

KU 0238 pozycja wydawnictw naukowych
Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie

© Wydawnictwa AGH, Kraków 2006
ISBN 83-7464-102-9

Redaktor Naczelny Uczelnianych Wydawnictw
Naukowo-Dydaktycznych AGH: *Jan Sas*

Z-ca Redaktora Naczelnego: *Beata Barszczewska-Wojda*

Komitety Naukowy UWND AGH:
Kazimierz Jeleń (przewodniczący),
Edward Fraś,
Tadeusz Sawik,
Ryszard Uberman,
Adam Paweł Wojda,
Mariusz Ziółko

Recenzent: *prof. zw. dr hab. inż. Ryszard Tadeusiewicz*

Redakcja techniczna: *Beata Barszczewska-Wojda*

Redakcja: *Małgorzata Koch*

Projekt okładki i strony tytułowej: *Jadwiga Kostrzevska*

Skład komputerowy: „Andre”, tel. 012 423 10 10

Druk i oprawa: „DjaF”, tel. 012 636 32 40, www.djaf.pl

Redakcja Uczelnianych Wydawnictw Naukowo-Dydaktycznych AGH
al. Mickiewicza 30, 30-059 Kraków
tel. 012 617 32 28, tel./fax 012 636 40 38
e-mail: redakcja@wydawnictwoagh.pl
www.WydawnictwoAGH.pl

MACIEJ KOSTRZEWSKI

**BAYESOWSKA ANALIZA
FINANSOWYCH SZEREGÓW CZASOWYCH
MODELOWANYCH PROCESAMI DYFUZJI**



Spis treści

Wstęp	7
1. Procesy dyfuzji stosowane w analizie danych finansowych	15
1.1. Uwagi wstępne	15
1.2. Procesy dyfuzji – definicje i własności	16
1.3. Stochastyczne równania różniczkowe Itô – definicje i własności	24
1.4. Dyskretna aproksymacja rozwiązań SRR	29
1.5. Wybrane procesy dyfuzji Itô	34
1.6. Uwagi końcowe	45
2. Aproksymacja gęstości przejścia procesów Itô	47
2.1. Uwagi wstępne	47
2.2. Schematy dyskretyzacji	48
2.3. Metoda zmiennych pośrednich	50
2.4. Metoda wielomianów Hermite’a	54
2.5. Metoda rozwiązań równań Fokkera, Plancka i Kołmogorowa	67
2.6. Uwagi końcowe	74
3. Wnioskowanie bayesowskie dla procesów Itô	75
3.1. Uwagi wstępne	75
3.2. Ogólne zasady wnioskowania bayesowskiego	76
3.3. Bayesowskie porównanie konkurencyjnych modeli	78
3.4. Estymacja, predykcja oraz porównanie modeli zadanych przez wybrane SRR	81
3.5. Metody numeryczne wykorzystywane we wnioskowaniu bayesowskim	91
3.6. Bayesowska estymacja parametrów przy zastosowaniu metody zmiennych pośrednich	101
3.7. Uwagi końcowe	111
4. Modelowanie stopy procentowej FED	112
4.1. Uwagi wstępne	112

4.2. Prezentacja danych	113
4.3. Wyniki estymacji nieznanych parametrów modeli	115
4.4. Wyniki porównania modeli	121
4.5. Prognoza poziomu stopy procentowej FED	124
4.6. Uwagi końcowe	127
5. Modelowanie indeksu WIG20	128
5.1. Uwagi wstępne	128
5.2. Prezentacja danych	129
5.3. Estymacja nieznanych parametrów	131
5.4. Porównanie modeli	135
5.5. Prognoza poziomu indeksu WIG20	138
5.6. Wycena europejskich opcji kupna	139
5.7. Uwagi końcowe	145
Zakończenie	147
Literatura	153

*Pracę dedykuję żonie Jadzi oraz córkom Joasi i Urszulce
z podziękowaniem za pomoc, cierpliwość i wyrozumiałość*

Wstęp

W nowoczesnej ekonometrii powszechne jest przekonanie o losowym charakterze zjawisk finansowych. Proces stochastyczny, będący rodziną zmiennych losowych określoną względem ustalonej osi czasu, jest obecnie najważniejszym narzędziem opisu pojawiających się wraz z upływem czasu zdarzeń losowych, w tym finansowych szeregów czasowych.

W 1906 roku A.A. Markow rozpoczął badania nad zdarzeniami zależnymi, wśród których wyodrębnił schemat doświadczeń powiązanych w łańcuch bez tzw. efektów następstwa. Wyniki jego pracy doprowadziły do rozwoju klasy procesów stochastycznych, posiadających własność braku pamięci, zwanych od jego nazwiska procesami Markowa (por. Fisz 1967; Jermakow 1976; Sobczyk 1996). Istotny wpływ na rozwój teorii procesów Markowa mieli A.N. Kołmogorow i W. Feller. Stochastyczne procesy Markowa można klasyfikować zależnie od tego, czy oś czasu oraz przestrzeń stanów (tzn. zbiór możliwych wartości procesu) są zbiorami dyskretnymi, tzn. o przeliczalnej liczbie elementów, czy ciągłymi, tzn. zbiorami o nieprzeliczalnej liczbie elementów, np.: przedziałami liczbowymi (por. Bailey 1990). W niniejszej monografii do opisu finansowych szeregów czasowych wykorzystywane są procesy o ciągłej osi czasu i nieprzeliczalnej przestrzeni stanów nazywane potocznie procesami ciągłymi.

Wśród ciągłych procesów Markowa na szczególną uwagę zasługują procesy dyfuzji. Ich nazwa wywodzi się stąd, że początkowo procesy te traktowane były jako matematyczne modele ruchu pojedynczej cząsteczki w zjawisku dyfuzji, tzn. w zjawisku przenikania cząsteczek jednej substancji w obręb drugiej na skutek chaotycznego ruchu molekuł (por. Wentzell 1980). R. Brown w 1827 roku wykorzystał do opisu ruchu cząsteczki w zjawisku dyfuzji (nazwanego od jego nazwiska ruchem Browna) pewien proces stochastyczny, który nazywany jest obecnie procesem Browna. Nad precyzyjną matematyczną teorią opisującą ruch Browna pracowali m.in. A. Einstein (por. Einstein 1905, 1906), M. Smoluchowski (por. Smoluchowski 1906) i N. Wiener (por. Wiener 1923). Dzięki pracom tego ostatniego, proces ruchu Browna stał się jednym z najważniejszych procesów losowych, wykorzystywanym w coraz to szerszych obszarach zastosowań (por. Gajek i Kałużska 1996; Platen 1999, Weron i Weron 1999). W literaturze przedmiotu zamiennie stosuje się pojęcie ruchu Browna i procesu Wienera (*Wiener process*). Jest to uzasadnione, gdyż definicja procesu Wienera w standardowej postaci

jest równoważna procesowi Browna (por. Lipcer i Szirajew 1981; Neftci 2000). Procesy dyfuzji stosowane są również do opisu zupełnie innych zjawisk, np.: są wykorzystywane do prognozy koncentracji szkodliwych zanieczyszczeń, służą do opisu procesu utleniania się krzemu, pojawiają się również jako procesy graniczne w biologicznych modelach dyskretnych opisujących zmiany w czasie liczby osobników określonego gatunku czy też zmiany koncentracji genów w populacji (por. Korbicz i Zacharewicz Zgurowski 1991). Zwiększające się zainteresowanie zagadnieniami kształtowania się cen aktywów finansowych oraz chęć lepszego ich opisu wymusiły poszukiwanie nowych, coraz bardziej wyrafinowanych narzędzi do tego celu. W konsekwencji do modelowania zjawisk zachodzących na rynkach finansowych wykorzystano procesy dyfuzji. Proces nazywany obecnie procesem Wienera po raz pierwszy został wykorzystany do modelowania procesów finansowych przez L. Bacheliera, który w swojej pracy doktorskiej z 1900 roku zaproponował opis dynamiki ceny akcji z paryskiej giełdy za pomocą tzw. arytmetycznego ruchu Browna (por. Weron i Weron 1999).

W literaturze przedmiotu znanych jest kilka, nie zawsze równoważnych, podejść do określania pojęcia procesu dyfuzji, prowadzących do różnych klas procesów (por. Wentzell 1980; Rogers i Williams 1987). W monografii przyjęto, wydaje się najbardziej rozpowszechnioną, definicję tego procesu – podobnie jak w pracach: Arnold (1974), Rozanow (1974), Kowalenko, Kuzniecowa i Szurienkow (1989), Schuss (1989), Bailey (1990), Kloeden i Platen (1992), Gajek i Kałuska (1996) oraz Sobczyk (1996).

Proces Wienera stanowi podstawę pojęcia stochastycznej całki Itô (por. Itô 1942, 1944, 1946, 1951), wykorzystywanej m.in. do budowy modeli matematycznych zadanych poprzez stochastyczne równania różniczkowe SRR (*stochastic differential equations – SDE*) Itô. Należy nadmienić, iż samo pojęcie stochastycznych równań różniczkowych zostało po raz pierwszy wprowadzone w 1934 roku przez S.N. Bernsteina. Dzięki wynikom, jakie osiągnęli P. Lévy (por. Lévy 1937, 1939, 1948), K. Itô (por. Itô 1942, 1944, 1946, 1951), I.V. Girsanov (Girsanov 1962) i inni, teoria stochastycznych równań różniczkowych stanowi rozwiniętą i bardzo ważną gałąź współczesnej teorii procesów stochastycznych. Znalazła ona liczne zastosowania, między innymi w nowoczesnej teorii finansów.

W literaturze przedmiotu zamiennie stosuje się określenie SRR Itô oraz SRR z zaburzeniem gaussowskim, a rozwiązania takich układów równań nazywa się procesami Itô. W niniejszej pracy, często dla prostoty zapisu, stosuje się określenie SRR, mając na myśli wyłącznie SRR Itô. Wykorzystywane w praktyce procesy dyfuzji są rozwiązaniami SRR Itô postaci (por. Arnold 1974):

$$\begin{cases} dX_t = a(X_t, t) dt + b(X_t, t) dW_t \\ X_{t_0} = x_0 \end{cases},$$

gdzie $a(\cdot)$ nazywana jest funkcją dryfu (*drift function*), natomiast $b(\cdot)$, a czasem $b^2(\cdot)$, funkcją zmienności lub dyfuzji (*volatility/diffusion function*). Dla podkreślenia przyna-

leżności do obydwu klas, przyjęto nazywać je procesami dyfuzji Itô (*Itô diffusion processes*) (por. Rogers i Williams 1987; Øksendal 2000). Do tejże klasy jednowymiarowych procesów ograniczono rozumowanie w prezentowanej monografii. Należy nadmienić, że ze względu na wspomniany związek procesów dyfuzji z procesami Itô w literaturze przedmiotu, często mówiąc o procesach dyfuzji, ma się na myśli rozwiązanie SRR i na odwrót – choć takie ujęcie nie jest precyzyjne, jest powszechnie stosowane (por. Dacunha-Castelle i Florens-Zmirou 1986; Florens-Zmirou 1989; Pedersen 1995b; Musiela i Rutkowski 1998; Eraker 1998; Jensen i Poulsen 1999; Poulsen 1999a; Aït-Sahalia 1999, 2002, 2003, 2004; Roberts i Stramer 2001; DiPietro 2001; Egorov, Li i Xu 2003; Elerian, Chib i Shephard 2001; Sørensen 2002; Nicolau 2002; Gutiérrez, Nafidi i Gutiérrez Sánchez 2004; Kostrzewski 2004). W tym właśnie sensie przyjęto w pracy stosować zamiennie pojęcia procesu dyfuzji, procesu Itô oraz procesu dyfuzji Itô.

W monografii wykorzystuje się pojęcie modelu w sensie matematycznego sposobu opisu zjawiska oraz w rozumieniu modelu statystycznego (przestrzeni statystycznej), a w szczególności statystycznego modelu bayesowskiego. Model matematyczny zadany jest w pracy poprzez układ stochastycznych równań różniczkowych lub wprost poprzez proces stochastyczny. Właściwe znaczenie i konkretny sens pojęcia modelu wynika z kontekstu.

Prace F. Blacka i M. Scholesa oraz R.C. Mertona z roku 1973 wywarły istotny wpływ na liczne działy ekonomii¹⁾. W pracach tych wyznaczono wzory na cenę europejskiej opcji kupna na akcję przy założeniu, że jej wartość opisywana jest poprzez proces dyfuzji Itô, a dokładnie poprzez tzw. geometryczny ruch Browna. Wspomniane rezultaty w dużej mierze przyczyniły się do stosowania na masową skalę procesów dyfuzji w zagadnieniach ekonomicznych. Obecnie procesy te wykorzystuje się m.in. w zagadnieniach zarządzania ryzykiem; do konstrukcji optymalnych strategii inwestycyjnych, modelowania struktury czasowej instrumentów bazowych (*underlying instrument*) oraz wyceny instrumentów pochodnych (*derivative security, derivative instrument*) (por. Hull 1993, 1998; Lamberton i Lapeyre 1996; *Instrumenty...* 1997; Crawford i Sen 1998; Musiela i Rutkowski 1998; Wilmott 1998; Weron i Weron 1999; Øksendal 2000; Jajuga i Jajuga 2000; Brigo i Mercurio 2001; Butler 2001; Filipović 2001; Janicki i Izydorczyk 2001; Jakubowski, Palczewski, Rutkowski i Stettner 2003).

Wzrastająca rola procesów dyfuzji Itô w zastosowaniach finansowych wymusiła podjęcie tematu estymacji parametrów funkcji dryfu i funkcji zmienności, stając się równocześnie główną inspiracją prezentowanej pracy. Dla większości rozwiązań SRR pomimo wiedzy o istnieniu gęstości przejścia nie można podać jej jawnego wzoru, a tym samym wzoru na funkcję wiarygodności (*likelihood function*) (por. Friedman

¹⁾ Odzwierciedleniem tego faktu było przyznanie nagrody Nobla w 1997 roku w dziedzinie nauk ekonomicznych dla M. Scholesa i R.C. Mertona (F. Black zmarł w 1995 roku i stąd nagroda nie została mu przyznana).

1975; Stroock i Varadhan 1979). Brak jawnego przepisu na funkcję wiarygodności stanowi główną trudność w przeprowadzaniu wnioskowania statystycznego. Z własności Markowa, obowiązującej dla procesów dyfuzji, funkcja wiarygodności jest iloczynem gęstości przejścia, stąd w literaturze przedmiotu podejmuje się tematykę przybliżania tych gęstości jako sposobu na aproksymację funkcji wiarygodności. Problem estymacji parametrów SRR nie jest nowym zagadnieniem. Obszerne zestawienie niebayesowskich metod estymacji można znaleźć w rozprawie doktorskiej O. Eleriana z 1999 roku. Znanych jest również kilka prac podsumowujących wybrane podejścia do zagadnienia estymacji: Jensen i Poulsen (1999), Ait-Sahalia (1999), DiPietro (2001), Sørensen (2002), Nicolau (2002). Zagadnienie estymacji parametrów procesu Itô na gruncie bayesowskim zostało poruszone w następujących pozycjach: Eraker (1998), Jones (1998), Elerian (1998, 1999), Elerian, Chib i Shephard (2001), DiPietro (2001), Roberts i Stramer (2001), Kostrzewski (2004, 2006a, b).

Celem monografii jest przedstawienie procesów dyfuzji oraz ich wykorzystanie do modelowania dynamiki finansowych szeregów czasowych, a w szczególności opracowanie i zastosowanie podejścia bayesowskiego do estymacji, predykcji oraz porównania mocy wyjaśniającej konkurencyjnych modeli. Podstawową tezę pracy jest stwierdzenie, że procesy dyfuzji, będąc przykładem ciągłych procesów stochastycznych, są ważnym i godnym uwagi narzędziem opisu dynamiki zjawisk finansowych obserwowanych w dyskretnych momentach czasu. Wnioskowanie bayesowskie wydaje się ciekawym, pożytecznym i jednocześnie obiecującym ujęciem wspomnianego modelowania. Dodatkowo brak rozważań tej tematyki w literaturze polskiej wskazuje na konieczność poszerzenia zakresu badań oraz podjęcia próby wykorzystania procesów dyfuzji do modelowania polskich danych finansowych.

Istota estymacji bayesowskiej polega na zastosowaniu twierdzenia Bayesa w celu wyznaczenia rozkładu a posteriori na podstawie założonej struktury a priori i stosownej dla rozpatrywanych modeli funkcji wiarygodności. Estymacja bayesowska łączy w ten sposób wstępną wiedzę badacza o parametrach modelu, posiadaną przed wglądem w dane, z informacją, jaką niesie próba. Rozkład a priori umożliwia tym samym w sposób jawny i sformalizowany włączenie intuicyjnych, subiektywnych informacji spoza próby; jest on miarą stopnia przekonania badacza (*degree of belief*), co do szans poszczególnych wartości parametrów (por. Zellner 1971; Box i Tiao 1973; DeGroot 1981; Osiewalski 1991, 2001; Szreder 1994; Męczarski 1998; Pipień 2000; Bernardo i Smith 2002; Pajor 2003b). Atrakcyjność podejścia bayesowskiego jest związana również z tym, iż rozkład a posteriori odzwierciedla informację o nieznanach parametrach w sposób czysto probabilistyczny i wolny od aproksymacji asymptotycznych, co pociąga za sobą możliwość pełnego wnioskowania w przypadku małych prób. Bayesowska prognoza realizacji badanego zjawiska opiera się na gęstości rozkładu warunkowego dla wektora prognozowanego przy znanym wektorze obserwacji, tzn. na gęstości predyktywnej. Gęstość tę wykorzystuje się m.in. do wyceny europejskich opcji kupna (por. Bauwens i Lubrano 1997; Osiewalski i Pipień 2003; Pajor 2003a, b). Wnioskowa-

nie bayesowskie dostarcza probabilistycznego narzędzia porównywania mocy wyjaśniających konkurujących ze sobą modeli (por. Zellner 1971; Osiewalski i Steel 1993). Daje również możliwość tzw. łączenia wiedzy (*bayesian pooling approach*) polegającego na ważonym uśrednieniu informacji płynącej z poszczególnych modeli (por. Osiewalski 2001). Zastosowaniom wnioskowania bayesowskiego towarzyszą problemy generowania realizacji z wielowymiarowych rozkładów a posteriori, których często jedynym rozwiązaniem jest aproksymacja tych rozkładów za pomocą metod Monte Carlo łańcuchów Markowa (*Markov Chain Monte Carlo*, MCMC) (por. O'Hagan 1994; Gamberman 1997; Krzykowski 2000; Osiewalski 2001; Pajor 2003b).

