The use of entropy in risk measurement on the WGPW

Adrianna Mastalerz-Kodzis[1]

Abstract

In the literature covering the fields of financial mathematics, portfolio theory or econometrics there are a great number of different methods of measurement of risk of exchange securities. The article describes an unconventional method of risk measurement – entropy. It shows that risk measurement carried out with the use of entropy gives results comparable to those reached while using classic methods.

Key words

Stochastic process, entropy, Hurst exponent, risk.

1 Introduction

Stochastic analysis is used in mathematical models, which can find use in finance, in models with continuous - time. It is assumed that, the dynamics of the process of the price change of financial instruments (stocks, stock exchange indexes and exchange rates) are described with stochastic differential equations Itô. Stochastic processes are solutions of equations. Their trajectories approximating real time series, that is diagrams of price change of financial instruments in time.

The main reasons for dealing with this issue is an intention of getting in-depth understanding of dynamically changing behaviour of financial instruments market and putting forward a proposal of unconventional methods of risk measurement for the purpose of conducting analyses. The objective of the article is to define entropy as risk measurement method and to compare this method with classic ones.

The article consists of four parts. The first part is the introduction. In the next chapter modelling of financial series with stochastic processes is shortly discussed. The third chapter deals with entropy, its definition, properties, interpretations and examples of calculations. The last chapter is of empirical character. On the basis of selected WGPW (Warsaw Stock Market) items the risks were assessed and compared. Then finally, conclusions were drawn after conducted analyses.

2 Stock exchange series treated as realisations of stochastic processes

In the literature, prices of financial instruments are modelled with stochastic processes [see 1,5]. These processes are defined by calculus of probability using random variable. Prices are random variables dependent on parameter; most often time is the parameter – t.

There is a given calculus of probability (Ω, Σ, P). Let us assume that T is a subset of set of real numbers and $t \in T$.

[1] Dr Adrianna Mastalerz-Kodzis, Faculty of Mathematics, University of Economics in Katowice adamast@ae.katowice.pl

Stochastic process X(t) is a family of random variables $\{X_t(\omega), t \in T, \omega \in \Omega\}$ dependent on the parameter t and specified within calculus of probability (Ω, Σ, P).

For an unspecified infinitesimal event $\omega \in \Omega$ stochastic process X(t) is the function of variable t determined for all $t \in T$.

Stochastic process $X(t), t \in T$ can be called **a stationary process** when for every n, for every subset $\{t_1, t_2, \dots, t_n : t_1 < t_2 < \dots < t_n\} \subset T$ and for every h for which $t_i + h \in T$ $(i = 1,2,\dots,n)$ the following equation is true:

$$F_{t_1,t_2,\dots,t_n}(x_1, x_2, \dots, x_n) = F_{t_1+h,t_2+h,\dots,t_n+h}(x_1, x_2, \dots, x_n). \tag{1}$$

3 Definition of entropy and its interpretation, examples of calculating entropy

The term **"entropy"** is met in many fields of science, e.g. thermodynamics, probability mathematics, theory of information, theory of dynamical systems (which includes theory of stochastic processes, ergodic theory, topological dynamics).

The first to use the term entropy was in 1854 a German physicist Rudolf Clausius dealing with thermodynamics. However, in this article we focus on the entropy with its roots in theory of information. This theory was developed more than fifty years ago and was put forward in a book titled "Mathematical theory of communication". The book consisted of the articles by Claud Elwood Shannon published previously in "Bell System Technical Journal".

Entropy understood in terms of theory of information, as introduced by C.E. Shannon, is measure of indeterminacy, chaotic behaviour, degree of disorder. It is measure of indeterminacy of experiment (test), whose result is not explicit. Entropy is also called function of probabilities of results of experiments. Entropy of random variable is characterised by uncertainty (randomness) of a priori results (before experiment). In cybernetics entropy is measure of how chaotic a set is. The more probable are states of the set the bigger the measure is. When one of the states is statistically distinguished the measure gets smaller. In physics it characterises condition of a set of material substances. The more probable condition is the bigger entropy.

In literature one can come across topological entropy for dynamical systems. R.L. Adler, A.G. Konheim and M.H. McAndrew introduced it in 1965. Entropy is then numerical parameter of dynamical system and characterises the speed of "mixing" of points in the process of conversion.

3.1 Definition of entropy

Below there is going to be shown the Shannon definition of entropy of random variable for discrete variable as well as continuous [see 3, 4, 5].

Entropy of discrete random variable X is H(X) defined as follows:

$$H(X) = -\sum_{i=1}^{n} p_i \log_a p_i, \tag{2}$$

where x_i has probability p_i.

