmetric portfolio". Tsaing (1972) suggests that "the inclusion of a higher moment is desirable, in this order and that under identical mean-variance criterion; the degree of positive skewness should be used to determine preferences. The use of higher moments enriches the various utility functions used in portfolio allocation".

Among the various studies on this issue, "*Skewness and the Bubble*" by Conrad, Dittmar, and Ghysels (2007) deserves our attention. The authors explore the possibility of higher moments in the distribution of returns to be significant in the explanation of securities returns. On the basis of data for the period between 1965 and 2005, they evaluate individual variances of securities, their skewness, and kurtosis. They find a significant negative relation between skewness and return – securities with positive or less negative skewness have lower returns in the next months. The authors come to the conclusion of a positive relation between kurtosis and return. It turns out that the sensitivity to skewness is different depending on the

production sector where securities are acquired. For example, the sensitivity to positive skewness is higher for the sectors of hardware, software production, semiconductor industry. There is an interesting conclusion, supported by practical results, that variance and skewness decrease with the increase of kurtosis. According to the authors, this emphasizes again the necessity for the evaluation of the link between return and higher moments to be done simultaneously. The results also show that shares with high variance and significant skewness have lower subsequent return, while higher kurtosis is related to higher subsequent returns.

The necessity of using higher moments was realized longtime ago. The reasons were both theoretical and purely practical. The common ground between the two types of reasons is the need to work with distributions different from normal ones, irrespectively of the influences having caused the respective distribution.

Karl Pearson (1895) pays attention to gamma distribution as a model of skewness¹. Pareto (1897) was also interested in asymmetric distributions, because part of the distributions in

the economy are asymmetric. As Groeneveld, R. A. and Meeden, G. (1984; 391) point out, "despite the many fields where asymmetric distributions are encountered and work "the concept and measurement of skewness remains imprecise."

The main question raised in almost all studies is what measures with different construction really determine and how to interpret the obtained results. It is obvious that there will be differences between the different measures of skewness, but what is important in the interpretation of the obtained results is to take into account the specific characteristics of those used in each particular case.

Unlike for the skewness indicator, the understanding of kurtosis has gone through serious evolution. The development is in three main directions. The first is related to the essence of the measure and its cognitive sense – what the presence of kurtosis means, what indeed is measured by the coefficient. The second direction is related to developing measures of kurtosis with different characteristics. The third direction is related to the first two and consists in the application of kurtosis in theory and practice.

* In case any of my readers may be unfamiliar with the term "kurtosis" we may define mesokurtic as "having β_2 equal to 3," while platykurtic curves have $\beta_2 < 3$ and leptokurtic > 3 . The important property which follows from this is that is that platykurtic curves have shorter "tails" than the



normal curve of error and leptokurtic longer "tails." I myself bear in mind the meaning of the words by the above *memoria technica*, where the first figure represents platypus, and the second kangaroos, noted for "lepping," though, perhaps, with equal reason they should be hares!

Figure 2.

¹ The chronology is from Seier, E. "Celebrating 100 years of Kurtosis 1905-2005".

During the period between 1906 and 1910, articles in the journal *Biometrika* contain comparisons between frequency distributions for very large populations in different fields and for normal distribution, using kurtosis and skewness.

Student, on p. 160 in his study "*Errors of Routine Analysis*" (1927), provides the following description of kurtosis, as shown on Figure 2.

An assumption is often made that, if for the same value of the argument $x = 0$, the density function is higher, then the respective distribution will have higher kurtosis. The term "*density crossing*" is also introduced as a sufficient condition for a distribution to have higher kurtosis β_2 than another distribution [Dyson, 1943, Funican, 1963]. It is presumed that, if two functions of probability distributions density with equal variance are crossed twice on each side about zero (the centre of the distribution), one will have higher kurtosis than the other.

It is most frequently accepted that kurtosis is a measure of whether the distribution is peakier or blunter (flattened) than the respective normal distribution. The higher kurtosis distribution reveals a trend of the peak being near the mean, it decreases fast and has heavy tails. Higher kurtosis means that a larger part of variance is due to the rarely met extreme deviations, contrary to medium deviations with high frequency. The distribution with a low value of kurtosis shows a trend of having rather a flat peak close to the arithmetical mean than a sharp peak.

