

275/2003

**Raport Badawczy**

**RB/72/2003**

**Research Report**

**Stochastyczny model  
strumienia nakładów  
na edukację**

**M. Berezinski, D. Wagner**

**Instytut Badań Systemowych  
Polska Akademia Nauk**

**Systems Research Institute  
Polish Academy of Sciences**



# **POLSKA AKADEMIA NAUK**

## **Instytut Badań Systemowych**

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 8373578

fax: (+48) (22) 8372772

Kierownik Pracowni zgłaszający pracę:  
Prof. dr Roman Kulikowski

Warszawa 2003

## Spis treści

### Streszczenie

1. Wprowadzenie .....	1
2. Dyskusja założeń, ogólna charakterystyka i skonstruowanie modelu .....	5
2.1. Charakterystyka zakłóceń impulsowych i ich serii w czasie .....	5
2.2. Rozkład liczby zakłóceń w funkcji wielkości ich amplitudy .....	11
2.3. Czasowy rozkład zakłóceń w serii .....	14
3. Wstępna analiza danych empirycznych .....	18
4. Zakończenie i wnioski .....	23

### Bibliografia



# STOCHASTYCZNY MODEL STRUMIENIA NAKŁADÓW NA EDUKACJĘ

Mirosław Bercziński, Dariusz Wagner  
Instytut Badań Systemowych PAN  
01-447 Warszawa, ul. Nowelska 6

**Streszczenie:** W pracy przeanalizowano teoretyczne założenia przyjmowane w stochastycznych modelach strumienia zakłóceń długofalowych procesów inwestycyjnych (na przykładzie edukacji). Stwierdzono i uzasadniono nieadekwatność założeń o jednorodności, czystej przypadkowości i stochastycznej niezależności zakłóceń do rzeczywistości, prowadzących do powszechni stosowanego modelu poissonowskiego. Stwierdzono też, że niezasadne jest przymywanie założenia o normalności rozkładu wielkości zakłóceń. Skonstruowano model, w którym przyjęto, że silne zakłócenia procesu występują seriami, odstępy czasu między seriami mają rozkład Pareto zaś wielkości zakłóceń podlegają rozkładowi hiperbolicznemu. Wyprowadzono dokładne formuły matematyczne stanowiące model oraz znaleziono ich przybliżenia analityczne, szczególnie wygodne dla przeprowadzania obliczeń. Podstawowe założenia i wyniki modelu zostały wstępnie sprawdzone w oparciu o dane rzeczywiste dotyczące jednej z uczelni wyższych. Potwierdzają one zasadność zaproponowanego modelu. Chociaż model nie ma bezpośrednio charakteru markowskiego, to jednak dobrze odwzorowuje stochastyczną zależność występowania serii zakłóceń.

## 1. Wprowadzenie

Ważnym czynnikiem kształtowania społeczeństwa informacyjnego jest doskonalenie dotychczasowych i wprowadzanie nowych metod edukacji. Edukacja przedszkolna, szkolna na poziomach podstawowym, gimnazjalnym, średnim oraz wyższa, wspomagane różnymi formami kształcenia uzupełniającego i doskonalącego (studia podyplomowe, kursy specjalistyczne, kursy doskonalenia zawodowego itp.) są elementami ogólnonarodowego programu aktywizacji wiedzy jako kluczowego czynnika rozwoju osobowego obywateli i społeczeństwa jako systemowej całości. Realizowanie takiej strategii wiąże się z koniecznością systematycznego przeznaczania na ten cel co najmniej niezbędnej minimalnej części dochodu narodowego, części dochodów przedsiębiorstw oraz części dochodów indywidualnych i racjonalnego nimi gospodarowania. Ponieważ procesy te trzeba rozpatrywać w bardzo długich horyzontach czasowych, a przewidywalność przyszłości maleje w miarę jego wydłużania się, więc ilościowo-jakościowa struktura strumieni pieniężnych, których składowymi będą nakłady na poszczególne lata, jest obciążony dużą niepewnością. Musi ona być brana pod uwagę przy podejmowaniu wszelkich decyzji dotyczących finansowania procesów edukacyjnych, zarówno indywidualnych, jak i zbiorowych.

Niepewność strumienia nakładów można odwzorowywać w różnoraki sposób. Jeżeli przyjąć, że mamy do czynienia z niepewnością typu probabilistyczno-statystycznego, to matematycznym modelem tego strumienia jest proces stochastyczny  $X(t) = \{X_t(\gamma), t \in T\}$ , w którym wielkości nakładów przewidywanych na poszczególne lata są realizacjami zmiennych losowych  $X_t(\gamma)$  zależnych od parametru  $t \in T$ , gdzie  $T$  - horyzont czasowy, określonych na pewnej przestrzeni probabilistycznej  $(\Gamma, \mathcal{F}, P)$ . Założymy, że przestrzeń ta jest dana i że są spełnione wszystkie warunki uprawniające do posługiwania się pojęciem prawdopodobieństwa.

Wiadomo, że pełną charakterystykę procesu stochastycznego  $X(t)$  daje jedynie zbiór wszystkich skończone wymiarowych rozkładów jego wartości we wszystkich możliwych układach chwil, i że rozkład  $n$ -wymiarowy jest scharakteryzowany przez dystrybucję  $n$ -wymiarowej zmiennej losowej,  $[X(t_1), X(t_2), \dots, X(t_n)]$

$$F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n) = P\{X(t_1) < x_1, X(t_2) < x_2, \dots, X(t_n) < x_n\}. \quad (1)$$

Każdy inny opis procesu stochastycznego jest uproszczeniem, którego forma powinna wynikać z rozwiązywanego zadania oraz z jakościowo-ilościowej struktury zbioru danych.

Dla skonstruowania modelu strumienia nakładów pieniężnych na edukację zauważamy, że z punktu widzenia teorii procesów stochastycznych może on być traktowany jako jednorodny strumień zdarzeń, tzn. jako proces losowy z czasem ciągłym, przyjmujący wartości całkowitoliczbowe, nieujemne. Istotnie, finansowe zasilanie procesu edukacyjnego nie jest na ogół aktem jednorazowym, lecz procesem rozłożonym w całym horyzoncie czasowym tego przedsięwzięcia. W określonych momentach czasu właściciel kapitału pieniężnego przekazuje lub przeznaczają na edukację określone sumy pieniędzy. Mamy więc do czynienia ze specyficzną formą inwestowania w człowieka. Proces ten jest sprzężony z procesem zwrotu z inwestycji, zazwyczaj przesuniętym w czasie względem początku procesu finansowania i zachodzącym w znacznie dłuższym horyzoncie czasowym niż sam proces edukacyjny. Oba te procesy podlegają zakłóceniom. Są dwa rodzaje zakłóceń: wewnętrzne i zewnętrzne w stosunku do nich. Zakłócenia wewnętrzne mają swoje źródło w łonie samych procesów formowania strumieni nakładów i gospodarowania nimi, a źródłem zakłóceń zewnętrznych jest otoczenie. Przyczyny zakłóceń mogą być obiektywne bądź subiektywne. Przyczyny obiektywne występują niezależnie od woli decydentów, natomiast przyczyny subiektywne tkwią w celowo ukierunkowanych działaniach decydentów, w ich posunięciach organizacyjnych, doświadczeniu, wiedzy itp.

W teorii finansowania procesów edukacyjnych istnieje wiele stochastycznych modeli strumienia nakładów, (zob. np.: Runggaldier 1997; Embrechts, Klüppelberg i Mikosch, 2001). Wszystkie je można podzielić na dwie kategorie. Pierwszą stanowią modele czysto teoretyczne. Podstawową słabością większości z nich jest nadmierne upraszczanie rzeczywistości, wynikające z przyjmowania w nich takich założeń, które pozwalają wprowadzić przeniesić na grunt teorii finansowania procesów edukacyjnych stosunkowo proste modele stochastyczne, ale zbyt daleko odbiegają od realiów, by mogły być użyteczne w praktyce. Drugą kategorię modeli stanowią modele czysto empiryczne, powstałe w wyniku statystycznej obróbki danych o rzeczywistych rozkładach wielkości nakładów finansowych na edukację w czasie. Przydatność tych modeli jest również ograniczona, jako że opisują one konkretne specyficzne sytuacje, które zaistniały w przeszłości i których na ogół nie można ekstrapolować w przyszłość. Ważną cechą modeli obu kategorii jest przyjmowanie, że w typowych warunkach zakłócenia strumienia nakładów finansowych na edukację są w pełni kontrolowane oraz na bieżąco korygowane i z tego powodu nie stanowią zagrożenia dla rytmiki strumienia pieniężnego zasilania i zwrotu z tej inwestycji. Przyjęcie tego założenia pozwoliło przenieść na grunt inżynierii finansowej niemal cały dorobek klasycznej teorii rachunku prawdopodobieństwa, statystyki matematycznej i teorii procesów stochastycznych. Szczególnie popularne stały się modele

