EQUILIBRIUM
1 (2) 2009
ISSN 1689-765X

Marcin Fałdziński

Szacowanie prawdopodobnej maksymalnej straty przy zastosowaniu teorii wartości ekstremalnych na przykładzie indeksów giełdowych

Słowa kluczowe: *teoria wartości ekstremalnych, stopy zwrotu, prawdopodobna maksymalna strata, ryzyko giełdowe*

Abstrakt: Prawdopodobna maksymalna strata (*Probable Maximum Loss*, PML) jest miarą wywodzącą się z rynku ubezpieczeń, gdzie stosuje się ją do portfeli ubezpieczeniowych. Twórca miary PML Wilkinson [1982] użył ją do estymacji klasycznych metod statystycznych. Natomiast dobrze znana reguła 20–80 mówi, że 20% roszczeń jest odpowiedzialna za więcej niż 80% całej sumy odszkodowań w dobrze zdefiniowanym portfelu. W związku z tym, że właśnie ekstremalne wydarzenia powodują zdecydowaną większość całej sumy odszkodowań, zdecydowano się na zastosowanie teorii wartości ekstremalnych. Celem artykułu jest oszacowanie prawdopodobnej maksymalnej straty dla indeksów giełdowych. Stopy zwrotu jak ogólnie wiadomo cechują się występowaniem wartości ekstremalnych. Podstawowym zadaniem analizy jest określenie i sprawdzenie użyteczności miary PML dla indeksów giełdowych przy użyciu teorii wartości ekstremalnych.

Wprowadzenie

Prawdopodobna maksymalna strata (*Probable Maximum Loss*, PML) jest miarą wywodzącą się z rynku ubezpieczeń, gdzie stosuje się ją do portfeli ubezpieczeniowych. Jest ona związana ze znaną regułą 20–80, która głosi, że w dobrze zdefiniowanym portfelu 20% roszczeń jest odpowiedzialnych za więcej niż 80% całej sumy odszkodowań. Celem artykułu jest oszacowanie prawdopodobnej maksymalnej straty dla indeksów giełdowych, traktowanych jako portfele papierów wartościowych. Szeregi stóp zwrotu, jak ogólnie wiadomo, cechują się występowaniem wartości ekstremalnych.

Teoria wartości ekstremalnych

Ostatnimi laty w finansach, ubezpieczeniach, informatyce i w innych dziedzinach coraz częściej do estymacji wysokich kwantyli stosuje się teorię wartości ekstremalnych. Centralne Twierdzenie Graniczne jest podstawą dla klasycznej teorii statystyki, natomiast fundamentem teorii wartości ekstremalnych jest twierdzenie graniczne dla maksimów Fishera i Tippetta[1] z 1928 roku. Konsekwencją twierdzenia Fishera i Tippetta jest uzyskanie uogólnionego rozkładu wartości ekstremalnych (*Generalized Extreme Value Distribution*, GEV) zdefiniowanego poniżej.

Definicja 1. Uogólniony rozkład wartości ekstremalnych (*Generalized Extreme Value Distribution*, GEV)[2]

Dystrybuanta rozkładu GEV, czyli H_γ jest definiowana jako:

$$H_\gamma(x) = \begin{cases} \exp\left[-(1+\gamma x)^{-1/\gamma}\right] & \gamma \neq 0 \\ \exp\left[-\exp(-x)\right] & \gamma = 0 \end{cases} \tag{1}$$

gdzie $1+\gamma x > 0$.

Często dystrybuanta rozkładu GEV przedstawiana jest w poniższej postaci:

$$H_\psi(x) = H_{\gamma,\mu,\sigma}(x) = \begin{cases} \exp\left\{-\left(1+\gamma\,\dfrac{x-\mu}{\sigma}\right)^{-1/\gamma}\right\} & \gamma \neq 0 \\ \exp\left\{-\exp\left(\dfrac{x-\mu}{\sigma}\right)\right\} & \gamma = 0 \end{cases} \tag{2}$$

gdzie $\mu \in \mathrm{R}$ $(\mu = d_n)$ jest parametrem położenia, natomiast $\sigma > 0$ $(\sigma = c_n)$ jest parametrem skali. Parametr γ standardowo w statystyce nazywany jest parametrem kształtu, natomiast w teorii wartości ekstremalnych ten parametr nazywany jest indeksem wartości ekstremalnych (*Extreme Value Index,* EVI). Parametr EVI ma bardzo istotną z punktu widzenia zastosowań w finansach interpretację: im wyższa wartość bezwzględna tego parametru, tym grubsze ogony rozkładu.

