PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1133 — 2006

Inwestycje finansowe i ubezpieczenia
– tendencje światowe a polski rynek

Krzysztof Jajuga

Wrocław University of Economics

APPLICATION OF EXTREME VALUE ANALYSIS IN PORTFOLIO ANALYSIS

1. Introduction

Portfolio analysis is one of the most important areas in modern finance. One the one hand, the birth of portfolio theory was one of the very few milestones in the scientific history of finance. On the other hand, portfolio theory is the main background of risk analysis and management. There are at least several approaches proposed in portfolio theory. The classical approach is well-known Markowitz theory, generalized by James Tobin to include risk-free instruments. Of course, the non-classical approaches are very often applied as well. Among the most often used non-classical approaches it is worth to mention the "safety-first" approaches, were the main stress is put on the avoiding large losses in a portfolio. The good description of these approaches is given in [1].

In this paper we propose the other possible ways to introduce "safety-first" concept in portfolio analysis. These proposals are based on Extreme Value Theory, being relative new statistical tool. The conceptual considerations are preceded by a synthetic presentation of Extreme Value Theory. In addition, some illustrative examples are given.

2. Extreme value theory – the most important facts

There are many different approaches that were developed in statistical methods. Classical statistical methods can be described as "mean-oriented" methods, aiming at the modeling of mean (expected value) of the statistical distribution. However, financial risk analysis is often interested in rare events, which lead to exceptionally

high losses. In this case, the suitable approach is the one, which aims at modeling the tails of the distribution – these models are concentrated on the analysis of "extreme" values. This approach is called Extreme Value Theory. It is briefly presented in this chapter. We discuss two cases:
- univariate extreme value theory;
- multivariate extreme value theory.

2.1. Univariate extreme value theory

Univariate extreme value theory consists of two types of strongly linked approaches:
- analysis of distribution of extremum;
- analysis of conditional excess distribution.

Analysis of distribution of extremum

Analysis of distribution of extremum is the analysis of the random variable, defined as maximum (or minimum) of the set of random variables. For simplicity, let us concentrate on the distribution of maximum (the distribution of minimum can be obtained in the obvious way from the distribution of maximum). The most important result for the distribution of maximum is given in the so-called Fisher-Tippet theorem (see [1]). In this theorem one considers the limiting distribution for the normalized (by location and scale parameter) maximum:

$$\lim_{n\to\infty} P\left(\frac{X_{n:n}-b_n}{a_n}\le x\right)=G(x), \tag{1}$$

$$X_{n:n}=\max(X_1,X_2,\ldots,X_n). \tag{2}$$

It can be proved that this limiting distribution belongs to the family of the so-called Generalized Extreme Value Distributions (GEVD), where a distribution function is given as:

$$G(x)=\exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \tag{3}$$

$$1+\xi\sigma^{-1}(x-\mu)>0.$$

Generalized Extreme Value Distribution has three parameters:
- location parameter, denoted by μ;
- scale parameter, denoted by σ;
- shape parameter, denoted by ξ.

Shape parameter reflects the fatness of tails of GEVD distribution – the higher value of this parameter, the fatter tails.

The family of Generalized Extreme Value Distributions contains three types of distributions; they differ with respect to the shape parameter, namely:
- Fréchet distribution – positive value of shape parameter;
- Weibull distribution – negative value of shape parameter;

– Gumbel distribution – the value of shape parameter converges to zero.

In financial problems one usually encounters Fréchet distribution as limiting distribution for maximum. Here the underlying observations come from the fat-tailed distribution, such as Pareto distribution, stable distribution (including Cauchy), etc.

Analysis of conditional excess distribution

Here we consider the so-called conditional excess distribution (also called conditional tail distribution), given as:

$$F_u(y) = P(X - u \le y | X > u) = \frac{F(u+y) - F(u)}{1 - F(u)},$$
$$0 \le y < x_0 - u; x_0 = \sup_x(x : F(x) < 1). \tag{4}$$

This distribution is given that the underlying random variable takes value from the tail. It depends on the choice of threshold u.

It can be proved (see [1]) that the conditional excess distribution can be approximated by the so-called Generalized Pareto Distribution (GPD), linked to Generalized Extreme Value Distribution. The distribution function of GPD is given as:

$$F_u(y) = 1 - (1 + \xi y / \beta)^{-1/\xi}, \tag{5}$$
$$\beta = \sigma + \xi(u - \mu). \tag{6}$$

This distribution has two parameters, namely:

- shape parameter, denoted by ξ (the same as in GEVD distributions);
- generalized parameter, denoted by β.

This generalized parameter depends on all three parameters given in GEVD distribution, as well as on the threshold u.