oparte na założeniu statystycznej niezależności i jednakowości rozkładów zmiennych losowych reprezentujących odpowiednie elementy rozważanego zagadnienia finansowego, co pozwoliło wykorzystywać w szerokim zakresie prawa małych i wielkich liczb oraz graniczne twierdzenia rachunku prawdopodobieństwa, dotyczące stochastycznej zbieżności ciągów zmiennych losowych i ciągów dystrybuant oraz rozkładów sum zmiennych losowych niezależnych o jednakowych rozkładach prawdopodobieństwa i skończonych wariancjach. Jak wiadomo, centralne miejsce w tych twierdzeniach zajmuje rozkład normalny. Nic więc dziwnego, że dominuje on w modelach strumieni finansowania przedsięwzięć edukacyjnych. Nie jest naszą intencją krytykowanie tego podejścia, gdyż niewątpliwie przyniosło ono wiele niezwykle wartościowych wyników dotyczących gospodarowania finansami w sferze edukacji. Prawdą jest jednak, że w miarę weryfikacji budowanych modeli stopniowo i coraz wyraźniej ujawniały się i wciąż ujawniają te ich aspekty, które istotnie odbiegały i odbiegają od realiów praktyki edukacyjnej, a które dla tej praktyki były i są ważne. Na przykład, okazało się, że za pomocą modeli opartych na założeniu normalności lub asymptotycznej normalności rozkładów zmiennych losowych charakteryzujących przypadkowe zakłócenia strumieni nakładów finansowych w różnych horyzontach czasowych nie daje się wystarczająco dobrze odwzorować występowania groźnych w skutkach zakłóceń o dużych amplitudach, pojawiających się nagle i niespodzianie. Będziemy je nazywać zakłóceniami impulsowymi. Zazwyczaj zakłócenia takie występują rzadko, nie trwają długo, ale ich skutki mogą być bardzo poważne. Co więcej, ich znaną cechą jest to, że rzadko kiedy występują one w odosobnieniu. Na ogół pojawiają się one w seriach, których czas trwania jest stosunkowo krótki w porównaniu z czasem trwania spokojnych faz długofalowego procesu finansowania inwestycji edukacyjnych. Z uwagi na niezwykle destrukcyjną naturę zakłóceń impulsowych nie mogą być one pomijane w matematycznych opisach strumieni finansowych. Tradycyjne przedstawianie tych strumieni jako stacjonarnych ciągów jednorodnych punktowych zdarzeń przypadkowych bez następstw lub z ograniczonym następstwem, a w szczególności jako procesów rekurencyjnych, jest więc wprawdzie wygodnym matematycznie, ale zbyt daleko idącym uproszczeniem rzeczywistości. Istnieje zatem potrzeba opracowania takiej koncepcji modelowania strumieni nakładów finansowych, która - zachowując prostotę modelu - pozwoli odwzorować ich prawdziwą strukturę, a zwłaszcza niepunktowy charakter zakłóceń impulsowych.

W świetle tej potrzeby celowe wydaje się wyróżnienie dwóch rodzajów przypadkowości w strumieniach nakładów finansowych. Pierwszy z nich będziemy wiązać z zakłóceniami monotonicznymi, występującymi często i mającymi niewielkie amplitudy. Drugi będzie odpowiadał silnym zakłóceniom impulsowym o dużych amplitudach, pojawiającym się z nagłą, rzadko i z reguły w krótkich seriach o losowej długości. Zakłócenia pierwszego rodzaju mają charakter masowy, co - wraz z przypadkowym działaniem czynników zewnętrznych w stosunku do wewnętrznych uwarunkowań strumienia nakładów - tworzy warunki dla wzajemnego znoszenia się dodatnich i ujemnych odchyień częstości zakłóceń od ich faktycznych prawdopodobieństw, sprzyjając ujawnianiu się rzeczywistych mechanizmów i prawidłowości tego strumienia. Jest tak, ponieważ w tym przypadku działa prawo wielkich liczb, które jest koniecznym warunkiem przejawiania się prawidłowości statystycznych, kształtujących się pod wpływem wewnętrznej logiki strumienia nakładów. Prawidłowości tych nie można wykryć w żadnej pojedynczej obserwacji strumienia nakładów, ale dają się one dostrzec w masie obserwacji i scharakteryzować za pomocą wartości średnich. Inaczej jest w przypadku zakłóceń impulsowych. Ponieważ nie mają one charakteru masowego, więc leżą poza obszarem stosowności

prawa wielkich liczb. Ze względu na ich unikatowość do ich opisu nie można stosować klasycznych reguł rachunku prawdopodobieństwa, o ile nie przyjmie się stosownych założeń upraszczających. O dopuszczalności takich założeń zawsze musi decydować przeznaczenie modelu.

Jak już wspomnieliśmy, tradycyjne podejście do modelowania probabilistyczno-statystycznego mechanizmu procesów finansowania inwestycji edukacyjnych zakłada, że zakłócenia, które oddziałują na ten proces są punktowymi zdarzeniami jednorodnymi zachodzącymi w czasie ciągłym oraz że strumień zakłóceń jest procesem przyjmującym nieujemne wartości całkowitoliczbowe. Przy tych założeniach strumień zakłóceń można opisać na dwa równoważne sobie sposoby: albo (1) poprzez określenie rozkładu prawdopodobieństwa zmiennej losowej, której realizacjami są liczby zakłóceń w różnych przedziałach czasu, położonych w dowolny sposób względem punktu przyjętego za początek osi czasu, albo też (2) poprzez określenie rozkładu prawdopodobieństwa zmiennej losowej, której realizacjami są długości odstępów czasu między zakłóceniami. W pierwszym przypadku z reguły zakłada się, że strumień zakłóceń jest procesem bez następstw (liczby zakłóceń w rozłącznych przedziałach osi czasu są realizacjami niezależnych zmiennych losowych) i stacjonarnym (probabilistyczny mechanizm strumienia nie zmienia się w czasie). W drugim przypadku opis strumienia zakłóceń polega na określeniu rodziny funkcji rozkładów prawdopodobieństwa wielowymiarowej zmiennej losowej, której realizacjami są długości odstępów czasu między zakłóceniami, obliczane względem dowolnie wybranego punktu odniesienia, przy czym z reguły przyjmuje się, że te przedziały są wzajemnie niezależne (tzn. że ma się do czynienia ze strumieniem z ograniczonym następstwem). Najpowszechniej stosowanym modelem strumienia zakłóceń jest proces najprostsz, który zakłada, że strumień zakłóceń jest regularny [funkcja wiodąca strumienia, czyli wartość oczekiwana liczby zakłóceń w przedziale czasu  $<0, t)$  jest ciągła], bez następstw, zwyczajny (prawdopodobieństwo wystąpienia w małym przedziale czasu więcej niż jednego zakłócenia jest wielkością pomijalnie małą w porównaniu z długością tego przedziału) i stacjonarny. Należy pamiętać, że jeżeli modelem strumienia zakłóceń jest proces najprostsz, to liczba zakłóceń obserwowanych w określonym przedziale czasu ma rozkład Poissona o stałej intensywności, natomiast długości odstępów czasu między zakłóceniami mają rozkład wykładniczy. Zachodzi również relacja odwrotna, tzn. wykładniczy rozkład długości tych odstępów implikuje poissonowskość rozkładu liczby zakłóceń.

Podstawowym źródłem poznania własności strumieni zakłóceń jest logiczna analiza mechanizmów ich powstawania i oddziaływania na długofalowe procesy finansowe oraz dane empiryczne. Bezpośrednio dostępnych danych empirycznych jest jednak wciąż zbyt mało i są one trudno dostępne, ale już nawet te, które dotychczas zebrano i opublikowano od dawna podważają zasadność wielu założeń przyjmowanych w matematycznych modelach strumieni finansowych, nie tylko w sferze edukacji (zob., np., Pareto, 1897; Bachelier, 1900; Gibrat, 1932; Mandelbrot, 1958, 1959, 1966, 1982; Mandelbrot i Taylor, 1967; Fama, 1965, 1970; Samuelson, 1967; Mirowski, 1990, 1996). W szczególności wskazują na niegaussowski charakter rozkładu wielkości zakłóceń, czego konsekwencją jest również niegaussowskość wielu obserwowanych zmiennych finansowych i ekonomicznych, asymetria ich rozkładów itp. Ujawniają też coraz więcej niedostatków procesu najprostszego, jako modelu strumienia zakłóceń. Przede wszystkim okazuje się, że aproksymacja strumienia czynników zakłócających strumień nakładów za pomocą strumienia niezależnych zdarzeń punktowych jest zbyt daleko idącym uproszczeniem. Jesz-



cze większym uproszczeniem jest przyjmowanie, że odstępy czasu między zakłóceniami są realizacjami zmiennych losowych o rozkładzie wykładniczym. Zakłócenia ani nie są zdarzeniami punktowymi, ani też nie występują pojedynczo. Jest wręcz przeciwnie. Mają one wyraźną tendencję do pojawiania się seriami, co świadczy o ich silnej stochastycznej zależności.

Chociaż rozkład serii zakłóceń w czasie jest losowy, to jednak koncentracja zakłóceń impulsowych w seriach świadczy o tym, że ich pojawianiem się nie rządzi czysta przypadkowość i że nie są one w pełni niezależnie od siebie. Dotychczasowe doświadczenia uczą (Fama, 1965, 1970; Mandelbrot, 1966; Markowitz 1952, 1959; Mirowski, 1990, 1996), że nie ma podstaw do odrzucenia hipotezy o statystycznej niezależności długości przedziałów czasu między seriami zakłóceń impulsowych. Zdając sobie sprawę, że brak podstaw do odrzucenia hipotezy statystycznej nie oznacza, że jest ona prawdziwa, utrzymamy ją jednak, ale równocześnie przyjmiemy, że rozkład serii zakłóceń w czasie podlega prawu Pareto. Myślą przewodnią proponowanego modelu jest wykorzystanie wprowadzonej przez Mandelbrota (1958, 1966, 1982) a nawiązującej do paretowskiego prawa rozkładu dochodów i dobrze udokumentowanej merytorycznie oraz empirycznie - również w procesach ekonomicznych i finansowych - idei samopodobieństwa. Jej wyróżnikiem w odniesieniu do struktury zakłóceń strumienia nakładów w długim horyzoncie czasu jest potwierdzony doświadczalnie fakt, że zakłócenia te występują w seriach, które też nie rozkładają się w czasie zupełnie przypadkowo, lecz mają skłonność do grupowania się itd. Tej właściwości nie można odwzorować za pomocą żadnego z procesów stochastycznych opartych na założeniu statystycznej niezależności zdarzeń. Za pomocą tych procesów można opisać jedynie rozkłady zakłóceń typowych, jednorodnych. Istotne znaczenie dla kształtowania strumienia nakładów i zarządzania nim mają jednak zakłócenia impulsowe, którym odpowiada jeden z ogonów rozkładu prawdopodobieństwa. Modelowanie tych ogonów ma więc duże znaczenie praktyczne.

Celem pracy jest zaproponowanie analitycznego modelu strumienia zakłóceń procesu nakładów finansowych w długim horyzoncie czasowym, odwzorowującego seryjność występowania zakłóceń impulsowych.

