

Katarzyna Brzozowska-Rup, Wiesław Dziubdziela
Wyższa Szkoła Handlowa im. Bolesława Markowskiego w Kielcach
e-mail: brzozows@poczta.fm, wiedziu@poczta.fm,

ESTYMACJA “INDEKSU OGONA” WYBRANYCH SZEREGÓW FINANSOWYCH ZA POMOCĄ ENTROPII RENYI’EGO.

Streszczenie: W pracy przedstawiono wstępne wyniki estymacji indeksu ogona rozkładów prawdopodobieństwa w wybranych szeregach finansowych. Zakładając rozkład Pareto, otrzymano za pomocą entropii Renyi’ego estymator indeksu ogona w przypadku stóp zwrotu indeksu. Analiza indeksu WIG wskazuje, że szereg czasowy jego zwrotów ma nieskończone momenty, natomiast w przypadku zwrotów indeksu DJ częstość pojawiania się dużych zmian jest znacznie mniejsza.

Słowa kluczowe: rozkłady z grubymi ogonami, rozkład Pareto, entropia Renyi’ego, estymacja jądrowa.

WSTĘP

Bogata empiryczna analiza stóp zwrotu finansowych szeregów czasowych pozwala stwierdzić, że dominujące do niedawna modele bazujące na rozkładzie normalnym, w wielu przypadkach są dużym nadużyciem. Większość szeregów finansowych wykazuje: tendencje do skupiania danych, długoterminową zależność, asymetrię (częściej występują zwroty dodatnie), a prawdopodobieństwo pojawienia się skrajnych wartości, czy bardzo dużych zmian jest wyraźnie większe niż w przypadku rozkładu Gaussa. Wykresy kwantyli (QQ plot) pokazują, że duże zmiany występują znacznie częściej niż przewiduje to rozkład normalny, natomiast mniej jest średnich zmian (wartości odchylających się od średniej od 0,5 do 2,5 odchyłeń standardowych). W związku z powyższym konieczne jest poszukiwanie jakościowo nowych modeli.

Na podstawie przeprowadzonych badań okazuje się, że w wielu zjawiskach wartości ekstremalne pojawiają się zgodne z rozkładami potęgowymi. Celem pracy jest wykorzystanie rozkładów z grubymi ogonami do analizy finansowych szeregów czasowych. Została podjęta próba estymacji ciężkości ogona rozkładów szeregu zwrotów: dziennych notowań indeksu WIG i DJ przy założeniu, że dane pochodzą z rozkładu Pareto.

W literaturze przedmiotu definiowanych jest wiele różnych estymatorów indeksu, między innymi estymator: Hilla, Pickandsa, Dekkersa-Einmahla-de Haana, niestety otrzymane na ich podstawie oszacowania pozostawiają wiele do życzenia. W artykule zaprezentujemy próbę oszacowania indeksu ogona za pomocą entropii Renyi’ego.

GRAFICZNA IDENTYFIKACJA ROZKŁADU POTĘGOWEGO.

Zmienna losowa X pochodzi z rozkładu potęgowego, jeżeli jej funkcja gęstości spełnia zależność:

$$\lim_{x \rightarrow \infty} \frac{f(x)}{L(x)x^{-\alpha-1}} = 1, \quad (2.1)$$

gdzie: $L(x)$ jest wolno zmieniającą się funkcją, tzn. $\forall t > 0 \lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1$,

a α jest stałą większą od zera, nazywaną indeksem ogona. Graficzna metoda wykrywania rozkładu potęgowego jest bezpośrednią konsekwencją własności funkcji przeżycia tego typu rozkładów. Zauważmy:

$$\bar{F}(x) = P(X > x) = 1 - F(x) = \int_x^{\infty} p(x)dx \approx \frac{L(x)}{\alpha} x^{-\alpha} \quad (2.2)$$

po obustronnym zlogarytmowaniu związku (2.2) otrzymujemy równanie prostej:

