Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
2013, t. 47, nr 2

Rafał Czyżycki
Uniwersytet Szczeciński

Estymacja jądrowa w modelowaniu rozkładu stopy zwrotu

Streszczenie, *Rozkład stopy zwrotu z inwestycji jest jednym z najczęściej analizowanych i wykorzystywanych pojęć na rynku kapitałowym. Często przy korzystaniu w różnego typu badaniach z funkcji gęstości stopy zwrotu przyjmuje się klasyczne założenie, że dana stopa zwrotu charakteryzuje się rozkładem normalnym. W artykule zaprezentowano możliwości wykorzystania do modelowania stóp zwrotu estymacji jądrowej wchodzącej w skład nieklasycznych metod estymacji.*

Słowa kluczowe: *stopa zwrotu, funkcja gęstości, metody nieklasyczne, estymacja jądrowa*

1. Wprowadzenie

Stopa zwrotu oprócz ryzyka jest jednym z najważniejszych pojęć w teorii i praktyce finansów. Jest ona podstawową determinantą m.in. przy podejmowaniu decyzji inwestycyjnych. Od poziomu prawdopodobieństwa osiągnięcia określonego zysku lub poniesienia ewentualnej straty inwestorzy najczęściej uzależniają decyzję nabycia lub zbycia określonych papierów wartościowych. Przy szacowaniu wielkości takiego prawdopodobieństwa, w praktyce bardzo często przyjmuje się założenie o normalności rozkładu stóp zwrotu. Założenie takie jest również często przyjmowane w różnego typu modelach opisujących funkcjonowanie rynku kapitałowego czy też mających wspomagać decyzje inwestycyjne. Występuje ono w modelu wyceny opcji Blacka-Scholsa, w modelu wyceny aktywów finansowych CAMP lub w przypadku wyznaczania wartości narażonej na ryzyko (VaR). Założenie o normalnym rozkładzie stóp zwrotu przyjmuje się przede wszystkim w celu przyspieszenia, uproszczenia i ułatwienia dokonywania określonych obliczeń. Jednak istotne odstępstwo rzeczywistych stóp zwrotu od przyjętych w założeniach może spowodować szereg negatywnych konsekwencji. Może m.in. być podstawą do zakwestionowania wiarygodności, a przez to aplikacyjności wielu technik, metod i modeli wykorzystywanych do analiz, diagnoz i prognoz rynku kapitałowego.

2. Badanie normalności stóp zwrotu

Stopę zwrotu najczęściej definiuje się jako arytmetyczną lub logarytmiczną stopę zwrotu. W przypadku analizowania arytmetycznej stopy zwrotu R_t jej wartość wyznacza się za pomocą formuły:

$$R_t = \frac{P_t - P_{t-1} + D_t}{P_{t-1}}, \qquad (1)$$

natomiast logarytmiczną stopę zwrotu wyznaczyć można jako:

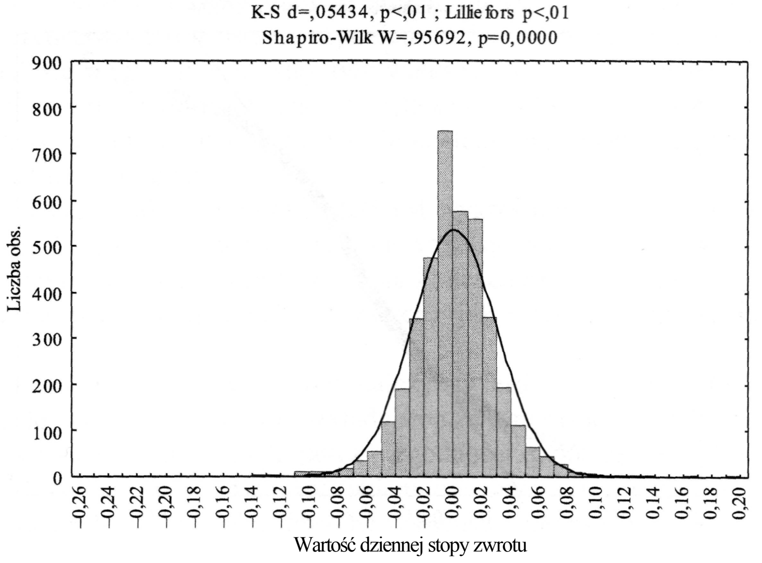
$$R_t^* = \ln \frac{P_t + D_t}{P_{t-1}}, \qquad (2)$$

gdzie:

P_t - cena papieru wartościowego w okresie *t*,
P_{t-1} - cena papieru wartościowego w okresie *f—l*,
D_t - wartość wypłaconej dywidendy w okresie *t*.

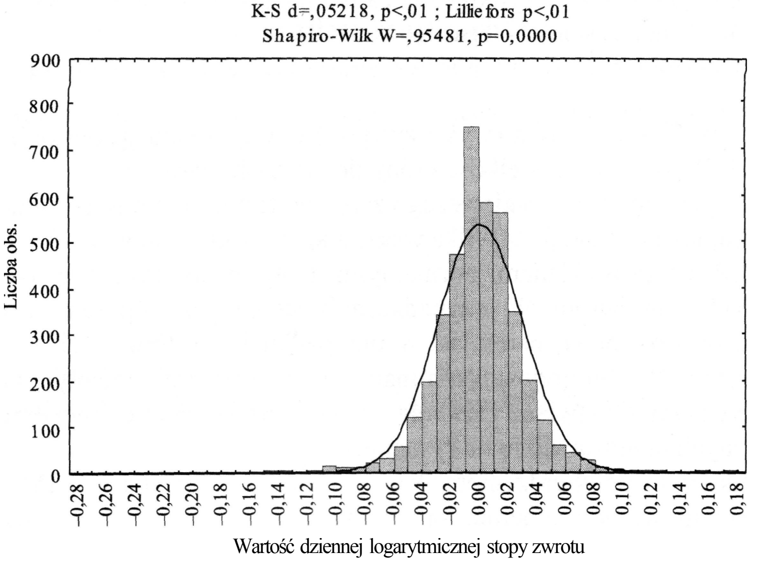
W zależności od przyjętego horyzontu czasowego można analizować dzienne, tygodniowe, miesięczne czy roczne stopy zwrotu. Na potrzeby artykułu analizę ograniczono tylko do dziennych stóp zwrotu. Ponadto ze względu na ograniczenia wydawnicze zaprezentowane wyniki dotyczą tylko spółki Kombinat Górniczo-Hutniczy Miedzi (KGHM), której wyboru dokonano przede wszystkim ze względu na dużą jej popularność wśród inwestorów na Giełdzie Papierów Wartościowych (GPW) w Warszawie. Badanie obejmuje okres od 11.07.1997 r. (drugie notowanie KGHM na giełdzie, czyli moment, dla którego można obliczyć pierwszy raz dzienną stopę zwrotu) do 30.04.2013 r., czyli w sumie 3961 notowań.

Badanie normalności stóp zwrotu polega na postawieniu dwóch hipotez: hipotezy zerowej, która mówi o zgodności dystrybuanty rozkładu danej stopy zwrotu z dystrybuantą rozkładu normalnego, oraz hipotezy alternatywnej do hipotezy zerowej zakładającej, że dystrybuantą rozkładu stopy zwrotu nie jest dystrybuantą rozkładu normalnego.