## **2. Ogólna charakterystyka modelu**

### **2.1. Rozkład zakłóceń impulsowych i ich serii w czasie**

Idea proponowanego modelu polega na przyjęciu, że na strumień nakładów oddziałuje strumień zakłóceń przypadkowych, niejednorodnych ani pod względem wielkości ich amplitud, ani też skutków, które powodują. W zależności od tego jakiego rodzaju zakłócenia dominują, w strukturze wypadkowego strumienia kształtują się dwie naprzemiennie występujące fazy: stosunkowo długa faza monotoniczna, w której funkcjonuje statystyczne prawo wielkich liczb, oraz znacznie od niej krótsza faza impulsowa, w której działanie tego prawa ulega zawieszeniu. Z punktu widzenia kształtowania struktury strumienia nakładów, a także zarządzania nim, szczególnie ważne jest poznanie i matematyczne opisanie własności zakłóceń o charakterze impulsowym, a zwłaszcza ich rozkładu w czasie oraz rozkładu wielkości ich amplitud. Jak już powiedzieliśmy, nie tylko z wewnętrznej logiki strumienia nakładów inwestycyjnych w sferze edukacji oraz ich interakcji z otoczeniem, ale też z ich obserwacji wynika, że oddziałujące na nie zakłócenia impulsowe na ogół nie występują odrębnie, lecz w seriach, których czas trwania jest na

ogół mały w porównaniu z czasem trwania całego procesu. Przyjęcie, że wystąpienie takiej serii jest zdarzeniem punktowym, bezwymiarowym, znacznie wprawdzie upraszcza matematyczny opis ciągu zakłóceń, ale równocześnie zacięra prawdziwy obraz rzeczywistości. Ponieważ w warunkach względnej stabilności otoczenia serie zakłóceń impulsowych występują stosunkowo rzadko, więc badanie mechanizmów ich pojawiania się oraz wywieranych przez nie skutków należy w gruncie rzeczy do statystyki zdarzeń ekstremalnych i wiąże się z koniecznością rozpatrywania ogonów rozkładów prawdopodobieństwa (Gumbel, 1962; Leadbetter, Lindgren i Rootzen, 1986; Kotz i Nadarajah, 2000). Ale w podstawowym nurcie statystyki zdarzeń ekstremalnych też zakłada się, że zdarzenia te są statystycznie niezależne (teoria zdarzeń ekstremalnych zależnych znajduje się dopiero w początkowym stadium rozwoju).

Dotychczasowe modele strumienia zakłóceń procesu nakładów inwestycyjnych - opierając się na założeniu o jednorodności i punktowości zakłóceń oraz ich statystycznej niezależności i - na ogół - wykładniczości rozkładu rozdzielających je odstępów czasu - nie odwzorowują zjawiska ich seryjności. Stąd konieczność zmiany sposobu widzenia systemowej struktury strumienia zakłóceń. W proponowanym modelu nie będziemy więc uważać go za proces czysto punktowy, lecz spojrzymy nań jako na proces, w którym zakłócenia mają określony czas trwania. Ponieważ krzywa rozkładu wykładniczego dość dobrze aproksymuje rzeczywisty rozkład długości odstępów czasu między zakłóceniami tylko w środkowej części zakresu zmienności ich amplitud - podczas gdy istotne znaczenie ma przede wszystkim informacja o zachowaniu się zakłóceń ekstremalnych, zwłaszcza impulsowych - więc odrzucamy założenie o wykładniczości. Przyjmujemy, że odstęp czasu między zakłóceniami są realizacjami zmiennej losowej o rozkładzie Pareto. Jest to rozkład ciągły, dwuparametrowy, o gęstości

$$f(t) = \begin{cases} \frac{\alpha\beta^\alpha}{t^{\alpha+1}}, & t \geq \beta, \\ 0, & t < \beta, \end{cases} \quad (2)$$

gdzie  $\alpha > 0$ ,  $\beta > 0$ . O wyborze tego rozkładu zdecydował fakt, że - jak uczy teoria i doświadczenie - jego dolny ogon pozwala odwzorować fakt częstego występowania zdarzeń, których statystyczne charakterystyki leżą znacznie poniżej średniej, zaś górny - rzadkiego pojawiania się zdarzeń o charakterystykach znacznie przekraczających tę średnią. Mając na uwadze tę właściwość zastanówmy się nad mechanizmem powstawania serii zakłóceń impulsowych.

W tym celu oznaczmy przez  $T_n$  zmienną losową, której realizacjami są chwile pojawiania się zakłóceń i rozważmy trzy kolejno po sobie następujące zakłócenia. Załóżmy, że są znane chwile  $t_{n-1}$  i  $t_{n+1}$  pojawienia się pierwszego i trzeciego z nich, natomiast nie jest znana chwila  $t_n$  ( $t_{n-1} < t_n < t_{n+1}$ ) pojawienia się trzeciego. Odcinki czasu  $t_n - t_{n-1}$  oraz  $t_{n+1} - t_n$  mają długość losową. Załóżmy, że między pojawieniem się pierwszego i trzeciego zakłócenia upłynął czas  $t = t_{n+1} - t_{n-1}$ , który jest realizacją zmiennej losowej  $T = T_{n+1} - T_{n-1}$  i niech  $\hat{t} = t_n - t_{n-1}$  będzie realizacją zmiennej losowej  $\hat{T} = T_n - T_{n-1}$ . Zważmy, że:

1. Gdyby przyjąć, że odstępy czasu między zakłóceniami są niezależne i mają rozkład geometryczny, to zmienna  $T_n$  miałaby na odcinku  $t_{n-1} < T_n < t_{n+1}$  rozkład jednostajny. Taki model jest jednak nie do przyjęcia, ponieważ rzeczywiste upływy czasu między chwilami pojawiania się zakłóceń impulsowych są bardzo zróżnicowane i nie mają rozkładu jednostajnego.
2. Gdyby założyć, że odstępy czasu między zakłóceniami są niezależne i podlegają rozkładowi dwumianowemu, to – w długim horyzoncie czasu - rozkład ten można by aproksymować za pomocą rozkładu normalnego o wartości oczekiwanej równej w przybliżeniu średniej arytmetycznej wyników obserwacji, tj.

$$m = \frac{t_{n-1} + t_{n+1}}{2},$$

przy czym trzeba by przyjąć, że zakłócenia rozkładają się prawie równomiernie w czasie. Ten postulat jest również nie do przyjęcia, gdyż zdecydowanie przeczy rzeczywistości.

3. Jeżeli jednak przyjąć, że odstępy czasu między zakłóceniami są niezależne i mają rozkład Pareto, to prawdopodobieństwo zdarzenia, że zmienna losowa  $\hat{T}$  przyjmie wartość  $\hat{t}$  pod warunkiem, że zmienna losowa  $T$  przyjmie wartość  $t$  wyrazi się wzorem

$$P(\hat{T} = \hat{t} | T = t) = \frac{\alpha^2 \hat{t}^{-(\alpha+1)} (t - \hat{t})^{-(\alpha+1)}}{2\alpha t^{-(\alpha+1)}} = \left(\frac{\alpha}{2}\right) \hat{t}^{-(\alpha+1)} \left(1 - \frac{\hat{t}}{t}\right)^{-(\alpha+1)}. \quad (3)$$

Wynika z niego, że rozkład zmiennej losowej  $T_n$  ma w chwilach  $t_{n-1} + 1$  i  $t_{n+1} - 1$  dwa ostre wysoki o takich samych amplitudach, niezależnych od  $t$ . Wskutek tego zakłócenie leżące wewnątrz przedziału będzie ciążyło ku zakłóceniu występującym w jednym z jego końców, co prowadzi do odwzorowania serii złożonej z dwóch zakłóceń (rys. 1c). W dokładnie taki sam sposób można odwzorować mechanizm powstawania dłuższych dłuższych serii.

Prawdopodobieństwo tego, że długość przedziału oddzielającego dwa kolejne zakłócenia impulsowe będzie niemniejsza od pewnej ustalonej wartości progowej  $t$  wynosi

$$P\{T \geq t\} = \int_t^{\infty} f(u) du, \quad (4)$$

czyli

$$P\{T \geq t\} = \alpha \beta^{\alpha} \int_t^{\infty} \frac{du}{u^{\alpha+1}}. \quad (5)$$

Po wykonaniu obliczeń dostajemy

$$P\{T \geq t\} = \begin{cases} \left(\frac{\beta}{t}\right)^\alpha, & t \geq \beta, \\ 1, & t < \beta. \end{cases} \quad (6)$$

Zauważmy, że ponieważ

$$\log\left(\frac{\beta}{t}\right)^\alpha = -\alpha \log t + \alpha \log \beta, \quad (7)$$

więc w logarytmicznym układzie współrzędnych prostokątnych  $\beta$  jest parametrem skali. Nie zmniejszając ogólności rozważań można więc przyjąć, że  $\beta = 1$ . Wtedy funkcja gęstości (2) przyjmie postać

$$f(t) = \begin{cases} \alpha t^{-\alpha-1}, & t \geq 1, \\ 0, & t < 1, \end{cases} \quad (8)$$

a prawdopodobieństwo wystąpienia przedziału bez zakłóceń, mającego długość co najmniej  $t$ , wyniesie

$$P\{T \geq t\} = \begin{cases} 1 t^{-\alpha}, & t \geq 1, \\ 1, & t < 1. \end{cases} \quad (9)$$

Z uwagi na niejednorodność zakłóceń rozsądnie jest przyjąć, że przedziały czasu wolne od zakłóceń są sumami skończonej liczby krótszych odcinków, będących realizacjami różnych zmiennych losowych paretońskich. Skoro tak, to gęstość prawdopodobieństwa zmiennej losowej będącej sumą  $r$  odcinków jest kompozycją  $r$  funkcji gęstości charakteryzujących poszczególne składowe. Oznaczmy tę gęstość przez  $f_r(t)$ . Jeżeli  $r = 2$ , to

$$f_2(t) = \int_1^{t-1} f(u)f(t-u)du, \quad (t \geq 2), \quad (10)$$

czyli, po uwzględnieniu (7),

$$f_2(t) = \alpha^2 \int_1^{t-1} u^{-(\alpha+1)}(t-u)^{-\alpha+1} du. \quad (11)$$