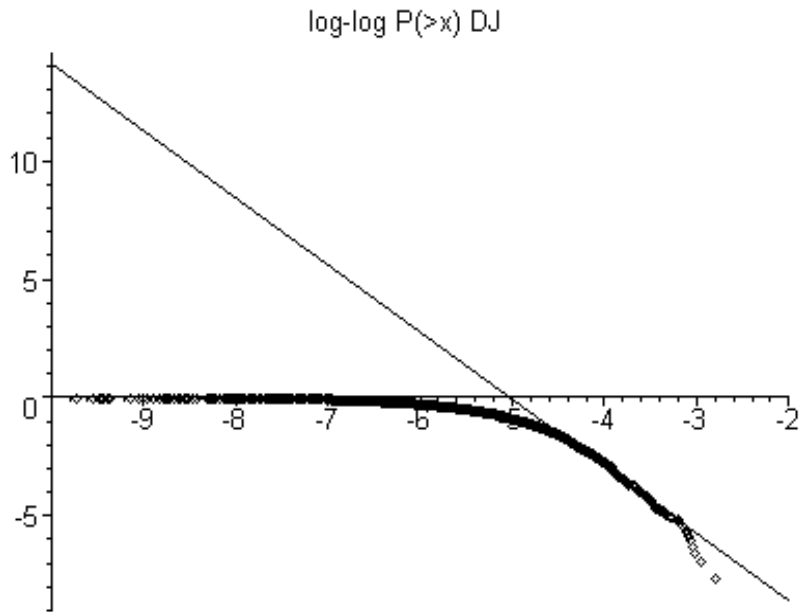
$$\ln(\bar{F}(x)) = -\alpha \ln x + \ln L(x) \quad (2.3)$$

Z powyższych rozważań wynika, że pierwszym krokiem badania szeregów pod kątem rozkładów z grubymi ogonami jest obserwacja empirycznej funkcji przeżycia w skali podwójnie logarytmicznej (Rysunek 1). Jeżeli dane pochodzą z rozkładu potęgowego, to wykres punktów $(x, y) = (\ln(x_{(i)}), \ln(\bar{F}(x_{(i)})))$ układa się wzdłuż prostej o współczynniku kierunkowym $(-\alpha)$. W celu wyznaczenia empirycznej funkcji przeżycia, należy uporządkować obserwacje od największej do najmniejszej $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$ (statystyki porządkowe), wtedy

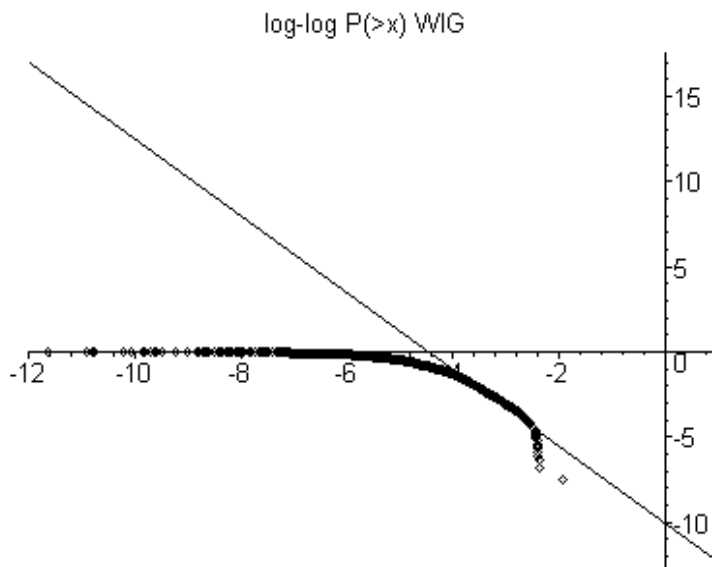
$$\bar{F}(x_{(i)}) = P(X \geq x_{(i)}) = \frac{i}{n} \quad (2.4)$$

Niestety metoda jest bardzo wrażliwa zarówno na wielkość próby, jak i na liczbę obserwacji uwzględnionych w modelu regresji. Z przeprowadzonych badań wynika, że właściwe zachowanie ogona widoczne jest dopiero dla bardzo dużych (co do wartości bezwzględnej) obserwacji. W praktyce oznacza to, że do otrzymania wiarygodnego estymatora ciężkości ogona, należy posłużyć się bardzo dużą liczbą obserwacji, wybierając tylko wartości skrajne.