W porównaniu z literaturą przedmiotu nowością w prezentowanej monografii jest opracowanie i zastosowanie metod wnioskowania bayesowskiego dla procesów dyfuzji Itô z wykorzystaniem najnowszych wyników teoretycznych dotyczących aproksymacji gęstości przejścia. Aproksymacji funkcji przejścia dokonano za pomocą wielomianów Hermite'a oraz poprzez rozwiązania równań Fokkera, Plancka i Kołmogorowa na podstawie wyników prac Aït-Sahalia (2003, 2004) oraz Egorov, Li i Xu (2003). Ilustrację empiryczną wspomnianych metod przeprowadzono na szeregu stóp procentowych banku centralnego USA oraz na szeregu dziennych wartości indeksu WIG20. Dodatkowo omówiono i zilustrowano przykładem empirycznym opartym na danych symulowanych zastosowanie metody zmiennych pośrednich do bayesowskiej estymacji parametrów SRR. Wokół tej metody oraz jej uogólnień koncentrują się wspomniane już wcześniej prace z zakresu wnioskowania bayesowskiego: Eraker (1998), Jones (1998), Elerian (1998, 1999), Elerian, Chib i Shephard (2001), Roberts i Stramer (2001), Kostrzewski (2004). Ponadto rozważono zagadnienie wnioskowania bayesowskiego dla nieautonomicznych SRR, rzadko rozpatrywanych w teorii i praktyce. W prezentowanej pracy podjęto problem bayesowskiej wyceny opcji z wykorzystaniem nieautonomicznego modelu EOU (*Extended Ornstein-Uhlenbeck model*). Należy w tym miejscu wspomnieć, że M. DiPietro w rozprawie doktorskiej pod tytułem *Bayesian Inference for Discretely Sampled Diffusion Processes with Financial Applications* (por. DiPietro 2001), w odróżnieniu od niniejszej pracy, nie rozważał zagadnienia bayesowskiego porównania konkurencyjnych modeli oraz problemu bayesowskiej prognozy. W prezentowanej monografii aproksymację gęstości przejścia oparto na pewnym uogólnieniu wzoru stosowanego w przytoczonej rozprawie, co było konieczne ze względu na rozpatrywanie modelu nieautonomicznego EOU. Również metoda rozwiązań równań FPK (Fokkera, Plancka i Kołmogorowa) wykorzystywana tutaj dla autonomicznych SRR nie była przedmiotem rozważań M. DiPietro. Należy dodać, że istotny wysiłek przygotowania prezentowanej monografii koncentrował się wokół rozwiązywania problemów numerycznych, towarzyszących zastosowaniom wnioskowania bayesowskiego.

Pomimo to że problem estymacji parametrów procesów dyfuzji nie jest nowym zagadnieniem, wciąż cieszy się dużym zainteresowaniem przynoszącym w ostatnich latach ciekawe wyniki teoretyczne. Niegasnąca popularność tej tematyki, widoczna w światowej literaturze, może świadczyć o nietrywialności rozważanego w pracy problemu.

Autor prezentowanej pracy jest pracownikiem Wydziału Matematyki Stosowanej Akademii Górniczo-Hutniczej w Krakowie. Monografia jest rozprawą doktorską napisaną pod kierunkiem prof. zw. dr. hab. Jacka Osiewalskiego, a obronioną dnia 4 lipca 2006 roku na Wydziale Zarządzania Akademii Ekonomicznej w Krakowie. Praca składa się z pięciu rozdziałów. Rozdział pierwszy zawiera szereg definicji i własności rozważanych w monografii procesów stochastycznych. Ograniczono się wyłącznie do najważniejszych charakterystyk, odsyłając po szczegóły do odpowiedniej literatury. W rozdziale tym wspomniano o numerycznych aspektach rozwiązań SRR. Rozważono zagadnienie schematów dyskretyzacji i ich zastosowania w aproksymacji trajektorii oraz charakterystyk rozkładów procesów. Ponadto zdefiniowano i szczegółowo omówiono zadane poprzez stochastyczne równania różniczkowe modele matematyczne: Vasička (V), Blacka i Scholesa (BS), Coxa, Ingersolla i Rossa (CIR), uogólniony model Ornsteina i Uhlenbecka (EOU) oraz Brennana i Schwartza (BSBS). Do tych modeli ograniczono rozumowanie w kolejnych rozdziałach pracy.

Rozdział drugi poświęcony jest aproksymacji funkcji wiarygodności dla procesów dyfuzji Itô. Z własności Markowa rozwiązań SRR funkcja wiarygodności jest iloczynem gęstości przejścia, stąd poprzez aproksymację tych gęstości otrzymuje się aproksymację funkcji wiarygodności.

Na podstawie najnowszych wyników teoretycznych omówiono trzy metody przybliżania wartości gęstości przejścia:

- 1) metodę zmiennych pośrednich,
- 2) metodę przybliżania szeregiem wielomianów Hermite'a,
- 3) metodę rozwiązań równań Fokkera, Plancka i Kołmogorowa.

W rozdziale trzecim przedstawiono ogólne zasady wnioskowania bayesowskiego. Zdefiniowano strukturę rozkładu a priori oraz określono rozkłady a posteriori i rozkłady predyktywne dla omawianych w pracy modeli. Rozważone zostało również zagadnienie porównania konkurencyjnych modeli. Znaczną część rozdziału poświęcono metodom numerycznym będącym często jedynymi narzędziami obliczania charakterystyk wielowymiarowych rozkładów a posteriori i predyktywnych. Ponadto omówiono zagadnienie estymacji parametrów za pomocą metody zmiennych pośrednich, ukazując jej niewielką przydatność w zastosowaniach.

W rozdziałach czwartym i piątym przedstawione zostały wyniki badań empirycznych przeprowadzonego wnioskowania bayesowskiego dla finansowych szeregów czasowych. W rozdziale czwartym modelowaniem objęto szereg miesięcznych stóp procentowych banku centralnego USA (FED – *Federal Reserve System*), a w rozdziale piątym szereg dziennych wartości indeksu WIG20 z zamknięcia sesji. W obydwu rozdziałach przeprowadzono estymację parametrów, porównanie konkurencyjnych modeli oraz prognozę dalszego poziomu analizowanego szeregu czasowego. Ponadto w rozdziale piątym dokonano wyceny europejskiej opcji kupna na indeks WIG20.

Słowa podziękowania i wdzięczności za inspirację, cenne uwagi, olbrzymi optymizm i wiarę we mnie składam na ręce Promotora prof. dr. hab. Jacka Osiewalskiego.

Dziękuję koleżankom i kolegom z Wydziału Matematyki Stosowanej Akademii Górniczo-Hutniczej w Krakowie oraz z Katedry Ekonometrii Akademii Ekonomicznej w Krakowie za pomoc i wyrozumiałość. W sposób szczególny chciałbym podziękować dr Annie Pajor za jej wskazówki pomocne w tworzeniu prezentowanej monografii. Dziękuję również Recenzentom mojej pracy doktorskiej prof. zw. dr. hab. inż. Ryszardowi Tadeusiewiczowi i dr. hab. Ryszardowi Domanowi za cenne uwagi i sugestie kierunku dalszej pracy naukowej oraz zachętę do opublikowania rozprawy.

1. Procesy dyfuzji

stosowane w analizie danych finansowych

1.1. Uwagi wstępne

W 1906 roku A.A. Markow rozpoczął badania nad zdarzeniami zależnymi, wśród których wyodrębnił schemat doświadczeń powiązanych w łańcuchach bez tzw. efektów następstwa (por. Doob 1953; Fisz 1967; Jermakow 1976; Feller 1978; Sobczyk 1996). Wyniki jego pracy doprowadziły do rozwoju pewnej klasy procesów stochastycznych nazywanych od jego nazwiska procesami Markowa. Istotny wpływ na rozwój teorii procesów Markowa mieli A.N. Kołmogorow i W. Feller. Wśród procesów Markowa określonych na ciągłej osi czasu i o nieprzeliczalnej przestrzeni stanów wyróżnione zostały procesy dyfuzji. Procesy te po raz pierwszy zostały użyte do modelowania ruchu pojedynczej cząsteczki w procesie przenikania cząsteczek jednej substancji w obręb drugiej na skutek chaotycznego ruchu molekuł. R. Brown w 1827 roku wykorzystał do opisu tego właśnie ruchu (tzw. ruchu Browna) pewien proces stochastyczny, który od jego imienia nazywany jest obecnie procesem Browna. Nad precyzyjną matematyczną teorią opisującą ruch Browna pracowali m.in. A. Einstein (por. Einstein 1905), M. Smoluchowski (por. Smoluchowski 1906) i N. Wiener (por. Wiener 1923). Szczególnie dzięki pracom Wienera proces ruchu Browna stał się jednym z najważniejszych procesów stochastycznych, wykorzystywanym w coraz to szerszych obszarach zastosowań (por. Gajek i Kałuszka 1996; Platen 1999; Weron i Weron 1999). W literaturze przedmiotu bardzo często zamiennie stosuje się pojęcie ruchu Browna i procesu Wienera (*Wiener process*). Jest to uzasadnione, gdyż definicja procesu Wienera w standardowej i zarazem najpopularniejszej formie jest równoważna procesowi Browna (por. Lipcer i Szirijajew 1981; Neftci 2000). Ponadto proces Wienera stanowi podstawę pojęcia stochastycznej całki Itô (por. Itô 1942, 1944, 1946, 1951), wykorzystywanej do budowy modeli zadanych poprzez stochastyczne równania różniczkowe¹⁾ SRR (*stochastic differential equations, SDE*) Itô. Do modelowania procesów finanso-

¹⁾ Termin ten po raz pierwszy wprowadził S.N. Bernstein w 1934 roku. W niniejszej pracy ograniczono rozważania dotyczące stochastycznych równań różniczkowych wyłącznie do SRR Itô.

wych proces Wienera po raz pierwszy został wykorzystany przez L. Bacheliera, który w swojej pracy doktorskiej z 1900 roku zaproponował opis dynamiki ceny akcji z paryskiej giełdy za pomocą tzw. arytmetycznego ruchu Browna (por. Weron i Weron 1999).

Celem niniejszego rozdziału jest przedstawienie wykorzystywanych w dalszej części pracy pojęć oraz omówienie podstawowych własności rozważanych procesów stochastycznych. Zaprezentowane zostaną procesy Markowa z czasem ciągłym i nieprzeliczalną przestrzenią stanów oraz ich szczególne przypadki, jakimi są procesy dyfuzji. Ponadto przedstawiony zostanie związek pomiędzy procesami dyfuzji oraz procesami Itô. Należy zaznaczyć, że rozważania podjęte w monografii zawężone zostaną do procesów dyfuzji Itô, tzn. procesów dyfuzji będących zarazem rozwiązaniami pewnych SRR Itô.

Ponieważ w ogólnej sytuacji nieznane są jawne wzory²⁾ rozwiązań SRR Itô, powstaje więc problem ich aproksymacji. W niniejszym rozdziale opisane zostaną wybrane schematy dyskretyzacji określające równania różnicowe, których rozwiązania uważane są za aproksymację rozwiązań SRR Itô. Ponadto opisane zostaną wykorzystane w dalszej części pracy modele: Vasička, Blacka i Scholesa, Coxa, Ingersolla i Rossa, Brennana i Schwartza oraz uogólniony model Ornsteina i Uhlenbecka.

1.2. Procesy dyfuzji – definicje i własności

Niech $(\Omega, \mathfrak{F}, P)$ oznacza przestrzeń probabilistyczną zupełną³⁾, gdzie Ω jest przestrzenią zdarzeń elementarnych, inaczej przestrzenią prób, $\mathfrak{F}$ jest σ -algebrą, natomiast P jest miarą probabilistyczną. Dodatkowo dla potrzeb niniejszej pracy przyjęto niżej wymienione definicje

Definicja 1.1

Procesem stochastycznym X jest funkcja dwóch zmiennych⁴⁾ $X: \mathbb{T} \times \Omega \rightarrow \mathbb{R}$, taka, że $X(t, \cdot): \Omega \rightarrow \mathbb{R}$ jest zmienną losową dla każdego $t \in \mathbb{T} \subset \mathbb{R}$.

²⁾ To znaczy takie, które nie zawierają w swojej postaci całek stochastycznych.

³⁾ Przestrzeń $(\Omega, \mathfrak{F}, P)$ jest przestrzenią zupełną wtedy i tylko wtedy, gdy $\forall A \subset \Omega: A \subset B \in \mathfrak{F}: P(B) = 0 \Rightarrow A \in \mathfrak{F}$ (innymi słowy, dowolny podzbiór zbioru miary zero jest mierzalny). Dla przestrzeni zupełnych, jeśli ξ jest zmienną losową oraz istnieje funkcja η taka, że $\xi \neq \eta$ na podzbiórze zbioru miary zero, wówczas η jest zmienną losową i z punktu widzenia teorii miary funkcje ξ i η są nierozróżnialne. Dowolną przestrzeń probabilistyczną można w prosty sposób rozszerzyć do przestrzeni probabilistycznej zupełnej (por. Gichman i Skorochod 1968; Lipcer i Szirajew 1981).

⁴⁾ W pracy przyjęto następujące oznaczenia: $\mathbb{R}$ – zbiór liczb rzeczywistych, $\mathbb{Z}$ – zbiór liczb całkowitych, $\mathbb{N}$ – zbiór liczb naturalnych.

Zmienne losowe definiujące proces stochastyczny indeksowane są elementami zbioru $\mathbb{T}$ interpretowanego jako czas. W literaturze przedmiotu rozważa się procesy z czasem dyskretnym, jeśli $\mathbb{T} = \mathbb{Z}$ lub $\mathbb{T} = \mathbb{N}$, oraz procesy z czasem ciągłym, jeśli $\mathbb{T}$ jest przedziałem liczbowym (najczęściej: $\mathbb{T} = [t_0, T]$, $\mathbb{T} = [t_0, \infty)$ lub $\mathbb{T} = (-\infty, \infty)$). Funkcja $X(\cdot, \omega): \mathbb{T} \rightarrow \mathbb{R}$ nazywana jest trajekcją, ścieżką lub realizacją procesu. Wśród rozpatrywanych w literaturze procesów rozróżnia się te, których zbiór wartości (stanów) jest zbiorem przeliczalnym lub nieprzeliczalnym. W monografii ograniczono rozważania do procesów ciągłych, tzn. procesów z czasem ciągłym, o nieprzeliczalnym zbiorze wartości. Dla wygody zapisu przyjęto oznaczenia na proces stochastyczny $\{X_t\}_{t \in \mathbb{T}}$ lub X .

Wśród procesów stochastycznych na szczególną uwagę zasługują procesy Markowa. Ich istota oparta jest na tzw. własności Markowa, zgodnie z którą przyszła realizacja procesu Markowa zależy wyłącznie od obecnego (teraźniejszego) stanu procesu, a nie od całej przeszłości⁵⁾. Formalnie procesy te określa się w następujący sposób:

Definicja 1.2

Proces stochastyczny X jest procesem Markowa, jeśli:

$$\begin{aligned} \forall n \in \mathbb{N}, \forall \{t_1, t_2, \dots, t_n, t_{n+1}\} \subset \mathbb{T}: t_1 < t_2 < \dots < t_n < t_{n+1}, \\ \forall \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}, \forall B \in \mathcal{B}(\mathbb{R}) \end{aligned}$$

zachodzi związek:

$$\begin{aligned} P(X_{t_{n+1}} \in B | X_{t_n} = x_n, X_{t_{n-1}} = x_{n-1}, \dots, X_{t_1} = x_1) = \\ = P(X_{t_{n+1}} \in B | X_{t_n} = x_n), \end{aligned}$$

gdzie $\mathcal{B}(\mathbb{R})$ oznacza σ -algebrę zbiorów borelowskich określonych na zbiorze liczb rzeczywistych.

Ważnym pojęciem w teorii procesów Markowa jest prawdopodobieństwo przejścia – inaczej funkcja przejścia (*transition probability/function*). Skończenie wymiarowe rozkłady procesu Markowa $\{X_t\}_{t \in [t_0, \infty)}$ są jednoznacznie wyznaczone przez rozkład początkowy $P_0(B) = P(X_{t_0} \in B)$ i prawdopodobieństwo przejścia $P(t_0, x; t, B)$ (por. Kowalenko, Kuzniecowa i Szurienko 1989). Prawdopodobieństwo to definiowane jest w następujący sposób:

⁵⁾ Poglądowo: „przyszłość nie zależy od przeszłości, gdy znana jest teraźniejszość” (por. Sobczyk 1996).

Definicja 1.3

Prawdopodobieństwem przejścia procesu X ze stanu x w chwili s do zbioru $B \in \mathcal{B}(\mathbb{R})$ w chwili t nazywana jest funkcja, przyporządkowująca czwórce (x, s, B, t) wartość:

$$P(B|x, s, t-s) := \begin{cases} P(X_t \in B | X_s = x) & \text{dla } t-s > 0 \\ I_B(x) & \text{dla } t-s = 0 \end{cases},$$

$$\text{gdzie } I_B(x) = \begin{cases} 1 & \text{dla } x \in B \\ 0 & \text{dla } x \notin B \end{cases}.$$

Dla ustalonych $x, s \leq t$: $(\cdot|x, s, t-s)$ jest miarą probabilistyczną określoną na σ -algebrze zbiorów borelowskich.

W monografii przyjęto założenie, że istnieją gęstości $p(x_1, t_1; x_2, t_2; \dots; x_k, t_k)$ skończenie wymiarowych zmiennych losowych postaci $(X_{t_1}, \dots, X_{t_k})$ dla $t_i \in \mathbb{T}$, $x_i \in \mathbb{R}$, $i \in \{1, \dots, k\}$, $k = 1, 2, \dots$. W rezultacie $P(\cdot|x, s, t-s)$ dla $s < t$ jest rozkładem absolutnie ciągłym o gęstości $p(\cdot|x; s, t-s)$. Funkcja⁶⁾ $p(\cdot|x; s, t-s)$ nazywana jest gęstością lub funkcją przejścia (*transition density or function*) procesu X . Z definicji gęstości wynika:

$$\forall B \in \mathcal{B}(\mathbb{R}): P(B|x, s, t-s) = \int_B p(y|x; s, t-s) dy.$$

Wśród procesów Markowa wyróżnia się procesy jednorodne i niejednorodne (*homogeneous and inhomogeneous Markov process*). Dla procesu jednorodnego prawdopodobieństwo przejścia nie zależy od wielkości s , a zależy jedynie od przyrostu czasu $t-s$. Natomiast dla procesu niejednorodnego istotna jest wartość zarówno s , jak i t . W języku gęstości przejścia własność jednorodności i niejednorodności definiuje się w następujący sposób:

Definicja 1.4

Proces Markowa X nazywany jest procesem jednorodnym, jeśli:

$$p(y|x; s, t-s) = p(y|x; t-s) \text{ dla } t-s > 0;$$

w przeciwnym przypadku proces jest niejednorodny.

⁶⁾ Gęstość przejścia oznacza się również poprzez $p(\cdot, t|x; s)$. Ponadto jeśli gęstość ta nie zależy od stanu x lub któregośkolwiek z momentów czasu, to odpowiednie oznaczenie pomija się w zapisie.

Dla procesów Markowa gęstość wektora $(X_{t_1}, \dots, X_{t_{n+1}})$ dla $t_1 < t_2 < \dots < t_n < t_{n+1}$ spełnia następujący związek:

$$\begin{aligned} p(x_1, t_1; x_2, t_2; \dots; x_{n+1}, t_{n+1}) &= \\ &= p(x_{n+1}, t_{n+1} | x_1, t_1; \dots; x_n, t_n) p(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = \\ &= p(x_{n+1} | x_n, t_n, t_{n+1} - t_n) p(x_1, t_1; x_2, t_2; \dots; x_n, t_n). \end{aligned}$$

W konsekwencji:

$$p(x_1, t_1; x_2, t_2; \dots; x_{n+1}, t_{n+1}) = p(x_1, t_1) \prod_{k=2}^{n+1} p(x_k | x_{k-1}; t_{k-1}, t_{k-1} - t_k) \quad (1.1)$$

gdzie $p(x_1, t_1)$ jest gęstością zmiennej X_{t_1} . Z powyższego związku można otrzymać postać gęstości próbkowej, prowadzącej do funkcji wiarygodności (*likelihood function*) dla parametrycznej klasy procesów Markowa. Gęstość próbkowa jest iloczynem gęstości przejścia:

$$p(x_1, \dots, x_{n+1} | \theta) = p(x_1, t_1; \theta) \prod_{k=2}^{n+1} p(x_k | x_{k-1}; t_{k-1}, t_{k-1} - t_k; \theta) \quad (1.2)$$

gdzie $\theta \in \Theta$ jest wektorem parametrów oraz $(t_1, t_2, \dots, t_n, t_{n+1})' : t_1 < t_2 < \dots < t_n < t_{n+1}$ jest wektorem momentów czasu, w których zaobserwowano wartości procesu $X: (x_1, \dots, x_{n+1})'$.