Podstawiając  $u = t \cdot z$  sprowadzamy tę całkę do postaci

$$f_3(t) = 2\alpha^2 t^{-(2\alpha+1)} \int_{1/t}^{1/2} z^{-(\alpha+1)} (1-z)^{-(\alpha+1)} dz. \quad (12)$$

Aby obliczyć jej wartość skorzystajmy z rozwinięcia funkcji  $(1-z)^{-(\alpha+1)}$  w szereg Taylora. Ponieważ

$$(1-z)^{-(\alpha+1)} = \sum_{s=1}^{\infty} \binom{\alpha+s}{s} z^s,$$

więc

$$f_2(t) = 2\alpha^2 t^{-(2\alpha+1)} \int_{1/t}^{1/2} z^{-(\alpha+1)} \left[ \sum_{s=0}^{\infty} \binom{\alpha+s}{s} z^s \right] dz. \quad (13)$$

Po zamianie kolejności całkowania i sumowania a następnie scałkowaniu szeregu wyraz po wyrazie i uporządkowaniu otrzymamy

$$f_2(t) = 2\alpha t^{-(\alpha+1)} \sum_{s=0}^{\infty} \frac{\alpha}{\alpha-s} \binom{\alpha+s}{s} \left[ \left(\frac{1}{t}\right)^s - \left(\frac{1}{2}\right)^s \left(\frac{2}{t}\right)^s \right]. \quad (14)$$

Zauważmy teraz, że

$$\begin{aligned} \frac{\alpha}{\alpha-s} \binom{\alpha+s}{s} &= \frac{\alpha}{\alpha-s} \frac{\alpha! (\alpha+1) \dots (\alpha+s)}{\alpha! s!} = \\ &= \frac{(-\alpha) (\alpha+1) \dots (\alpha+s)}{\alpha-s s!} = \\ &= \frac{(-\alpha)}{\alpha-s} \cdot \frac{(-\alpha+1)(-\alpha+2) \dots (-\alpha+s-1)}{(-\alpha+1)(-\alpha+2) \dots (-\alpha+s-1)} \cdot \frac{(\alpha+1) \dots (\alpha+s)}{s!}. \end{aligned}$$

Stosując symbol  $(a)_s = a(a+1) + \dots + (a+s-1)$  nadamy temu wyrażeniu postać

$$\frac{\alpha}{\alpha-s} \binom{\alpha+s}{s} = \frac{(-\alpha)_s (\alpha+1)_s}{(-\alpha+1)_s s!}.$$

Jeżeli przyjmiemy oznaczenie

$$A_s = \frac{(-\alpha)_s (\alpha+1)_s}{(-\alpha+1)_s s!},$$

to

$$f_2(t) = 2\alpha t^{-(\alpha+1)} \sum_{s=0}^{\infty} A_s \left(\frac{1}{t}\right)^s - 2^{\alpha+1} \alpha t^{-\alpha} \sum_{s=0}^{\infty} A_s \left(\frac{1}{2}\right)^s. \quad (15)$$

Szeregi występujące w pierwszym i drugim składniku tego wyrażenia są szeregami hipergeometrycznymi (zob., np., Kamke, 1959; Graham, Knuth i Patashnik, 1996)  $F(-\alpha, 1 + \alpha; 1 - \alpha, \frac{1}{t})$  oraz  $F(-\alpha, 1 + \alpha; 1 - \alpha, \frac{1}{2})$ . Ponieważ

$$F\left(-\alpha, \alpha + 1; -\alpha + 1, \frac{1}{2}\right) = 2^{-\alpha} \frac{\Gamma(1 - \alpha) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(-\alpha + \frac{1}{2}\right) \Gamma(1)}, \quad (16)$$

gdzie  $\Gamma(x) = \int_0^{\infty} e^{-u} u^{x-1} du$  i ponieważ  $\Gamma(1) = 1$ , więc ostatecznie

$$f_2(t) = 2\alpha t^{-(\alpha+1)} \left[ F\left(-\alpha, \alpha + 1; -\alpha + 1, \frac{1}{t}\right) - 2^{\alpha+1} t^{-\alpha} \frac{\Gamma(1 - \alpha) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(-\alpha + \frac{1}{2}\right)} \right]. \quad (17)$$

Wróćmy jeszcze do formuły (14). Można jej nadać postać

$$f_2(t) = 2\alpha t^{-(\alpha+1)} \left[ 1 - \left(\frac{4}{t}\right)^{\alpha} \frac{\Gamma(1 - \alpha) \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(-\alpha + \frac{1}{2}\right)} - \sum_{k=1}^{\infty} \frac{\alpha}{s - \alpha} \binom{\alpha + s}{s} \left(\frac{1}{t}\right)^s \right].$$

Zauważmy, że dla dużych wartości  $t$  ma miejsce następujące przybliżenie

$$f_2(t) \approx 2\alpha t^{-(\alpha+1)}. \quad (18)$$

W przypadku, gdy przedziały czasu wolne od zakłóceń są sumami więcej niż dwóch odcinków losowych, wtedy dla znalezienia gęstości rozkładu zmiennej losowej charakteryzującej długość całego przedziału trzeba powyższą procedurę kompozycji rozkładu powtórzyć rekurencyjnie odpowiednią liczbę razy. Kontynuując to rozumowanie otrzymamy, że jeżeli odcinków tych jest  $r$ , to dla dużych wartości  $t$

$$f_r(t) \approx r\alpha t^{-(\alpha+1)}. \quad (19)$$

Ta postać formuły jest szczególnie wygodna dla przeprowadzania obliczeń.

## 2.2. Rozkład liczby zakłóceń w funkcji wielkości amplitudy

Znamienną cechą zakłóceń impulsowych jest to, że ich liczba maleje w miarę wzrostu stopnia ich ekstremalności. Np., największy procentowy odsiew studentów w czasie egzaminów ma miejsce na pierwszym roku studiów, mniejszy na drugim, jeszcze mniejszy na trzecim itd.; łatwość przejścia do wyższej grupy dochodowej jest tym większa, im większy dochód już się posiada; stopień podatności kursów walutowych na zakłócenia jest tym większy, im mniej stabilna jest waluta, itp. Z tego rodzaju zjawiskiem mamy również do czynienia w przypadku strumienia zakłóceń procesu długofalowego finansowania inwestycji, np. edukacyjnych. Zakłóceń drobnych, w zasadzie nie wywierających większego wpływu na przebieg procesu jest bardzo dużo, ale w miarę zwiększania się amplitudy zakłóceń ich liczba oraz częstość występowania gwałtownie zmniejszają się. Jeżeli przyjąć, że względna liczba zakłóceń maleje proporcjonalnie do stopnia ich ekstremalności, to podstawowym matematycznym modelem tej zmienności jest hiperbola, określona wzorem:

$$h(x) = \frac{A}{(x + \alpha)^k}, \quad (20)$$

gdzie  $x \geq 0$ ,  $x$  - wielkość amplitudy zakłócenia,  $A$  - stała dodatnia,  $\alpha$  - dowolnie mała stała dodatnia, grająca rolę czynnika skalującego,  $k$  - stała dodatnia. Ponieważ wartość  $\alpha$  jest mała, więc wartość ta nie wpływa na kształt hiperboli w obszarze dużych wartości amplitud zakłóceń.

Przez odpowiedni dobór wartości stałej  $A$  można nadać funkcji (10) interpretację gęstości prawdopodobieństwa. W tym celu należy rozwiązać względem  $A$  równanie

$$\int_0^{\infty} h(x) dx = 1, \quad (21)$$

czyli

$$\int_0^{\infty} \frac{A}{(x + \alpha)^k} dx = 1. \quad (22)$$

Po scałkowaniu i wykonaniu stosownych przekształceń otrzymamy

$$A = (k - 1)\alpha^{k-1}.$$

Wobec tego, funkcja

$$h(x) = \frac{(k - 1)\alpha^{k-1}}{(x + \alpha)^k} \quad (23)$$

jest gęstością prawdopodobieństwa dla rozkładu hiperbolicznego.

Obliczmy prawdopodobieństwo tego, że amplituda zakłócenia przyjmie wartość równą lub przekraczającą ustaloną wartość progową  $x$ , tzn.  $P\{X \geq x\}$ . Mamy

$$P\{X \geq x\} = \int_x^{\infty} h(z) dz, \quad (24)$$

czyli

$$P\{X \geq x\} = (k-1)\alpha^{k-1} \int_x^{\infty} \frac{dz}{(z+\alpha)^k}. \quad (25)$$

Po scałkowaniu otrzymamy

$$P\{X \geq x\} = \frac{(k-1)\alpha^{k-1}}{(x+\alpha)^{k-1}}. \quad (26)$$

Wartość oczekiwana zmiennej losowej o rozkładzie hiperbolicznym wynosi

$$E(X) = \int_0^{\infty} xh(x) dx, \quad (27)$$

czyli

$$E(X) = (k-1)\alpha^{k-1} \int_0^{\infty} \frac{x}{(x+\alpha)^k} dx. \quad (28)$$

Po wykonaniu całkowania i odpowiednich przekształceniach otrzymamy

$$E(X) = \frac{\alpha}{k-2}. \quad (29)$$

Wariancję rozkładu oblicza się według wzoru

$$D^2(X) = \int_0^{\infty} x^2 h(x) dx, \quad (30)$$

czyli

$$D^2(X) = (k-1)\alpha^{k-1} \int_0^{\infty} \frac{x^2}{(x+\alpha)^k} dx. \quad (31)$$

Po wykonaniu obliczeń otrzymując się



$$D^2(X) = \frac{2\alpha^2}{(k-1)(k-2)}. \quad (32)$$

Wobec tego odchylenie standardowe wyraża się wzorem

$$\sigma(X) = \alpha \sqrt{\frac{2}{(k-1)(k-2)}}. \quad (33)$$

Zauważmy, że dla  $k = 2$  wartość oczekiwana, a dla  $k = 1$  i  $k = 2$  odchylenie standardowe rozpatrywanego rozkładu przyjmują wartość nieskończoną. Zauważmy też, że dla ustalonego  $k$  współczynnik  $\alpha$  - z uwagi na małość swych wartości - nie wpływa ani na kształt krzywej rozkładu prawdopodobieństwa w obszarze dużych wartości amplitud, ani też na stosunek wartości amplitud do wartości średniokwadratowych, ale wpływa na ten rozkład w obszarze małych amplitud.