Another definition of kurtosis is that the latter represents a degree of peakedness of the distribution, defined as a shape of the fourth central moment of the distribution. There are several methods of its representation. For example, it is marked as β_2 (Abramovitz and Stegun 1972; 928) or α_4 (Kenney and Keeping 1951; 27; Kenney and Keeping 1961; 99-102),

$$\text{where } \beta_2 = \frac{\mu_4}{\mu_2^2}, \quad (3)$$

and μ_i indicates the i -th central moment.

The same authors also present the following version of the formula:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3, \quad (4)$$

which is more frequently used as it measures the kurtosis with respect to normal distribution. By subtracting the number 3, the zero value is obtained for a normal kurtosis distribution.

Van Zwet W.R. (1964) introduced for a class of symmetric distributions an ordering \leq_S defined by $F \leq_S G$, if $R_{F,G}(x) = G^{-1}(F(x))$ is convex for $x > m_F$, where m_F is the symmetry point of F . It follows from the assumption of symmetry of distributions that $R_{F,G}(x)$ is convex for $x > m_F$, if it is concave for $x < m_F$. $F \leq_X G$ is valid, if the random variable X with distribution F can be attributed to the random variable Y with distribution G through increasing concave-convex function with respect to the median. Van Zwet defines kurtosis as an ordering of symmetric distributions and says that we should not be representing it by a single measure. He proposes a method of ordering of two distributions according to skewness. The author's idea found further advance and Loh, Oja, Lawrence orderings are known in the literature. The concave-convex functions are applied in the development of portfolio analysis methods.

Chissom B. S. (1970), experimenting with adding and subtracting cases, shows what is, in his opinion, the correct interpretation of the kurtosis coefficient (meaning the measure based on the fourth central moment). He used three distributions for this purpose – approximately normal, rectangular, and bimodal.

In his conclusion, he emphasizes that: "It is important to remember that kurtosis is dependent on the distribution peak and tails, and a major emphasis must be placed on the tails of the distribution in the determination of the fourth moment" [1970; 22].

In 1970, a discussion was started on the question whether the measures of kurtosis reveal the presence of bimodality. At the basis of this discussion was the study of Richard Darlington "Is Kurtosis Really "Peakedness?", published in *The American Statistician*. According to the author, the term "peakedness" is mistaken and there is a better term to describe kurtosis: "bimodality of the distribution" [1970; 19]. To prove this statement, Darlington uses the deviations z instead of the original values of X . In this way, the initial formula of the kurtosis coefficient

$$k = \frac{N^{-1} \sum (x - m)^4}{s^4} \quad (5)$$

is transformed into

$$k = N^{-1} \sum z^4. \quad (6)$$

According to Darlington, this means that k can be interpreted as "a measure of the degree, to which the values of z^2 are grouped around their mean with value 1; higher grouping (concentration), means a lower k . As z is equal to +1 or -1, when $z^2 = 1$, k can be interpreted as a measure of the degree, to which the z -values of the distribution are grouped around +1 and -1. The best characteristic to describe such grouping is "bimodality" [1970; 20]. In his publication, David K. Hildebrand (1971) also pays attention to this idea. He presents a family of double gamma distributions, which are bimodal and have a kurtosis coefficient between -2 and +3. It turns out, however, that bimodal distributions can have high kurtosis, when the modes are not near the values $z = \pm 1$. According to J. J. A. Moors,

this is the reason why the idea of Darlington has not received the attention it deserves. Moors (1986; 283) expresses the opinion that "kurtosis measures the variation around two values $\mu \pm \sigma$, which is a reverse measure of concentration in these two points. High kurtosis can be present in two situations: a) concentration of probability mass near μ (corresponding to peaked single-modal distribution) and b) concentration of probability mass at the tails of distributions. The existence of these two possibilities explains the confusion in the interpretation of kurtosis."