Stosując przedstawioną metodę dla stóp zwrotu indeksu DJ otrzymaliśmy $\hat{\alpha} = -2,84$, natomiast w przypadku stóp zwrotu oszacowany parametr $\hat{\alpha} = -2,26$.



Rysunek 1. Funkcje przeżycia szeregów stóp zwrotu indeksów DJ w skali podwójnie logarytmicznej wraz z dopasowanymi funkcjami regresji dla wartości skrajnych.



Rysunek 2. Funkcje przeżycia szeregów stóp zwrotu indeksów WIG w skali podwójnie logarytmicznej wraz z dopasowanymi funkcjami regresji dla wartości skrajnych

ESTYMACJA FUNKCJI GĘSTOŚCI ROZKŁADU.

W teorii statystyki istnieje wiele sposobów estymacji nieznannej postaci funkcji gęstości, niestety część z nich nie może być użyta w przypadku rozkładów z grubymi ogonami, np. metoda krzywych Pearsona - zakłada istnienie czterech pierwszych momentów zwykłych i centralnych.

W naszych badaniach zastosujemy estymację funkcji gęstości za pomocą jądra. Wyboru jądra dokonujemy w taki sposób, aby otrzymany estymator gęstości uwzględniał istnienie grubszych ogonów niż w rozkładzie normalnym. Definiujemy jądra za pomocą pewnej modyfikacji (tak, aby spełnione były poniższe warunki) funkcji gęstości rozkładu Laplace'a i rozkładu Cauchy'ego.

Estymatorem jądrowym gęstości $f(x)$ nazywamy funkcję postaci:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (3.1)$$

gdzie $K(x)$ jest odwzorowaniem spełniającym następujące warunki:

$$1. \int_{\mathbb{R}} |K(x)|^2 dx < \infty,$$

$$2. \int_{\mathbb{R}} K(x) dx = 1,$$

$$3. K(x) = K(-x) \text{ dla } x \in (-\infty, \infty),$$

Okazuje się, że jakość estymacji jądrowej w większym stopniu zależy od wyboru szerokości pasma, h (bandwidth) niż od kształtu jądra. Zbyt małe wartości parametru h (gdy h dąży do zera) powodują, że szukany estymator gęstości jest nieregularny, „najeżony pikami”, z dużym prawdopodobieństwem zawiera wahania przypadkowe. Natomiast w przeciwnym przypadku, gdy szerokość pasma jest zbyt duża estymator jest co prawda gładki, ale może okazać się za bardzo wygładzony, a przez to tracić istotne własności poszukiwanego rozkładu. Szerokość pasma można traktować jako stopień/parametr wygładzania estymowanej funkcji gęstości, ale chociaż metodom jego wyboru poświęcono wiele prac, nadal nie jest to zadanie proste.

Jako wstępne kryterium wyboru szerokości pasma zastosowaliśmy metodę graficzną – wizualna ocena zgodności kształtu histogramu empirycznego i szacowanej gęstości.

RELACJA MIĘDZY ENTROPIĄ A INDEKSEM OGONA ROZKŁADU.

Pojęcie entropii występuje w wielu dziedzinach nauki, między innymi takich jak termodynamika czy teoria informacji - powoduje to różne sposoby

definiowania tej wielkości. Niemniej w każdym przypadku entropia jest pewną miarą niepewności – nieporządku systemu dynamicznego (przy czym nieporządek niekoniecznie musi oznaczać nieład przestrzenny struktury). Entropia jest funkcją częstotliwości występowania określonych konfiguracji.

W naszych badaniach posłużymy się definicją entropii zaproponowaną przez Renyi'ego, wówczas ogólny wzór na entropię rzędu β ma postać:

$$R_\beta(f) = \frac{1}{1-\beta} \ln \int_R f^\beta(x) dx, \quad \beta > 0, \beta \neq 1 \quad (4.1)$$

lub w przypadku dyskretnym:

$$\hat{R}_\beta(p) = \frac{1}{1-\beta} \ln \sum_{i=1}^k p_i^\beta \quad (4.2)$$

Wielkość k określa liczbę pojawiających się mikrostanów, β jest parametr skalowania (zooming parameter).