Największe zagrożenie bezpieczeństwa inwestowanych pieniędzy powstanie wtedy, gdy amplituda zakłóceń impulsowych,  $A(t)$ , osiągnie lub przekroczy najwyższą dozwoloną wartość  $A^*(t)$  subiektywnie ustaloną przez właściciela kapitału. Ponieważ wartość ta może zmieniać się w zależności od różnych okoliczności (kondycja finansowa właściciela kapitału, stabilność gospodarki, techniczna i ekonomiczna gospodarność procesu inwestycyjnego, sytuacja na rynku finansowym itp.) więc jej matematycznym modelem może być zmienna losowa ciągła lub dyskretna. Tak więc warunkiem rozstrzygającym, czy miała lub ma miejsce sytuacja najgorsza, czy też nie, jest zachodzenie nierówności stochastycznej  $A(t) \geq A^*(t)$ . Skorzystać z niej można jednak tylko w odniesieniu do przeszłości lub teraźniejszości. Tymczasem dla właściciela kapitału znacznie ważniejsza byłaby możliwość określenia, jakie jest prawdopodobieństwo zdarzenia, że w określonej przyszłej chwili wystąpi zakłócenie, którego wielkość przekroczy ustaloną dla tej chwili normę, czyli  $P\{A(t) \geq A^*(t)\}$ . Stawiamy więc hipotezę zerową, że amplituda zakłócenia  $A(t) \geq A^*(t)$  i wprowadzamy skrócony zapis  $H_0(A \geq A^*)$ .

Dla zweryfikowania tej hipotezy jest potrzebny zbiór danych ilustrujących zmienność amplitudy zakłóceń procesu finansowego w przeszłości. Załóżmy, że dysponujemy takim zbiorem i że liczy on  $n$  elementów. Gdyby hipoteza  $H_0$  była prawdziwa, to średnia liczba zdarzeń  $A$  w zbiorze danych powinna być znacznie większa od  $A^*(t)$ . Sensowny jest więc następujący test hipotezy  $H_0$ : jeżeli średnia liczba zakłóceń impulsowych zarejestrowanych w zbiorze danych jest mniejsza od wartości  $A^*(t)$ , to hipotezę odrzucamy. W przeciwnym przypadku stwierdzamy, że nie ma podstaw do odrzucenia hipotezy i badamy funkcję mocy testu jako funkcję prawdziwej wartości amplitudy zakłócenia  $A(t)$ . Innymi słowy, sprawdzamy, jakie jest prawdopodobieństwo odrzucenia hipotezy  $H_0(A \geq A^*)$ . Maksymalna wartość tej funkcji czyli maksymalne prawdopodobieństwo odrzucenia hipotezy, gdy jest ona prawdziwa, nazywa się poziomem istotności testu. Praktyczną wartość mają tylko testy o małej wartości poziomu istotności. Należy pamiętać, że wybór poziomu istotności jest decyzją subiektywną. Dlatego każdą hipotezę staty-

styczną należy testować wielokrotnie, aby w jak największej mierze uwolnić wynik od elementu subiektywizmu.

Jeżeli jest znana gęstość prawdopodobieństwa  $a(t)$  zmiennej  $A(t)$ , to

$$P\{A(t) \geq A^*\} = \int_{A^*}^{\infty} a(u) du.$$

Umiejętność prawidłowego wykonania procedury testowania hipotezy o wielkości amplitudy zakłócenia procesu nakładów inwestycyjnych na naukę i wszelkich innych długofalowych procesów inwestycyjnych jest niezwykle ważna, ponieważ w wyniku jej wykonania inwestor otrzymuje statystyczną informację dotyczącą przypadku, który w jego subiektywnej ocenie uchodzi za najgorszy. Jest więc w stanie podjąć z wyprzedzeniem takie działania, które pozwolą mu uniknąć tego przypadku lub zminimalizują jego skutki.

### 2.3. Czasowy rozkład zakłóceń w serii

Założmy, że rozpatrywany horyzont czasowy został podzielony na kolejno po sobie następujące i całkowicie wypełniające go interwały czasowe  $T_1, T_2, \dots, T_n$  o jednakowej długości  $\hat{T}$ . Wybór długości odcinka zależy od celu i stopnia szczegółowości badania. W wielu z tych odcinków proces finansowania inwestycji będzie przebiegał w sposób zgodny z zamierzonym, ale w niektórych z nich zostanie zakłócony. Oznaczmy symbolem  $t_s$  chwilę, w której pojawiło się silne zakłócenie procesu, dające początek serii innych, podobnych zakłóceń, których amplitudy też przekraczają z góry ustalony poziom krytyczny. Założmy, że rozkład liczby zakłóceń w każdej serii jest przypadkowy. Oznaczmy przez  $\delta$  czas trwania serii. W przedziałach czasu między dwiema kolejnymi seriami prawdopodobieństwo wystąpienia zakłócenia impulsowego jest równe zero.

Dla obliczenia rozkładu liczby zakłóceń impulsowych w pojedynczej serii, a więc w odcinku czasu  $\delta$  założmy, że odcinek ten został podzielony na  $n$  małych podprzedziałów o jednakowej długości  $\Delta t$ ; wobec tego  $n = \frac{\delta}{\Delta t}$ . Założmy też, że każdy z

tych podprzedziałów jest na tyle krótki, że zasadniczo może w nim wystąpić nie więcej niż jedno zakłócenie impulsowe. Prawdopodobieństwo pojawienia się dwóch lub większej liczby takich zakłóceń w tym podprzedziale jest pomijalnie małe w porównaniu z prawdopodobieństwem wystąpienia co najwyżej jednego zakłócenia. W każdym podprzedziale  $\Delta t$  wystąpienie lub brak zakłócenia impulsowego mogą być traktowane jako wynik doświadczenia bernoulliańskiego. Jeżeli pojawienie się zakłócenia będziemy uważać za sukces, to prawdopodobieństwo tego zdarzenia będzie równe  $\mu \Delta t$ , gdzie  $\mu$  jest stałą dodatnią. Niech  $N(n\Delta t)$  oznacza liczbę zakłóceń impulsowych w odcinku  $\delta$ . Jeżeli przyjąć, że doświadczenia bernoulliańskie w kolejnych podprzedziałach  $\Delta t$  są niezależne, to liczba zakłóceń impulsowych pojawiających się w zdyskretyzowanym odcinku czasu

$\delta$  ma rozkład dwumianowy o parametrach  $n = \frac{\delta}{\Delta t}$  i  $p = \mu \Delta t$ . Ponieważ interesuje nas

rozkład prawdopodobieństwa liczby zakłóceń impulsowych w serii, gdy  $\Delta t \rightarrow 0$  (tzn. gdy liczba podprzedziałów odcinka  $\delta$  dąży do nieskończoności,  $n \rightarrow \infty$ ), więc

$$\begin{aligned} \lim_{n \rightarrow \infty} P\{N(t) = k\} &= \lim_{n \rightarrow \infty} \binom{n}{k} (\mu \Delta t)^k (1 - \mu \Delta t)^{n-k} = \\ &= \frac{(\mu t)^k}{k!} e^{-\mu t}, \end{aligned} \quad (34)$$

gdzie  $k = 0, 1, 2, \dots$  (zob., np., Fisz, 1968; Plucińska i Pluciński, 2000). Otrzymaliśmy więc, że liczba zakłóceń impulsowych w serii jest realizacją zmiennej losowej o rozkładzie Poissona.

Oznaczmy przez  $\theta$  odstęp czasu, jaki upłynął od chwili pojawienia się pierwszego zakłócenia impulsowego w serii do końca interwału czasowego  $\hat{T}$ , w którym się ona zaczęła (seria zakłóceń nie musi mieścić się całkowicie w jednym interwale czasu). Jeżeli długość odcinka  $\theta$  nie przekracza długości odcinka  $\delta$ , to prawdopodobieństwo zdarzenia, że w odcinku  $\theta$  nie pojawi się ani jedno zakłócenie impulsowe wyraża się wzorem

$$p(0, \theta) \stackrel{\text{df}}{=} P\{N(\theta) = 0\} = e^{-\mu \theta}. \quad (35)$$

Jeżeli jednak  $\theta$  jest większe niż  $\delta$ , ale nie większe niż  $\hat{T}$ , to prawdopodobieństwo tego, że w przedziale  $\delta$  proces finansowania nie zostanie zaburzony przez zakłócenie impulsowe wynosi

$$p(0, \delta) \stackrel{\text{df}}{=} P\{N(\delta) = 0\} = e^{-\mu \delta}. \quad (36)$$

Rozpatrzmy teraz dwa pierwsze interwały czasowe,  $T_1$  i  $T_2$ . Niech  $q(z, \tau)$  oznacza prawdopodobieństwo wystąpienia zakłócenia w przedziale czasu  $\tau$ . Na podstawie (35) prawdopodobieństwo wystąpienia zakłócenia impulsowego w interwale  $T_1$  opisują formuły

$$q(z, \theta) = 1 - p(0, \theta) = 1 - e^{-\mu \theta} \quad (37)$$

i

$$q(z, \delta) = 1 - p(0, \gamma) = 1 - e^{-\mu \delta}. \quad (38)$$

Analogicznie do formuł (35) i (37), jeżeli  $\theta \leq \delta$ , to prawdopodobieństwo niewystąpienia bądź wystąpienia w interwale  $T_2$  zakłóceń impulsowych rozważanego procesu wyraża się odpowiednio formułami

$$p(0, \delta - \theta) = e^{-\mu(\delta - \theta)}, \quad (39)$$

$$q(z, \delta - \theta) = 1 - e^{-\mu(\delta - \theta)}. \quad (40)$$

Jeżeli jednak  $\theta > \delta$ , to - przez analogię do formuł (36) i (38) - prawdopodobieństwa te wynoszą odpowiednio

$$p(0, \theta) = 1, \quad (41)$$

$$q(z, \theta) = 0. \quad (42)$$

Z punktu widzenia praktyki finansowej ważna jest umiejętność oszacowania średniej liczby zakłóceń impulsowych w interwale  $\hat{T}$ . Aby wyznaczyć tę średnią rozpatrzmy zbiór wszystkich możliwych wersji pary interwałów,  $T_1$  i  $T_2$ , otrzymanych przez przesuwanie chwili rozpoczęcia się serii zakłóceń w każdym interwale o odcinek  $\theta$ . Załóżmy, że zbiór ten liczy  $M$  elementów. Oznaczmy przez  $l_1$  i  $l_2$  liczby zakłóceń impulsowych odpowiednio w pierwszym i drugim z tych interwałów, a przez  $l$  ich łączną liczbę w obu. Jeżeli przez  $q(z, T_1)$  oznaczymy prawdopodobieństwo pojawienia się zakłóceń impulsowych w interwale  $T_1$ , a przez  $q(z, T_2)$  w interwale  $T_2$  (patrz formuły (37), (38), (40) i (42)), to łączna liczba tych spośród  $M$  wszystkich interwałów, w których pojawiły się takie zakłócenia wynosi  $M[q(z, T_1) + q(z, T_2)]$ . Wobec tego średnia liczba zakłóceń impulsowych w interwale należącym do rozpatrywanego zbioru wynosi

$$\frac{l}{M[q(z, T_1) + q(z, T_2)]}, \text{ a ponieważ } \frac{l}{M} = \mu\delta, \text{ więc jest ona równa } \frac{\mu\delta}{q(z, T_1) + q(z, T_2)}.$$