According to David Rupert [1987; 1] "kurtosis is often regarded as a measure of the weight of the distribution tails with respect to normal distribution. According to other authors, it measures the narrowness of the peak near the centre of the distribution". According to Rupert, the interpretation of kurtosis is too short and usually the attempts to make it more understandable are unsuccessful. On the same page, he indicates that "The fundamental problem, as Bickel and Lehmann (1975) noticed, is that there is no agreement on what kurtosis measures". Some authors (e.g. Kendall and Buckland 1971 and Levin 1984) state that kurtosis differentiates "the narrowness of the central peak" from "flatness". Darlington emphasizes that the opposite of "peaky" is bimodal. The flattened distribution is between these two extremes. But the terms "narrowness of the central peak", "flatness", and "bimodality" understate the dependency of kurtosis on the behavior of tails. Chissom draws attention on the fact that "a major emphasis must be placed on the tails of the distribution in the determination of the fourth moment".

Johnson and Kotz accept that "Kurtosis is a measure of the departure from normality depending on the relative frequency of values either near the mean or far from it, with respect to those located at intermediate distance from the mean" [1985; 22]. Alternative measures of

kurtosis are proposed, many of them developed to measure only the narrowness of the central peak or only the weight of tails, while according to Rupert, kurtosis characterizes both.

Using Hampel influence function (1968, 1974), Rupert develops further the discussion by Darlington. He carries out comparative analysis between the following three measures of kurtosis:

- Classical $k(F) = \mu_4(F) = \frac{\mu_4(F)}{\sigma^4(F)}$.
- Hogg measure (1974), which is supposed to measure the weight of tails on the basis of "final values", using the mean of tails:

$$\bar{U}_p(F) = \int_{q_{1-p}(F)}^{\infty} x dF(x) , \tag{7}$$

mean measuring the weight of the upper tail

$$\bar{L}_p(F) = \int_{-\infty}^{q_p(F)} x dF(x) , \tag{8}$$

mean measuring the weight of the lower tail.

As it is indicated

$$Q_p(F) = \frac{\bar{U}_p(F) - \bar{L}_p(F)}{\bar{U}_{0,5}(F) - \bar{L}_{0,5}(F)} \tag{9}$$

is invariant with respect to the location and scale, and it measures the weight of tails. In the quoted study, Hogg (1974) uses the second quantiles of the function F with the condition that $0 < p \leq 0.5$. This measure is found to be unstable as the upper means are sensitive to outliers, although not as much as the classical measure based on the fourth moment.

- according to Rupert, in order for the measure of kurtosis to be stable, it must be built on the basis of the ratio of two stable functions from the type:

$$R_{\eta,p}(F) = \frac{\{q_{1-p}(F) - q_p(F)\}}{\{q_{1-\eta}(F) - q_\eta(F)\}} \text{ for } 0 < p < \eta < 0.5. \tag{10}$$

Hampel influence function provides a quantitative understanding of kurtosis. A comparative analysis of the abovementioned three measures of kurtosis was carried out using this function. An attempt has been made to establish what is the relative importance of the weight of tails and the narrowness of the central peak in each one of them.

The following conclusions reached by Rupert by using the influence function, are of special importance for both theory and practice:

- kurtosis measures both the narrowness of the central peak and the weight of tails. There are no pure measures only of the narrowness of the central peak or only of the weight of tails;
- the location of outliers is as important as their frequencies;
- disturbances in the centre have much lower influence than those at the end of tails. The measure of kurtosis based on the fourth moment k is mostly a measure of the behavior of tails and less so of the narrowness of the central peak;
- the three compared measures have common characteristics.

J. J. A. Moors (1988) proposes a measure of kurtosis built on the basis of "octiles", i.e. quantiles dividing the row in eight equal parts.

The formula of the measure is the following:

$$\frac{[(E_7 - E_5) + (E_3 - E_1)]}{(E_6 - E_2)} , \tag{11}$$

where E_i ($i = 1, 2, 3, 5, 6, 7$) are, respectively, first, second, third, fifth, sixth, and seventh quantiles.

Moors indicates the following advantages of this measure of kurtosis:

- it can be used even when moments are nonexistent;
- it is not influenced by (outliers) tails of the distribution;
- it is easily calculated (can be determined even graphically).