Celem pracy jest analiza ogonów rozkładów finansowych szeregów czasowych. Przypomnijmy, funkcja gęstości rozkładu Pareto wyraża się wzorem:

$$f(x) = \begin{cases} \alpha \theta^\alpha x^{-(\alpha+1)} & x > \theta \\ 0 & x \leq \theta \end{cases} \quad (4.3)$$

gdzie: $\theta > 0$, $\alpha > 0$ - jest miarą grubości ogona rozkładu. Grubość ogona określa prawdopodobieństwo pojawienia się obserwacji ekstremalnych, im mniejsze α tym grubszy ogon, a tym samym więcej jest takich wartości.

Podstawowe charakterystyki rozkładu Pareto mają postać:

Wartość oczekiwana:

$$E(X) = \frac{\alpha}{\alpha-1} \cdot \theta, \quad \text{dla } \alpha > 1 \quad (4.4)$$

Wariancja:

$$D^2(X) = \frac{\alpha}{(\alpha-1)(\alpha-2)} \cdot \theta^2, \quad \text{dla } \alpha > 2 \quad (4.5)$$

Z powyższych relacji otrzymujemy naturalną interpretację wykładnika α , a mianowicie określa maksymalny rząd momentu skończonego.

Do ustalenia wstępnych wniosków przyjmijmy $\beta = 2$ (quadratic entropy), wówczas entropia Renyi'ego, dla rozkładu Pareto wynosi :

$$R_2(f) = -\ln \int_{\theta}^{+\infty} \alpha^2 \theta^{2\alpha} x^{-2(\alpha+1)} dx = -\ln \frac{\alpha^2}{1+2\alpha} \theta^{-1} \quad (4.6)$$

\Updownarrow

$$\exp(-R_2(f)) = \frac{\alpha^2}{1+2\alpha} \theta^{-1} \quad (4.7)$$

Zdefiniujmy nową zmienną $T = \theta \exp(-R_2(f))$, wówczas równość (4.7) przekształca się w proste równanie kwadratowe:

$$\alpha^2 - (1 + 2\alpha)T = 0 \quad (4.8)$$

$$\Delta = 4T^2 + 4T$$

Zauważmy, że T przyjmuje wartości większe od 0, skąd wynika, że równanie (4.8) posiada dwa pierwiastki rzeczywiste, które określają związek między entropią a wykładnikiem α :

$$\alpha_1(T) = T - \sqrt{T^2 + T} \quad (4.9)$$

$$\alpha_2(T) = T + \sqrt{T^2 + T} \quad (4.10)$$

W definicji rozkładu Pareto wykładnik $\alpha > 0$, stąd jedynym rozwiązaniem jest pierwiastek $\alpha_2(T)$.

Powyższe obliczenia pozwalają określić estymator parametru α wzorem: $\hat{\alpha}(T) = T + \sqrt{T^2 + T}$. Teraz problem sprowadza się do oszacowania wartości entropii rzędu 2.

ESTYMACJA ENTROPII RENYI'EGO.

Wzór na entropię Renyi'ego rzędu β zmiennej losowej pochodzącej z rozkładu o funkcji gęstości f_X , oraz zauważmy, że można go wyrazić za pomocą wartości oczekiwanej zmiennej $f_X^{\beta-1}(X)$:

$$R_\beta(f) = \frac{1}{1-\beta} \ln \int_R f_X^\beta(x) dx = \frac{1}{1-\beta} \ln E_X[f_X^{\beta-1}(X)] \quad (5.1)$$

Jak wyżej rozważamy szczególnie przypadek, gdy parametr $\beta = 2$, wówczas:

$$R_2(f) = -\ln E_X[f_X(X)] \quad (5.2)$$

Wzór, z którego będziemy korzystać estymując entropię, otrzymujemy podstawiając w relacji (5.2) estymator jądrowy (3.1) funkcji f_X , stąd:

$$R_2(f) = -\ln E_X[f_X(X)] = -\ln \left[\frac{1}{n^2 h} \sum_{j=1}^n \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \right] \quad (5.3)$$