The family of Generalized Pareto Distributions contains three types of distributions; they differ with respect to the shape parameter. These are:

- Pareto distribution – positive value of shape parameter;
- Pareto type II distribution – negative value of shape parameter;
- Exponential distribution – the value of shape parameter converges to zero.

Therefore, in finance, when the underlying distribution has fat tails, we have Pareto distribution characterizing behavior in the tail.

2.2. Multivariate extreme value theory

In the multivariate case, we can apply two different approaches:

- analysis of distribution of vector of maxima;
- tail dependence analysis.

Both approaches are based on the so-called copula analysis, to be presented firstly.

Copula analysis

This approach has a crucial importance in multivariate analysis. Its main practical use comes from the fact, that it performs the analysis of multivariate distributions by separate analysis of marginal univariate distributions and the analysis of dependence between components of the random vector. Therefore the analysis of dependence is "separated" from the analysis of marginal distributions.

One of the most important facts in copula analysis is the so-called Sklar theorem [5], in which the multivariate distribution function is represented as a copula function linking the univariate distribution functions, given in the following formula (m is the number of variables):

$$H(x_1,...,x_m)=C(F_1(x_1),...,F_m(x_m)). \tag{7}$$

Here copula function reflects the dependence between the components of a random vector. There are very many possible copula functions. The popular family contains the so-called Archimedean copulas, defined on the base of strictly decreasing and convex function, called generator. In the bivariate case it is given as:

$$C(u_1,u_2)=\psi^{-1}(\psi(u_1)+\psi(u_2)), \tag{8}$$

$$\psi:[0;1]\to[0;\infty), \tag{9}$$

$$\psi(1)=0. \tag{10}$$

Copula functions are used in two approaches presented above.

Analysis of distribution of vector of maxima

This is a direct generalization of univariate case. Here we consider a vector of maxima, each maximum taken for single variable. The main result is obtained for the limiting distribution of normalized maxima, given as:

$$\lim_{n\to\infty} P\left(\frac{X_{n:n}^1-b_n^1}{a_n^1}\le x^1,...,\frac{X_{n:n}^m-b_n^m}{a_n^m}\le x^m\right)=G(x^1,...,x^m). \tag{11}$$

It can be proved (e.g. [3]) that this limiting distribution can be presented in the following form:

$$G(x^1,...,x^m)=C(G_1(x^1),...,G_m(x^m)). \tag{12}$$

Here, the multivariate distribution of maxima, called Multivariate Extreme Value Distribution (MEVD), is represented according to Sklar theorem, and:

- univariate distributions belong to the family of Generalized Extreme Value Distributions, therefore they are Fréchet, Weibull or Gumbel distributions;
- the copula function denoted by, is the so-called Extreme Value Copula (EVC), to be presented below.

As one can see, Multivariate Extreme Value Distribution is obtained by applying Extreme Value Copula to univariate Generalized Extreme Value Distributions.

Extreme Value Copula is the copula satisfying the following relation:

$$C(u_1{}^t,...,u_m{}^t)=C^t(u_1,...,u_m), \quad t>0. \tag{13}$$

In the bivariate case Extreme Value Copula can be represented in the following form:

$$C(u_1,u_2)=\exp\left\{\log(u_1u_2)A\left(\frac{\log u_1}{\log(u_1u_2)}\right)\right\}. \tag{14}$$

Here A is a convex function satisfying the following relations:

$$A(0)=A(1)=1, \tag{15}$$

$$\max(w,1-w)\le A(w)\le 1. \tag{16}$$

Three often applied bivariate Extreme Value Copulas are:

1. Gumbel copula, where:

$$C(u_1,u_2)=\exp[-(\log u_1{}^{\theta}+\log u_2{}^{\theta})^{1/\theta}], \tag{17}$$

$$A(w)=[w^{\theta}+(1-w)^{\theta})]^{1/\theta}, \quad \theta\in[1;\infty). \tag{18}$$

2. Gumbel II copula, where:

$$C(u_1,u_2)=u_1u_2\exp[\theta(\log u_1\log u_2)/(\log u_1+\log u_2)], \tag{19}$$

$$A(w)=\theta w^2-\theta w+1, \quad \theta\in[0;1]. \tag{20}$$

3. Galambos copula, where:

$$C(u_1,u_2)=u_1u_2\exp[((\log u_1)^{-\theta}+(\log u_2)^{-\theta})^{-1/\theta}], \tag{21}$$

$$A(w)=1-[w^{-\theta}+(1-w)^{-\theta}]^{-1/\theta}, \quad \theta\in[0;\infty). \tag{22}$$

The presented copulas are one parameter functions. This parameter can be interpreted as dependence parameter.