W klasie procesów Markowa na szczególną uwagę zasługują procesy dyfuzji. Procesy te służą m.in. do opisu zmian położenia cząstki zanurzonej w cieczy i poruszającej się pod wpływem zderzeń z cząsteczkami ośrodka (por. Gajek i Kałuszka 1996). W literaturze przedmiotu znanych jest kilka, nie zawsze równoważnych, podejść do określania wspomnianych procesów (por. Rogers i Williams 1987). W monografii przyjęto następującą definicję (por. Arnold 1974; Rozanow 1974; Kowalenko, Kuzniecow i Szurienkow 1989; Schuss 1989; Bailey 1990; Kloeden i Platen 1992; Gajek i Kałuszka 1996; Sobczyk 1996):

Definicja 1.5

Proces Markowa $\{X_t\}_{t \in \mathbb{T}}$ o ciągłych trajektoriach⁷⁾ nazywany jest procesem dyfuzji, jeśli dla dowolnych s i t należących do zbioru $\mathbb{T}$ takich, że $s < t$ oraz $\varepsilon > 0$, prawdopodobieństwo przejścia $P(\cdot | x, s, t - s)$ spełnia następujące warunki:

⁷⁾ Ściślej, prawie wszędzie ciągłych, tzn. ciągłych z prawdopodobieństwem jeden. Innymi słowy, proces stochastyczny $\{X_t\}_{t \in \mathbb{T}}$ ma realizacje ciągłe z prawdopodobieństwem jeden, jeśli zbiór $\Omega_0 = \{\omega: \text{odwzorowanie } t \rightarrow X_t(\omega) \text{ jest ciągłe w } \mathbb{T}\}$ jest mierzalny i $P(\Omega_0) = 1$. Zgodnie z konwencją przyjętą w literaturze przedmiotu podawane w niniejszej pracy własności procesów są prawdziwe z dokładnością do zbioru miary zero.

$$(i) \quad \lim_{t \rightarrow s^+} \frac{1}{t-s} \int_{|y-x| > \varepsilon} p(y|x; s, t-s) dy = 0,$$

$$(ii) \text{ istnieje funkcja } a: \quad \lim_{t \rightarrow s^+} \frac{1}{t-s} \int_{|y-x| \leq \varepsilon} (y-x) p(y|x; s, t-s) dy = a(s, x),$$

$$(iii) \text{ istnieje funkcja } b: \quad \lim_{t \rightarrow s^+} \frac{1}{t-s} \int_{|y-x| \leq \varepsilon} (y-x)^2 p(y|x; s, t-s) dy = b(s, x).$$

Przy czym funkcja a nazywana jest funkcją dryfu lub współczynnikiem przenoszenia (*drift coefficient*), natomiast b funkcją lub współczynnikiem dyfuzji (*diffusion function/coefficient*).

Warto zauważyć, że jeśli istnieje dodatnia liczba $\varepsilon > 0$ taka, że:

$$\lim_{t \rightarrow s^+} \frac{1}{t-s} E(|X_t - X_s|^{2+\varepsilon} | X_s = x) = 0,$$

wówczas warunek (i) definicji 1.5 jest spełniony, natomiast warunki (ii) oraz (iii) równoważne są następującym granicom:

$$\lim_{t \rightarrow s^+} \frac{1}{t-s} E(X_t - X_s | X_s = x) = a(s, x) \quad (1.3)$$

$$\lim_{t \rightarrow s^+} \frac{1}{t-s} E((X_t - X_s)^2 | X_s = x) = b(s, x) \quad (1.4)$$

Z warunku (i) powyższej definicji wynika, że prawdopodobieństwo wystąpienia dużych zmian procesu X w krótkiej chwili jest małe, tzn. (por. Gajek i Kałuszka 1996; Sobczyk 1996):

$$\lim_{t \downarrow s} \frac{1}{t-s} P(|X_t - X_s| > \varepsilon | X_s = x) = 0 \quad \text{dla } \varepsilon > 0.$$

Jeśli zachodzi warunek (1.3), to⁸⁾:

$$E(X_t - X_s | X_s = x) = a(s, x)(t-s) + o(t-s) \quad \text{dla } t \rightarrow s^+.$$

⁸⁾ Zapis $f(x) = o(g(x))$ dla $x \rightarrow x_0$ oznacza, że $f(x)/g(x) \xrightarrow{x \rightarrow x_0} 0$. Natomiast zapis $f(x) = O(g(x))$ w otoczeniu punktu x_0 oznacza, że wyrażenie $f(x)/g(x)$ jest ograniczone w otoczeniu x_0 .

Innymi słowy, średnia zmiana wartości procesu dyfuzji w małym przedziale czasu jest w przybliżeniu wprost proporcjonalna do wielkości $t - s$ i współczynnika dryfu a .

Jeśli natomiast zachodzi wzór (1.4), to:

$$E\left((X_t - X_s)^2 \mid X_s = x\right) = b(s, x)(t - s) + o(t - s) \text{ dla } t \rightarrow s^+.$$

Zatem średni kwadrat przyrostu procesu w małym przedziale czasu jest w przybliżeniu wprost proporcjonalny do wartości $t - s$ i współczynnika dyfuzji b .

Poniżej przedstawione zostały twierdzenia (w jednej z możliwych postaci) opisujące ważną własność procesu dyfuzji, mówiącą o tym, że jego gęstość przejścia jest rozwiązaniem równań Fokkera, Plancka i Kołmogorowa (FPK), tzn. równań (1.5) i (1.6) (por. Fisz 1967; Gichman i Skorochod 1968; Rozanow 1974; Arnold 1974; Lipcer i Szirajew 1981; Schuss 1989; Kowalenko, Kuzniecowa i Szurienkow 1989; Sobczyk 1996).

Twierdzenie 1.1

Jeżeli $\{X_t\}_{t \in [t_0, T]}$ jest procesem dyfuzji z ciągłymi funkcjami a i b takim, że zbieżności z warunków (i) – (iii) definicji 1.5 są jednostajne dla $s \in [t_0, T]$, a ponadto gęstość przejścia $p(\cdot | x; s, t - s)$ procesu X jest ciągła względem s oraz istnieją i są ciągłe względem s pochodne cząstkowe:

$$\frac{\partial p(y | x; s, t - s)}{\partial x} \text{ i } \frac{\partial^2 p(y | x; s, t - s)}{\partial x^2},$$

wówczas spełnione jest równanie:

$$\frac{\partial p(y | x; s, t - s)}{\partial s} + a(s, x) \frac{\partial p(y | x; s, t - s)}{\partial x} + \frac{1}{2} b(s, x) \frac{\partial^2 p(y | x; s, t - s)}{\partial x^2} = 0 \quad (1.5)$$

Równanie różniczkowe (1.1) nazywane jest retrospektywnym równaniem Kołmogorowa (lub pierwszym równaniem Kołmogorowa) (*backward Kolmogorov equation*).

Twierdzenie 1.2

Jeżeli $\{X_t\}_{t \in [t_0, T]}$ jest procesem dyfuzji takim, że zbieżności z warunków (i) – (iii) definicji 1.5 są jednostajne dla s i x , a ponadto dla gęstości przejścia $p(\cdot | x; s, t - s)$ procesu X istnieją i są ciągłe następujące pochodne:

$$\frac{\partial p(y | x; s, t - s)}{\partial t}, \frac{\partial}{\partial y} \{a(t, y) p(y | x; s, t - s)\}, \frac{\partial^2}{\partial y^2} \{b(t, y) p(y | x; s, t - s)\},$$

wówczas dla ustalonych s oraz x takich, że $s < t$, spełnione jest prospektywne równanie Kołmogorowa (nazywane również drugim równaniem Kołmogorowa lub równaniem Fokkera i Plancka) (*forward Kolmogorov equation*):

$$\frac{\partial p(y|x; s, t-s)}{\partial t} + \frac{\partial}{\partial y} \{a(t, y) p(y|x; s, t-s)\} - \frac{1}{2} \frac{\partial^2}{\partial y^2} \{b(t, y) p(y|x; s, t-s)\} = 0 \quad (1.6)$$

Jednym z ważniejszych procesów dyfuzji, który odgrywa istotną rolę w teorii stochastycznych równań różniczkowych, jest proces Wienera (ruch Browna). Dla tego procesu funkcje dryfu i dyfuzji wynoszą odpowiednio: $a(s, x) = 0$ i $b(s, x) = 1$. Jego definicja jest następująca⁹⁾:

Definicja 1.6

Procesem Wienera nazywany jest proces stochastyczny $W: \Omega \times [0, \infty) \rightarrow \mathbb{R}$ taki, że:

- (i) $P(W_0 = 0) = 1$,
- (ii) przyrosty procesu $W_t - W_s, W_u - W_t$ dla $s < t < u$ są niezależne,
- (iii) przyrosty $W_{t+h} - W_t$ mają rozkład normalny o wartości oczekiwanej 0 i wariancji h .

Proces Wienera jest jednorodnym procesem dyfuzji o następującej gęstości przejścia¹⁰⁾:

$$p(y|x; s, t-s) = \frac{1}{\sqrt{2\pi(t-s)}} \exp \left\{ -\frac{(y-x)^2}{2(t-s)} \right\}.$$

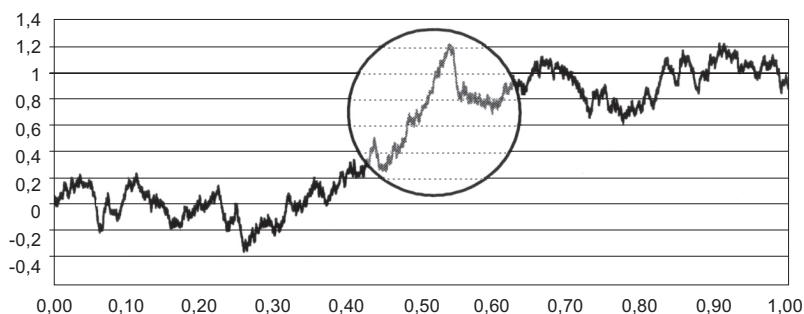
Ponadto można pokazać, że proces ten ma realizacje ciągłe, prawie wszędzie nieróżniczkowalne o wahanu nieskończonym¹¹⁾ na każdym skończonym przedziale liczbowym. Wspomniane wyżej własności pociągają za sobą dużą nieregularność trajektorii procesu widoczną na rysunkach¹²⁾ 1.1 i 1.2.

⁹⁾ W ten sposób zdefiniowany proces stochastyczny w literaturze przedmiotu nazywany jest czasami standardowym procesem Wienera (*standard Wiener process*) dla odróżnienia od definicji, w której wartość oczekiwana i wariancja z punktu (iii) definicji 1.6 wynoszą odpowiednio αh i $\sigma^2 h$ (por. Kowalenko, Kuzniecowa i Szurienkow 1989; Sobczyk 1996).

¹⁰⁾ Jest to gęstość rozkładu normalnego o wartości oczekiwanej x i wariancji $t-s$.

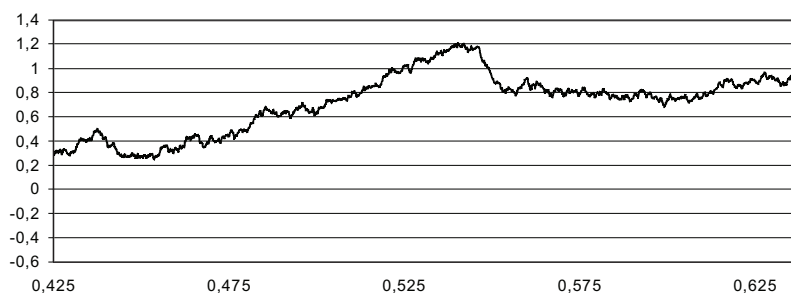
¹¹⁾ Jeśli dla f określonej w przedziale $[a, b]$ kres górny sum $\sum_{i=1}^n \left| f(x_{i+1}^{(n)}) - f(x_i^{(n)}) \right|$ jest skończony dla dowolnego podziału $a = x_1 < x_2 < \dots < x_{n+1} = b$, wówczas mówi się, że funkcja f jest funkcją o wahanu skończonym. Jeśli kres górny jest nieskończony, wówczas mówi się, że funkcja ma wahanie nieskończone.

¹²⁾ W literaturze przedmiotu znanych jest kilka metod konstrukcji procesu Browna np.: konstrukcja Levy'ego-Ciesielskiego, aproksymacja poprzez proces błędzenia przypadkowego oraz aproksymacja ruchu Browna otrzymana przez sumowanie odpowiednich przyrostów (por. Janicki 1996). Ostatnia z wymienionych metod została wykorzystana w niniejszej pracy – odpowiada ona zastosowaniu schematu Eulera omówionego w podrozdziale 1.4.



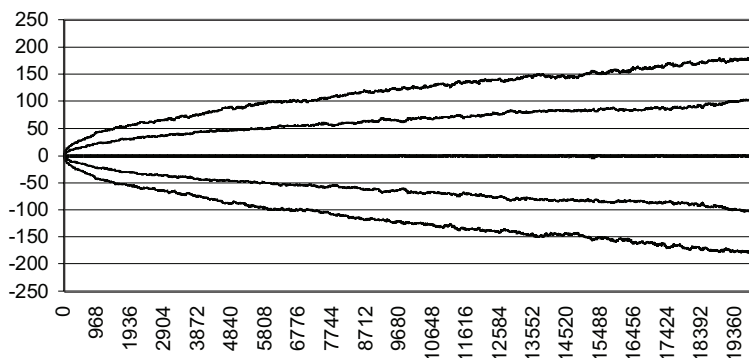
Rys. 1.1. Przykładowa trajektoria procesu Wienera

Źródło: Opracowanie własne



Rys. 1.2. Przykładowa trajektoria procesu Wienera – powiększenie obszaru wyszczególnionego na rysunku 1.1

Źródło: Opracowanie własne



Rys. 1.3. Linie kwantylowe rzędu 0,1, 0,25, 0,5, 0,75 i 0,9 procesu Wienera

Źródło: Opracowanie własne

Rysunek 1.3 przedstawia linie kwantylowe procesu Wienera¹³⁾. Widoczne jest zwiększające się rozproszenie procesu wraz ze wzrostem czasu.

¹³⁾ W ustalonej chwili czasu wartość procesu Wienera określa stosowna zmienna losowa, dla której można wyznaczyć kwantyle rozkładu (por. definicja 1.12).

1.3. Stochastyczne równania różniczkowe Itô – definicje i własności

Niech $(\Omega, \mathfrak{F}, P)$ będzie przestrzenią probabilistyczną zupełną wyposażoną w filtrację $\{\mathfrak{F}_t\}_{t \geq 0}$ zgodną z procesem Wienera $\{W_t\}_{t \geq 0}$, tzn.:

Definicja 1.7

Filtracją zgodną z procesem Wienera $\{W_t\}_{t \geq 0}$ nazywa się rodzinę σ -algebr $\{\mathfrak{F}_t\}_{t \geq 0}$ spełniającą następujące warunki:

- (i) $\mathfrak{N} \subset \mathfrak{F}_s \subset \mathfrak{F}_t \subset \mathfrak{F}$, $0 \leq s < t$, gdzie $\mathfrak{N}$ jest rodziną zbiorów miary zero;
- (ii) zmienna losowa W_t jest $\mathfrak{F}_t$ -mierzalna dla każdego $t \geq 0$;
- (iii) $\forall 0 \leq s < t$ zmienna losowa $W_t - W_s$ jest niezależna od σ -algebry $\mathfrak{F}_s$.

Podstawowym terminem, na którym opiera się teoria SRR Itô, jest pojęcie całki stochastycznej Itô $\int_{t_0}^T b(s, \omega) dW_s$ (por. Itô 1942, 1944, 1946, 1951). Całkę tę określa się

dla zbioru funkcji $M_2[t_0, T]$, tzn. dla funkcji $b: [t_0, T] \times \Omega \rightarrow \mathbb{R}$ mierzalnych względem układu zmiennych (t, ω) (tzn. względem σ -algebry $\mathcal{B}([t_0, T]) \times \mathfrak{F}$), nieantycypujących dla każdego $t \in [t_0, T]$ (tzn. $b(t, \cdot)$ jest mierzalna względem σ -ciała $\mathfrak{F}_t$ dla każdego $t \in [t_0, T]$) oraz z prawdopodobieństwem jeden zachodzi: $\int_{t_0}^T |b(t, \omega)|^2 dt < \infty$ (gdzie całkę przy ustalonym ω oblicza się w sensie Riemanna).

Dla tak określonego zbioru funkcji definiuje się pojęcie całki Itô. Całkę tę wprowadza się najpierw dla funkcji schodkowych (inaczej skokowych)¹⁴⁾ w następujący sposób:

$$\int_{t_0}^T b(t, \omega) dW_t(\omega) = \sum_{k=0}^{n-1} b(t_k, \omega) [W_{t_{k+1}}(\omega) - W_{t_k}(\omega)] \quad (1.7)$$

gdzie $t_0 < t_1 < \dots < t_n = T$ jest stosownym podziałem odcinka całkowania. Następnie z faktu, że dla dowolnej funkcji $b \in M_2[t_0, T]$ istnieje ciąg funkcji schodkowych $b_n \in M_2[t_0, T]$

taki, że z prawdopodobieństwem jeden zachodzi $\lim_{n \rightarrow \infty} \int_{t_0}^T |b(t, \omega) - b_n(t, \omega)|^2 dt = 0$

¹⁴⁾ To znaczy dla funkcji, dla których istnieje podział odcinka $[t_0, T]$: $t_0 < t_1 < \dots < t_n = T$ taki, że $b(t, \omega) = b(t_i, \omega)$ dla $t \in [t_i, t_{i+1})$, $i = 0, 1, \dots, n-1$.

(por. Gichman i Skorochod 1968), całkę stochastyczną Itô z funkcji b definiuje się jako granicę (wg prawdopodobieństwa) całek dla funkcji schodkowych, tzn.:

$$\int_{t_0}^T b(t, \omega) dW_t(\omega) := \lim_{n \rightarrow \infty} \int_{t_0}^T b_n(t, \omega) dW_t(\omega) \quad (1.8)$$

Należy nadmienić, że granica nie zależy od doboru ciągu b_n .

Całki Itô nie można definiować jako całki Riemanna–Stieltjesa, ze względu na nieskończone wahanie trajektorii procesu Wienera (por. Karatzas i Shreve 1988). Ponadto całka Itô jest równa granicy (wg prawdopodobieństwa) sum aproksymacyjnych całki Riemanna–Stieltjesa, dla których punkty pośrednie są lewymi krańcami podziału odcinka $[t_0, T]$. Dobór punktów pośrednich jest istotny, gdyż prowadzi do różnych wyników, tzn. całek stochastycznych¹⁵⁾ (por. Protter 1990). W szczególności przyjęcie środków przedziałów jako punktów pośrednich określa tzw. całkę Stratonovicha (por. Stratonovich 1965, 1966), natomiast przyjęcie prawych krańców definiuje tzw. całkę wsteczną (por. Schuss 1989). Szczegółowe informacje dotyczące definicji i własności całki Itô można znaleźć m.in. w następujących pracach: Gichman i Skorochod (1968), Ikeda i Watanabe (1981), Lipcer i Szirajew (1981), Karatzas i Shreve (1988), Kloeden i Platen (1992), Sobczyk (1996), Øksendal (2000) oraz Brzeźniak i Zastawniak (2002).

Kolejnym ważnym terminem wykorzystywanym w pracy jest pojęcie procesu Itô (por. Lipcer i Szirajew 1981).