BADANIA EMPIRYCZNE

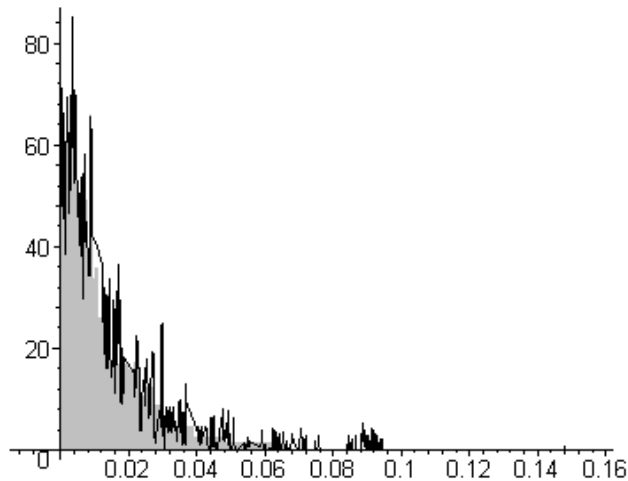
W naszych badaniach ograniczyliśmy się do szacowania ciężkości prawych ogonów rozkładów. Estymując funkcję gęstości za pomocą jądra, obliczenia powtórzyliśmy dla kilku szerokości pasma.

Poniższe tabele i wykresy przedstawiają wyniki naszej analizy przeprowadzonej dla:

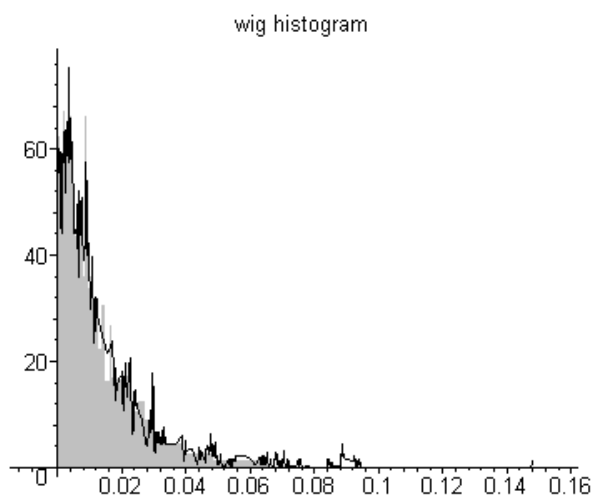
- a** Dziennych notowań indeksu DJ, kurs zamknięcia (okres obserwacji: 02.10.1990-0.3.03.2006, łącznie 4078 danych, 2128 dodatnich stóp zwrotu),
- b** Dziennych notowań indeksu WIG, kurs zamknięcia (okres obserwacji: 23.04.1991- 0.3.03.2006, 3271, 1728 dodatnich stóp zwrotu),

Zanim przystąpiliśmy do analizy danych rzeczywistych, dokonaliśmy estymacji funkcji gęstości dla szeregu wygenerowanego z rozkładu Pareto ze znanym parametrem α , generowaliśmy szereg $\{u_t\}_{t=1}^{1000}$, gdzie $U_t = u_{\min} (1-x)^{-1/\alpha}$, a zmienna losowa X jest zmienną o rozkładzie jednostajnym na przedziale $[0, 1)$. W przypadku, gdy $\alpha = 2$, zarówno z jądrem Laplace'a jak i Cauchy'ego wartość wyestymowanej grubości ogona pokrywa się z wartością rzeczywistą, o ile szerokość pasma nie jest zbyt duża.