The mentioned characteristics are a result of the use of different positional means, which are known not to be calculated from all values of the indicator. This is the reason why they are preferred in case of strongly deviating values.

During the same year, Balanda and MacGillivray [1988; p. 111] defined kurtosis as **“location and scale-free movement of probability mass from the shoulders of a distribution into its centre and tails and we admit that it and can be formalized in many ways”**. This definition is accepted as one of the most successful and is frequently quoted, as it takes into account the values and their frequencies not only in the centre of the distribution, but in its tails and shoulders. According to Balanda and MacGillivray, though moments play an important role in statistical inference they are very poor indicators of distributional shape. This is the reason why researchers look for different constructions of the measure of kurtosis. According to Balanda and MacGillivray, most studies concentrate on measuring kurtosis in symmetric distributions and less attention is paid to asymmetric distributions, including the connection between skewness and kurtosis. Obviously, this problem deserves more attention in the future. The opinion of the authors is that the formalization of kurtosis should be pursued in the partial ordering of distributions on the basis of Van Zwet’s concept (1964).

Two years later, in 1990, the same authors broadened Van Zwet’s criterion of asymmetric distributions. They introduced a coherent structure of ordering and measures, which do not require the presence of symmetry. This is made on the basis of the spread function.

During the same year, Hosking (1990) defined L-kurtosis, which is calculated according to the formula:

$$\tau = L_4/L_2, \quad (12)$$

where L are moments which, according to Hosking, represent summary statistics for probability distributions and data samples.

The theory also includes procedures like order statistics of Gini’s mean difference and provides promising innovations like measures of skewness and kurtosis, as well as new evaluation methods of the parameters of several distributions.

Groeneveld R. A. (1998) checks the degree of sensitivity of measures proposed by Groeneveld and Meeden to the shape of the symmetric distribution (1984). These measures are built on the basis of quantiles, defined in the following way [1998; p. 325]:

“For a continuous random variable X with distribution function F , and continuous density f , symmetric about zero, we focus on the distribution of $|X|$, which has density function $F_*(x) = 2F(x) - 1$ ”. The measures are defined by the expression representing a quantile measure of skewness:

$$\begin{aligned} \gamma_2(p, F) &= \frac{F_*^{-1}(1-p) + F_*^{-1}(p) - 2F_*^{-1}(0,5)}{F_*^{-1}(1-p) - F_*^{-1}(p)} = \\ &= \frac{F^{-1}(1 - (p/2)) + F^{-1}((1+p)/2) - 2F^{-1}(0,75)}{F^{-1}(1 - (p/2)) - F^{-1}((1+p)/2)}, \\ &0 < p < 0,5. \end{aligned} \quad (13)$$

Groeneveld R. A. demonstrates that, in essence, the measure is a ratio of differences between two positive distances and the sum of those distances.

$$\gamma_2(p, F) = \frac{[F_*^{-1}(1-p) - F_*^{-1}(0,5)] - [F_*^{-1}(0,5) - F_*^{-1}(p)]}{[F_*^{-1}(1-p) - F_*^{-1}(0,5)] + [F_*^{-1}(0,5) - F_*^{-1}(p)]} = \frac{d_2(p) - d_1(p)}{d_2(p) + d_1(p)} \quad (14)$$

In his study, Groeneveld defines several problems to be solved

- establishing to what extent $\gamma_2(p, F)$ is an appropriate measure of kurtosis;
- using the influence function to compare how $\gamma_2(p, F)$ and β_2 reflect the distributional shape;
- finding out to what extent the Shapiro-Wilk test is appropriate for measuring the difference from the normal distribution;

The answer to the first problem is that for appropriate values of p , $\gamma_2(p, F)$ is a useful measure of kurtosis.

The answer to the second problem is that the measure based upon moments gives higher influence to the weight of tails and relatively lower influence to the mass in the centre of the distribution. "The influence function for β_2 in the normal case shows that it is strongly affected by the small displacements of the mass towards the tails of the distribution. This effect significantly decreases with the quantile measure" [1998; p. 329].

The simulation carried out by the author allowed him to establish that the Shapiro-Wilk test is a sensitive measure of deviation from the normal distribution for small values of p , $\gamma_2(p, F)$.