Definicja 1.8

Proces stochastyczny o ciągłych trajektoriach $\{X_t\}_{t \in [t_0, T]}$ nazywany jest procesem Itô względem procesu Wienera W_t i filtracji $\mathfrak{F}_t$, jeśli istnieją funkcje a i b mierzalne względem σ -algebry $\mathcal{B}[t_0, T] \times \mathfrak{F}$, będące procesami nieantycypującymi, dla których zachodzą (z prawdopodobieństwem jeden) następujące warunki:

$$(i) \int_{t_0}^T |a(t, \omega)| dt < \infty,$$

$$(ii) \int_{t_0}^T b^2(t, \omega) dt < \infty,$$

¹⁵⁾ Dobierając punkty pośrednie w zależności od $\lambda \in [0, 1]$ w postaci $(1-\lambda)t_j^{(n)} + \lambda t_{j+1}^{(n)}$ dla $i = 0, 1, \dots, n-1$, gdzie $a = t_1^{(n)} < t_2^{(n)} < \dots < t_{n+1}^{(n)} = b$ stanowią n -ty podział odcinka $[a, b]$ taki, że $\max \left\{ |t_{i+1}^{(n)} - t_i^{(n)}| : i = 1, \dots, n-1 \right\} \rightarrow 0$ dla $n \rightarrow \infty$, otrzymuje się, przechodząc do granicy, $\int_{t_0}^T W(s, \omega) dW(s, \omega) = \frac{1}{2}W(T, \omega)^2 - \frac{1}{2}W(t_0, \omega)^2 + \left(\lambda - \frac{1}{2}\right)(T - t_0)$, co potwierdza zależność granicy od doboru punktu pośredniego.

(iii) $X_t(\omega) = X_{t_0}(\omega) + \int_{t_0}^t a(s, \omega) ds + \int_{t_0}^t b(s, \omega) dW_s(\omega)$ dla $t \in [t_0, T]$, przy czym $\int_{t_0}^t a(s, \omega) ds$ jest całką Riemanna liczoną na każdym ustalonym elemencie $\omega \in \Omega$, natomiast $\int_{t_0}^t b(s, \omega) dW_s(\omega)$ jest całką Itô.

Warunek (iii) definicji 1.8 zapisuje się często w formie skróconej:

$$dX_t = a(t) dt + b(t) dW_t \quad (1.9)$$

Ponadto, jeśli proces X spełnia równanie (1.9), to mówi się, że ma on stochastyczną różniczkę Itô dX_t . Często w zastosowaniach wykorzystuje się fakt istnienia różniczki stochastycznej dla transformacji procesu X . Własność tę określa poniższe twierdzenie zwane wzorem Itô:

Twierdzenie 1.3

Jeśli funkcja $u: [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ posiada ciągłe pochodne cząstkowe $\frac{\partial u}{\partial t}(t, x)$, $\frac{\partial u}{\partial x}(t, x)$ i $\frac{\partial^2 u}{\partial x^2}(t, x)$, a ponadto X ma różniczkę stochastyczną postaci (1.9), to proces stochastyczny $Y_t(\omega) = u(t, X_t(\omega))$ ma różniczkę stochastyczną postaci:

$$dY_t = \left(\frac{\partial u}{\partial t}(t, X_t) + a(t) \frac{\partial u}{\partial x}(t, X_t) + \frac{1}{2} b^2(t) \frac{\partial^2 u}{\partial x^2}(t, X_t) \right) dt + b(t) \frac{\partial u}{\partial x}(t, X_t) dW_t \quad (1.10)$$

Ważną klasę procesów stochastycznych stanowią rozwiązania stochastycznych równań różniczkowych (SRR) Itô:

$$\begin{cases} dX_t = a(t, X_t) dt + b(t, X_t) dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.11)$$

Zapis (1.11) jest uproszczoną formą układu równań (tzw. postaci całkowej¹⁶⁾):

$$\begin{cases} X_t(\omega) = X_{t_0}(\omega) + \int_{t_0}^t a(s, X_s(\omega)) ds + \int_{t_0}^t b(s, X_s(\omega)) dW_s(\omega) \\ X_{t_0}(\omega) = x_0(\omega) \end{cases} \quad (1.12)$$

¹⁶⁾ Układ równań (1.12) ma właściwy sens matematyczny w odróżnieniu od (1.11). Jednak z powodu prostoty zapisu i łatwości interpretacji w literaturze przedmiotu powszechnie stosuje się zapis uproszczony (1.11).

W pracy pod pojęciem stochastycznego równania różniczkowego rozumiane jest SSR w sensie Itô.

Jeśli funkcje a i b nie zależą od zmiennej t , wówczas mówi się o autonomicznych SRR. W literaturze przedmiotu funkcja a nazywana jest funkcją lub współczynnikiem dryfu lub przenoszenia, funkcja b (lub b^2) nazywana jest funkcją lub współczynnikiem dyfuzji lub zmienności. Rozwiązanie powyższych układów określa się w następujący sposób:

Definicja 1.9

Niech funkcje a i b będą jak w definicji 1.8. Ponadto niech zmienna losowa x_0 będzie mierzalna względem $\mathcal{F}_t$ oraz niezależna od $W_t - W_{t_0}$ dla $t > t_0$. Wówczas proces X jest silnym, inaczej mocnym, rozwiązaniem (lub po prostu rozwiązaniem) układu (1.11), gdy X_t jest $\mathcal{F}_t$ mierzalną zmienną losową spełniającą (z prawdopodobieństwem jeden) układ (1.12) dla $t \in [t_0, T]$.

W teorii stochastycznych równań różniczkowych wyróżnia się dwa pojęcia rozwiązań: mocne (inaczej silne) i słabe. Mocne rozwiązanie (por. definicja 1.9) jest procesem, który przy zadanej (dowolnej) przestrzeni probabilistycznej i procesie Wienera spełnia układ (1.12). Natomiast słabym rozwiązaniem SRR jest proces, dla którego istnieje przestrzeń probabilistyczna i proces Wienera, dla których spełniony jest układ (1.12). W kolejnych rozdziałach monografii rozważane będą wyłącznie mocne rozwiązania. Szczegóły dotyczące słabych rozwiązań można znaleźć m.in. w następujących pracach: Ikeda i Watanabe (1981), Karatzas i Shreve (1988).

Należy zauważyć, że dla każdego rozwiązania układu (1.11) istnieje stochastycznie równoważne mu rozwiązanie ciągłe – tzn. rozwiązanie o ciągłych trajektoriach – do którego zwykle w literaturze przedmiotu zawęża się rozumowanie (por. Gichman i Skorochod 1968). Przegląd wielu znanych SRR wraz z ich rozwiązaniami znajduje się w pracy Kloeden i Platen (1992).

Poniżej przytoczone zostanie podstawowe twierdzenie (w jednej z możliwych form) gwarantujące istnienie i jedność silnego rozwiązania stochastycznego równania różniczkowego oraz wyjaśniające związek pomiędzy rozwiązaniami SRR a procesami dyfuzji (por. Arnold 1974; Sobczyk 1996).

Twierdzenie 1.4

Niech funkcje a i b będą określone i mierzalne na zbiorze $[t_0, T] \times \mathbb{R}$ oraz niech istnieje stała $K > 0$ taka, że dla $x, y \in \mathbb{R}$ i $t \in [t_0, T]$:

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K|x - y| \quad (\text{warunek Lipschitza}) \quad (1.13)$$

$$a^2(t, x) + b^2(t, x) \leq K(1 + x^2) \quad (\text{warunek wzrostu}) \quad (1.14)$$

Ponadto niech X_{t_0} będzie zmienną losową, dla której zachodzi $E(X_{t_0}^2) < \infty$.

Wówczas istnieje dokładnie jedno¹⁷⁾ mocne rozwiązanie układu (1.11) o ciągłych trajektoriach. Rozwiązanie to jest procesem Markowa z rozkładem początkowym $P_{t_0}(B) = P(X_{t_0} \in B)$ i rozkładem przejścia $P(\cdot | x, s, t-s)$ (por. definicja 1.3), dla którego $E\left(\sup_{t \in [t_0, T]} X_t^2\right) < \infty$. Przy czym, jeśli współczynniki SRR Itô nie zależą od t , tzn. $a(t, x) = a(x)$ i $b(t, x) = b(x)$, to rozwiązanie X jest jednorodnym procesem Markowa.

Jeśli dodatkowo funkcje $a(t, x)$, $b(t, x)$ są ciągłe względem $t \in [t_0, T]$, to rozwiązanie X jest procesem dyfuzji w przedziale $[t_0, T]$ z funkcją dryfu $a(t, x)$ i funkcją dyfuzji $b^2(t, x)$ (por. definicja 1.5).

Cytowane twierdzenie jest prawdziwe przy założeniu silnego warunku Lipschitza (zwanego globalnym) – często niespełnionego w praktyce. Okazuje się, że warunek ten można zastąpić przez lokalny warunek Lipschitza:

$$\begin{aligned} \forall N > 0 \exists K_N : [t \in [t_0, T], |x| \leq N, |y| \leq N \Rightarrow \\ \Rightarrow |a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K_N |x - y|]. \end{aligned}$$

Znane są również inne modyfikacje założeń twierdzenia 1.4 (por. Karatzas i Shreve 1988; Ikeda i Watanabe 1981; Kloeden i Platen 1992).

Naturalne jest pytanie, kiedy proces dyfuzji jest rozwiązaniem SRR Itô. Przy stosownych założeniach można pokazać, że istnieje proces Wienera oraz układ równań (1.11), którego rozwiązaniem jest zadany proces dyfuzji – jest to wówczas rozwiązanie w sensie słabym (por. Prohorov i Rozanov 1969; Kowalenko, Kuzniecowa i Szurienkow 1989; Kloeden i Platen 1992).

Ze względu na zastosowanie SRR Itô do opisu ruchu cząsteczki w procesie dyfuzji (por. Øksendal 2000) oraz ze względu na tezę twierdzenia 1.4 wraz z podanymi po nim uwagami, przyjęto utożsamiać¹⁸⁾ procesy dyfuzji z rozwiązaniami SRR. W niniejszej pracy brane są pod uwagę procesy dyfuzji będące rozwiązaniami SRR. Procesy te, dla uwypuklenia przynależności do obydwu klas, określa się jako procesy dyfuzji Itô (por. Øksendal 2000; Janicki i Izydoreczek 2001) i stąd uzasadnione jest stosowanie wymiennie dla rozważanych w monografii procesów pojęcia procesu dyfuzji, procesu Itô oraz rozwiązania SRR.

¹⁷⁾ Jeśli istnieje inny proces ξ będący również silnym rozwiązaniem (1.11), to zachodzi:

$$P\left(\sup_{t \in [t_0, T]} |X_t - \xi_t| > 0\right) = 0.$$

¹⁸⁾ W pracy L. Arnolda z 1974 roku zwrócono uwagę, że większość rozważanych i istotnych w praktyce procesów Itô będących rozwiązaniami układów (1.11) jest procesami dyfuzji i na odwrót. Stąd te dwie klasy procesów zdefiniowane w odmienny sposób zasadniczo uznaje się za identyczne.

1.4. Dyskretna aproksymacja rozwiązań SRR

Ponieważ w wielu przypadkach nie można wskazać jawnego wzoru na rozwiązanie SRR, pomimo wiedzy o jego istnieniu i jedyności, zmusza to do wnioskowania o jego własnościach poprzez badanie jego przybliżenia. Użyteczna w takich sytuacjach okazuje się teoria dotycząca analizy numerycznej stochastycznych równań różniczkowych. Wśród bogatej literatury poświęconej numerycznym aspektom SRR można wymienić m.in. następujące prace: Kloeden i Platen (1992), Janicki i Weron (1994), Janicki (1996), Platen (1999), Cyganowski, Kloeden i Ombach (2001), Higham (2001) oraz Janicki i Izydoreczyk (2001). W niniejszym podrozdziale zaprezentowane zostaną wybrane algorytmy umożliwiające wygenerowanie trajektorii „bliskiej” rzeczywistej realizacji badanego procesu. Pojęcie „bliskiej” definiowane jest w analizie numerycznej poprzez odpowiedni sposób zbieżności. Umiejętność wygenerowania przybliżonej trajektorii rozwiązania SRR umożliwia w szczególności graficzną interpretację przebiegu procesu oraz wyznaczenie linii kwantylowych, ilustrujących rozkład procesu w danej chwili.

Podstawowe schematy dyskretyzacji

Pojęcie dyskretyzacji procesu (modelu) opisującego zmieniające się w czasie zjawisko polega na dyskretyzacji osi czasu, tzn. polega na wprowadzeniu skończonego podziału przedziału czasu, a następnie zdefiniowaniu w punktach podziału skończonej rodziny zmiennych losowych (dyskretnego procesu stochastycznego) aproksymujących w ustalonym sensie proces ciągły. Określony w ten sposób proces dyskretny nazywany jest numeryczną lub dyskretną aproksymacją procesu ciągłego.

Już pierwsze prace oparte na symulacyjnych eksperymentach i badaniach teoretycznych (por. Clements i Anderson 1973; Wright 1974; Clark i Cameron 1980) pokazały, że nie wszystkie klasyczne i heurystyczne dyskretyzacje SRR zbiegają w użytecznym sensie do rozwiązania SRR.

Efektywność schematów dyskretyzacji określa się poprzez sposób i rząd zbieżności do prawdziwego procesu. Rozróżnia się dwa podstawowe typy zbieżności, które opisują aproksymację trajektorii lub aproksymację wybranych charakterystyk procesu. Poniżej przytoczone zostaną definicje określające rodzaj zbieżności aproksymacji Y (otrzymanej poprzez dyskretyzację SRR) do procesu X będącego rozwiązaniem równania Itô.

Definicja 1.10

Aproksymujący proces Y zbiega w silnym sensie z rzędem $v \in (0, \infty]$ w chwili T do procesu X , jeśli istnieją stałe $K < \infty$ i $\Delta_0 > 0$ takie, że

$$E(|X_T - Y_T|) \leq K \Delta^v$$

dla dowolnego kroku dyskretyzacji $\Delta \in (0, \Delta_0)$.

W wielu zagadnieniach praktycznych nie jest wymagana znajomość zachowania się trajektorii procesu, wystarcza natomiast znajomość aproksymacji rozkładu procesu w pewnym punkcie osi czasu, czy też aproksymacji jego momentów tzn. $E(g(X_T))$.

Definicja 1.11

Aproksymacja Y zbiega w słabym sensie z rzędem $v \in (0, \infty]$ w chwili T do procesu X , jeśli dla dowolnego wielomianu g istnieją $E(g(X_T))$, $E(g(Y_T))$ oraz stałe $K < \infty$ i $\Delta_0 > 0$ takie, że:

$$\left| E(g(X_T)) - E(g(Y_T)) \right| \leq K \Delta^v$$

dla dowolnego kroku dyskretyzacji $\Delta \in (0, \Delta_0)$.

W literaturze przedmiotu najczęściej rozpatrywanymi schematami dyskretyzacji są schemat Eulera (zastosowany po raz pierwszy przez Maruyame w 1955 roku i stąd nazywany również schematem Eulera–Maruyamy) oraz schemat Milsteina (por. Milstein 1974, 1978). Poniżej podane zostaną definicje i własności wspomnianych schematów.

Niech proces $\{X_t\}_{t \in [t_0, T]}$ będzie rozwiązaniem układu równań:

$$\begin{cases} dX_t = a(t, X_t) dt + b(t, X_t) dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.15)$$

Ponadto niech $N > 0$ będzie liczbą naturalną, względem której definiuje się podział odcinka $[t_0, T]$: $t_0 = \tau_0 < \tau_1 < \dots < \tau_N = T$. Dla prostoty rozumowania zakłada się równy podział, tzn. $\Delta \equiv \tau_{k+1} - \tau_k = \frac{T - t_0}{N}$.

Dyskretyzując układ równań (1.15) względem schematu Eulera, otrzymujemy proces stochastyczny $Y = \{Y_{k\Delta}, k = 0, \dots, N\}$ spełniający następujący związek:

$$\begin{cases} Y_{k\Delta} = Y_{(k-1)\Delta} + a((k-1)\Delta, Y_{(k-1)\Delta})\Delta + b((k-1)\Delta, Y_{(k-1)\Delta})(W_{k\Delta} - W_{(k-1)\Delta}) \\ Y_0 = x_0 \end{cases} \quad (1.16)$$

Należy nadmienić, że gęstość przejścia p_Y procesu Y jest postaci:

$$p_Y(\cdot | y; k\Delta, \Delta) = \phi(\cdot | y + a(k\Delta, y)\Delta; b^2(k\Delta, y)\Delta) \quad (1.17)$$

gdzie $\phi(\cdot | m, v)$ jest gęstością rozkładu normalnego, dla którego m i v są odpowiednio wartością oczekiwaną i wariancją. Funkcję tę często wykorzystuje się jako aproksymację gęstości przejścia procesu X .

Najczęściej spotykaną praktyką dookreślenia procesu Y na cały przedział $[t_0, T]$ jest wykorzystanie do tego celu interpolacji liniowej:

$$Y_t = Y_{(k-1)\Delta} + (Y_{k\Delta} - Y_{(k-1)\Delta})(t - (k-1)\Delta)/\Delta, \quad t \in [(k-1)\Delta, k\Delta).$$

Poniżej przytoczone zostanie twierdzenie (w jednej z możliwych form) opisujące sposób zbieżności rozwiązania równania różnicowego (1.16) do rozwiązania SSR (1.15) (por. Milstein 1978; Talay 1984; Kloeden i Platen 1992; Janicki i Izydorczyk 2001).

Twierdzenie 1.5

Niech $E\left(\left|X_{t_0}\right|^2\right) < \infty$. Ponadto niech dla dowolnych $s, t \in [t_0, T]$ oraz $x, y \in \mathbb{R}$ istnieje stała $K > 0$ niezależna od Δ taka, że spełnione są następujące związki:

$$|a(t, x) - a(t, y)| + |b(t, x) - b(t, y)| \leq K|x - y|,$$

$$|a(t, x)| + |b(t, x)| \leq K(1 + |x|),$$

$$|a(t, x) - a(s, x)| + |b(t, x) - b(s, x)| \leq K(1 + |x|)|t - s|^{1/2}.$$

Wówczas¹⁹⁾:

$$E|X_T - Y_T| \leq C\Delta^{1/2}, \quad \text{gdzie } C \text{ jest stałą niezależną od } \Delta,$$

ponadto proces Y jest zbieżny jednostajnie w sensie średniokwadratowym do X , gdy $\Delta \rightarrow 0$:

$$E|X_t - Y_t|^2 = O(\Delta) \quad \text{dla } t \in [0, T].$$

W przypadku autonomicznego SRR, jeśli funkcje $a(x)$ i $b(x)$ są czterokrotnie różniczkowalne, wówczas schemat Eulera jest schematem rzędu pierwszego w sensie słabym²⁰⁾.

Z twierdzenia 1.5 wynika, że aproksymacja Eulera zbiega w silnym sensie do rozwiązania równania Itô z rzędem $\nu = 0,5$. Warto nadmienić, że dla deterministycznych równań różniczkowych schemat ten jest rzędu $\nu = 1$ w silnym sensie²¹⁾.

¹⁹⁾ Założenia twierdzenia można osłabić (por. Kloeden i Platen 1992; Pedersen 1995b).

²⁰⁾ Przy odpowiedniej modyfikacji założeń tezę tę da się uogólnić na równania nieautonomiczne (por. Kloeden i Platen 1992).

²¹⁾ W przypadku deterministycznym silną zbieżność rzędu $\nu \in (0, \infty]$ definiuje się w następujący sposób: $\exists K < \infty, \Delta_0 > 0 : |X_T - Y_T| \leq K\Delta^\nu$ dla dowolnego kroku dyskretyzacji $\Delta \in (0, \Delta_0)$.