The logarithm basis can be any positive number. For the purpose of the analyses presented below a = 2. The bigger H(X) is the greater risk. For discrete variable entropy is positive. Entropy H(X) = 0 means that distribution is constant (only one random variable is possible), so there is no indeterminacy and risk. The bigger H(X) is the greater the risk gets.

It is also possible to come up with an interpretation of entropy, which takes into account the probability of occurring random events. Entropy is measure of indeterminacy, so the less probable a given event is the greater entropy, the greater risk. The more probable a given event (more frequent occurrence of a given random variable) is the smaller entropy gets, the smaller risk. In the common understanding of the problem entropy is ascending function of probability of occurring of an event.

Entropy of continuous random variable X is H(X) defined by the following formula:

$$H(X) = -\int_{-\infty}^{+\infty} f(x)\log_2 f(x)dx, \tag{3}$$

where f(x) is the function of density of distribution.

From the above formula one can conclude that for random variable with normal distribution entropy is calculated like this:

$$H(X) = \frac{1}{2}\log_2\left(2\pi e\sigma^2\right), \tag{4}$$

while for logarithmic – normal distribution like this:

$$H(X) = \frac{1}{2\ln 2}(2R+1) + \frac{1}{2}\log_2\left(2\pi\sigma^2\right), \tag{5}$$

where R is the average random variable, and σ^2 is variance of the variable.

In the case of variable with constant distribution within closed range <a, b> entropy is dependent only on the length of the range and is described by the following formula:

$$H(X) = \log_2(b-a) \tag{6}$$

Entropy of continuous random variable can be positive as well as negative. The bigger H(X) is the greater the risk gets.

3.2 Examples of calculating entropy for discrete and continuous random variables

- Discrete random variable X assumes value x with the probability of 1. Then:
- $H(X) = -1 \cdot \log_2 1 = 0$
- Discrete random variable X assumes value x_1 with the probability of ½ and x_2 also with the probability of ½. Then:
- $H(X) = -\left(\frac{1}{2}\log_2\frac{1}{2} + \frac{1}{2}\log_2\frac{1}{2}\right) = -\left(\log_2\frac{1}{2}\right) = \log_2\left(\frac{1}{2}\right)^{-1} = \log_2 2 = 1$
- Continuous random variable X has a constant distribution within interval <5, 9>. Then its density function looks like this:

$$f(x) = \begin{cases} 0 & \text{dla} \quad x \le 5 \\ 1/4 & \text{dla} \quad x \in (5,9] \\ 0 & \text{dla} \quad x > 9 \end{cases}$$

Then:

$$H(X) = -\int_{-\infty}^{+\infty} f(x)\log_2 f(x)dx = -\left[0 + \int_5^9 (1/4)\log_2(1/4)dx + 0\right] = -\int_5^9 (1/4)\log_2(1/4)dx =$$

$$= \left[-1/4 \cdot \log_2(1/4) \cdot x\right]_5^9 = -1/4 \cdot \log_2(1/4)(9-5) = -\log_2(1/4) = \log_2(1/4)^{-1} = \log_2 4 =$$

$$\overset{\text{formula 6}}{=} \log_2(9-5) = 2$$

4 Comparison of selected measures of risk on the basis of price quotations of stocks and WGPW indexes

4.1 Empirical analysis of selected stock exchange indexes

The quotations of selected indexes of the Warsaw Stock Market (WGPW) in he period 31.12.1997 – 06.02.2004 were used in the research.

In succession calculated were:

- The expected rate of return of index R,
- Variance of rates of return σ^2,
- Entropy H(X) of the index with the assumption that the distribution of the rates of return is normal,
- H(X) of the index with the assumption that the distribution of the rates of return is logarithmic – normal.

The results of the calculations are shown in Table 1. Indexes are put according to ascending variance.

Table 1. Indexes according to ascending variance.

INDEX	**Expected rate of return R**	**Variance σ^2**	**H(X) normal distribution**	**H(X) logarithmic – normal distribution**
WIRR	9,6653E-05	0,000187134	-4,144725613	-4,144586172
MIDWIG	0,00036574	0,000196882	-4,108094618	-4,107566969
WIG FOOD PRODUCTS	0,00028424	0,000242407	-3,958044463	-3,957634389
WIG - PL	0,00041331	0,000269132	-3,882602472	-3,882006194
WIG	0,00041568	0,000269155	-3,882541164	-3,881941467
WIG CONSTRUCTION	0,00012227	0,000283018	-3,846313076	-3,846136672
WIG BANKS	0,00075874	0,000302571	-3,798123169	-3,797028534
WIG 20	0,000274187	0,00040812	-3,582264204	-3,581868635
WIG TELEKOMUNICATIONS	9,9631E-06	0,000688211	-3,205335113	-3,205320739
NIF	-7,1359E-05	0,000708299	-3,184581258	-3,184684207