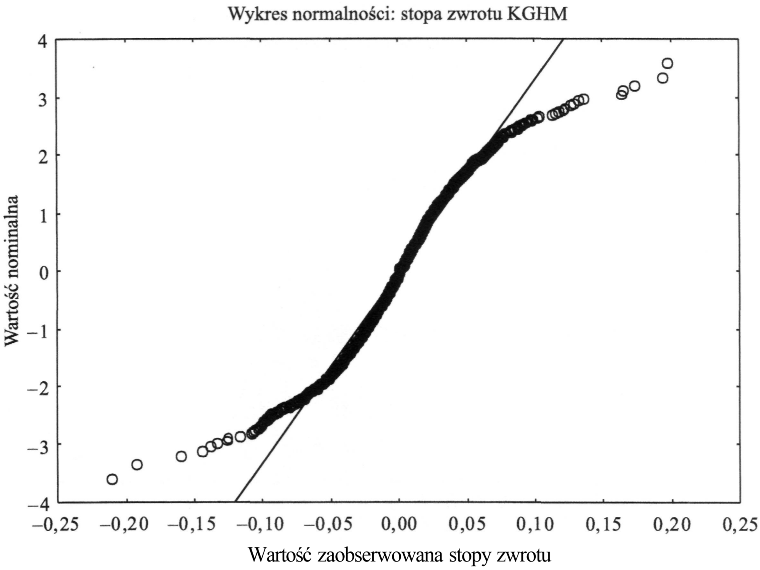
Rysunek 1. Histogram dziennej stopy zwrotu dla KGHM w okresie od 11.07.1997 roku do 30.04.2013 roku wraz z oszacowaną teoretyczną linią rozkładu normalnego

Źródło: opracowanie własne.



Rysunek 2. Histogram dziennej logarytmicznej stopy zwrotu dla KGHM w okresie od 11.07.1997 roku do 30.04.2013 roku wraz z oszacowaną teoretyczną linią rozkładu normalnego

Źródło: opracowanie własne.



Rysunek 3. Wykres normalności dla dziennej stopy zwrotu akcji KGHM
w okresie od 11.07.1997 roku do 30.04.2013 roku

Źródło: opracowanie własne.

Hipotezy powyższe weryfikuje się za pomocą odpowiednich testów zgodności, do których najczęściej zalicza się:

- test normalności Kołmogorowa-Smirnowa (K-S) z poprawką Lilleforsa oraz
- test W Shapiro-Wilka (wykorzystywany w tym celu pakiet obliczeniowy StatisticalOPL ogranicza wielkość próby do 5000 obserwacji).

Dla przyjętego okresu badawczego zarówno test normalności Kołmogorowa-Smirnowa, jak i test Shapiro-Wilka wskazują, że rozkłady dziennych stóp zwrotu dla KGHM nie są rozkładami normalnymi. Dotyczy to arytmetycznych i logarytmicznych stóp zwrotu. W przypadku arytmetycznych stóp zwrotu występuje asymetria prawostronna, natomiast w przypadku logarytmicznych stóp zwrotu - lewostronna. W obu przypadkach mamy do czynienia z wyraźnym rozkładem leptokurtycznym. Omawiane rozkłady wraz z wartościami odpowiednich testów statystycznych przedstawiono na rysunkach 1 i 2.

Dodatkowym potwierdzeniem tego, że stopy zwrotu dla KGHM[1] nie podlegają rozkładowi normalnemu, są wyniki analizy tzw. wartości odstających. Narzędziem ułatwiającym taką analizę jest m.in. wykres normalności, na którym

[1] Przeprowadzone badania świadczą, że jest to zjawisko występujące w przypadku zdecydowanej większości spółek notowanych na giełdzie.

linia prosta obrazuje wartości dla rozkładu normalnego, natomiast poszczególne punkty przedstawiają wartości rzeczywiście zaobserwowane. Im bardziej punkty „układają" się wzdłuż linii prostej, tym większa jest zgodność danego rozkładu z rozkładem normalnym, natomiast im większe są odchylenia punktów od prostej, tym większe jest prawdopodobieństwo, że analizowana zmienna podlega rozkładowi innemu niż normalny.

W przypadku stóp zwrotu akcji KGHM wykres normalności przedstawiono na rysunku 3. Analizując ten wykres, zauważa się występowanie tzw. grubych ogonów, czyli wartości ekstremalnych na końcach przedziału zmienności. Dotyczy to zarówno minimalnych (ujemnych), jak i maksymalnych (dodatnich) stóp zwrotu.