Dyskretyzując układ (1.15) względem schematu Milsteina, otrzymuje się proces Y spełniający układ:

$$\begin{cases} Y_{k\Delta} = Y_{(k-1)\Delta} + a((k-1)\Delta, Y_{(k-1)\Delta})\Delta + b((k-1)\Delta, Y_{(k-1)\Delta})(W_{k\Delta} - W_{(k-1)\Delta}) + \\ \quad + \frac{1}{2}b((k-1)\Delta, Y_{(k-1)\Delta}) \frac{\partial b((k-1)\Delta, x)}{\partial x} \Big|_{x=Y_{(k-1)\Delta}} \left[(W_{k\Delta} - W_{(k-1)\Delta})^2 - \Delta \right] \\ Y_0 = x_0. \end{cases} \quad (1.18)$$

Należy zauważyć, że jeśli $b(t, x) = b(t)$, a stąd $\frac{\partial b(t, x)}{\partial x} = 0$, to schemat Milsteina jest równoważny schematowi Eulera. Natomiast w przeciwnym razie, gęstość przejścia procesu Y jest postaci (por. Elerian 1998):

$$p_Y(y_{(k+1)\Delta} | y; k\Delta, \Delta) = \frac{\exp(-\lambda/2)}{|A|\sqrt{2\pi}} (z)^{-1/2} \exp\left(-\frac{z}{2}\right) \cosh(\sqrt{\lambda}z) \mathbf{I}_{z \geq 0},$$

gdzie:

$$z = \frac{y_{(k+1)\Delta} - B(k\Delta, y)}{A(k\Delta, y)},$$

$$\lambda = \frac{1}{\Delta} \left(\frac{\partial b}{\partial x}(k\Delta, y) \right)^{-2},$$

$$A = \frac{\Delta}{2} b \frac{\partial b}{\partial x},$$

$$B = -\frac{b}{2} \left(\frac{\partial b}{\partial x} \right)^{-1} + y_{k\Delta} + a\Delta - A.$$

Zbieżność rozwiązania równania różnicowego (1.18) do rozwiązania SSR (1.15) rozpatrywanego schematu opisuje następujące twierdzenie (por. Kloeden i Platen 1992; Janicki i Izydorczyk 2001):

Twierdzenie 1.6

Niech $E\left(\left|X_{t_0}\right|^2\right) < \infty$. Ponadto niech dla dowolnych $s, t \in [t_0, T]$ oraz $x, y \in \mathbb{R}$ istnieje stała $K > 0$ niezależna od Δ , taka że spełnione są następujące związki:

$$\begin{aligned}
|a^*(t, x; \theta) - a^*(t, y; \theta)| &\leq K|x - y|, \\
|b(t, x; \theta) - b(t, y; \theta)| &\leq K|x - y|, \\
|Lb(t, x; \theta) - Lb(t, y; \theta)| &\leq K|x - y|, \\
|a^*(t, x; \theta)| + |La(t, y; \theta)| &\leq K(1 + |x|), \\
|b(t, x; \theta)| + |Lb(t, x; \theta)| &\leq K(1 + |x|), \\
|L^2b(t, x; \theta)| &\leq K(1 + |x|), \\
|a^*(t, x; \theta) - a^*(s, x; \theta)| &\leq K(1 + |x|)|t - s|^{1/2}, \\
|b(t, x; \theta) - b(s, x; \theta)| &\leq K(1 + |x|)|t - s|^{1/2}, \\
|Lb(t, x; \theta) - Lb(s, x; \theta)| &\leq K(1 + |x|)|t - s|^{1/2},
\end{aligned}$$

gdzie $a^* = a - \frac{1}{2}bb'$, $L = b \frac{\partial}{\partial y}$.

Wówczas:

$$E|X_T - Y_T| \leq C\Delta, \text{ gdzie } C \text{ jest stałą niezależną od } \Delta.$$

Ponadto proces Y jest zbieżny jednostajnie w sensie średniokwadratowym do X , gdy $\Delta \rightarrow 0$:

$$E|X_t - Y_t|^2 = O(\Delta^2) \text{ dla } t \in [0, T].$$

W przypadku autonomicznego SRR, jeśli funkcje $a(x)$ i $b(x)$ są sześciokrotnie ciągle różniczkowalne, wówczas schemat Milsteina jest schematem rzędu drugiego w sensie słabym.

Z twierdzenia 1.6 wynika, że dyskretyzacja Milsteina zbiega w silnym sensie do rozwiązania równania Itô z rzędem $\nu = 1$.

Możliwe jest wskazanie sposobów dyskretyzacji o wyższych rzędach zbieżności niż rozpatrywane tutaj schematy Eulera i Milsteina. Uzyskuje się je poprzez uwzględnienie kolejnych wyrazów wzoru Taylora dla SRR Itô (por. Wagner i Platen 1978; Komori, Saito i Mitsui 1994; Saito i Mitsui 1996; Kloeden i Platen 1992; Platen 1999). Zagadnienie to nie będzie jednak rozważane w prezentowanej monografii.

W zastosowaniach często rozważa się SRR, których rozwiązania są procesami przyjmującymi wartości dodatnie, jednakże aproksymacja rozwiązania otrzymana w wyniku dyskretyzacji może przyjmować wartości ujemne. Wówczas często spotykaną praktyką jest transformacja wyjściowego układu równań (np.: poprzez funkcję logarytm) przy użyciu wzoru Itô (por. twierdzenie 1.3) do procesu o wartościach rzeczywistych, do którego dopiero stosuje się schemat dyskretyzacji. Unika się wówczas sytuacji, w której nieujemny proces jest przybliżany procesem przyjmującym wartości ujemne.

Linie kwantylowe

Umiejętność wygenerowania trajektorii procesu Itô lub jej aproksymacji za pomocą schematów dyskretyzacji daje możliwość opisu procesu za pomocą tzw. linii kwantylowych (por. Janicki i Izydorczyk 2001). Pojęcie to definiuje się w następujący sposób:

Definicja 1.12

Dla ustalonego $p \in (0, 1)$ oraz procesu stochastycznego $\{X_t\}_{t \in [t_0, T]}$ linią p -kwantylową nazywana jest deterministyczna funkcja $q: [t_0, T] \rightarrow \mathbb{R}$ spełniająca następujący związek:

$$P(X_t \leq q_p(t)) = p \text{ dla } t \in [t_0, T].$$

Obliczenie wartości $q_p(t)$ odpowiada wyznaczeniu kwantyla rzędu p zmiennej losowej X_t . W niniejszej pracy przyjęto następujący algorytm oszacowania p -tego kwantyla (por. Janicki i Izydorczyk 2001):

1. Uporządkowanie w sposób niemalejący wygenerowanej R -elementowej próbki zmiennej X_t lub jej aproksymacji Y_t .
2. Wyznaczenie indeksu, tj. funkcji $k = k(p)$ zdefiniowanej w następujący sposób:

$$k(p) = \begin{cases} pR, & \text{gdy } pR \in \mathbb{N}, \\ \lfloor pR \rfloor + 1, & \text{gdy } pR \notin \mathbb{N}. \end{cases}$$

Za p -ty kwantyl przyjmuje się $k(p)$ -ty element uporządkowanej próbki.

Zastosowanie schematów dyskretyzacji do aproksymacji trajektorii procesu oraz wykorzystanie linii kwantylowych do graficznej ilustracji rozkładów rozpatrywanych w pracy procesów dyfuzji znajduje się w podrozdziale 1.5. Ponadto schemat Eulera stanowi podstawę jednej z prezentowanych w pracy metod aproksymacji gęstości przejścia omówionej w podrozdziale 2.3.

1.5. Wybrane procesy dyfuzji Itô

Celem niniejszego podrozdziału jest zaprezentowanie rozpatrywanych w dalszej części monografii procesów dyfuzji. Jak wspomniano wcześniej, rozważania ograni-

czono wyłącznie do procesów dyfuzji Itô, tzn. procesów dyfuzji będących rozwiązaniami SRR. W pracy analizowanych będzie pięć modeli: Vasička, Blacka i Scholesa, Coxa, Ingersolla i Rossa, Brennana i Schwartza oraz uogólniony model Ornsteina i Uhlenbecka. Każdy z modeli zadany jest poprzez stosowne stochastyczne równania różniczkowe.

Model Vasička (V) (por. Vasiček 1977)

Model zadany jest poprzez liniowe autonomiczne SRR postaci²²⁾:

$$\begin{cases} dX_t = (\alpha - \beta X_t)dt + \sigma dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.19)$$

gdzie o parametrach α , β i σ zakłada się, że są dodatnie. Rozwiązanie układu (1.19) jest następującej postaci:

$$X_t = X_{t_0} e^{-\beta(t-t_0)} + \frac{\alpha}{\beta} \left(1 - e^{-\beta(t-t_0)}\right) + \sigma \int_{t_0}^t e^{-\beta(t-u)} dW_u \quad (1.20)$$

Rozkład przejścia procesu jest rozkładem normalnym o gęstości zadanej wzorem:

$$p_X^{(V)}(x | x_0; t_0, \Delta) = \phi(x; m, v) \quad (1.21)$$

gdzie: $\Delta = t - t_0$, $m = \frac{\alpha}{\beta} + \left(x_0 - \frac{\alpha}{\beta}\right) e^{-\beta\Delta}$ i $v = \sigma^2 \frac{1 - e^{-2\beta\Delta}}{2\beta}$.

Warto zauważyć, że rozważany proces jest jednorodnym procesem Markowa. Ponadto rozwiązanie układu (1.19) jest procesem powracającym do średniej²³⁾ równej α/β , wielkość ta nazywana jest długoterminowym średnim poziomem procesu (*mean reversion level, long run equilibrium value*). Parametr β utożsamiany jest z szybkością, z jaką proces zmierza do średniej równej α/β , natomiast parametr σ lub σ^2 nazywany jest parametrem zmienności. Można zauważyć, że funkcja dryfu procesu jest dodatnia, jeśli wartość procesu jest poniżej wielkości α/β , oraz ujemna, jeśli wartość procesu

²²⁾ W literaturze przedmiotu rozważa się często równoważną do (1.19) postać SRR:

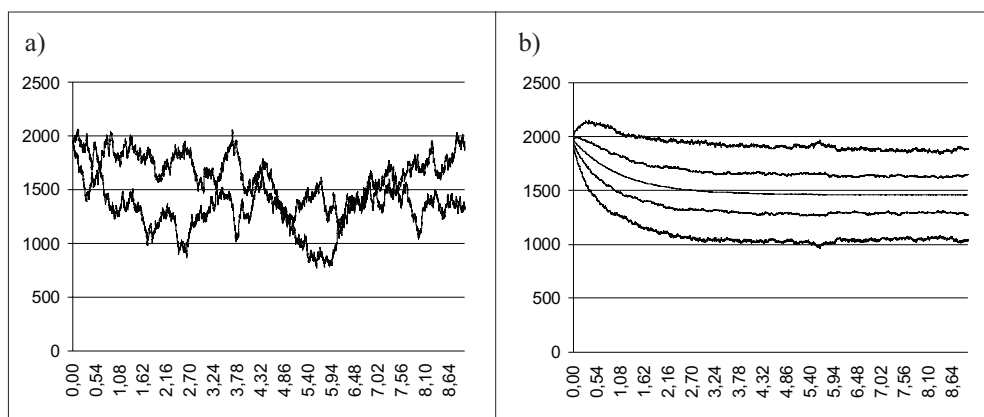
$$\begin{cases} dX_t = \beta(\mu - X_t)dt + \sigma dW_t \\ X_{t_0} = x_0 \end{cases}, \text{ gdzie } \mu = \frac{\alpha}{\beta}.$$

²³⁾ Proces ma własność powracania do średniej, jeśli istnieje i jest skończona granica: $\lim_{t \rightarrow \infty} E(X_t | X_0 = x_0)$ (por. Brigo i Mercurio 2001).

przekracza α/β . Innymi słowy, proces dąży do wartości średniej w każdym momencie czasu. Jeśli $\Delta \rightarrow \infty$, to graniczną gęstością przejścia jest $\phi\left(x; \frac{\alpha}{\beta}, \frac{\sigma^2}{2\beta}\right)$.

W literaturze przedmiotu proces zadany wzorem (1.20) nazywany jest powracającym do średniej procesem Ornsteina i Uhlenbecka (*Ornstein–Uhlenbeck mean reverting process*). Pomimo istotnej wady, tzn. przyjmowania wartości ujemnych, jest on powszechnie wykorzystywany do modelowania zjawisk finansowych – zwłaszcza do opisu chwilowych (krótkoterminowych) stóp procentowych (*short-term rates, spot rates, instantaneous interest rates*) (por. Weron i Weron 1999).

Rysunek 1.4 przedstawia dwie przykładowe trajektorie oraz linie kwantylowe procesu zadanego wzorem (1.20). Ilustrację graficzną wykonano przy następujących założeniach²⁴: $\alpha = 1350,9942$, $\beta = 0,9241$, $\sigma^2 = 129864$, $x_0 = 1960,57$, $T = 9$, $p \in \{5\%, 25\%, 50\%, 75\%, 95\%\}$ (por. definicja 1.12), przy czym linie kwantylowe wyznaczono na podstawie 1000 trajektorii. Widoczna jest zbieżność lini kwantylowych odpowiadająca zbieżności do rozkładu granicznego.



Rys. 1.4. Wybrane trajektorie (a) oraz linie kwantylowe (b) dla rozwiązania modelu V

Źródło: Opracowanie własne

Model Blacka i Scholesa (BS) (por. Black i Scholes 1973)

Model zadany jest poprzez liniowe autonomiczne SRR postaci:

$$\begin{cases} dX_t = \beta X_t dt + \sigma X_t dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.22)$$

²⁴) Przyjęte wartości parametrów są wartościami oczekiwanymi a posteriori prezentowanymi w tabeli 5.2.

gdzie o parametrze β zakłada się, że przyjmuje wartości z całego zbioru liczb rzeczywistych, natomiast σ jest stałą dodatnią. Rozwiązaniem układu (1.22) jest tzw. geometryczny ruch Browna (*geometric Brownian motion*) o następującej postaci:

$$X_t = X_{t_0} \exp \left\{ \left(\beta - \frac{1}{2} \sigma^2 \right) (t - t_0) + \sigma (W_t - W_{t_0}) \right\} \quad (1.23)$$

Ze wzoru (1.23) wynika, że proces przyjmuje wartości większe od zera, jeśli $x_0 > 0$. Rozkład przejścia procesu jest rozkładem lognormalnym o gęstości zadanej wzorem:

$$p_X^{(BS)}(x | x_0; t_0, \Delta) = \frac{1}{\sigma x \sqrt{2\pi\Delta}} \exp \left\{ - \frac{\left[\ln(x) - \left(\ln(x_0) + \left(\alpha - \sigma^2 / 2 \right) \Delta \right) \right]^2}{2\sigma^2 \Delta} \right\} \quad (1.24)$$

przy czym wartość oczekiwana i wariancja rozkładu przejścia wynoszą odpowiednio:

$$E(X_{t_0+\Delta} | X_{t_0} = x_0) = x_0 e^{\beta\Delta},$$

$$Var(X_{t_0+\Delta} | X_{t_0} = x_0) = x_0^2 e^{2\beta\Delta} (e^{\sigma^2\Delta} - 1).$$

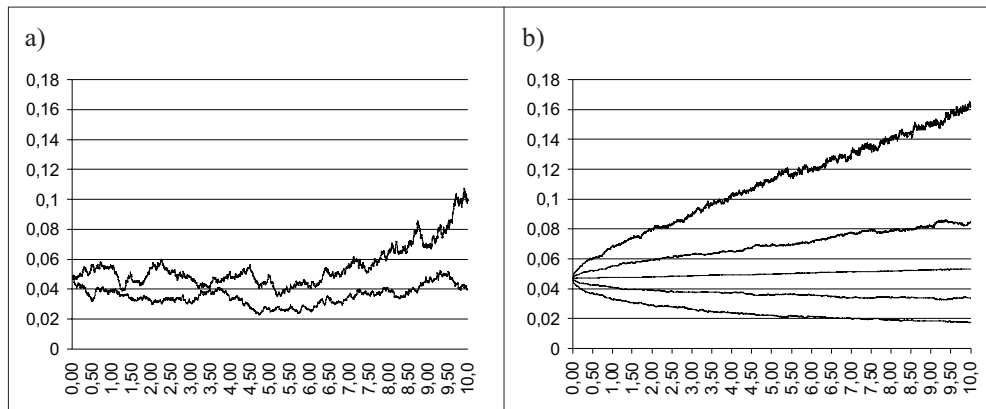
Warto zauważyć, że rozważany proces jest jednorodnym procesem Markowa. Ponadto dla $\beta \leq 0$ rozwiązanie układu (1.22) jest procesem powracającym do średniej równej warunkowi początkowemu x_0 , gdy $\beta = 0$, oraz 0 gdy $\beta < 0$. Parametr β nazywany jest parametrem dryfu, współczynnikiem wzrostu lub oczekiwaną stopą zwrotu (*growth rate, appreciation rate*), natomiast σ lub σ^2 nazywany jest parametrem zmienności procesu. Jeśli $\Delta \rightarrow \infty$ i $\beta > 0$, to zarówno wartość oczekiwana, jak i wariancja rozkładu przejścia rosną w sposób nieograniczony do nieskończoności.

Geometryczny ruch Browna po raz pierwszy został wykorzystany do opisu dynamiki papierów wartościowych w pracach M.F.M. Osborna z 1959 roku i P.A. Samuelsona z 1965 roku (por. Weron i Weron 1999). W 1973 roku Black i Scholes wyznaczyli wzór na cenę europejskiej opcji kupna na akcję niewypłacającą dywidendy przy założeniu, że dynamika akcji opisana jest równaniem (1.22).

Rysunek 1.5 przedstawia dwie przykładowe trajektorie oraz linie kwantylowe procesu zadanego wzorem (1.23). Ilustrację graficzną wykonano przy następujących założeniach²⁵⁾: $\beta = 0,03720$, $\sigma^2 = 0,04812$, $x_0 = 0,0468$, $T = 10$, $p \in \{5\%, 25\%, 50\%, 75\%, 95\%\}$

²⁵⁾ Przyjęte wartości parametrów są wartościami oczekiwanymi a posteriori prezentowanymi w tabeli 4.2.

(por. definicja 1.12), przy czym linie kwantylowe wyznaczono na podstawie 1000 trajektorii. Widoczne jest rosnące rozproszenie rozkładu warunkowego procesu wraz z wydłużaniem się horyzontu czasu.