The results are expected and logical, taking into account the averaging procedure included in the measure based upon moments. Raising to the fourth power in the numerator assigns higher weight to large deviations from the centre of the distribution and they obviously compensate the higher frequencies of values in the centre and close to it.

In 2002, sample g-kurtosis was defined, based on Geary's test of normality (1936) for symmetric distributions – $\hat{\tau}/\hat{\sigma}$.

$$g = \hat{w} = 13.29(\ln \hat{\sigma} - \ln \hat{\tau}), \quad (15)$$

$$\text{where } \hat{\tau} = (1/n) \sum_{i=1}^n |x_i - \bar{x}| \quad (16)$$

Seier and Bonett (2003) define two families of measures: $E(g(z))$, where g is a function of the standardized variable z , and a is a standardized variable selected in such a way that the kurtosis for normal distribution equals 3.

$$g(z) = ab^{-|z|}, \quad 2 \leq b \leq 20 \quad (17)$$

$$\text{and } g(z) = a[1-|z|^b], \quad 0.2 \leq b \leq 1 \quad (18)$$

It is indicated on these measures that they attribute higher weight to the central peak in comparison to the weight of tails, unlike the measure proposed by Karl Pearson. As a consequence, the normality tests based on these measures are more sensitive to the distribution peak. It is also established that the two measures satisfy Van Zwet's ordering and are closely related to L-kurtosis and to the measure based upon quantiles.

Samuel Kotz and Edith Seier (2007) turn again to the quantile measure of kurtosis introduced by Groeneveld. By replacing the difference between third and first quartile in the numerator $[Q_3(p) - Q_1(p)]$ with the difference between the second and first quartile $[Q_2 - Q_1(p)]$ and using an influence function, the authors explore the link between measures of kurtosis based upon moments and those based upon quantiles. The question of the ratio between the centre and tails of the distribution is again discussed. Five distributions, characterized as single mode distributions with heavy tails, are considered for this purpose.

Conclusion

During the last hundred years, there has been evolution from the more general definition of Pearson on distributions with kurtosis as relatively more or less flat-peaked than the normal curve to the idea of kurtosis as "location and scale-free movement of probability mass from the shoulders of a distribution into its centre and tails". Between these two ways of understanding, two other approaches have found their place. The first one, still frequent nowadays, including in Bulgarian-language literature, is that kurtosis represents the narrowness of the distribution's central peak. In this case, the necessary attention is not paid to the distribution tails. The other extreme interpretation is related to the exaggerated role of tails and the bimodality of the distribution.

The main problem that the different authors dealing with kurtosis are trying to find a decision to, is which is the most appropriate measure. Part of the authors accept measures constructed on the basis of moments (second and fourth central) as good enough – even if, according to a number of studies, these measures assign higher weight to tails, including to outliers, than to values in the centre or near the centre. For this reason, there are measures proposed on the basis of quantiles in many different versions. According to the author of the present paper, they have less cognitive value because the positional means do not account for the values of the variable. They are only a function of frequencies and use smaller quantity of information. Undoubtedly, they are not influenced by outliers, but it is not always possible and advisable to overlook the presence of the latter. The measure of kurtosis is a summarized numerical characteristic of the narrowness of the distribution peak, which must account for the values of the variable in the centre and at the tails of the distribution simultaneously with frequencies. Such is the measure based on moments. It is especially important to link the

measure and its cognitive sense with the behavior of the units of the respective population and the causes underlying the specific distribution. It is obvious that because of the different logic behind the construction of individual measures, they would have different values for the same distribution. But the most important question is what information is conveyed by each measure, how it will be interpreted and how it can be used in analytical activities.

In parallel with theoretical quest for cognitive sense and the most appropriate measure of kurtosis, it has been used in practice, in the solution of problems of different type. A number of researchers use the presence or absence of kurtosis together with skewness as an indicator of deviation or coincidence with the normal distribution. Other researchers use kurtosis (usually together with skewness) to solve problems in the field of medical science, engineering sciences, finance, etc.