The following conclusions can be drawn from the research done for the period specified above:

- The order of indexes according to ascending variance is the same as the order in the case of entropy with normal distribution,
- (it is the consequence of formula 4, entropy with the assumption of normal distribution is monotonic function of variance),
- The order of indexes according to ascending variance (in this example) is also the same in the case of entropy for logarithmic – normal distribution;
- (in spite of the fact that in the case of this distribution entropy is dependent on both expected rate of return and variance of rate of return (see formula 5)).

4.2 Comparison of measures of risk of selected stocks quoted on WGPW

Quotations of selected stocks on WGPW in the period 16.04.1991 – 08.06.2005 were used in the research.

Four measures of risk mentioned in example 4.1 and also fractal measure in the form of Hurst exponent were calculated [see 2]. Hurst exponent deals with the length of the memory in the series. Its value is always in the range (0, 1). The greater the value the longer the memory is, the lesser risk of change of value of the series. In the case of most stock exchange series this exponent is within the range (0.45; 0.65). For the sake of calculations was used the method of graduated range described e.g. in [2]. The results of calculating are shown in table 2. The stocks are set according to ascending variance.

Table 2. Stocks according to ascending variance.

	KROSNO	ŚLĄSKA FABRYKA KABLI	PRÓCHNIK	TONSIL
Expected rate of return R	0,001632	0,001179	0,000655	0,000546
Variance σ^2	0,001382	0,001806	0,002708	0,004439
H(X) - normal distribution	-2,70243	-2,50959	-2,21709	-1,86074
H(X) - logarithmic – normal distribution	-2,70007	-2,50789	-2,21615	-1,85996
H – Hurst exponent	0,6242	0,6328	0,6185	0,6070

The following conclusions can be drawn from the calculations in table 2:

- For the selected stocks quoted on WPGW the order according to ascending variance is the same as in the case of ascending entropy both for normal distribution and logarithmic – normal distribution of rate of return;
- It is the consequence of the selecting for the purpose of research very long series of data (more than three thousand). In the case of long series distribution of rate of return approximates normal distribution,
- Correlation between the value of entropy and Hurst exponent is negative, which is understandable, in the light of the fact that the greater entropy the higher degree of uncertainty (greater risk); the smaller Hurst exponent the greater risk. Bearing the above in mind it is obvious that the least risky stock among those researched is KROSNO. The most risky one is TONSIL,
- In the case when distributions of empirical rates of return approximate normal distribution the risk measured with variance and entropy (of normal distribution) will introduce the same order in the group of researched securities.

Hurst exponent is calculated in this case for the whole researched period of time. It is assumed then that in the researched period the series was stationary (its properties were not subject to change in time). Empirical studies prove, however, that stock exchange series are not stationary, so it is necessary to study their local properties and not these of global nature. Moreover, literature often described studies of distributions of empirical rates of return and it shows that from practical point of view they are far from being normal distribution.

Not in all cases the order according to ascending variance is just like the order according to ascending entropy and decreasing Hurst exponent. Such a case is presented in example 4.3.

4.3 Non-stationary series of WPGW stocks

The global Hurst exponent, calculated above, gives information concerning the behaviour of series in the whole researched period of time. What about its stability in time? The question is relevant because when different periods of time (for example following years) are analysed, the exponent changes its value (table 3). The analysis was begun starting with the year 1995 so as to get series of data of comparable length.

Table 3. Hurst exponent in following years for the selected companies.

	KROSNO	***PRÓCHNIK***	***ŚLĄSKA FABRYKA KABLI***	***TONSIL***
1995	*0,5268*	*0,4527*	*0,4915*	*0,4567*
1996	*0,4994*	*0,3736*	*0,4430*	*0,5423*
1997	*0,5321*	*0,4615*	*0,6323*	*0, 4132*
1998	*0,5670*	*0,2972*	*0,7057*	*0,6007*
1999	*0,5740*	*0,5807*	*0,6350*	*0,6445*

Hurst exponent of the series of exchange quotation changes in time. So it would be advisable to treat it as a local parameter rather than a global one. On the basis of a comparison, of empirical and theoretical values of the exponent for different lengths of time periods it is possible to show a long-term dependence of data.

In the same way the value of entropy changes. The results of calculations are in table 4.