Rys. 1.5. Wybrane trajektorie (a) oraz linie kwantylowe (b) dla rozwiązania modelu BS

Źródło: Opracowanie własne

Model Coxa, Ingersolla i Rossa (CIR) (por. Cox, Ingersoll i Ross 1985)

Model zadany jest poprzez nieliniowe autonomiczne SRR postaci²⁶⁾:

$$\begin{cases} dX_t = (\alpha - \beta X_t)dt + \sigma\sqrt{X_t}dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.25)$$

gdzie o wielkościach x_0 , α , β i σ zakłada się, że są dodatnie. Ponadto, aby rozwiązanie przyjmowało wyłącznie wartości dodatnie, zakłada się dodatkowo, że $2\alpha > \sigma^2$. Rozwiązaniem SRR (1.25) jest tzw. proces pierwiastka kwadratowego (*square-root process*) o następującej postaci:

$$X_t = \frac{\alpha}{\beta} + \left(X_{t_0} - \frac{\alpha}{\beta} \right) e^{-\beta(t-t_0)} + \sigma \int_{t_0}^t e^{-\beta(t-u)} \sqrt{X_u} dW_u \quad (1.26)$$

²⁶⁾ W literaturze przedmiotu rozważa się często reparametryzację układu postaci:

$$\begin{cases} dX_t = \beta(\mu - X_t)dt + \sigma\sqrt{X_t}dW \\ X_{t_0} = x_0 \end{cases}, \text{ gdzie } \mu = \frac{\alpha}{\beta}.$$

Rozkład przejścia procesu zadany jest poprzez następującą gęstość:

$$p_X^{(CIR)}(x | x_0; t_0, \Delta) = c e^{-u-v} \left(\frac{v}{u} \right)^{q/2} I_q(2\sqrt{uv}) \quad (1.27)$$

gdzie:

$$c = \frac{2\beta}{\sigma^2(1 - e^{-\beta\Delta})},$$

$$u = cx_0 e^{-\beta\Delta},$$

$$v = cx,$$

$$q = \frac{2\alpha}{\sigma^2} - 1,$$

$I_q(\cdot)$ – zmodyfikowana funkcja Bessela²⁷⁾ pierwszego rodzaju, rzędu q :

$$\begin{aligned} I_q(z) &= \left(\frac{z}{2} \right)^q \sum_{k=0}^{\infty} \frac{(z^2/4)^k}{k! \Gamma(q+k+1)} = \\ &= \frac{1}{\pi} \int_0^{\pi} e^{z \cos \varphi} \cos(q\varphi) d\varphi - \frac{\sin(q\pi)}{\pi} \int_0^{\infty} e^{-z \cosh(t) - qt} dt \end{aligned} \quad (1.28)$$

Warunkowa wartość oczekiwana i wariancja zmiennej losowej $X_{t_0+\Delta}$ przy założeniu, że $X_{t_0} = x_0$, dane są następującymi wzorami:

$$E(X_{t_0+\Delta} | X_{t_0} = x_0) = x_0 e^{-\beta\Delta} + \frac{\alpha}{\beta} (1 - e^{-\beta\Delta}),$$

$$Var(X_{t_0+\Delta} | X_{t_0} = x_0) = x_0^2 \left(\frac{\sigma^2}{\beta} \right) (e^{-\beta\Delta} - e^{-2\beta\Delta}) + \frac{\alpha\sigma^2}{2\beta^2} (1 - e^{-\beta\Delta})^2.$$

Warto zauważyć, że rozważany proces jest jednorodnym procesem Markowa. Ponadto rozwiązanie układu (1.25) jest procesem powracającym do średniej równej

²⁷⁾ Zmodyfikowane funkcje Bessela pierwszego i drugiego rodzaju rzędu $q \in \mathbb{R}$ są rozwiązaniami równania różniczkowego $x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - (x^2 + q^2)y = 0$ (por. Weisstein 2003).

α/β . Wielkość α/β nazywana jest długoterminowym średnim poziomem wartości procesu. Parametr β utożsamiany jest z szybkością, z jaką proces zmierza do średniej α/β , natomiast parametr σ lub σ^2 nazywany jest parametrem zmienności. Jeśli $\Delta \rightarrow \infty$, to gęstość przejścia zbiega do gęstości rozkładu gamma: $\frac{\omega^v}{\Gamma(v)} x^{v-1} e^{-\omega x}$, gdzie $\omega = 2\beta/\sigma^2$, $v = 2\alpha/\sigma^2$, dla którego wartość oczekiwana i wariancja wynoszą odpowiednio: $\frac{\alpha}{\beta}$ i $\frac{\alpha\sigma^2}{2\beta^2}$.

Można zauważyć, że gęstość przejścia zmiennej losowej $2cX_t$ jest gęstością niecentralnego rozkładu $\chi^2(\cdot; 2q + 2, 2u)$ (*noncentral chi-squared distribution*), gdzie $2q + 2$ jest liczbą stopni swobody (przyjmującą wartości ze zbioru dodatnich liczb rzeczywistych), natomiast $2u$ jest parametrem niecentralności. Gęstość $\chi^2(\cdot; \kappa, \lambda)$ wyraża się wzorem (por. Brigo i Mercurio 2001):

$$P_{\chi^2(\kappa, \lambda)}(z) = \sum_{i=0}^{\infty} \frac{e^{-\lambda/2} (\lambda/2)^i}{i!} P_{\chi^2(\kappa+2i)}(z) \quad (1.29)$$

gdzie $P_{\chi^2(\kappa+2i)}(\cdot)$ oznacza gęstość rozkładu χ^2 z $\kappa + 2i$ stopniami swobody.

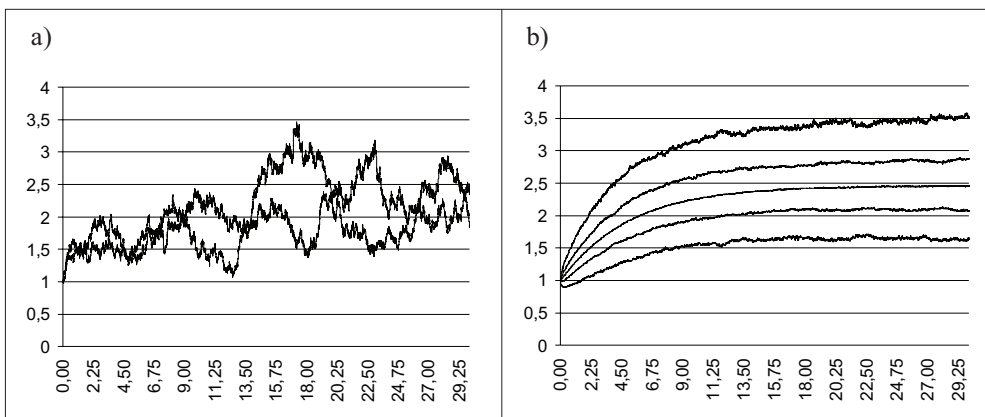
Pomimo jawnego wzoru na gęstość przejścia (1.27), niemożliwe jest dokładne wyliczenie wartości funkcji – ze względu na postać wzorów (1.28) i (1.29) konieczne jest zastosowanie algorytmów numerycznych.

W pracy przyjęto dodatkowe ograniczenie zbioru możliwych parametrów modelu CIR poprzez warunek $\alpha > \frac{3}{4}\sigma^2$, zapewniający całkowalność stosowanej aproksymacji gęstości przejścia (por. rozdz. 2).

Proces pierwiastka kwadratowego jest powszechnie wykorzystywany do modelowania zjawisk finansowych – w szczególności do opisu krótkoterminowych stóp procentowych.

Rysunek 1.6 przedstawia dwie przykładowe trajektorie oraz linie kwantylowe procesu zadanego wzorem (1.26). Ilustrację graficzną wykonano przy następujących założeniach²⁸⁾: $\alpha = 0,5$, $\beta = 0,2$, $\sigma^2 = 0,05$, $x_0 = 1$, $T = 30$, $p \in \{5\%, 25\%, 50\%, 75\%, 95\%\}$ (por. definicja 1.12), przy czym linie kwantylowe wyznaczono na podstawie 1000 trajektorii. Widoczna jest zbieżność lini kwantylowych odpowiadająca zbieżności rozkładu warunkowego wraz z upływem czasu do rozkładu granicznego.

²⁸⁾ Przyjęte wartości parametrów są identyczne jak w przykładzie empirycznym przedstawionym w podrozdziale 3.6.



Rys. 1.6. Wybrane trajektorie (a) oraz linie kwantylowe (b) dla rozwiązania modelu CIR

Źródło: Opracowanie własne

Uogólniony model Ornsteina i Uhlenbecka (EOU) (por. Egorov, Li i Xu 2003)

Model zadany jest poprzez nieliniowe i nieautonomiczne SRR postaci:

$$\begin{cases} dX_t = -\beta X_t dt + \sigma e^{\sigma_1 t} dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.30)$$

gdzie o β i σ_1 zakłada się, że są parametrami przyjmującymi wartości z całego zbioru liczb rzeczywistych, natomiast σ jest stałą dodatnią. Rozwiązaniem układu (1.30) jest proces dany następującym wzorem:

$$X_t = X_{t_0} e^{-\beta(t-t_0)} + \sigma e^{-\beta t} \int_{t_0}^t e^{(\beta+\sigma_1)u} dW_u \quad (1.31)$$

Rozkład przejścia procesu jest rozkładem normalnym o gęstości zadanej wzorem:

$$p_X^{(EOU)}(x | x_0; t_0, \Delta) = \phi(x; m, v) \quad (1.32)$$

gdzie:

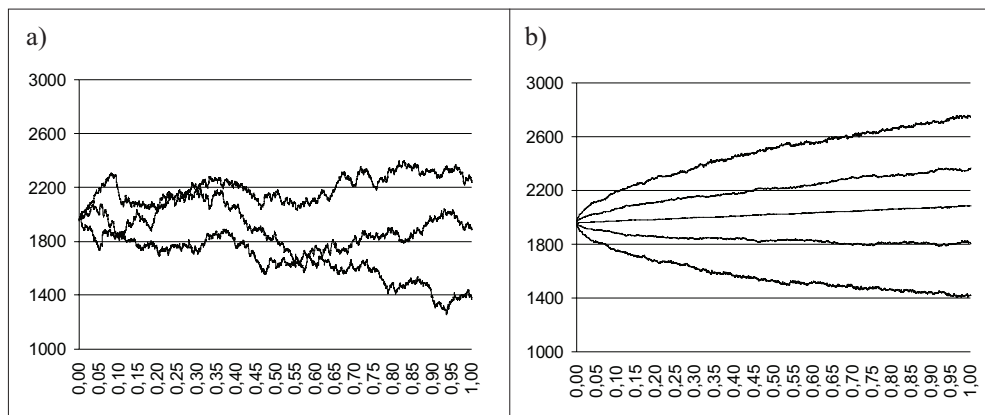
$$m = x_0 \exp(-\beta\Delta),$$

$$v = \begin{cases} \frac{\sigma^2}{2(\beta + \sigma_1)} \left[\exp(2\sigma_1(t_0 + \Delta)) - \exp(2(-\beta\Delta + \sigma_1 t_0)) \right] & \text{dla } \beta + \sigma_1 \neq 0 \\ \sigma^2 \exp(-2\beta(t_0 + \Delta)) \Delta & \text{dla } \beta + \sigma_1 = 0 \end{cases}.$$

Warto zauważyć, że rozważany proces nie jest jednorodnym procesem Markowa. Ponadto, jeśli $\beta \geq 0$, to rozwiązanie układu (1.30) jest procesem powracającym do średniej równej zero. Wówczas parametr β interpretowany jest jako prędkość, z jaką proces powraca do średniej. Parametr σ lub σ^2 nazywany jest parametrem zmienności, natomiast parametr σ_1 odzwierciedla wpływ czasu na funkcję zmienności. Należy zauważyć, że proces (1.31) może przyjmować wartości ujemne. Autorowi nieznane są zastosowania modelu EOU w praktyce.

W pracy przyjęto dodatkowe ograniczenia $\sigma_1 \in [-10, 10]$ i $\beta + \sigma_1 \geq -1$ zbioru możliwych parametrów modelu EOU. Założenia te gwarantują zbieżność stosowanej aproksymacji gęstości przejścia dla $\Delta \leq 0,5$ oraz pozwalają uniknąć problemów numerycznych w wykorzystywanych w pracy algorytmach.

Rysunek 1.7 przedstawia trzy przykładowe trajektorie oraz linie kwantylowe procesu zadanego wzorem (1.31). Ilustrację graficzną wykonano przy następujących założeniach²⁹⁾: $\beta = -0,0620$, $\sigma_1 = -0,0832$, $\sigma^2 = 177408$, $x_0 = 1960,57$, $T = 1$, $p \in \{5\%, 25\%, 50\%, 75\%, 95\%\}$ (por. definicja 1.12), przy czym linie kwantylowe wyznaczono na podstawie 1000 trajektorii.



Rys. 1.7. Wybrane trajektorie (a) oraz linie kwantylowe (b) dla rozwiązania modelu EOU

Źródło: Opracowanie własne

Model Brennana i Schwartza (BSBS) (por. Brennan i Schwartz 1979,1980)

Model zadany jest poprzez liniowe i autonomiczne SRR postaci:

$$\begin{cases} dX_t = (\alpha - \beta X_t)dt + \sigma X_t dW_t \\ X_{t_0} = x_0 \end{cases} \quad (1.33)$$

²⁹⁾ Przyjęte wartości parametrów są wartościami oczekiwanymi a posteriori prezentowanymi w tabeli 5.2.

gdzie o parametrach α i β zakłada się, że przyjmują wartości z całego zbioru liczb rzeczywistych, natomiast σ jest stałą dodatnią. Rozwiązaniem układu (1.33) jest proces dany następującym wzorem:

$$X_t = X_{t_0} \exp \left\{ - \left(\beta + \frac{\sigma^2}{2} \right) (t - t_0) + \sigma (W_t - W_{t_0}) \right\} + \alpha \int_{t_0}^t \exp \left\{ - \left(\beta + \frac{\sigma^2}{2} \right) (t - u) + \sigma (W_t - W_u) \right\} du \quad (1.34)$$

Postać gęstości przejścia procesu w ogólnym przypadku nie jest znana (por. Cox 1996; Aït-Sahalia 1999). Wiadomo natomiast, że:

$$E(X_{t_0+\Delta} | X_{t_0} = x_0) = \begin{cases} x_0 e^{-\beta\Delta} + \frac{\alpha}{\beta} (1 - e^{-\beta\Delta}) & \text{dla } \beta \neq 0 \\ x_0 + \alpha\Delta & \text{dla } \beta = 0 \end{cases},$$

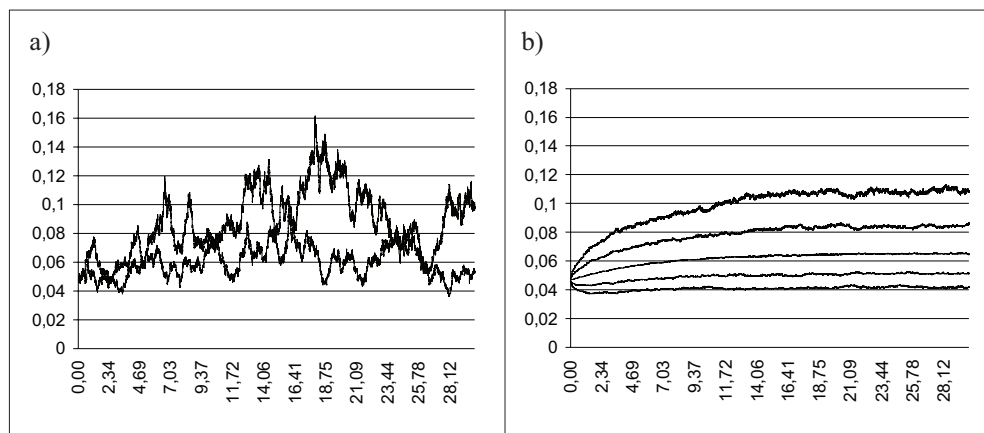
$$E(X_{t_0+\Delta}^2 | X_{t_0} = x_0) =$$

$$\begin{aligned} & \left[e^{-(2\beta-\sigma^2)\Delta} \left\{ \frac{2\alpha \left(x_0 - \frac{\alpha}{\beta} \right)}{(\beta-\sigma^2)} e^{(\beta-\sigma^2)\Delta} - \frac{2\alpha \left(x_0 - \frac{\alpha}{\beta} \right)}{(\beta-\sigma^2)} + \frac{2\alpha^2}{\beta(2\beta-\sigma^2)} \left(e^{(2\beta-\sigma^2)\Delta} - 1 \right) + x_0^2 \right\} \right. \\ & \qquad \qquad \qquad \text{dla } \beta \neq 0, \beta \neq \sigma^2, \beta \neq \frac{\sigma^2}{2}. \\ & = \left[\frac{4\alpha^2\Delta}{\sigma^2} - \left(\frac{4\alpha x_0}{\sigma^2} - \frac{8\alpha^2}{\sigma^4} \right) \left(e^{-\sigma^2\Delta/2} - 1 \right) \right. \\ & \qquad \qquad \qquad \text{dla } \beta = \frac{\sigma^2}{2} \neq 0 \\ & \qquad \qquad \qquad \left. e^{-\sigma^2\Delta} \left\{ \frac{2\alpha^2}{\sigma^4} \left(e^{\sigma^2\Delta} - 1 \right) - \frac{2\alpha^2\Delta}{\sigma^2} + 2\alpha x_0\Delta + x_0^2 \right\} \right. \\ & \qquad \qquad \qquad \text{dla } \beta = \frac{\sigma^2}{2} \neq 0 \\ & \qquad \qquad \qquad \left. e^{\sigma^2\Delta} \left\{ -\frac{2\alpha x_0}{\sigma^2} \left(e^{-\sigma^2\Delta} - 1 \right) - \frac{2\alpha^2}{\sigma^4} \left((\sigma^2\Delta + 1) e^{-\sigma^2\Delta} - 1 \right) + x_0^2 \right\} \right] \text{ dla } \beta = 0 \end{aligned}$$

Rozważany proces jest jednorodnym procesem Markowa. Ponadto, jeśli $\beta > 0$ lub $\left(\beta < 0 \text{ i } x_0 = \frac{\alpha}{\beta} \right)$, to rozwiązanie układu (1.33) jest procesem powracającym do średniej równej $\frac{\alpha}{\beta}$. Parametr β interpretowany jest wówczas jako prędkość, z jaką proces powraca do średniej. Parametr σ lub σ^2 nazywany jest parametrem zmienności. Proces (1.34) przyjmuje wyłącznie wartości dodatnie, jeśli $x_0 > 0$. Należy zwrócić uwagę, że jeśli $\alpha = 0$, wtedy model BSBS jest równoważny modelowi BS wykorzystywanemu do opisu dynamiki, m.in. cen akcji. Ponadto model BSBS stosowany jest do modelowania krótkoterminowych stóp procentowych.

W pracy przyjęto dodatkowe ograniczenie zbioru możliwych parametrów modelu $\alpha \in (0, 1)$ umożliwiające uniknięcie numerycznych problemów w procedurach wykorzystywanych w pracy. Należy zauważyć, że przy takim założeniu model BS nie jest szczególnym przypadkiem modelu BSBS.

Rysunek 1.8 przedstawia dwie przykładowe trajektorie oraz linie kwantylowe dyskretnej aproksymacji procesu zadanego wzorem (1.34). Do aproksymacji wykorzystano schemat Milsteina (por. wzór (1.18)). Ilustrację graficzną wykonano przy następujących założeniach³⁰⁾: $\alpha = 0,01150$, $\beta = 0,15974$, $\sigma^2 = 0,04851$, $x_0 = 0,0468$, $T = 30$, $p \in \{5\%, 25\%, 50\%, 75\%, 95\%\}$ (por. definicja 1.12), przy czym linie kwantylowe wyznaczono na podstawie 1000 trajektorii. Widoczna jest zbieżność lini kwantylowych wraz z upływem czasu



Rys. 1.8. Wybrane trajektorie (a) oraz linie kwantylowe (b) dla aproksymacji rozwiązania modelu BSBS

Źródło: Opracowanie własne

³⁰⁾ Przyjęte wartości parametrów są wartościami oczekiwanymi a posteriori prezentowanymi w tabeli 4.2.

W tabeli 1.1 zestawione zostały podstawowe własności omówionych powyżej modeli.

Tabela 1.1
Podstawowe własności rozważanych modeli

Model	SRR	Rozwiązanie SRR			
		Znana gęstość przejścia	Jednorodność	Dodatniość	Powracanie do średniej
V	$dX_t = (\alpha - \beta X_t)dt + \sigma dW_t$	Tak	Tak	Nie	Tak
BS	$dX_t = \beta X_t dt + \sigma X_t dW_t$	Tak	Tak	Tak	Tak dla $\beta \leq 0$
CIR	$dX_t = (\alpha - \beta X_t)dt + \sigma \sqrt{X_t} dW$	Tak	Tak	Tak	Tak
EOU	$dX_t = -\beta X_t dt + \sigma e^{\sigma t} dW_t$	Tak	Nie	Nie	Tak dla $\beta \geq 0$
BSBS	$dX_t = (\alpha - \beta X_t)dt + \sigma X_t dW_t$	Nie	Tak	Tak	Tak dla $\beta > 0$ lub $\left(\beta < 0 \text{ i } x_0 = \frac{\alpha}{\beta} \right)$

Źródło: Opracowanie własne.