Each statistical characteristic has certain cognitive sense and its significance should not be overstated. This is fully valid for skewness and kurtosis coefficients. They are related to the shape of the distribution and are meaningful mainly with one-peak distributions.

In reality, there are cases when distributions are not only with positive or negative kurtosis, but are also asymmetric. This issue is almost not developed in the literature. Each situation requires special attention to reveal the causes underlying such distribution.

Researchers pay attention mainly to the construction of measures of skewness and kurtosis and their application. There is insufficient work on the issue of interpretation and cognitive sense of the respective measures. The most exhaustive work on these issues is done in the field of finance, biology, and some technical sciences, which explains what has been said in a

part of the previous statement. In economics, as a rule, higher moments are unduly overlooked.

Literature

1. Borodachev, N. A., (1950), Main questions for the theory of preciseness of production. Moscow.
2. Borodachev, N. A., (1946), Justifying the methodology of calculation of assumptions and mistakes in sized cinematic chains, Moscow.
3. Byrne, P., and S. Lee, (1997), "Real estate portfolio analysis under conditions of non-normality: The case of NCREIF", *Journal of Real Estate Portfolio Management*.
4. Chissom, Brad S., (1970), "Interpretation of the Kurtosis Statistics", *The American Statistician*, Vol. 34, No. 4, pp. 19-22.
5. Conrad, J., R. Dittmar, and E. Ghysels, (2007), "Skewness and the Bubble", Preliminary Report.
6. Cox, D. R., D. Oakes (1988). *Analysis of survival data*, Moscow.
7. Darlington, R. B., (1970), "Is Kurtosis Really Peakedness?" *The American Statistician*, 24, pp. 19-22.
8. Dyson, F. J., (1943), "A Note on Kurtosis" *Journal of the Royal Statistical Society*, 106, 4, pp. 360-361.
9. Edith Seier- and Bonett, D. G., (2003), "Two Families of Kurtosis Measures", *Metrika*, 58, 1, pp. 59-70.
10. Finucan, H. M., (1964), "A Note on Kurtosis," *Journal of the Royal Statistical Society. Series B*, 26, 1, pp. 111-112.
11. Gatev, K., (1986). *General Theory of Statistics*, Sofia.
12. Geary, R. C., (1936), "Moments of the ratio of the mean deviation to the standard deviation for normal samples". *Biometrika*, 28, pp. 295-307.
13. Groeneveld, R. A., and G. Meeden, (1984), "Measuring Skewness and Kurtosis", *The Statistician*, 33, pp. 391-399.
14. Hampel, F. R., (1968), "Contributions to the theory of robust estimation. Ph.D. dissertation."
15. Hildebrand, David K., (1971), "Kurtosis Measures Bimodality?", *The American Statistician*, 25, No. 1, pp. 42-43.
16. Hogg, R. V., (1974), "Adaptive robust procedures: A partial review and some suggestions for future applications and theory", *J. Am. Stat. Assoc.*, 69, No. 348, pp. 909-927.
17. Hosking, J. R. M., (1986), *The theory of probability weighted moments*, Research Rep. RC 12210, p. 160, IBM Research Division, Yorktown Heights, NY.
18. Johnson, N. I., and S. Kotz, (1985), *Encyclopedia of Statistical Sciences*. Vol. 5, New York: John Wiley & Sons, Inc.
19. Kaloyanov, T., (1998). On the cognitive sense of skewness and kurtosis coefficients, *Alternatives*, 10/11.
20. Kenney, J. F., and E. S. Keeping, (1962), "Kurtosis" §7.12 in *Mathematics of Statistics*, Pt. 1, 3rd ed. Princeton, NJ: Van Nostrand, pp. 102-103.
21. MacGillivray, H. L., K. P. Balanda, (1988), "The relationships between skewness and kurtosis", *Austral. J. Statist.* 30 (3), pp. 319-337.