[1] R. A. Fisher, L. H. C. Tippet, *Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample*, Proc. Cambridge Phil. Soc. 24 (2), 1928, s. 163–190.

[2] P. Embrechts, C. Klüppelberg, T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin 2003.

W dalszej części tej pracy pośrednio będziemy korzystać z metody Peaks over Threshold, dlatego też zostanie przedstawiony krótki schemat tej metody[3].

Wybieramy wartość progową u dla danego szeregu zmiennych $X_1,\ldots,X_n$ (i.i.d.) pochodzących z nieznanej dystrybuanty F. Niech N_u będzie liczbą obserwacji przekraczających u ($X_{i_1},\ldots,X_{i_{N_u}}$), czyli $Y_j = X_{i_j} - u \geq 0$. Dopasowujemy rozkład $G_{\gamma,\sigma}$ do przekroczeń $Y_1,\ldots,Y_{N_u}$, aby otrzymać oszacowanie parametrów γ i σ. Realizacje zmiennej losowej X mieszczą się w przedziale od 0 do u, stąd estymacja F w tym przedziale nie sprawia problemu. Postać rozkładu $G_{\gamma,\sigma}$ otrzymujemy dzięki drugiemu podstawowemu twierdzeniu w teorii wartości ekstremalnych.

Twierdzenie 1 Pickands-Balkemy-de Haana[4]

Dla rozkładów danych rzeczywistą dystrybuantą F warunkowy rozkład przekroczeń $F_u(y)$, dla dużej wartości u jest dobrze aproksymowany za pomocą $F_u(y) \approx G_{\gamma,\sigma}(y),\ u \to \infty$, gdzie

$$G_{\gamma,\sigma}(y) = \begin{cases} 1 - \left(1 + \dfrac{\gamma}{\sigma} y\right)^{-1/\gamma} & \text{jeżeli } \gamma \neq 0 \\ 1 - e^{-y/\sigma} & \text{jeżeli } \gamma = 0 \end{cases} \tag{3}$$

dla $y \in \left[0, (x_F - u)\right]$ jeżeli $\gamma \geq 0$ i $y \in \left[0, -\sigma/\gamma\right]$ jeżeli $\gamma < 0$, gdzie $G_{\gamma,\sigma}$ jest uogólnionym rozkładem Pareto.

Prawdopodobna maksymalna strata

Przydatną miarą oceny ryzyka nie tylko w ubezpieczeniach jest prawdopodobna maksymalna strata (Probable Maximum Loss, PML). Wilkinson[5] zaproponowała, aby ustalić PML równe $(1+\theta)E[M_n]$, albo $E[M_n] + \theta\sqrt{Var[M_n]}$, gdzie θ jest określonym współczynnikiem. Według Cebrian, Denuit i Lambert[6] PML można również otrzymać, rozwiązując równanie:

---

[3] J. A. McNeil, F. Frey, *Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach*, „Journal of Empirical Finance" 7 (2000), s. 271–300.

[4] J. Pickands, *Statistical Inference Using Extreme Order Statistics,* „Annals of Statistics" 2 (5) 1975, s. 119–131, A. A. Balkema, L. de Haan, *Residual Life Time at Great Age*, „Annals of Probability" 2 (5) 1974, s. 792–804.

[5] M.E. Wilkinson, *Estimating probable maximum loss with order statistics*, Proceedings of the Casualty Actuarial Society 1982, s. 195–209..

[6] A.C. Cebrian, M. Denuit, P. Lambert, *Generalized Pareto fit to the society of actuaries' large claims database*, „North American Actuarial J", 3/2003, s. 18–36..

$$P\left[M_n \leq PML_\varepsilon\right] = 1 - \varepsilon \tag{4}$$

dla małego $\varepsilon > 0$, gdzie M_n jest szeregiem maksimów. Takie sformułowanie oznacza, że PML jest wysokim kwantylem maksimów próby losowej długości n. Równanie (4) pokazuje jednocześnie związek między PML a Value-at-Risk. PML możemy traktować jako wartość zagrożoną, ale wyznaczoną jedynie dla samych ekstremów. W związku z tym, że M_n przekroczy PML tylko w ε przypadkach, jest mało prawdopodobne, że pojedyncza wartość będzie większa niż PML. W związku z tym otrzymujemy:

$$PML_\varepsilon = F_{M_n}^{-1}(1 - \varepsilon) \tag{5}$$

tak że PML jest obliczany jako $1-\varepsilon$ kwantyl rozkładu maksymalnych strat w określonym przedziale czasu. W dalszej części zostanie przedstawione podejście oparte na aproksymacji przekroczeń.