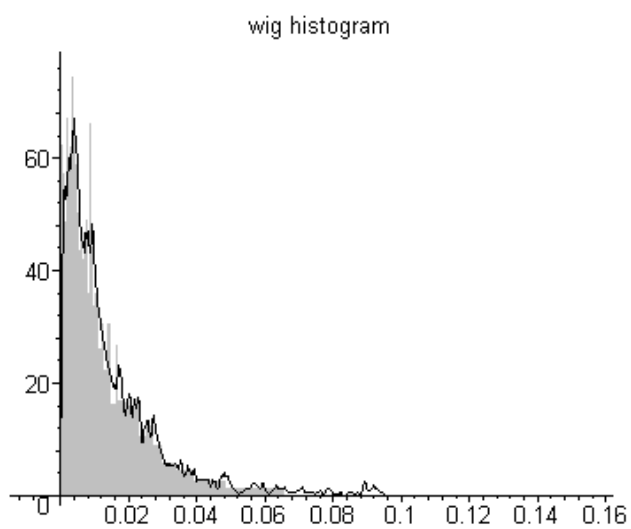
wig histogram



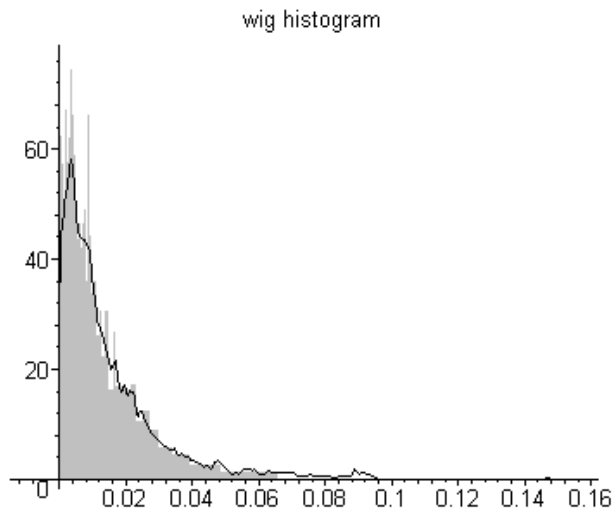
Rysunek 3 Estymacja jądrem Laplace'a szerokość pasma $2 \cdot h$



Rysunek 4. Estymacja jądrem Cauchy'ego szerokość pasma $2 \cdot h$

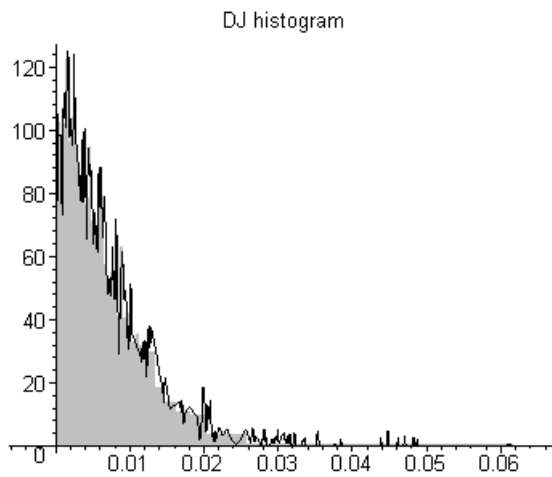


Rysunek 5. Estymacja jądrem Laplace'a szerokość pasma $10 \cdot h = 0,036$

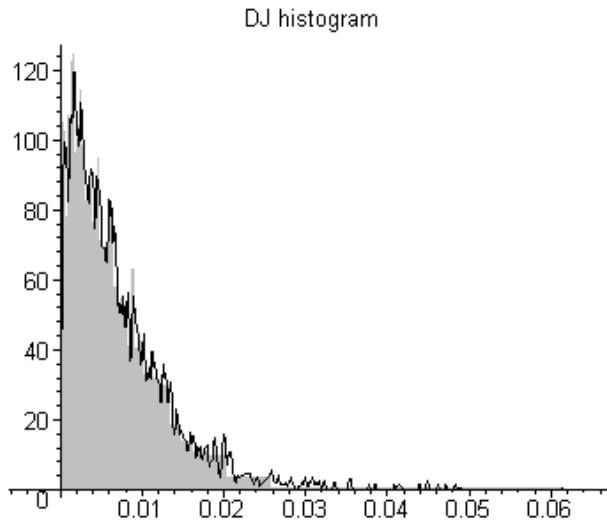


Rysunek 6. Estymacja jądrem Cauchy'ego szerokość pasma $10 \cdot h = 0,036$

Poniższe rysunki przedstawiają wybrane efekty estymacji funkcji gęstości szeregów zwrotów notowań indeksu DJ za pomocą jądra Laplace'a, $L_{(s/\sqrt{2}, 0)}$.