1.6. Uwagi końcowe

W niniejszym rozdziale przedstawione zostały podstawowe definicje i własności procesów stochastycznych, do których ograniczono rozważania dalszej części pracy. Omówiono procesy określone na ciągłej osi czasu i o nieprzeliczalnej przestrzeni stanów, a ściślej – procesy dyfuzji, procesy Itô oraz ich szczególny przypadek, jakim są procesy dyfuzji Itô, tzn. procesy dyfuzji będące rozwiązaniami SRR Itô.

Ponieważ w ogólnej sytuacji nie istnieje jawny wzór na rozwiązanie SRR, w wielu sytuacjach pozostaje wyłącznie jego numeryczna aproksymacja. W rozdziale omówiono podstawowe schematy dyskretyzacji oraz zdefiniowano pojęcie linii kwantylowych, jako metody aproksymacji i graficznej ilustracji rozwiązań SRR.

Omówiono pięć specyfikacji SRR definiujących modele: Vasička, Blacka i Scholesa, Coxa, Ingersolla i Rossa, Brennana i Schwartza oraz uogólniony model Ornsteina i Uhlenbecka. Rozwiązania każdego z nich, z wyjątkiem rozwiązania układu określającego model EOU, są procesami jednorodnymi. Na odpowiednich podzbiorach parametrów każdy z modeli posiada własność powracania do średniej. Dzięki odpowiednim założeniom rozważane procesy stochastyczne przyjmują wartości dodatnie. Wyjątkiem jest powracający do średniej proces Ornsteina i Uhlenbecka oraz uogólniony proces Ornsteina i Uhlenbecka. Ponadto dla każdego z procesów, z wyjątkiem rozwiązania

SRR określającego model BSBS, znana jest jawna postać gęstości przejścia. Jednym z głównych problemów rozważanych w monografii jest zagadnienie aproksymacji gęstości przejścia. Wybór modeli, dla których istnieje jawny przepis na tę gęstość, umożliwił porównanie wyników wnioskowania bayesowskiego, uzyskanych w kolejnych rozdziałach pracy, na podstawie aproksymowanej i prawdziwej gęstości.

Warto podkreślić, że modele V, CIR i BSBS są najczęściej stosowane do opisu stóp procentowych, natomiast za pomocą modelu BS wyjaśniana jest dynamika cen akcji. Z punktu widzenia teorii nie ma przeciwwskazań, by modele te stosować do opisu całkiem innych zjawisk – co zresztą jest powszechną praktyką w szeroko rozumianej literaturze z zakresu ekonomii (i nie tylko). Opisane w niniejszym rozdziale modele będą wykorzystywane w dalszej części pracy do opisu zarówno dynamiki wybranej stopy procentowej, jak i wartości indeksu giełdowego (por. rozdz. 4 i 5).

2. Aproksymacja gęstości przejścia procesów Itô

2.1. Uwagi wstępne

U podstaw prezentowanego w monografii wnioskowania statystycznego leży zagadnienie wyznaczenia funkcji wiarygodności, a ściślej jej aproksymacji. Problem ten, z własności silnych rozwiązań równań Itô, sprowadza się do zagadnienia aproksymacji gęstości przejścia, gdyż funkcja wiarygodności jest ich iloczynem (por. wzór (1.2)). W literaturze znanych jest wiele prób przybliżania tej funkcji. Opierając się na najnowszym wynikach teoretycznych, w monografii rozważono trzy metody:

- 1) zmiennych pośrednich,
- 2) wielomianów Hermite'a,
- 3) rozwiązań równań Fokkera, Plancka i Kołmogorowa.

Należy podkreślić, że celem pracy nie jest porównanie wymienionych metod, a wyłącznie konstrukcja na ich podstawie narzędzi umożliwiających przeprowadzenie bayesowskiego wnioskowania statystycznego dla procesów Itô.

Naturalnym podejściem do badania własności rozwiązań SRR jest, podobnie jak w przypadku deterministycznych równań różniczkowych, dyskretyzacja równań Itô. Dla otrzymanego w wyniku dyskretyzacji równania różnicowego, a ściślej jego rozwiązania, wyznacza się gęstość przejścia, oczekując jednocześnie, że będzie ona naśladowała gęstość wyjściowego procesu zadanego przez SRR tak, by wyniki wnioskowania statystycznego nie różniły się znacznie od tych opartych na rzeczywistej funkcji wiarygodności. W niniejszym rozdziale podany zostanie jednak przykład równania Itô, dla którego estymator metody największej wiarygodności parametrów, wyznaczony z równań różnicowych, mieć będzie nieoczekiwane słabe własności. W literaturze przedmiotu znane są różne udoskonalenia metod (schematów) dyskretyzacji (por. Kloeden i Platen 1992; Pedersen 1995a, b; Eraker 1998; Jones 1998; Kim, Shephard i Chib 1998; Elerian, Chib i Shephard 2001; Roberts i Stramer 2001). Spośród nich opisana zostanie metoda oparta na wprowadzeniu danych pośrednich (brakujących). Podejście to można traktować jako pomost pomiędzy procesami dyskretnymi a ciągłymi. Szczegółowo opisane zostaną też inne metody, oparte na zastosowaniu wielomianów Hermite'a do autonomicznych i nieautonomicznych SRR. Ostatnia część rozdziału poświęco-

na zostanie metodzie wyznaczenia aproksymacji gęstości przejścia jako rozwiązania równań różniczkowych Fokkera, Plancka i Kołmogorowa.

Zastosowanie metody zmiennych pośrednich zaprezentowane zostanie w podrozdziale 3.6, w którym poddano analizie wygenerowane trajektorie logarytmu procesu pierwiastka kwadratowego. Podejście do aproksymacji funkcji wiarygodności zarówno za pomocą wielomianów Hermite’a, jak i aproksymacji opartej na metodzie rozwiązań równań FPK, zostanie wykorzystane w rozdziałach 4 i 5 niniejszej pracy do bayesowskiego wnioskowania o nieznanych charakterystykach rzeczywistych szeregów czasowych stopy procentowej FED i indeksu WIG20.

2.2. Schematy dyskretyzacji

W literaturze przedmiotu przez pojęcie metod dyskretyzacji rozumie się metody aproksymacji gęstości przejścia wyznaczone ze schematów dyskretyzacji (por. podrozdz. 1.4). Dla większości SRR nie są znane jawne postacie gęstości przejścia. Najczęściej rozważanym sposobem radzenia sobie z takim problemem jest zastąpienie prawdziwej gęstości przejścia poprzez jej odpowiednik uzyskany w wyniku zastosowania schematu dyskretyzacji rozważanego SRR. Tak skonstruowana aproksymacja funkcji wiarygodności stosowana jest do przeprowadzenia wnioskowania statystycznego.

Najczęściej rozważane w praktyce schematy dyskretyzacji to schemat Eulera oraz schemat Milsteina. Po zastosowaniu schematu dyskretyzacji Eulera do ogólnej postaci SRR:

$$\begin{cases} dX_t = a(t, X_t; \theta) dt + b(t, X_t; \theta) dW_t \\ X_{t_0} = x_0 \end{cases}$$

otrzymujemy aproksymację gęstości przejścia ze stanu $x_{(k-1)\Delta}$ w chwili $(k-1)\Delta$ (gdzie $k \in \{0, 1, 2, \dots\}$) do stanu $x_{k\Delta}$ po czasie $\Delta > 0$ w postaci gęstości rozkładu normalnego:

$$\begin{aligned} p_X^E \left(x_{k\Delta} \mid x_{(k-1)\Delta}, (k-1)\Delta, \Delta; \theta \right) &= \\ &= \phi \left(x_{k\Delta} \mid x_{(k-1)\Delta} + a \left((k-1)\Delta, x_{(k-1)\Delta}; \theta \right) \Delta; b^2 \left((k-1)\Delta, x_{(k-1)\Delta}; \theta \right) \Delta \right). \end{aligned}$$

Natomiast po zastosowaniu schematu dyskretyzacji Milsteina aproksymacja gęstości przejścia ze stanu $x_{(k-1)\Delta}$ w chwili $(k-1)\Delta$ do stanu $x_{k\Delta}$ po czasie Δ , przy założeniu, że $b'_x \neq 0$, ma postać (por. Elerian 1998):

$$\begin{aligned} p_X^M \left(x_{k\Delta} \mid x_{(k-1)\Delta}, (k-1)\Delta, \Delta; \theta \right) &= \\ &= \frac{\exp(-\lambda/2)}{|A|\sqrt{2\pi}} (z_{k\Delta})^{-1/2} \exp\left(-\frac{z_{k\Delta}}{2}\right) \cosh(\sqrt{\lambda} z_{k\Delta}) \mathbf{I}_{z_{k\Delta} > 0}, \end{aligned}$$

$$\text{gdzie: } z_{k\Delta} = \frac{x_{k\Delta} - B}{A}, \quad \lambda = \frac{1}{\Delta(b'_x)^2}, \quad A = \frac{bb'_x\Delta}{2}, \quad B = -\frac{b}{2b'_x} + x_{(k-1)\Delta} + a\Delta - \frac{bb'_x\Delta}{2}.$$

Natomiast jeśli $b'_x = 0$ (tj. gdy $b(t, x; \theta) = b(t; \theta)$), to schemat Milsteina staje się schematem Eulera.

Należy zwrócić uwagę na popularność zastosowań metod dyskretyzacji spowodowaną w dużej mierze bardzo intuicyjnym ich charakterem oraz łatwością implementacji. Podejście takie jeszcze do końca lat 80. poprzedniego wieku było często w sposób bezkrytyczny wykorzystywane w literaturze i praktyce (por. Lo 1988). Natomiast obecnie stosuje się je głównie do porównania wyników różnych metod aproksymacji (por. np.: Pedersen 1995a, b; Elerian 1998; Ait-Sahalia 1999; Jensen i Poulsen 1999; Poulsen 1999a, b; Elerian, Chib i Shephard 2001; DiPietro 2001; Nicolau 2002, 2004). Warto w tym miejscu podać kilka faktów znanych z literatury przedmiotu, a odnoszących się do niebayesowskich reguł wnioskowania statystycznego, jako przestrożę przed naiwnym stosowaniem metod dyskretyzacji. Zakładając, że funkcja wiarygodności wyznaczona jest z prawdziwych gęstości przejścia zbudowanych na podstawie danych zaobserwowanych w równych odstępach czasu Δ , można pokazać, przy stosownych warunkach oraz rosnącej liczbie obserwacji, że estymator największej wiarygodności nieznanych parametrów procesu jest zgodny i asymptotycznie normalny (por. Billingsley 1961a, b; Dacunha-Castelle i Florens-Zmirou 1986; Barndorff-Nielsen i Sørensen 1994; Brand i Santa Clara 2002; Bingham i Kiesel 2004). Przy czym własności te nie zależą od wielkości Δ . W sytuacjach gdy układu SRR nie można rozwiązać w sposób jawny, często nie da się wyznaczyć funkcji wiarygodności, a w konsekwencji estymatora największej wiarygodności. Mając wówczas na uwadze zbieżność rozważanych w pracy schematów dyskretyzacji (por. twierdzenia 1.5 i 1.6), wydaje się sensowne zastosowanie metody największej wiarygodności dla procesu powstałego przez dyskretyzację SRR i wnioskowanie w ten sposób o nieznanych charakterystykach wyjściowego (ciągłego) procesu. Jeśli Δ jest „odpowiednio małe”, to naturalne jest oczekiwanie od tak wyznaczonego estymatora takich samych własności jak od „prawdziwego” estymatora największej wiarygodności. Okazuje się jednak, że estymacja oparta na danych dyskretnych, zarówno ze względu na dyskretyzację Eulera, jak i Milsteina, może być niezgodna i obciążona (por. Lo 1988; Florens-Zmirou 1989; Elerian 1998).

Poniżej przedstawiony zostanie przykład SRR, zaczerpnięty z pracy Lo (1988), dla którego takie podejście prowadzi do estymatora największej wiarygodności o słabych własnościach.

Przykład 2.1

W przykładzie rozważono następujące stochastyczne równanie różniczkowe¹⁾

$$dX_t = \beta X_t dt + \sigma X_t dW_t.$$

¹⁾ Jego rozwiązaniem jest geometryczny ruch Browna.

Ponieważ istnieje jawna postać funkcji przejścia dla tego równania, można wyznaczyć estymatory największej wiarygodności nieznanymi parametrów, które są estymatorami zgodnymi. Niestety własność ta nie zachowuje się przy wyznaczaniu estymatora największej wiarygodności na podstawie dyskretnej aproksymacji gęstości przejścia. Dyskretyzując powyższe równanie względem schematu Eulera, otrzymano proces określony w ustalonych momentach czasu (ściślej w krotnościach ustalonej wielkości Δ):

$$Y_{k\Delta} = Y_{(k-1)\Delta} + \beta Y_{(k-1)\Delta} \Delta + \sigma Y_{(k-1)\Delta} \xi_{k\Delta},$$

gdzie $\{\xi_{k\Delta}\}_{k=1}^{\infty}$ jest ciągiem niezależnych zmiennych losowych o takim samym rozkładzie normalnym $N(0, \Delta)$.

Estymatorami największej wiarygodności wektora parametrów $\theta = (\beta, \sigma^2)$ są:

$$\hat{\beta}_n = \frac{1}{n\Delta} \sum_{k=1}^n \left[\frac{Y_{k\Delta}}{Y_{(k-1)\Delta}} - 1 \right], \quad \hat{\sigma}_n^2 = \frac{1}{n\Delta} \sum_{k=1}^n \left[\frac{Y_{k\Delta}}{Y_{(k-1)\Delta}} - 1 - \hat{\beta}_n \Delta \right]^2,$$

gdzie n oznacza liczbę obserwacji, na podstawie której szacowane są parametry. Zachodzą wówczas następujące związki: $E(\hat{\beta}_n) = \beta$ i $E(\hat{\sigma}_n^2) = \sigma^2 \frac{n-1}{n} \neq \sigma^2$, oznaczające obciążoność estymatora $\hat{\sigma}_n^2$ (przy czym estymator ten jest asymptotycznie nieobciążony). Ponadto, przy ustalonym parametrze Δ i rosnącej liczbie obserwacji, prawdziwe są zależności:

$$\text{P-lim}_{n \rightarrow \infty} \hat{\beta}_n = \frac{1}{\Delta} [\exp(\beta\Delta) - 1] \neq \beta, \quad \text{P-lim}_{n \rightarrow \infty} \hat{\sigma}_n^2 = \frac{1}{\Delta} \exp(2\beta\Delta) [\exp(\sigma^2\Delta) - 1] \neq \sigma^2.$$

Innymi słowy, estymatory parametrów $\hat{\beta}_n$ i $\hat{\sigma}_n^2$ nie są zgodne. Oczywiście w pewnych sytuacjach oraz przy małym Δ niezgodność ta może okazać się nieistotna. Nasuwa się naturalne pytanie, czy przy ustalonym przedziale czasu $[0, T]$ i zwiększającej się częstotliwości pojawiania danych, tak by $\Delta = \frac{T}{n}$ (przy rosnącej liczbie obserwacji Δ maleje do zera), estymatory parametrów będą zgodne. W takim przypadku okazuje się, że estymator $\hat{\sigma}_n^2$ jest zgodny (ze zbieżnością według prawdopodobieństwa). Niestety nie można tego samego powiedzieć o estymatorze $\hat{\beta}_n$, który jest zbieżny w sensie słabym (wg rozkładów) do $\beta + \frac{\sigma}{\sqrt{T}} [W_1 - W_0]$.

Tym samym pokazano, że własności estymatorów największej wiarygodności parametrów rzeczywiście są słabsze, niż można było przypuszczać.

2.3. Metoda zmiennych pośrednich

Obecnie zaprezentowana zostanie metoda aproksymacji gęstości przejścia wykorzystująca schemat Eulera oraz ideę dolosowania (odtworzenia) brakujących/pośrednich danych.

Dane finansowe obserwowane są w dyskretnych momentach czasu i jako takie są powszechnie opisywane przez modele dyskretnie. W niniejszej pracy finansowy szereg czasowy traktowany jest jako realizacja procesu ciągłego, a ściślej procesu Itô. Ze względu na trudność, a w wielu przypadkach niemożliwość, podania analitycznej postaci tego procesu oraz odpowiadającej mu gęstości przejścia stosuje się jego dyskretną aproksymację. Sukces aproksymacji modeli ciągłych poprzez ich odpowiedniki dyskretnie jest często uzależniony od możliwości włączenia do badania danych o dowolnie wysokiej częstotliwości – w rzeczywistości jednakże nie ma możliwości obserwacji ścieżki procesu pomiędzy dwoma zarejestrowanymi punktami. Związany jest z tym problem danych pośrednich, określanych w literaturze przedmiotu również jako dane nieobserwowalne, brakujące, ukryte lub pomocnicze. Umiejętność poprawnego wygenerowania/odtworzenia brakującej ścieżki jest istotna dla zredukowania niedokładności aproksymacji (tj. zmniejszenia znaczenia nierealnego założenia $\Delta \rightarrow 0$) i może być uważana za użyteczne narzędzie łączące modele dyskretnie z ciągłymi. Celem niniejszego podrozdziału jest zaprezentowanie metody zmiennych pośrednich jako narzędzia aproksymacji nieznanego gęstości przejścia procesu Itô.

Dyskretyzacja ogólnej postaci stochastycznego równania różniczkowego:

$$\begin{cases} dX_t = a(t, X_t; \theta)dt + b(t, X_t; \theta)dW_t \\ X_{t_0} = x_0 \end{cases}$$

względem schematu Eulera prowadzi do nowego procesu stochastycznego Y spełniającego następujący układ równań:

$$\begin{cases} Y_{k\Delta} = Y_{(k-1)\Delta} + a((k-1)\Delta, Y_{(k-1)\Delta}; \theta)\Delta + b((k-1)\Delta, Y_{(k-1)\Delta}; \theta)(W_{k\Delta} - W_{(k-1)\Delta}) \\ Y_0 = y_0 \end{cases}$$

Proces ten definiuje się w dyskretnych momentach czasu²⁾ $k\Delta$ dla $k = 0, \dots, \infty$. Dla przejrzystości zapisu momenty czasu oznaczane będą symbolem t – w szczególności, jeśli t odpowiada momentowi $k\Delta$, wówczas $t+1$ utożsamiane jest z $(k+1)\Delta$. Gęstość przejścia dla Y jest gęstością rozkładu normalnego:

$$p_Y(\cdot | y_t; t, \Delta, \theta) = \phi(\cdot | y_t + a(t, y_t; \theta)\Delta; b^2(t, y_t; \theta)\Delta) \quad (2.1)$$

Na podstawie powyższej funkcji skonstruowana zostanie aproksymacja gęstości przejścia p_X procesu X . W celu przeprowadzenia dalszego rozumowania wprowadzone zostaną dodatkowe oznaczenia i definicje.

²⁾ Metodę zmiennych pośrednich można również definiować dla zmieniającej się odległości pomiędzy kolejnymi momentami czasu (por. Pedersen 1995a).

Niech $(x_0, x_1, \dots, x_T)'$ oznacza wektor kolejnych obserwacji wyjściowego procesu Itô w chwilach $\tau_0, \tau_1, \dots, \tau_T$ dla $\Delta = \tau_{i+1} - \tau_i$.