Ponieważ długości serii zakłóceń impulsowych mogą zmieniać się w sposób losowy, w zakresie od 0 (brak zakłóceń impulsowych w interwale) do  $\hat{T}$  (seria zakłóceń wypełnia cały interwał), więc konieczne jest jeszcze uśrednienie ostatniej wielkości względem  $\theta$  w tym zakresie. Załóżmy, że  $\theta$  zmienia się w sposób dyskretny, mogąc przyjmować  $m$  wartości ( $\theta_1 = 0, \theta_2 = 1, \dots, \theta_m = \hat{T}$ ). Skoro tak, to - po uwzględnieniu tej gradacji - liczba wszystkich możliwych par interwałów wyniesie  $mM$ . Niech  $q(z, T_1, \theta_j)$  będzie prawdopodobieństwem tego, że w interwale  $T_1$  pojawiła się seria zakłóceń impulsowych mająca długość  $\theta_j = j$ , a  $q(z, T_2, \theta_j)$  prawdopodobieństwem tego, że zdarzenie to miało miejsce w interwale  $T_2$ , ( $j = 1, 2, \dots, m$ ). Jeżeli  $L$  jest łączną liczbą zakłóceń zarejestrowanych we wszystkich  $mM$  parach interwałów, to ogólna liczba interwałów, w których pojawiły się zakłócenia impulsowe, wynosi

$$\begin{aligned} & M[q(z, T_1, 1) + q(z, T_1, 2) + \dots + q(z, T_1, m)] + M[q(z, T_2, 1) + q(z, T_2, 2) + \dots + q(z, T_2, m)] = \\ & = M \sum_{j=1}^m [q(z, T_1, j) + q(z, T_2, j)]. \end{aligned} \quad (43)$$

Z drugiej strony, zauważmy, że średnie liczby zakłóceń impulsowych w interwałach  $T_1$  i  $T_2$  wynoszą odpowiednio

$$q(z, T_1) = \frac{1}{m} \sum_{j=1}^m q(z, T_1, j) \quad (44a)$$

$$q(z, T_2) = \frac{1}{m} \sum_{j=1}^m q(z, T_2, j). \quad (44b)$$

Wobec tego

$$M \sum_{j=1}^m [q(z, T_1, j) + q(z, T_2, j)] = mM [\overline{q(z, T_1)} + \overline{q(z, T_2)}] \quad (45a)$$

czyli

$$\sum_{j=1}^m [q(z, T_1, j) + q(z, T_2, j)] = m [\overline{q(z, T_1)} + \overline{q(z, T_2)}]. \quad (45b)$$

Średnia liczba zakłóceń impulsowych  $Z$  w interwale czasowym, w którym takie zakłócenia wystąpiły, wynosi

$$Z = \frac{L}{mM [\overline{q(z, T_1)} + \overline{q(z, T_2)}]}, \quad (46a)$$

lub po uwzględnieniu, że  $mM = L\mu\delta$ ,

$$Z = \frac{\mu\delta}{\overline{q(z, T_1)} + \overline{q(z, T_2)}}. \quad (46b)$$

Zauważmy, że w miarę zagęszczania skali zmienności parametru  $\theta$ , mianownik wyrażenia (46b) można z coraz większą dokładnością aproksymować formułą

$$\overline{q(z, T_1)} + \overline{q(z, T_2)} \approx \int_0^{\delta} [q(z, T_1) + q(z, T_2)] d\theta. \quad (47)$$

Obliczmy wartość tej całki. Pamiętając o formułach (37) - (40) otrzymujemy

$$\int_0^{\delta} [q(z, T_1) + q(z, T_2)] d\theta = \int_0^{\delta} (1 - e^{-\mu\theta}) d\theta + \int_{\delta}^{\hat{T}} (1 - e^{-\mu\delta}) d\theta + \int_0^{\delta} (1 - e^{-\mu(\delta-\theta)}) d\theta. \quad (48)$$

Po wykonaniu całkowania i uporządkowaniu wyników dostajemy

$$\int_0^{\delta} [q(z, T_1) + q(z, T_2)] d\theta = \left(\delta - \frac{2}{\mu} + \hat{T}\right) + e^{-\mu\delta} \left(\delta + \frac{2}{\mu} - \hat{T}\right). \quad (49)$$

Wobec tego formuła (46b) przyjmuje postać

$$Z = \frac{\mu\delta}{\left(\delta - \frac{2}{\mu} + \hat{T}\right) + e^{-\mu\delta} \left(\delta + \frac{2}{\delta} - \hat{T}\right)} \quad (50a)$$

lub

$$Z = \frac{(\mu\delta)^2}{(\delta^2\mu - 2\delta + \mu\delta\hat{T}) + \mu e^{-\mu\delta} (\delta^2 + 2 - \delta\hat{T})} \quad (50b)$$

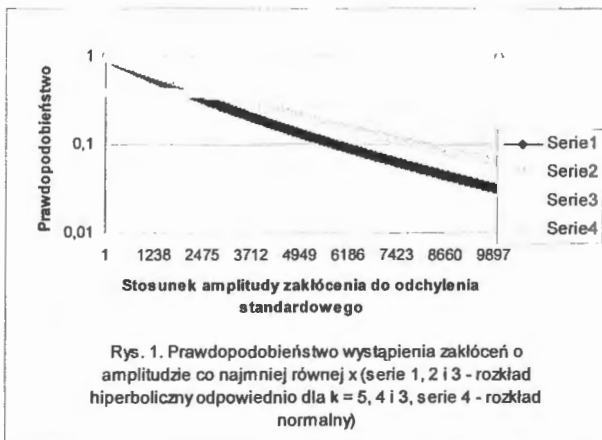
Widzimy więc, że średnia liczba zakłóceń impulsowych w interwale zależy od trzech wielkości: długości interwału  $\hat{T}$ , intensywności ich pojawiania się  $\mu$  oraz długości czasu trwania ich serii.

### 3. Wstępna analiza danych empirycznych

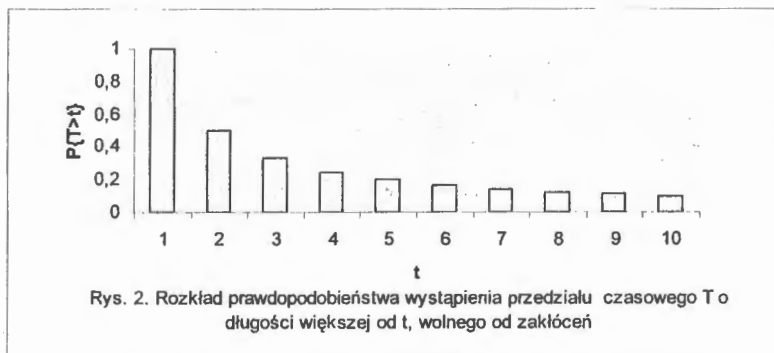
Dla przeanalizowania ewentualnej przydatności proponowanego modelu zakłóceń strumienia nadładów na edukację jest potrzebny długi szereg czasowy, którego elementami będą w jakikolwiek sposób wyrażone relacje między planowanymi a faktycznie poniesionymi nakładami w ustalonych okresach czasu. Autorzy pracy nie dysponowali tego rodzaju danymi dla uczelni krajowej. Skorzystano więc z danych dotyczących jednego z uniwersytetów zagranicznych (Kuwait University). Jego działalność została oficjalnie zainaugurowana w październiku 1966 r. i od tego czasu corocznie są publikowane bardzo szczegółowe statystyki charakteryzujące pełną działalność edukacyjną, finansową oraz administracyjną tej uczelni jako całości, a także poszczególnych fakultetów i działów wspomagających proces edukacyjny (KU, 2002). W oparciu o te dane skonstruowano zerowy szereg czasowy zawierający trzydzieści pięć elementów, w którym symbol 0 był przyporządkowywany tym latom, w których ogólne wydatkowane nakłady finansowe mieściły się w dopuszczalnych granicach tolerancji, zaś symbol 1 tym, w których nakłady te przekroczyły ustalone wcześniej limity. Uzyskano więc ciąg, w którym serie zer (okresy czasu, w których nie występowały impulsowe zakłócenia strumienia nakładów) przeplatały się z bardzo krótkimi w stosunku do nich seriami jedynek (okresów, w których miały miejsce silne zakłócenia strumienia nakładów). Długości odcinków czasu między zakłóceniami impulsowymi rozpatrzono (1) jako realizacje standaryzowanej zmiennej losowej gaussowskiej o gęstości

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2}}, \quad (51)$$

oraz (2) jako realizacje zmiennej paretowskiej - formuła (19) - dla  $n = 3, 4$  i  $5$ . Zmienność prawdopodobieństwa wystąpienia serii zakłóceń impulsowych jako funkcji stosunku długości serii  $x$  do odchylenia standardowego  $\sigma$  jest zilustrowana na rys. 1. Jak widać, krzywe paretowskie zdecydowanie lepiej odwzorowują prawidłowości rozpatrywanego szeregu czasowego niż krzywa gaussowska.