Na podstawie powyższych badań należy stwierdzić, że często przyjmowane a priori założenie o normalności rozkładu stóp zwrotu nie jest podejściem właściwym. Również weryfikowanie hipotez o zgodności rozkładu danej stopy zwrotu z innymi znanymi, klasycznymi rozkładami nie zawsze daje pożądane rezultaty. Oznacza to, że często jedynym sposobem estymacji nieznanego rozkładu badanej zmiennej losowej jest wykorzystanie metod nieklasycznych, wśród których coraz większym zainteresowaniem cieszy się estymacja jądrowa.

3. Pojęcie i możliwości wykorzystania estymacji jądrowej w analizie stóp zwrotu

Jądrowy estymator gęstości prawdopodobieństwa jednowymiarowej zmiennej losowej definiowany jest wzorem:

$$\hat{f}(x)\frac{1}{mh}\sum_{i=1}^{m}K\left(\frac{x-x_i}{h}\right), \tag{3}$$

gdzie:

m - liczebność próby losowej,

h - parametr wygładzania,

K -jądro funkcji, spełniające następujące warunki:

$$\int_{-\infty}^{\infty}K(x)dx=1, \tag{4}$$

$$K(x)=K(-x), \text{ dla każdego } x\in R, \tag{5}$$

$$K(0)\geq K(x), \text{ dla każdego } x\in R. \tag{6}$$

Przy wykorzystywaniu estymacji jądrowej do szacowania funkcji gęstości ważne znaczenie ma wybór postaci jądra oraz właściwe wyznaczenie parametru wygładzania. Spośród wielu znanych funkcji, które spełniają wyżej wskazane wymagania dla funkcji jądrowej, ostatecznego wyboru dokonuje się przede wszystkim na podstawie efektywności danej funkcji oraz dodatkowych własności, które dana funkcja posiada. Efektywność danego jadra jest to stopień, w jakim minimalizowana jest wartość scałkowanego błędu średniokwadratowego *MISE (Mean Intergrated Sqyare Error),* definiowanego jako:

$$MISE = \int_{-\infty}^{\infty} E\left(\left[\hat{f}(x) - f(x)\right]^2\right) dx. \qquad \mathbf{(7)}$$

Z tego punku widzenia najbardziej efektywne jest jądro Epanecznikowa, zdefiniowane jako:

$$K(x) = \begin{cases} \frac{3}{4}(1-x)^2 & \text{dla} \quad x \in [-1;1], \\ 0 & \text{dla} \quad x \in (-\infty;1) \cup (1;\infty), \end{cases} \qquad (8)$$

a inne, również często stosowane jądro, jakim jest jądro normalne o postaci:

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \qquad (9)$$

jest jedynie o 5% mniej efektywne od jądra Epanecznikowa. W przypadku innych popularnych funkcji jądra (np. dwuwagowy, trójkątny) spadek efektywności względem jądra Epanecznikowa również jest istotnie mały. Z tego też względu dokonując wyboru funkcji jadra, uwzględnia się przede wszystkim jego własności, takie jak klasa regularności, prostota obliczeń czy też inne istotne dla prowadzonych badań cechy jądra (np. jądro normalne ma pochodną dowolnego rzędu, ale jego całka nie jest wyrażalna wzorem analitycznym, a w przypadku innych jąder całka jest możliwa do obliczenia, ale pochodna nie istnieje w całej dziedzinie)[2]. Sam wybór jądra nie ma jednak dla jakości estymacji takiego znaczenia jak wybór parametru wygładzania *h.* Przyjęcie zbyt dużej wartości współczynnika wygładzania powoduje zbyt mocne wygładzenie estymowanej funkcji gęstości rozkładu prawdopodobieństwa, przyjęcie natomiast zbyt małej wartości spowoduje pojawienie się dużej liczby lokalnych ekstremów. Istnieje kilka sposobów wyznaczenia wartości parametru wygładzania *h.* Najprostszą z nich jest metoda przybliżona, która sprawdza się jednak tylko w tych przypadkach, kiedy estymowany rozkład jest zbliżony do rozkładu normalnego. W tym przypadku jednak