Jeśli N_u jest liczbą przekroczeń powyżej wartości progowej u i ciąg wartości progowych u_n spełnia warunek $\lim\limits_{n\to\infty} n\left(1 - F(u_n)\right) = \tau$ dla $k = 0,1,...$ to

$$\lim_{n\to\infty} P\left[N_u \leq k\right] = e^{-\tau} \sum_{s=0}^{k} \frac{\tau^s}{s!}.$$

Oznacza to, że pod warunkiem określonych wymagań liczba przekroczeń N_u ponad wartość progową u jest w przybliżeniu procesem Poissona. W takim przypadku można udowodnić, że rozkład maksimów M_n tych N_u przekroczeń może być określony za pomocą uogólnionego rozkładu wartości ekstremalnych $H(\gamma, \mu, \sigma)$, gdzie $\mu = \beta\gamma^{-1}\left(\lambda^\gamma - 1\right)$, $\sigma = \beta\lambda^\gamma$ i $\lambda = E\left[N\right]$. Używając tego rozkładu otrzymamy formułę na oszacowanie PML:

$$PML_\varepsilon = u + \left\{\left(-\frac{\lambda}{\ln(1-\varepsilon)}\right)^\gamma - 1\right\}\frac{\beta}{\gamma} \tag{6}$$

Standardowo przyjmuje się ε na poziomie 0,05 i 0,01. Autorzy przedstawionego tutaj podejścia pokazują, że prawdopodobna maksymalna strata liczona przy użyciu teorii wartości ekstremalnych okazała się zdecydowanie lepsza niż przy użyciu innych metod. Wynika to z tego, że teoria wartości ekstremalnych posiada narzędzia i podstawy teoretyczne do tego, aby mogła opisywać zachowania ekstremów. Jest to główne uzasadnienie, dlaczego EVT powinno być stosowane do problemów zarządzania ryzykiem w finansach czy ubezpieczeniach.

Analiza empiryczna

Podstawowym założeniem przyjętym przez autora tej pracy do analizy empirycznej, jest to że indeksy giełdowe możemy traktować jako jednoskładnikowe portfele. Ponadto przyjęto, że liczymy nie tylko prawdopodobną maksymalną stratę, ale również prawdopodobny maksymalny zysk. Polega to na tym, że w teorii wartości ekstremalnych rozważania na temat minimów są całkowicie równoważne tym na temat maksimów, dzięki spełnieniu następującej własności:

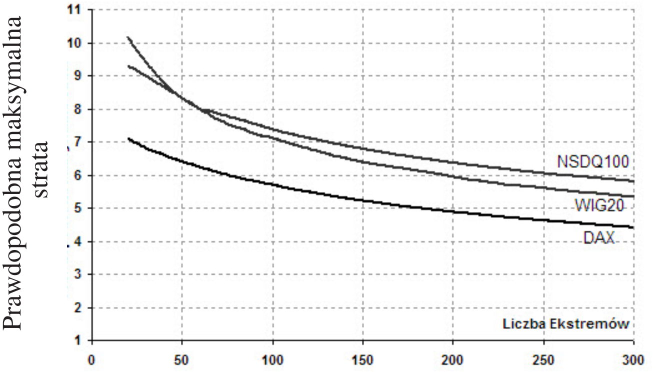
$$\min\left(X_1,\ldots,X_n\right) = -\max\left(-X_1,\ldots,-X_n\right). \tag{7}$$