Table 4. Changes of measures of risk in time.

		KROSNO	**PRÓCHNIK**	**ŚLĄSKA FABRYKA KABLI**	**TONSIL**
1995	**Expected rate of return R**	0,003142732	-0,001979811	0,001471539	-0,000375874
	Variance σ^2	0,00151332	0,001498681	0,001689823	0,001739262
	H(X) normal distribution	-2,636938002	-2,643949717	-2,557360616	-2,536558927
	H(X) logarithmic – normal distribution	-2,632403998	-2,646805981	-2,555237634	-2,537101199
1996	**Expected rate of return R**	0,002042001	-0,000847612	-0,000410689	0,000987557
	Variance σ^2	0,000468137	0,000875964	0,00039033	0,000822648
	H(X) normal distribution	-3,483295153	-3,031324815	-3,614413717	-3,076623006
	H(X) logarithmic – normal distribution	-3,480349167	-3,032547661	-3,615006217	-3,075198263
1997	**Expected rate of return R**	-0,003103267	0,003052794	0,002552028	-0,000852089
	Variance σ^2	0,000981608	0,00199327	0,001185633	0,001004305
	H(X) normal distribution	-2,949187019	-2,43822781	-2,812967622	-2,932698115
	H(X) logarithmic – normal distribution	-2,953664087	-2,43382356	-2,809285823	-2,933927419
1998	**Expected rate of return R**	-0,000799476	-0,001918049	-0,003750932	-0,003119846
	Variance σ^2	0,001337369	0,000952266	0,001870104	0,001830788
	H(X) normal distribution	-2,726097611	-2,971078394	-2,484237174	-2,499564204
	H(X) logarithmic – normal distribution	-2,72725101	-2,973845554	-2,489648625	-2,50406519

1999	**Expected rate of return R**	0,003267256	-0,00246793	0,000662172	-0,000222316
	Variance σ^2	0,000692958	0,001290531	0,000579887	0,001002907
	H(X) normal distribution	-3,20037619	-2,751814259	-3,328875298	-2,933702513
	H(X) logarithmic – normal distribution	-3,195662536	-2,755374729	-3,327919986	-2,934023248

Measuring risk with variances gives in the case of an analysis in time the same results as entropy. However, it is worth noticing that in the following years, the extend of risk changes. So does the order. For example in 1998 Próchnik was the most secure company while Śląska Fabryka Kabli was the most risky. In the next year 1999 the least risky company was Śląska Fabryka Kabli and the one most risky was Próchnik. The stock exchange series of these two companies are not stationary.

It can also happen that the order according to entropy for normal distribution will not be the same as in the case of logarithmic – normal distribution. This situation may occur when short series without normal distributions are analysed. Below there are examples of calculations done for the analysed companies in the period 2.11.1999 – 29.12.1999.

Table 5 shows that the order differs when entropies with different distributions are studied.

Table 5. Placing in order according to ascending entropy of normal distribution.

	KROSNO	**ŚLĄSKA FABRYKA KABLI**	**TONSIL**	**PRÓCHNIK**
H(X) normal distribution	-2,981813692	-2,84149601	-2,787433953	-2,761278436
H(X) logarithmic – normal distribution	-2,974756399	-2,840957673	-2,764834912	-2,78719735

Reference

[1] KIJIMA M.: *Stochastic processes witch applications to finance*. Chapman @ Hall/CRC, 2003

[2] MANDELBROT B. B.: *Fractals and Scaling in Finance. Discontinuity, Concentration, Risk*. Springer-Verlag, New York,1997.

[3] MICHALSKA E.: *Entropia jako miara ryzyka decyzji inwestycyjnych. Modelowanie Preferencji a Ryzyko '99*, Wydawnictwo Akademii Ekonomicznej w Katowicach, Katowice 1999, str.259-270

[4] MYNARSKI S.: *Elementy teorii systemów i informacj*i. AE, Kraków.

[5] SOBCZYK K.: *Stochastyczne równania różniczkowe*. Wydawnictwa Naukowo-Techniczne, Warszawa 1996.

Summary

The use of entropy in risk measurement on the WGPW

In the article selected unconventional measures of risk were discussed and then compared with classic measures. Measures introduced some order in a set of stock exchange securities. However, studying series in relation to time proved that measures of risk are also susceptible to change. It is worth analysing exchange quotations in not too long periods of time or using methods of non-stationary analysis.

Measures of risk described in the article, and not taken into account in classic exchange analyses, can be useful in any confrontation with those known in literatures concerning

measures of risk. They are also useful in the process of investment decisions and they can facilitate the process of risk management.