[2] P. Kulczycki, *Estymatory jądrowe w analizie systemowej,* Wyd. Naukowo-Techniczne, Warszawa 2005, s. 67-68.

należałoby się zastanowić nad celowością wykorzystywania w prowadzonych badaniach estymacji jądrowej. Inną metodą jest metoda podstawień *k*-tego rzędu, którą można wykorzystać tylko wówczas, kiedy mamy do czynienia z jednowymiarowym rozkładem zmiennej losowej i w dodatku przyjęta funkcja jądra jest różniczkowalna $(2k + 2)$-krotnie. Ponieważ w praktyce przy korzystaniu z tej metody przyjmuje się $k = 2$ (czyli korzysta się z metody 2. rzędu) wystarczy, aby dla funkcji jądra istniały tylko cztery pierwsze pochodne. Bardziej uniwersalną metodą jest np. metoda krzyżowego uwiarygodnienia, którą można stosować niezależnie od wymiaru analizowanej zmiennej losowej, ale w praktyce wykorzystuje się ją przede wszystkim w przypadku rozkładów wielowymiarowych. Ponadto niezależnie od sposobu wyznaczenia wartości parametru wygładzania, w celu lepszego dopasowania estymowanej funkcji gęstości do rozkładu analizowanej zmiennej, dokonuje się modyfikacji parametru wygładzania. Sprowadza się ona do tego, że przy wyznaczaniu wartości estymatora jądrowego, jądro wyznaczone dla każdej wartości /-tej zmiennej koryguje się o wartość parametru $S_i > 0$, wyznaczonego na podstawie wzoru:

$$S_i = \left(\frac{\hat{f}(x_i)}{\tilde{s}} \right)^{-c}, \tag{10}$$

gdzie c jest nieujemnym parametrem decydującym o intensywności modyfikacji, najczęściej przyjmującym wartość 0,5, $\hat{f}(x_i)$ -jest estymatorem jądrowym wyznaczonym na podstawie wzoru (1), a $\tilde{s}$ jest średnią geometryczną z tych estymatorów, otrzymaną jako:

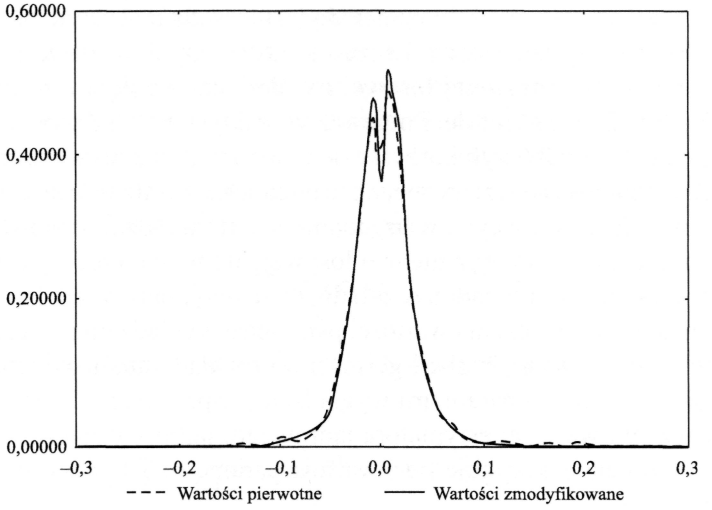
$$\tilde{s} = \exp\left(\frac{1}{m} \sum_{i=1}^{m} \ln(\hat{f}(x_i)) \right). \tag{ii}$$