Rysunek 7. Estymacja jądrem Laplace'a szerokość pasma $12 \cdot h$



Rysunek 8. Estymacja jądrem Laplace'a szerokość pasma $20 \cdot h$

Wyniki prezentowanej procedury szacowania ciężkości ogonów analizowanych szeregów przedstawiają tabele poniżej:

Tabela 1. Wyniki analizy dla szeregu generowanego z rozkładu Pareto z parametrem $\alpha = 2$:

jądro/pasmo	$h/4$	$h/2$	$h=0,018$	$10h$
Laplace'a	2	2	2	1,08
Cauchy	2	2	2	1,65

Tabela 2. Wyniki analizy szeregu dodatnich stop zwrotu indeksu WIG (estymacja gęstości dla szeregu dodatnich stop zwrotu)

jądro/pasmo	$h/2$	$h=0,0036$	$2h$	$5 \cdot h$	$10 \cdot h$
Laplace'a	1,019	0,974	0,949	0,926	0,908
Cauchy	1,052	0,988	0,948	0,908	0,875

Tabela 3. Wyniki analizy szeregu dodatnich stop zwrotu indeksu DJ (estymacja gęstości tylko dla dodatnich stop zwrotu)

jądro/pasmo	$h/2$	$h=0,001$	$5 \cdot h$	$10h$	$12 \cdot h$	$20 \cdot h$
Laplace'a	3,076	2,808	2,384	2,529	2,521	2,494
Cauchy	3,004	2,88	2,234	2,494	2,481	2,477

WNIOSKI.

Rozkłady potęgowe często nazywane są rozkładami z grubymi/ciężkimi ogonami - im grubszy ogon, tym więcej wartości ekstremalnych.

W wyniku przeprowadzonych badań na szeregu stóp zwrotu indeksu WIG okazało się, że estymator indeksu ogona w większości przypadków nie przekracza jedynki, a to oznaczałoby, że szereg pochodzi z rozkładu, w którym nie istnieje żaden moment skończony. Proces stochastyczny o nieskończonej wariancji, choć jest dobrze określony matematycznie, jest niesłychanie trudny do zastosowania. W przypadku układów finansowych nieskończona wariancja wprowadziłaby ogromne komplikacje przy szacowaniu ryzyka.

W przypadku analizy stóp zwrotu indeksu DJ wyniki nie są tak zaskakujące, estymatory otrzymane metodą entropii przyjmują wartości powyżej dwójki i pokrywają się z wartością estymatora $\hat{\alpha}$ otrzymaną metodą graficzną.

Zaprezentowano wstępne wyniki estymacji indeksu ogona rozkładu w zamiarach autorów leży dogłębne, teoretyczne zbadanie własności zaproponowanego estymatora, jak również wykorzystanie innych sposobów estymacji entropii, potrzebnej do oszacowania szukanego indeksu.

LITERATURA

- Domański Cz., Pruska K. (2000) „Nieklasyczne metody statystyczne”, PWE, Warszawa,
 Ombach J. (1996) „Wykłady z równań różniczkowych”, WUJ, Kraków,
 Reimann S, „An Elementary Model of Price Dynamice In A Financial Market. Distribution, Multiscaling & Entropy”, physics/0602097.
 Weron A., Weron R. (1999) „Inżynieria finansowa”, WNT, Warszawa,

Use Renyi's entropy for estimating tail index in financial time series.

Summary: Empirical analysis of financial time series indicates that there is a great demand for qualitative new models. Paper presents preliminary results for estimating tail index applying Renyi entropy. Empirical results show that exchange rate of WIG's time series are generated from distribution with significantly fatter tails than DJ's time series.

Key words: heavy-tailed distribution, Pareto distribution, Renyi entropy, Kernel estimation.