Tail dependence analysis

The main goal of tail dependence coefficients is to measure the dependence in tails of distribution. These coefficients, defined in the bivariate case, are:

- lower tail dependence coefficient:

$$\lambda_L=\lim_{u\to 0}P(X_2\le G_2{}^{-1}(u)\mid X_1\le G_1{}^{-1}(u)), \tag{23}$$

- upper tail dependence coefficient:

$$\lambda_U=\lim_{u\to 1}P(X_2> G_2{}^{-1}(u)\mid X_1> G_1{}^{-1}(u)). \tag{24}$$

The main idea behind tail dependence coefficient is based on the calculation of the probability that one variable takes value from the tail (lower or upper) given that the other variable takes value from the tail (lower or upper). The value in the tail is defined as a (lower or upper) quantile. This probability as taken as limiting probability given one goes with the probability in the tail to 0 (lower tail dependence) or to 1 (upper tail dependence).

We should notice that the tail dependence coefficients can be treated as the generalization of the approach based on the conditional excess distribution.

Both tail dependence coefficients can take values from the interval [0;1]. If tail dependence coefficient is equal to 0, we call this asymptotic independence. If tail dependence coefficient is higher than 0, we call this asymptotic dependence.

The important property of tail dependence coefficients is that they can be represented through copula functions. This is given in the following formulas:

$$\lambda_L = \lim_{u \to 0}[C(u,u)/u], \tag{25}$$

$$\lambda_U = \lim_{u \to 1}[(1-2u+C(u,u))/(1-u)]. \tag{26}$$

The bivariate approach can be generalized to multivariate case, by looking at the probability that some variables take value from the tail given the other variables has taken value from the tail.

3. Applications of Extreme Value Theory in portfolio analysis – some proposals

Now we give some proposals how the presented approaches of Extreme Value Theory can be applied in portfolio analysis. These proposals will be illustrated by some empirical examples. These examples were presented for the first time in Jajuga and Papla [4].

The proposals will be divided into two parts:
- applications of univariate extreme value theory;
- applications of multivariate extreme value theory.

Applications of univariate extreme value theory

In the univariate case, one can deal with two possible structures in portfolio analysis:
- distribution of the return of single stock;
- distribution of the return of a portfolio of stocks – here return of a portfolio is simply weighted average of returns of the individual stocks, where the weights are fractions of these stocks in a portfolio.

In each these two structures one can apply two main tools of univariate extreme value theory. In the analysis of the distribution of maximum, one looks at the distribution of maximal loss on a single stock or on the portfolio of stocks in the respective time period. Here loss is taken as negative return, understood as negative rate of return. Since the distribution of single stock return (same for loss) and sometimes also the distribution of portfolio return (same for loss) has fat tails, it seems that Fréchet distribution is likely distribution for maximal loss.

In the analysis of conditional excess distribution, one looks at distribution of loss on a single stock or on a portfolio of stocks, given that this loss exceeds some threshold. The threshold should be provided by investor, for example: if it set equal to 0, then investor looks at the distribution of loss (on single stock or portfolio of stocks), given that loss occurs. Of course, instead of 0, any other number, for example required rate of return, can be set.

Analysis of conditional excess distribution leads to the calculation of risk measure called Expected Shortfall (other names: Conditional Value at Risk, Expected Tail Loss). It is defined as:

$$ES = E(X - u | X > u). \tag{27}$$

Expected Shortfall is the expected value of the conditional excess distribution.

Example

In this example the logarithmic rate of returns for the following stock market indices are considered:

- indices of Warsaw Stock Exchange: WIG, WIG20, MIDWIG, TechWIG;
- indices of US market: DJIA, S&P 500;
- indices of EU market: DAX, FT-SE100.

The financial time series of logarithmic rates of return come from the period January 2, 1995-October 3, 2003. The stationarity of each time series was verified (positively) by the augmented Dickey-Fuller test.

For each time series Generalized Extreme Value Distributions were estimated by using maximum likelihood method. The results of the estimation are given in the table 1.

Table 1. The estimates of the parameters of Generalized Extreme Value Distributions

Data	Shape parameter	Location parameter	Scale parameter
WIG	0.374	0.040	0.012
WIG20	0.450	0.037	0.022
MIDWIG	0.604	0.033	0.011
TechWIG	0.147	0.066	0.012
DJIA	0.519	0.027	0.006
S&P 500	0.244	0.027	0.007
FT-SE 100	-0.048	0.031	0.006
DAX	-0.084	0.041	0.011

Source: own calculations.