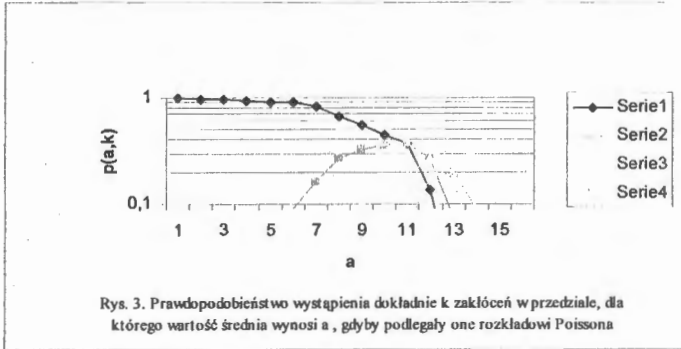
Rozkład serii zakłóceń impulsowych w czasie najwygodniej jest opisać w kategoriach odstępów czasu między nimi. Na rys. 2 pokazano jak zmienia się empiryczne prawdopodobieństwo zdarzenia, że odstęp czasu  $T$  jaki upłynie od chwili wystąpienia silnego zakłócenia strumienia nakładów do chwili pojawienia się następnego takiego zakłócenia będzie wynosił co najmniej  $t$  jednostek czasowych (zob. formuła (7)). Jak widać, krzywa tego rozkładu potwierdza paretowski bądź hiperboliczny charakter rozkładu zmiennej losowej



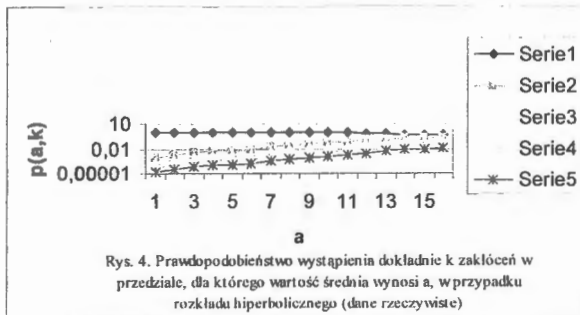
$T$ . Znacznie ważniejszy jest jednak wniosek, że kształt tej krzywej zdecydowanie przeczy możliwości przyjęcia, iż zmienna losowa  $T$  ma rozkład normalny. Tymczasem w dotychczasowych modelach strumieni nakładów – nie tylko w sferze edukacji – powszechnie

przyjmuje się jako fakt oczywisty, że długości odstępów czasu między zakłóceniami tych strumieni są realizacjami zmiennych losowych gaussowskich.

Zdecydowanie odrzucając jako niepoprawne teoretycznie i empirycznie założenie o normalności rozkładu zmiennej losowej  $T$  zastanówmy się czy i na ile dopuszczalne jest przyjmowanie, że matematycznym modelem procesu zakłóceń strumienia nakładów jest proces Poissona. Na rys. 3 przedstawiono rodzinę wykresów ilustrujących, jak musiałoby zmieniać się prawdopodobieństwo wystąpienia dokładnie  $k$  zakłóceń w przedziale czasu, gdyby proces ich pojawiania się był poissonowski o wartości średniej  $a$ . Są to krzywe



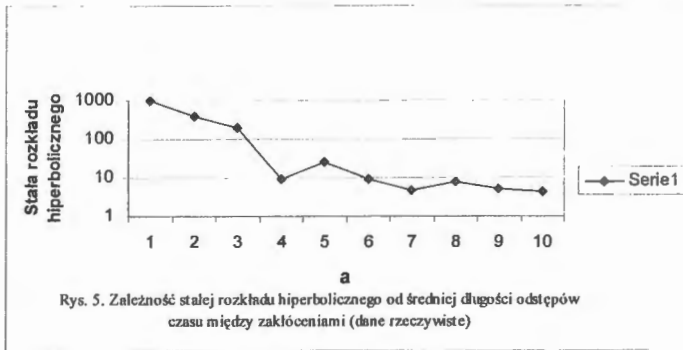
teoretyczne. Należy je porównać z kształtem krzywych empirycznych z rys. 4, które są zupełnie inne niż miałyby to miejsce w przypadku, gdyby proces zakłóceń miał charakter poissonowski. Charakter tych krzywych znów wskazuje na to, że hipoteza o parętowskim lub hiperbolicznym rozkładzie odstępów czasu między zakłóceniami jest bliższa rzeczywistości, niż jednoznacznie związana z procesem Poissona hipoteza o jego wykładniczości.



W modelu hiperbolicznym wprowadziliśmy stałą rozkładu  $\alpha$  ( $\alpha > 0$ ), mającą charakter parametru skali. Ważne znaczenie dla poznania rzeczywistych mechanizmów rządzą-



cych procesem zakłóceń strumienia nakładów jest identyfikacja postaci zależności między tą stałą a średnią wartością  $a$  długości odstępów czasu między zakłóceniami. Empiryczna zależność tych wielkości jest zilustrowana na rys. 5.



Zauważmy, że formuły (28) i (29), określające wartość oczekiwaną zmiennej losowej o rozkładzie hiperbolicznym, dotyczą sytuacji czysto teoretycznej, dopuszczającej zmienność amplitudy zakłóceń w zakresie od 0 do nieskończoności. W rzeczywistości taka idealizacja nigdy nie ma miejsca, bowiem każdy eksperyment i wszelkie wielkości w nim obserwowane są zawsze skończone. Amplituda zakłócenia może być bardzo duża, ale jest skończona. Zasadne jest więc przyjąć, że jej wielkość zmienia się od 0 do pewnej wartości  $x_{\max}$ . Z matematycznej postaci gęstości rozkładu hiperbolicznego (formuła (23)) wynika, że należy odrębnie rozpatrzeć przypadek, gdy  $k = 2$ . Wtedy funkcja gęstości rozkładu ma postać

$$h(x) = \frac{\alpha}{(x + \alpha)^2}, \quad (52)$$

zaś empiryczna wartość oczekiwana wyraża się wzorem:

$$\tilde{E}(X) = \alpha \int_0^{x_{\max}} \frac{xdx}{(x + \alpha)^2}. \quad (53)$$

Po podstawieniu  $u = x + \alpha$  i wykonaniu niezbędnych przekształceń otrzymamy:

$$\tilde{E}(X) = \alpha \int_{\alpha}^{x_{\max} + \alpha} \frac{du}{u} - \alpha^2 \int_{\alpha}^{x_{\max} + \alpha} \frac{du}{u^2}, \quad (54)$$

skąd, po scałkowaniu,

$$\tilde{E}(X) = \alpha \left[ \ln \frac{x_{\max} + \alpha}{\alpha} - \left( 1 + \frac{\alpha}{x_{\max} + \alpha} \right) \right]. \quad (55)$$

Zauważmy teraz, że dla  $k = 2$ , teoretyczne prawdopodobieństwo zdarzenia, że amplituda zakłócenia przyjmie wartość większą od ustalonej wartości  $x_{\max}$  wynosi:

$$P\{X \geq x_{\max}\} = \alpha \int_{x_{\max}}^{\infty} \frac{dz}{(z + \alpha)^2}, \quad (56)$$

tj.

$$P\{X \geq x_{\max}\} = \frac{\alpha}{x_{\max} + \alpha}. \quad (57)$$

Stwierdzamy więc, że jeżeli rzeczywiste wielkości amplitud zakłóceń mają rozkład hiperboliczny rzędu 2, to dla obliczenia empirycznej średniej wartości  $\tilde{E}(X)$  wystarczy znajomość teoretycznego prawdopodobieństwa przekroczenia przez amplitudę pewnej ustalonej wartości progowej.

W przypadku, gdy  $k \neq 2$  empiryczna wartość średnia amplitudy zakłócenia wyrazi się wzorem:

$$\tilde{E}(X) = (k-1)\alpha^{k-1} \int_0^{x_{\max}} \frac{x dx}{(x + \alpha)^k}. \quad (58)$$

Stąd

$$\tilde{E}(X) = (k-1)\alpha^{k-1} \left[ -\frac{1}{(k-2)(x + \alpha)^{k-2}} + \frac{\alpha}{(k-1)(x + \alpha)^{k-1}} \right]_0^{x_{\max}},$$

zaś po wykonaniu stosownych przekształceń otrzymujemy:

$$\tilde{E}(X) = \alpha \left\{ \left[ \left( \frac{\alpha}{x_{\max} + \alpha} \right)^{k-1} - 1 \right] - \frac{k-1}{k-2} \left[ \left( \frac{\alpha}{x_{\max} + \alpha} \right)^{k-2} - 1 \right] \right\}. \quad (59)$$

Wyrażenie to ma sens fizyczny tylko wtedy, gdy  $\tilde{E}(X) \geq 0$ . Ponieważ  $\alpha$  jest dowolnie małą stałą dodatnią, więc czynnik zawarty w nawiasach klamrowych musi mieć wartość nieujemną. Na podstawie danych empirycznych należy zatem tak dobrać wartość  $k$  (rzęd hiperboli), żeby  $\tilde{E}(X) \geq 0$ . Prowadzi to do nierówności:

$$\left( \frac{\alpha}{x_{\max} + \alpha} \right)^{k-1} - 1 \geq \frac{k-1}{k-2} \left[ \left( \frac{\alpha}{x_{\max} + \alpha} \right)^{k-2} - 1 \right],$$

czyli

$$\left( \frac{\alpha}{x_{\max} + \alpha} \right)^{k-2} \left[ \frac{\alpha}{x_{\max} + \alpha} - \frac{k-1}{k-2} \right] \leq \frac{1}{k-2}.$$

Ponieważ wyrażenie  $\left( \frac{\alpha}{x_{\max} + \alpha} \right)^{k-2}$  jest zawsze dodatnie, więc

$$\frac{\alpha}{x_{\max} + \alpha} - \frac{k-1}{k-2} \leq \frac{1}{k-2} \left( \frac{\alpha}{x_{\max} + \alpha} \right)^{-(k-2)}.$$

Wchodzą w grę dwie możliwości:  $0 < k < 2$  lub  $k > 2$ .