22. Mansfield, E., (1987). *Statistics for Business and Economics. Methods and Applications*. Third ed. W.W. Norton & Company, Inc. New York.
23. Markovitz, H. M., (1952), "Portfolio Selection", *The Journal of Finance*, 7.
24. Markovitz, H. M., (1959), "Portfolio Selection: Efficient Diversification of Investments", New York, John Wiley & Sons.
25. Moors, J. J. A., (1988), "A Quantile Alternative for Kurtosis." *The Statistician*, 37, 1, pp. 25-32.
26. Pareto, V., (1987). *Cours d'Economie Politique*. Rouge and Cie, Lausanne and Paris.
27. Roy, A. D., (1952), "Safety first and the holding of assets", *Econometrica*, 20, pp. 431-49.
28. Ruppert, D. (1987), "What is Kurtosis? An Influence Function Approach." *Amer. Statist.* 41, pp. 1-5.
29. Stefanov, I., A. Y. Totev, (1960). *Theory of statistics*, Sofia.
30. Student, (July 1927), "Errors of Routine Analysis" *Biometrika*, 19, pp. 151-164.
31. Tobin, J., (1958), "Liquidity preference as behavior towards risk", *Review of Economic Studies*, February, pp. 65-86.
32. Tobin, J., (1965), "The Theory of Portfolio selection" in F. H. Hahn and F. P. R. Brechling (eds.), *The Theory of Interest Rates*, London, Macmillan.
33. Tsiang, S. I., (1972), "The Rationale of the Mean-Standard Deviation Analysis, Skewness Preferences, and the Demand for Money," *The American Economic Review*, 62, pp. 354-371.
34. Tsonev, V. (1971). *Foundations of the representative study*, Sofia.
35. Van Zwet, W. R., (1964), "Convex Transformations of Random Variables". *Mathematics Centre Tract 7*,: Mathematisch Centrum, Amsterdam.
36. Venezkii, I. G., V. I. Venezkaia, (1979), *Main mathematical-statistical terms and formulas in economic analysis*, Moscow.
37. Zvi Bodie, Alex Kane, Alan Marcus. (2000), *Investments*, Naturela, S. **VIA**

Bulgarian Nonprofit Sector in the Light of the Structural-Operational Definition¹

Chief Assist. Prof. Denitsa Gorchilova, Ph.D.

*Department of Public Administration
and Regional Development, UNWE
e-mail: gorchilova@yahoo.com*

JEL: L31.

Summary: In the last 20 years serious efforts are made worldwide to clarify the scope, structure, finances, and activities of the nonprofit sector. A satellite account on the nonprofit institutions in the System of National Accounts is developed and tested in over 30 countries for that purpose. The account is based on the so called “structural-operational definition”, which serves as a point of reference when deciding whether to include or exclude different types of organizations in the nonprofit sector.

The Bulgarian nonprofit sector was resurrected quickly in the years of democratization but still is rather unknown. Its image is built on media coverage (which is frequently more on the negative side), on fragmentary surveys, and the statistics does not account for the full size of its role in the economic and social development. The aim of that text is to test the applicability of the structural-operational definition in Bulgaria and to check if there are the precedent conditions for the country to join the international efforts to specify the statistical image of the nonprofit sector.

Key words: structural-operational definition, Bulgarian nonprofit sector, associations, foundations, chitalishte (community centers).

Introduction

Despite the growing global presence of the civil society structures and the mounting interest from social researchers, politicians, and statisticians, the size and scope of nonprofit activities still remain almost invisible. Even in countries with long tradition in the statistical portrayal of the nonprofit presence and contribution, one rarely can find comprehensive and internationally comparative data for that type of organizations.

The **object** analyzed in that text are Bulgarian nonprofit organizations and the **subject** – the possibility to depict the whole variety they constitute in a statistically verified methodology. The **aim** of the analysis is to check whether is it possible to include the Bulgarian nonprofit organizations in the last years’ global efforts to represent them adequately in the System of National Accounts by creating a specialized satellite account. The research has several **tasks**: 1) to outline the faults in the present nonprofit contribution reporting by the national statistics; 2) to present in detail the structural-operational definition; 3) to retrospect historically the Bulgarian nonprofit sector development; and 4) to analyze the applicability of the structural-operational definition to the Bulgarian practice.

¹ The structural-operational definition is a specific instrument that tests whether any given organization can be accepted as a part of the nonprofit sector. The definition is developed by a team of the Johns Hopkins University Center for Civil Society Studies.