The analysis of the results in the table 1 leads to the following conclusions:

- in most cases we obtained Fréchet distribution (estimate of shape parameter is positive), which suggests that underlying returns are characterized by fat-tailed distribution;
- for FTSE-100 and DAX the estimate of shape parameter is negative but close to zero, which suggests either Weibull distribution or Gumbel distribution;
- in most cases, the Warsaw Stock Exchange indices have fatter tails than the other indices;
- the Warsaw Stock Exchange indices have larger estimates of location (related to expected return) and larger estimates of scale parameter (related to risk).

Applications of multivariate extreme value theory

In multivariate case, one considers vector of returns on many (or several) stocks. Similarly, one can consider vector of returns on portfolios of stocks (market indices). One can here apply both presented tools of multivariate extreme value theory. In the analysis of distribution of vector of maxima, one looks at the distribution of maximal losses of all considered stocks, where loss is taken as negative return, understood as negative rate of return. Here the limiting distributions for the components are univariate Generalized Extreme Value Distributions and the relation between the maxima is reflected through Extreme Value Copula.

In tail dependence distribution, performed for bivariate case, one looks at the probability (in limit) that one stock faces large losses, given that the other stock faces large losses. It is very often useful to study this for market indices rather than for individual stocks. This allows detecting the so-called contagion effect.

Example

We consider different pairs of stock market indices, where each pair is composed of WIG index and another index, studied in the previous example. In the first step we estimated the parameters of Generalized Extreme Value Distributions. In the second step we used empirical distribution functions obtained in the first step and estimated three copulas belonging to Extreme Value Copulas family: Gumbel, Gumbel II and Galambos. The best fit was obtained for Gumbel copula. The results are presented in the table 2.

Table 2. The estimates of Gumbel copula for stock indices

Bivariate data	Value of dependence parameter
WIG and WIG20	21.345
WIG and DJIA	14.862
WIG and FTSE-100	2.275
WIG and DAX	5.562

Source: own calculations.

As we know, the dependence parameter for Gumbel copula takes values higher or equal to 1. The higher value of this parameter, the stronger dependence between maximal losses of respective variables. The results given in this table indicate the strong dependence (it could have been expected) between stock indices of Warsaw Stock Exchange. It also shows stronger extreme dependence between Warsaw Stock Exchange and New York Stock Exchange than between Warsaw Stock Exchange and two large European exchanges.

Then for each pair we calculated the approximation of tail dependence by taking level of u (probability) close to 0 (for lower tail) or close to 1 (for upper tail). We present the results in the case of Frank copula, belonging to the family of Archimedean copulas. Due to the symmetry of Frank copula, we present the results for upper tail dependence. They are given in the table 3.

Table 3. Upper tail dependence for stock indices (Frank copula)

Probability	WIG and WIG20	WIG and DJIA	WIG and FTSE-100	WIG and DAX
0.9	0.546	0.200	0.195	0.198
0.95	0.368	0.147	0.106	0.108
0.99	0.103	0.086	0.023	0.023
0.999	0.011	0.072	0.002	0.002

Source: own calculations.

As we can see, one can suspect that only in the case of two pairs, WIG and WIG20, as well as WIG and DJIA, there are some indications of small asymptotic tail dependence.

These illustrative examples shows some usefulness of Extreme Value Theory in portfolio analysis.

Literature

[1] Elton E.J., Gruber M.J., Brown S.J., Goetzmann W.N., *Modern Portfolio Theory and Investment Analysis*, Wiley, New York 2003.

[2] Embrechts P., Klüppelberg C., Mikosch T., *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, Berlin 1997.

[3] Galambos J., *The Asymptotic Theory of Extreme Order Statistics*, Kreiger Publishing, Melbourne 1978.

[4] Jajuga K., Papla D., *Extreme Value Analysis and Copula*, [w:] P. Cizek, W. Hardle, R. Weron (red.), *Statistical Tools for Finance and Insurance*, Springer, Berlin 2005, s. 45-64.

[5] Sklar A., *Fonctions de repartition à n dimensions et leurs marges*, Publications de l'Institut de Statistique de l'Université de Paris, 8, Paris 1959, s. 229-231.

ZASTOSOWANIE ANALIZY WARTOŚCI EKSTREMALNYCH W ANALIZIE PORTFELOWEJ

Streszczenie

W artykule podano kilka propozycji zastosowania teorii wartości ekstremalnych w analizie portfelowej. Rozpatrzono dwa przypadki: przypadek jednowymiarowy i przypadek wielowymiarowy. Artykuł rozpoczyna syntetyczna prezentacja teorii wartości ekstremalnych, a w dalszej części zaprezentowano zastosowania praktyczne ilustrowane przykładami z rynku finansowego.