Przy uwzględnieniu występowania parametru modyfikującego S_i, estymator jądrowy dla jednowymiarowej zmiennej losowej definiowany jest wzorem:

$$\hat{f}(x) = \frac{1}{mh} \sum_{i=1}^{m} \frac{1}{S_i} K\left(\frac{x - x_i}{S_i h} \right). \tag{12}$$

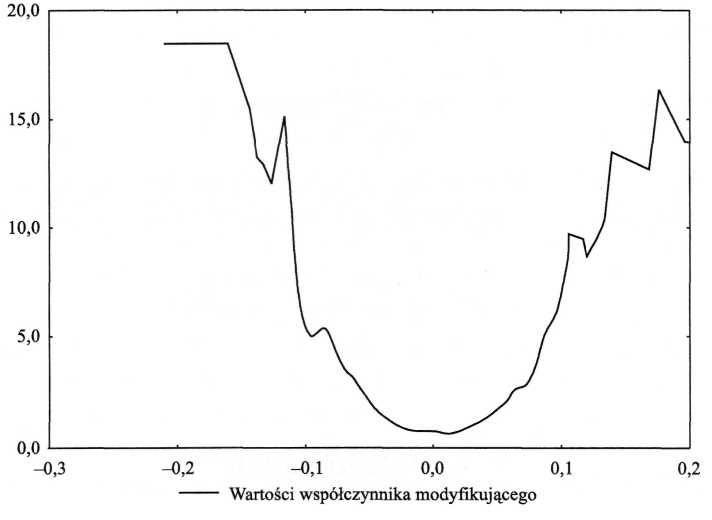
4. Wyniki badań

Szacując rozkład gęstości stóp zwrotu z akcji KGHM w okresie od 11.07.1997 r. do 30.04.2013 r., na podstawie wzoru (3), z jądrem normalnym (9), otrzymano rozkład, który przedstawiono na rysunku 4.



Rysunek 4. Wykres otrzymanej funkcji jądrowej rozkładu dziennych stóp zwrotu akcji KGHM w okresie od 11.07.1997 roku do 30.04.2013 roku bez współczynnika modyfikującego parametr wygładzający oraz z jego uwzględnieniem

Źródło: opracowanie własne.



Rysunek 5. Wartości współczynnika korygującego parametr wygładzania na tle wartości dziennej stopy zwrotu

Źródło: opracowanie własne.

Pomimo że przeciętna dzienna stopa zwrotu kształtowała się na poziomie +0,0953%, to w wyniku analizy otrzymanego wykresu gęstości, który istotnie różni się od rozkładu normalnego, można m.in. wskazać na dwie dominujące wartości stóp zwrotu: wśród ujemnych wartości najczęściej występowały stopy oscylujące wokół wartości -0,757%, natomiast najczęściej występującą dodatnią dzienną stopą zwrotu była stopa oscylująca wokół wartości +0,856%). W celu wyznaczenia wartości funkcji gęstości uwzględniającej modyfikacje parametru wygładzania, zgodnie ze wzorem (11) wyznaczono wartość $s \approx 0{,}000154$, a następnie na podstawie wzoru (10) wyznaczono odpowiednie wartości korygujące S_i. Zgodnie z założeniami teoretycznymi, duże wartości parametru korygującego miały miejsce w przypadku rzadko występujących stóp zwrotu, natomiast w obszarach stóp zwrotu o dużej częstości występowania, wartości współczynnika korygującego były zdecydowanie niższe. Rozkład wartości korygujących s. przedstawiono na rysunku 5, natomiast wartości funkcji jądrowej uwzględniające modyfikacje parametru wygładzania, w celu porównania z wartościami pierwotnie otrzymanymi, na rysunku 4. Wszystkie obliczenia i wykresy otrzymano, wykorzystując autorski program napisany w Microsoft Visual Studio Professional 2012.