Do badania użyto szeregów czasowych złożonych z 3000 obserwacji (dane dzienne: 07.11.1994–31.10.2006) zlogarytmowanych stóp zwrotu. Parametry szacowane były za pomocą metody największej wiarygodności. Rysunki 1 i 2 ukazują, jak kształtuje się prawdopodobna maksymalna strata (odpowiednio zysk) względem liczby ekstremów przyjętych do estymacji. Wraz ze wzrostem liczby ekstremów k branych do estymacji miara PML maleje. Takie kształty wykresów nie są żadnym zaskoczeniem, ponieważ wynikają z przedstawionych wcześniej założeń. W tabeli 1 i 2 pokazano oszacowania prawdopodobnej maksymalnej straty (odpowiednio zysku) indeksów giełdowych dla wybranych poziomów ekstremów. W ostatniej kolumnie możemy zobaczyć również minimalną (maksymalną) wartość danego indeksu giełdowego w całej próbie. Łatwo zauważyć, że często prawdopodobna maksymalna strata (odpowiednio zysk) jest mniejsza od wartości minimalnej (maksymalnej). Wynika to z założeń przyjętych do miary PML, dlatego też otrzymane oszacowania miary PML w tabeli 1 należy traktować jako średnią prawdopodobną maksymalną stratę dla określonej liczby ekstremów. Przykładowo zostanie zinterpretowany wynik z tabeli 1 indeksu WIG dla $k = 20$: Na poziomie $\varepsilon = 0{,}05$ średnia prawdopodobna maksymalna strata indeksu WIG jest równa 8,39 , przy założeniu 20 ekstremów.

Aby otrzymać teoretyczną prawdopodobną maksymalną stratę (odpowiednio zysk) należałoby skorzystać z asymptotycznego przedziału ufności (*asymptotic confidence interval*) przedstawionego w pracy McNeila[7]. Wartość miary PML zależy również od przyjętego ε , przy czym im mniejszy ε , tym większa jest miara PML. Aby prawdopodobna maksymalna strata miała jak najbardziej zbliżone wartości do tych rzeczywistych, to należy wybierać

[7] J. A. McNeil, *Calculating quantile risk measures for financial time series using extreme value theory*, Preprint, ETH, Zurych 1998.

Rysunek 1. Wykres liczby ekstremów k względem prawdopodobnej maksymalnej straty dla indeksów NSDQ100, DAX i WIG



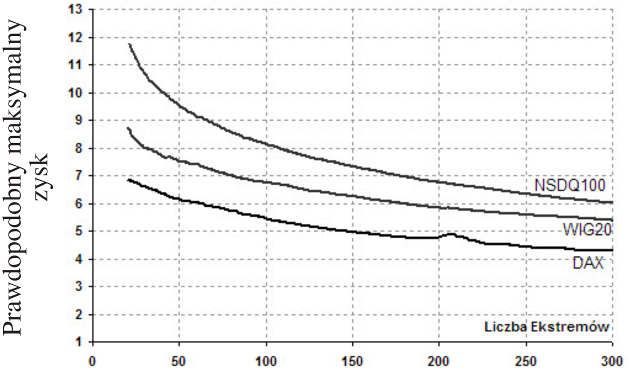
Źródło: opracowanie własne.

Tabela 1. Prawdopodobna maksymalna strata

Indeks	Prawdopodobna maksymalna strata				Minimalna wartość szeregu
	k=20	k=100	k=200	k=300	
WIG	8,39	6,04	5,06	4,54	10,28
WIG20	10,18	7,09	5,93	5,32	14,16
NSDQ100	9,30	7,36	6,35	5,79	10,37
DAX	7,53	5,69	4,87	4,41	8,87
STI	7,01	5,32	4,34	3,91	9,15
SSMI	6,19	4,82	4,06	4,00	7,33
KOSPI	9,89	7,39	6,35	5,76	12,8
HSI	9,50	6,45	5,32	4,72	14,74
ESTOXX	6,15	5,17	4,49	4,11	6,61
Indeks	Prawdopodobna maksymalna strata				Minimalna wartość szeregu
	k=20	k=100	k=200	k=300	
WIG	9,40	7,60	6,72	6,21	10,28
WIG20	11,25	8,71	7,59	6,93	14,16
NSDQ100	9,82	8,50	6,79	7,01	10,37
DAX	8,67	6,71	6,24	5,84	8,87
STI	8,01	7,32	6,80	6,59	9,15
SSMI	7,67	6,63	5,72	5,55	7,33
KOSPI	11,3	8,87	7,79	7,20	12,80
HSI	10,81	8,41	7,13	6,50	14,74
ESTOXX	6,78	5,98	5,70	6,08	6,61

Źródło: opracowanie własne.

Rysunek 2. Wykres liczby ekstremów k względem prawdopodobnego maksymalnego zysku dla indeksów NSDQ100, DAX i WIG



Źródło: opracowanie własne.