W pierwszym przypadku  $\frac{1}{k-2} < 0$ . Wobec tego musi być spełniony warunek  $\frac{\alpha}{x_{\max} + \alpha} - \frac{k-1}{k-2} \leq 0$ , z którego wynika, że  $k \leq 1 - \frac{\alpha}{x_{\max}}$ . Stwierdziliśmy więc, że jeżeli aproksymujemy empiryczny rozkład wielkości amplitud zakłóceń za pomocą hiperbolicznego rozkładu prawdopodobieństwa rzędu  $0 < k < 2$ , to rząd ten nie może być większy od  $1 - \frac{\alpha}{x_{\max}}$ .

Zauważmy ponadto, że również w tym przypadku dla oszacowania empirycznej wartości  $\tilde{E}(X)$  można skorzystać ze znajomości prawdopodobieństwa teoretycznego

$$P\{X \geq x_{\max}\} = \frac{\alpha}{x_{\max} + \alpha}$$

odpowiadającego rozkładowi hiperbolicznemu rzędu 2.

#### 4. Zakończenie i wnioski

Podstawowym celem pracy było przeanalizowanie teoretycznych założeń przyjmowanych w stochastycznych modelach strumieni zakłóceń długofalowych procesów inwestycyjnych, np. w dziedzinie edukacji. Przez zakłócenie rozumie się tutaj każde zaburzenie zaplanowanego procesu finansowego zasilania edukacji, przekraczające dopuszczalne przedziały tolerancji. Stwierdzono, że podstawowym mankamentem istniejących modeli jest przyjmowanie, że zakłócenia są zdarzeniami występującymi niezależnie od siebie i rozkładają się w czasie w sposób czysto przypadkowy, czego konsekwencją jest opisywanie strumienia zakłóceń za pomocą procesu Poissona. Z poissonowskiego modelu rozkładu liczby zakłóceń w czasie wynika, że rozkład prawdopodobieństwa zmiennej losowej, której realizacjami są długości odstępów czasu między zakłóceniami jest wykładniczy. Wykazano, że zarówno z teoretycznego punktu widzenia (rozdz. 2), jak i w świetle posiadanych materiałów empirycznych (rozdz. 3) są to założenia zbyt abstrakcyjne. Podobnie nierealne – chociaż wygodne teoretycznie – jest przyjmowanie, że wielkości amplitud zakłóceń mają rozkład normalny. I w tym przypadku zarówno przesłanki logiczne, jak i materiał empiryczny wskazują, że musi to być rozkład silnie niesymetrycz-

ny. Co więcej, model strumienia musi uwzględniać fakt niejednorodności i stochastycznej zależności zakłóceń, czego empirycznym wyrazem jest ich grupowanie się w serie.

Jest oczywiste, że stochastycznego modelu strumienia spełniającego te postulaty należałoby poszukiwać na gruncie teorii procesów Markowa. Z drugiej strony powstaje jednak pytanie, czy jest to jedyne alternatywne podejście w stosunku do dotychczasowych. Myślą przewodnią pracy było wykazanie, że dobrą aproksymację rzeczywistości, dostatecznie wiernie odwzorowującą zjawisko seryjności występowania impulsowych zakłóceń procesów finansowych oraz asymetrię rozkładu ich wielkości można otrzymać jeżeli zauważy się, że rzeczywiste ciągi zakłóceń tych procesów w długich okresach czasu wykazują empirycznie stwierdzaną cechę samopodobieństwa. Zbudowano więc model, w którym założono, że długości odstępów czasu między seriami zakłóceń mają rozkład typu paretońskiego, zaś wielkości zakłóceń mają rozkład hiperboliczny. Z uwagi na niejednorodność zakłóceń przyjęto, że przedziały czasu wolne od zakłóceń impulsowych są sumami skończonej liczby krótszych odcinków, będących realizacjami różnych zmiennych losowych paretońskich. Pokazano, że gęstość prawdopodobieństwa zmiennej losowej, której realizacjami są te odstępów, można aproksymować za pomocą bardzo wygodnej do przeprowadzania obliczeń formuły rekurencyjnej (19).

Zwrócono uwagę, że szczególnie trafnie powinien być określony rozkład prawdopodobieństwa wielkości amplitud zakłóceń, ponieważ jest on wykorzystywany do oszacowania prawdopodobieństwa wystąpienia tzw. najgorszego dla inwestora przypadku. Oszacowanie to dostarcza inwestorowi informacji, z której może skorzystać przy podejmowaniu decyzji o ewentualnym udziale i warunkach uczestniczenia w procesie finansowania inwestycji. Pokazano, że dokładność oszacowania średniej liczby silnych zakłóceń tworzących serię zależy od sposobu dyskretyzacji osi czasu, od ich intensywności oraz od długości czasu trwania serii zakłóceń i wyprowadzono stosowną formułę matematyczną (wzór (50b)).

Podstawowe postulaty matematyczne, wynikające z przeprowadzone w pracy logicznej analizy strumienia zakłóceń procesu inwestowania edukacji w długim horyzoncie czasowym, zostały zweryfikowane w oparciu o dane rzeczywiste dotyczące jednej z uczelni zagranicznych. Potwierdzona została słuszność przyjętych założeń. Wydaje się więc, że dalsze badania w tym kierunku powinny być prowadzone.

## Bibliografia

- Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, III-17, 21-86.
- Campbell G.A. (1923). Probability curves, showing poisson's exponential summation. *Bell System Technical Journal*, 2, 95-113.
- Embrechts P., Klüppelberg C., Mikosch T. (2001). Modelling extremal events for insurance and finance. Springer-Verlag, New York.
- Fama E.F. (1965). The behavior of stockmarket prices. *Journal of Business*, 38, 34-105.
- Fama E.F. (1970). Efficient capital markets: a review of theory and empirical work. *Journal of Finance*, 25, 383-417.
- Feller W. (1969). Wstęp do rachunku prawdopodobieństwa. T. II. PWN, Warszawa.

- Fisz M. (1968). Rachunek prawdopodobieństwa i statystyka matematyczna. PWN, Warszawa.
- Gibrat A. (1932). Les inégalités économiques. Sirey, Paris.
- Gniedenko B.W., Kofmogorow A.N. (1957). Rozkłady graniczne sum zmiennych losowych niezależnych. PWN, Warszawa.
- Graham R.L., Knuth D.E., Patashnik O. (1996). Matematyka konkretna. Wydawnictwo Naukowe PWN, Warszawa.
- Gumbel E. (1962). Statistics of extremes. Columbia University Press, New York.
- Kamke E. (1959). Differentialgleichungen. Lösungsmethoden und Lösungen. Teubner Verlagsgesellschaft, Leipzig.
- Kijima M. (1997). Markov processes for stochastic modeling. Chapman & Hall, London.
- Kotz S., Nadarajah S. (2000). Extreme value distributions. Imperial College Press, London.
- KU (2002). Kuwait University. A statistical review from 1966/67 to 1999/2000. Kuwait University Press, Kuwait.
- Kulikowski R., Libura M., Słomiński L. (1998). Wspomaganie decyzji inwestycyjnych. Instytut Badań Systemowych PAN, Warszawa.
- Lévy P. (1925). Calcul des probabilités. Gauthier-Villars, Paris.
- Leadbetter M.R., Lindgren G., Rootzen H. (1986). Extremes and related properties of random sequences and processes. Springer-Verlag, New York.
- Lintner J. (1965a). Security prices, risk, and maximal gains from diversification. *Journal of Finance*, **20**, 587-615.
- Lintner J. (1965b). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, **47**, 13-37.
- Mandelbrot B.B. (1958). Les lois statistiques macroscopiques du comportement (rôle de la loi de Gauss et des lois de Paul Lévy). *Psychologie Française*, **3**, 237-249.
- Mandelbrot B.B. (1959). Variables et processus stochastiques de Pareto-Lévy, et de la répartition des revenus. *Comptes Rendus Hebdomadaires des Sciences de l'Académie des Sciences*, **249**, 613-615, 2153-2155.
- Mandelbrot B.B. (1966). Nouveaux modèles de la variation des prix (cycles lents et changements instantanés). *Cahiers du Séminaire d'Économétrie*, **9**, 53-66.
- Mandelbrot B.B. (1982). The many faces of scaling: fractals, geometry of nature and economics. W: W.C. Schieve i P.M. Allen, eds., *Self-organization and dissipative structures*. Springer-Verlag, New York.
- Mandelbrot B.B., Taylor H.M. (1967). On the distribution of stock price differences. *Operations Research*, **15**, 1057-1062.
- Markowitz H.M. (1952). Portfolio selection. *Journal of Finance*, **7**, 77-91.
- Markowitz H.M. (1959). Portfolio selection: efficient diversification of investments. John Wiley, New York.
- Mirowski P. (1990). From Mandelbrot to chaos in economic theory. *Southern Economic Journal*, **57**, 289-307.
- Mirowski P. (1996). Mandelbrot's economics after a quarter century. *Fractals*, **3**, 181-196.
- Pareto V. (1897). Cours d'économie politique. Rouge & Cie, Lausanne & Paris.
- Parzen E. (1962). Stochastic processes. Holden-Day, San Francisco.
- Pietrow W.W. (1972). Summy niezawisomych słuczajnych wieliczin. Nauka, Moskwa.
- Plucińska A., Pluciński E. (2000). Rachunek prawdopodobieństwa. Statystyka matematyczna. Procesy stochastyczne. WNT, Warszawa.
- Runggaldier W.J., ed. (1997). Financial mathematics. Springer-Verlag, New York.

- Samuelson P.A. (1967). Efficient portfolio selection for Pareto-Lévy investments. *Journal of Financial and Quantitative Analysis*, **2**, 107-122.
- Sharpe W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. *Journal of Finance*, **29**, 425-442.
- Tobin J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies*, **25**, 65-86.
- Willett A.H. (1951). The economic theory of risk insurance. University of Pennsylvania Press, Philadelphia.
- Zołotariew W.M. (1986). Sowriemiennaja teorija summirowanija niezawisimych słuczajnych wieliczin. Nauka, Moskwa.