W tabeli 1 zaprezentowano wyniki porównania szacowanego prawdopodobieństwa poniesienia określonej dziennej straty na akcjach KHGM w analizowanym okresie, otrzymanego dla założenia normalności rozkładu stóp zwrotu

Tabela 1. Porównanie wartości dystrybuant: empirycznej, rozkładu normalnego oraz funkcji jądrowej dla wybranych stóp zwrotu KGHM w okresie od 11.07.1997 roku do 30.04.2013 roku

Stopa zwrotu	Dystrybuanta empiryczna	Dystrybuanta rozkładu normalnego	Dystrybuanta funkcji jądrowej bez modyfikacji parametru wygładzania*	Dystrybuanta funkcji jądrowej z modyfikacją parametru wygładzania*
-8,02% lub mniejsza	0,01	0,0037	0,0106	0,0107
-6,41% lub mniejsza	0,02	0,0160	0,0200	0,0195
-5,37% lub mniejsza	0,03	0,0358	0,0315	0,0307
-4,86% lub mniejsza	0,04	0,0512	0,0412	0,0399
-4,46% lub mniejsza	0,05	0,0666	0,0519	0,0501

* Wartość dystrybuanty wyznaczono, wykorzystując całkowanie numeryczne.

Źródło: opracowanie własne.

i z wykorzystaniem oszacowanej funkcji jądrowej. Na przykład w okresie od 11.07.1997 r. do 30.04.2013 r. w przypadku 1% notowań dzienna stopa zwrotu z akcji KGHM ukształtowała się na poziomie -8,02% lub mniejszym. W przypadku szacowania prawdopodobieństwa takiej straty przy założeniu, że stopa zwrotu podlega rozkładowi normalnemu, oszacowane prawdopodobieństwo wynosiłoby 0,37%, natomiast wykorzystując w tym celu funkcję jądrową, przewidywane prawdopodobieństwo jest zdecydowanie bliższe rzeczywistemu i kształtuje się na poziomie od 1,06%, w przypadku braku modyfikacji parametru wygładzania, do 1,07% w przypadku wykorzystania podejścia z modyfikacją parametru wygładzania funkcj i jądrowej.

5. Podsumowanie

Na podstawie przeprowadzonych w artykule rozważań można przyjąć dwa istotne wnioski:

1. Często przyjmowane przy prowadzeniu badań nad rynkiem kapitałowym wstępne założenie, że stopy zwrotu podlegają rozkładowi normalnemu, nie jest uzasadnione. Zaprezentowane w artykule wyniki dotyczą jedynie akcji KGHM, ale przeprowadzone przez autora inne badania potwierdzają, że brak normalności stóp zwrotu dotyczy wielu innych spółek notowanych na GPW w Warszawie.
2. W celu właściwej analizy rozkładu stóp zwrotu z akcji, w większości przypadków należy wykorzystywać nieklasyczne metody, do których należy m.in. zaprezentowana w artykule metoda estymacji jądrowej. W coraz większym zakresie sprzyja temu ciągły wzrost mocy obliczeniowej komputerów oraz, co ważniejsze, pojawianie się specjalistycznych, dedykowanych nieklasycznym metodom (w tym symulacyjnym), programów statystycznych.

Literatura

Klóska R., *Statystyczna analiza poziomu rozwoju społeczno-gospodarczego w Polsce w ujęciu regionalnym,* „Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu" 2012, nr 42.

Kulczycki R, *Estymatory jądrowe w analizie systemowej,* Wyd. Naukowo-Techniczne, Warszawa 2005.

Purczyński J., *Wykorzystanie symulacji komputerowych w estymacji wybranych modeli ekonometrycznych i statystycznych,* Wyd. US, Szczecin 2003.

Wand M.R, Jones M.C., *Kernel Smoothing,* Chapman and Hall, London 1995.

Rate of Return Distribution Modeled by the Kernel Function

Abstract *The rate of return is one of the most important concepts in the theory and practice of finance. It is often assumed that it is characterized by a normal distribution. This paper presents the possibility of using the kernel function in modeling the rate of return.*

Keywords: *rate of return, density function, non-classical methods, kernel smoothing*