Tabela 2. Prawdopodobny maksymalny zysk

Indeks	Prawdopodobna maksymalna strata				Maksymalna wartość szeregu
	k=20	k=100	k=200	k=300	
WIG	6,97	5,82	5,01	4,53	7,89
WIG20	8,71	6,74	5,82	5,37	13,7
NSDQ100	11,75	8,12	6,74	6,00	17,20
DAX	6,84	5,42	4,77	4,29	7,55
STI	8,67	5,37	4,37	4,05	12,87
SSMI	6,42	5,86	4,3	3,5	7,46
KOSPI	7,75	6,52	5,48	5,38	8,16
HSI	10,62	6,49	5,15	4,57	17,25
ESTOXX	6,08	5,20	6,59	3,99	7,07
Indeks	Prawdopodobna maksymalna strata				Maksymalna wartość szeregu
	k=20	k=100	k=200	k=300	
WIG	7,40	6,79	6,77	6,45	7,89
WIG20	10,96	8,23	7,33	6,77	13,70
NSDQ100	13,16	9,86	8,49	7,73	17,20
DAX	7,03	7,28	7,95	7,21	7,55
STI	9,71	7,22	6,57	6,38	12,87
SSMI	7,20	6,45	5,64	5,32	7,46
KOSPI	9,35	6,98	5,86	5,65	8,16
HSI	12,51	8,53	7,28	6,54	17,25
ESTOXX	6,47	6,50	6,59	6,21	7,07

Źródło: opracowanie własne.

bardzo małą liczbę ekstremów. Pozostaje jeszcze problem testowania lub/i sprawdzania dokładności prawdopodobnej maksymalnej straty. Takie narzędzia aktualnie nie są jeszcze rozwinięte, ale pewnymi propozycjami mogłyby być: analiza historyczna ekstremów lub przedział ufności.

Zakończenie

Wydarzenia ekstremalne w finansach, których wpływ na cały rynek jest bardzo duży, są bardzo trudne do prognozowania na długi okres do przodu. W takich przypadkach korzystanie z teorii wartości ekstremalnych jest czymś naturalnym, a wręcz można powiedzieć, że wskazanym. Wykorzystanie prawdopodobnej maksymalnej straty jako miary średnich największych strat dla indeksów giełdowych jest poprawne i możliwe do zastosowania. Okazuje się, że prawdopodobna maksymalna strata może być użytecznym narzędziem do analizy ryzyka lub/i pozyskiwania informacji w celach diagnostycznych, ale należy być świadomym wad zarówno teorii wartości ekstremalnych (problem wyboru liczby ekstremów), jak i samej miary PML.

Literatura

Balkema A. A., de Haan L., *Residual Life Time at Great Age*, „Annals of Probability" 2 (5) 1974, s. 792–804.

Beirlant J., Matthys G., *Extreme quantile estimation for heavy-tailed distributions*, August 2001, //www.gloriamundi.org/detailpopup.asp?ID=453055854, (01.02.2007).

Cebrian A. C., Denuit M., Lambert P., *Generalized Pareto fit to the society of actuaries' large claims database*, „North American Actuarial J." 3/2003, s. 18–36.

Embrechts P., Klüppelberg C., Mikosach T., *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin 2003.

Fisher R. A., Tippet L. H. C., *Limiting Forms fo the Frequency Distribution of the Largest or Smallest Member of a Sample*, Proc. Cambridge Phil. Soc., 24 (2) 1928, s. 163–190.

McNeil J. A., *Calculating quantile risk measures for financial time series using extreme value theory*, Preprint, ETH, Zurych 1998.

McNeil J. A., Frey F. *Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach*, „Journal of Empirical Finance„" 7/2000, s. 271–300.

Pickands J., *Statistical Inference Using Extreme Order Statistics*, „Annals of Statistics" 3 (1) 1975, s. 119–131.

Wilkinson M.E, *Estimating probable maximum loss with order statistics*, „Proceedings of the Casualty Actuarial Society" 1982, s. 195–209.

Estimation of the Probable Maximum Loss Based on Extreme Value Theory for Stock Returns

Summary

The probable maximum loss is a measure that originated on the insurance market, where it is applied for insurance portfolio analysis. This corresponds to the 20-80 rule, which states that in a well defined portfolio 20% of the individual claims is responsible for more than 80% of the total claim amount. The paper presented attempts to estimate the probable maximum loss for stock returns which are treated as portfolios of securities. The probable maximum loss proved to be a useful tool for risk analysis or/and other diagnostic purposes on capital markets, it should be noted; however, that it has its drawbacks as well.