Ponadto niech:

$$y = (y_0, y_1, \dots, y_T)' = (x_0, x_1, \dots, x_T)', \quad y^* = (y_0^*, y_1^*, \dots, y_{T-1}^*)' \text{ i } y_t^* = (y_{t,1}^*, \dots, y_{t,M}^*)',$$

gdzie $y_{t,1}^*, \dots, y_{t,M}^*$, to M punktów pośrednich pomiędzy zaobserwowanymi wielkościami y_t i y_{t+1} , przy czym przyjęto $y_{t,0}^* = y_t = x_t$, $y_{t,M+1}^* = y_{t+1} = x_{t+1}$ dla $t \leq T-1$. Dodatkowo zdefiniowano wektory $\hat{y}$ i $\hat{y}_{\setminus 0}$ określające wartości zaobserwowane y i dane pośrednie y^* , przy czym $\hat{y}_{\setminus 0}$ nie uwzględnia warunku początkowego:

$$\hat{y} = (\hat{y}_0, \hat{y}_1, \dots, \hat{y}_n)' = (y_0, y_{0,1}^*, \dots, y_{0,M}^*, y_1, \dots, y_{T-1,M}^*, y_T)',$$

$$\hat{y}_{\setminus 0} = (y_{0,1}^*, \dots, y_{0,M}^*, y_1, \dots, y_{T-1,M}^*, y_T)',$$

gdzie $n = T(M+1)+1$ oznacza wymiar wektora $\hat{y}$. Przez $y_{t(k,m)}^*$ oznaczono blok długości m , tzn. $y_{t(k,m)}^* = (y_{t,k}^*, y_{t,k+1}^*, \dots, y_{t,k+m-1}^*)$ dla $1 \leq m \leq M - k + 1$. Momenty czasu odpowiadające kolejnym punktom pośrednim wynoszą $\tau_{t,k}^* = t + k$. Należy zauważyć, że uwzględnienie punktów pośrednich zmniejsza krok dyskretyzacji $\Delta^* = \frac{\Delta}{M+1}$, poprawiając tym samym aproksymację procesu X poprzez jego dyskretny odpowiednik Y .

Przy odpowiednich założeniach (por. Pedersen 1995a, b) można zdefiniować ciąg gęstości $\{p_X^{(M)}\}_{M=0}^{\infty}$ aproksymujący gęstość przejścia p_X :

– dla $M = 0$ punktów pośrednich³⁾:

$$\begin{aligned} p_X^{(0)}(x|x_t; t, \Delta, \theta) &= p_Y(x|x_t; t, \Delta, \theta) = \\ &= \frac{1}{\sqrt{2\pi b^2(t, x_t; \theta)\Delta}} \exp\left(-\frac{(x - x_t - a(t, x_t; \theta)\Delta)^2}{2b^2(t, x_t; \theta)\Delta}\right) \end{aligned} \quad (2.2)$$

– dla $M \geq 1$ punktów pośrednich:

$$\begin{aligned} p_X^{(M)}(x|x_t; t, \Delta, \theta) &= p_Y^{(M)}(x|x_t; t, \Delta, \theta) = \\ &= \int_{\mathbb{R}^M} p_Y(x|y_{t,M}^*; \tau_{t,M}^*, \Delta^*, \theta) \left\{ \prod_{j=1}^M p_Y(y_{t,j}^* | y_{t,j-1}^*; \tau_{t,j-1}^*, \Delta^*, \theta) \right\} dy_{t,M}^*, \dots, dy_{t,1}^* \end{aligned} \quad (2.3)$$

³⁾ Dla $M = 0$ aproksymacja pokrywa się z gęstością otrzymaną wprost ze schematu Eulera.

Przy odpowiednich założeniach można pokazać, że aproksymacja $p_X^{(M)}(x|x_t; t, \Delta; \theta)$ zbiega do gęstości przejścia $p_X(x|x_t; t, \Delta, \theta)$ w sensie normy przestrzeni $L^1(\lambda)$, gdzie λ oznacza jednowymiarową miarę Lebesgue'a⁴⁾. Prawdą jest również, że zlogarytmowana aproksymacja funkcji wiarygodności zbiega do logarytmu rzeczywistej funkcji wiarygodności (wg prawdopodobieństwa), jeśli $M \rightarrow \infty$. Ponadto dla prawie wszystkich x zachodzi

$$p_X(x|x_t; t, \Delta, \theta) = \liminf_{M \rightarrow \infty} p_X^{(M)}(x|x_t; t, \Delta, \theta),$$

zatem jeśli wartości funkcji $p_X^{(M)}(x|x_t; t, \Delta; \theta)$ przy ustalonym x są zbieżne, to ich granicą jest wartość gęstości przejścia procesu X w tym samym punkcie (por. Pedersen 1995a). Ponadto w pracach Jensen (1986) i Pedersen (1995b) wykazano, przy odpowiednich założeniach, zgodność i asymptotyczną normalność estymatora największej wiarygodności $\hat{\theta}_n^{(M)}$ wyznaczonego z powyżej zdefiniowanej aproksymacji gęstości przejścia.

Zaprezentowana aproksymacja została zadana poprzez M -wymiarową całkę, która w ogólnym przypadku nie ma jawnego rozwiązania. Jednym ze sposobów wyznaczenia przybliżonej wartości funkcji $p_X^{(M)}$ jest zauważenie następującego związku:

$$p_X^{(M)}(x|x_t; t, \Delta; \theta) = E\left(p_Y\left(x|y_{t,M}^*; \tau_{t,M}^*, \Delta^*, \theta\right)\right).$$

Z prawa wielkich liczb powyższą wartość oczekiwaną można przybliżyć wyrażeniem:

$$\frac{1}{N} \sum_{n=1}^N p_Y\left(x|y_{t,M}^{*(n)}; \tau_{t,M}^*, \Delta^*, \theta\right) \quad (2.4)$$

gdzie $y_{t,M}^{*(n)}$ jest zmienną losową, generowaną rekurencyjnie według schematu:

$$y_{t,0}^{*(n)} = y_t = x_t,$$

$$y_{t,k}^{*(n)} = y_{t,k-1}^{*(n)} + a\left(\tau_{t,k-1}^*, y_{t,k-1}^{*(n)}; \theta\right)\Delta^* + b\left(\tau_{t,k-1}^*, y_{t,k-1}^{*(n)}; \theta\right)\sqrt{\Delta^*}U_k^{(n)}$$

dla $k=1, \dots, M$, $n=1, \dots, N$.

⁴⁾ f_n zbiega do f w sensie normy przestrzeni $L^1(\lambda)$, jeśli zachodzi $\int_{\mathbb{R}} |f_n - f| d\lambda \xrightarrow{n \rightarrow \infty} 0$.

Przy tym $\left\{U_k^{(n)}\right\}_{k=1, n=1}^{M-1, N}$ jest ciągiem niezależnych zmiennych losowych o standardowym rozkładzie normalnym. Zbieżność powyższego przybliżenia wyjaśnia następująca nierówność (por. Talay i Tubaro 1990):

$$E \left| \frac{1}{N} \sum_{n=1}^N p_Y \left(x | y_{t,M}^{*(n)}; \tau_{t,M}^*, \Delta^*, \theta \right) - p_X(x | x_t; t, \Delta, \Theta) \right| \leq C \left(\frac{1}{\sqrt{N}} + \frac{\Delta}{M} \right).$$

Problemem towarzyszącym zastosowaniom metody zmiennych pośrednich jest wybór optymalnego M , który zależy od konkretnej specyfikacji SRR. Czym większa wartość M , tym bardziej precyzyjne jest oszacowanie, ale jednocześnie zwiększa się w sposób istotny czas obliczeń numerycznych, stawiając zastosowanie metody zmiennych pośrednich pod dużym znakiem zapytania.

Praktyczne wykorzystanie zaprezentowanej metody przedstawione zostanie w podrozdziale 3.6, gdzie poddano analizie szeregi danych symulowanych, będące dyskretnymi trajektoriami procesu pierwiastka kwadratowego.

2.4. Metoda wielomianów Hermite’a

W pracach z lat 1999 i 2002 Y. Aït-Sahalia używając wielomianów Hermite’a, skonstruował ciąg funkcji zbieżny do prawdziwej, lecz niekoniecznie znanej, gęstości przejścia autonomicznych stochastycznych równań różniczkowych. A.V. Egorov, H. Li oraz Y. Xu w roku 2003 uogólnili teoretyczne wyniki Y. Aït-Sahalii na przypadek nieautonomiczny.

Pomysł Y. Aït-Sahalii został oparty na rozwinięciu funkcji w szereg funkcyjny Grama–Charliera typu A przy zastosowaniu wielomianów Hermite’a. Podstawowe twierdzenia gwarantujące zbieżność takiego szeregu przytoczono w dalszej części podrozdziału. Okazuje się, że założenia wspomnianych twierdzeń spełnia relatywnie „mały” zbiór odwzorowań (por. Cramer 1925, 1957). Powstaje więc problem sprawdzenia, czy gęstość przejścia SRR spełnia te założenia, co może być trudne, gdy nie znana jest jej analityczna postać. W niniejszym podrozdziale podane zostaną warunki sformułowane w języku funkcji dryfu i dyfuzji, przy których obserwowany w sposób dyskretny proces X można przekształcić do procesu Z , dla którego gęstość przejścia można rozwinąć w szereg wielomianów Hermite’a. Wspomniane przekształcenie jest jednoznaczne, co umożliwia transformację aproksymacji gęstości dla procesu Z do aproksymacji dla X . Większość stosowanych w praktyce modeli stochastycznych równań różniczkowych spełnia wymagane założenia z ewentualnym ograniczeniem przestrzeni parametrów procesu.

Wielomiany Hermite’a

W literaturze występują różne definicje wielomianów Hermite’a, nie zawsze równoważne (por. np.: Cramer 1925). Wydaje się zatem konieczne precyzyjne określenie używanego w pracy pojęcia wielomianu Hermite’a i przedstawienie jego podstawowych własności.

Definicja 2.1

Wielomianem Hermite'a stopnia v , gdzie $v \in \{0, 1, 2, \dots\}$, nazywamy funkcję

$$H_v: \mathbb{R} \rightarrow \mathbb{R}$$

taką, że:

$$H_v(x) = (-1)^v e^{\frac{x^2}{2}} \frac{d^v}{dx^v} e^{-\frac{x^2}{2}}.$$

Przykład 2.2

$$H_0(x) = 1,$$

$$H_1(x) = x,$$

$$H_2(x) = x^2 - 1,$$

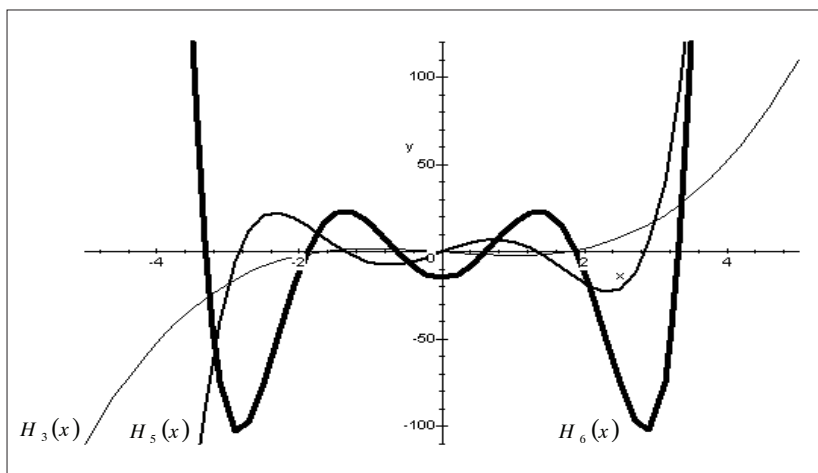
$$H_3(x) = x^3 - 3x,$$

$$H_4(x) = x^4 - 6x^2 + 3,$$

$\vdots$

$$H_v(x) = x^v - \frac{v(v-1)}{2 \cdot 1!} x^{v-2} + \frac{v(v-1)(v-2)(v-3)}{2^2 \cdot 2!} x^{v-4} - \dots$$

Na rysunku 2.1 przedstawiono wykresy wybranych wielomianów.



Rys. 2.1. Wykresy wielomianów $H_3(x)$, $H_5(x)$, $H_6(x)$

Źródło: Opracowanie własne

Rozważane wielomiany Hermite'a mają następujące własności⁵⁾:

$$H_{v+1}(x) = xH_v(x) - vH_{v-1}(x),$$

$$\frac{dH_v}{dx}(x) = vH_{v-1}(x),$$

$$\sum_{v=0}^{\infty} \frac{H_v(x)}{v!} t^v = e^{\frac{1}{2}t^2 + tx},$$

$$\frac{d^v \phi}{dx^v}(x) = (-1)^v H_v(x) \phi(x),$$

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) d\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} H_n(x) H_m(x) e^{-\frac{x^2}{2}} dx = \begin{cases} n! & n=m \\ 0 & n \neq m \end{cases}.$$

Z powyższych własności wynika, że $\left\{ \frac{1}{\sqrt{v!}} H_v(x) \right\}_{v=0}^{\infty}$ jest ciągiem wielomianów ortonormalnych w przestrzeni $L^2(\mu)$, gdzie μ jest jednowymiarową miarą gausowską⁶⁾. Z ortogonalności wielomianów wynika istnienie pewnej klasy funkcji, której każdy element f możemy rozwinąć w szereg funkcyjny wielomianów Hermite'a (zwany szeregiem Grama–Charliera typu A) o wyrazach zadanych w następujący sposób:

$$f(x) = \sum_{v=0}^{\infty} c_v \phi(x) H_v(x),$$

gdzie $c_v = \frac{1}{v!} \int_{-\infty}^{\infty} f(x) H_v(x) dx$.

Warto przytoczyć twierdzenia gwarantujące prawdziwość powyższego wzoru (por. Cramer 1925, 1957):

Twierdzenie 2.1

Jeśli f jest funkcją o ciągłej pochodnej, dla której istnieje całka:

$$\int_{-\infty}^{\infty} (f'(x))^2 e^{\frac{x^2}{4}} < \infty$$

⁵⁾ ϕ i Φ są odpowiednio gęstością i dystrybucją standardowego rozkładu normalnego.

⁶⁾ $L^2(\mu)$ jest zbiorem funkcji mierzalnych spełniających warunek $\int_{\mathbb{R}} |f(x)|^2 \phi(x) dx < \infty$.

oraz zachodzi: $\lim_{|x| \rightarrow \infty} f(x) = 0$, wówczas szereg $\sum_{v=0}^{\infty} c_v \phi(x) H_v(x)$ jest absolutnie i jednostajnie zbieżny do $f(x)$.

Twierdzenie 2.2

Jeśli f jest funkcją o wahanu ograniczonym⁷⁾ na każdym ograniczonym przedziale oraz istnieje całka:

$$\int_{-\infty}^{\infty} |f(x)| e^{\frac{x^2}{4}} < \infty,$$

wówczas szereg $\sum_{v=0}^{\infty} c_v \phi(x) H_v(x)$, zbiega do $\frac{1}{2} [f(x^+) + f(x^-)]$ dla $x \in \mathbb{R}$, gdzie

$f(x^-)$ i $f(x^+)$ oznaczają odpowiednio granicę lewostronną i prawostronną funkcji f w punkcie x . Zbieżność szeregu funkcyjnego jest jednostajna na każdym ograniczonym przedziale ciągłości funkcji f .

W niektórych przypadkach możliwa jest interpretacja początkowych współczynników c_v , którą ilustruje przykład 2.3.

Przykład 2.3

Niech ξ będzie zmienną losową, dla której istnieją wszystkie momenty centralne μ_v , natomiast f oznacza gęstość zmiennej $\frac{\xi - m}{\sigma}$, gdzie m i σ są odpowiednio wartością oczekiwaną i odchyleniem standardowym zmiennej ξ . Ponadto niech f spełnia założenia jednego z powyższych twierdzeń. Wówczas początkowe współczynniki rozwinięcia funkcji f w szereg wielomianów Hermite'a mają postać:

$$c_0 = 1, \quad c_1 = c_2 = 0, \quad c_3 = -\frac{\mu_3}{\sigma^3}, \quad c_4 = \frac{\mu_4}{\sigma^4} - 3,$$

gdzie μ_n jest momentem centralnym stopnia n zmiennej ξ . Wyrazy c_3 oraz c_4 to, odpowiednio, wartość przeciwna do współczynnika skośności oraz współczynnik ekscesu (spłaszczenia).

Ostatecznie sumę początkowych wyrazów szeregu funkcyjnego można uznać jako aproksymację nieznaną gęstości poprzez gęstość standardowego rozkładu normalnego oraz „poprawkę” uwzględniającą podstawowe własności rozkładu.

Rozwinięcie gęstości przejścia w szereg wielomianów Hermite'a

Jak już wspomniano, Y. Aït-Sahalia w pracach z roku 1999 i 2002 zastosował wielomiany Hermite'a do aproksymacji gęstości przejścia autonomicznych stocha-

⁷⁾ Na przykład funkcją różniczkowalną o ograniczonej pochodnej.

stycznych równań różniczkowych, natomiast A.V. Egorov, H. Li oraz Y. Xu w roku 2003 uogólnili to podejście na przypadek nieautonomiczny. Obecnie na podstawie tych prac zaprezentowane zostaną założenia oraz podstawowe wyniki aproksymacji gęstości przejścia za pomocą wielomianów Hermite’a.

Rozpatrywana będzie ogólna postać SRR:

$$\begin{cases} dX_t = a(X_t, t; \theta) dt + b(X_t, t; \theta) dW_t \\ X_0 = x_0 \end{cases} \quad (2.5)$$

gdzie:

$\theta = (\theta_1, \dots, \theta_k)$ – nieznaną wektor parametrów modelu, $\theta \in \Theta$, gdzie $\Theta \subset \mathbb{R}^k$ jest zbiorem zwartym,

$D_X = (x, \bar{x})$ – zbiór wartości procesu X ,

$a: D_X \times [0, \infty) \times \Theta \rightarrow \mathbb{R}$ – funkcja dryfu,

$b: D_X \times [0, \infty) \times \Theta \rightarrow \mathbb{R}$ – funkcja dyfuzji (zmienności).

Przyjmuje się, że $D_X = (-\infty, \infty)$ lub $D_X = (0, \infty)$ (warunki te spełnione są dla większości silnych rozwiązań SRR wykorzystywanych w literaturze przedmiotu do modelowania finansowych szeregów czasowych).

Jak wcześniej wspomniano, nie każdą funkcję da się rozwinąć w szereg wielomianów Hermite’a zbieżny do niej samej. Jednakże przy odpowiednich założeniach istnieje transformacja przekształcająca wyjściowy proces X do procesu Z , którego gęstość da się rozwinąć w szereg wielomianów Hermite’a. Transformacja ta odbywa się dwuetapowo, tzn. proces X przekształcany jest do procesu Y , a następnie Y do procesu Z . Spełnione muszą być następujące założenia:

Założenie 2.1

Funkcje $a(x, t; \theta)$ oraz $b(x, t; \theta)$ są nieskończenie wiele razy różniczkowalne ze względu na zmienną $x \in D_X$ i $t \in [0, \infty)$ oraz klasy C^2 względem $\theta \in \Theta$.

Założenie 2.2

1. Jeśli $D_X = (-\infty, \infty)$, to istnieje stała c taka, że $b(x, t; \theta) > c > 0$ dla wszystkich $(x, t, \theta) \in D_X \times [0, \infty) \times \Theta$.
2. Jeśli $D_X = (0, \infty)$, wówczas dla każdego $\xi > 0$ i $\hat{t} > 0$ istnieje stała c_ξ taka, że $b(x, t; \theta) \geq c_\xi > 0$ dla wszystkich $(x, t, \theta) \in [\xi, \infty) \times [0, \hat{t}] \times \Theta$. Ponadto jeśli $\lim_{x \rightarrow 0^+} b(x, t; \theta) = 0$ dla pewnego $t > 0$, wówczas dla dowolnego $\hat{t} > 0$ istnieją stałe $b(x, t; \theta) \geq \omega x^p$ takie, że $b(x, t; \theta) \geq \omega x^p$ dla wszystkich $(x, t, \theta) \in (0, \xi_0] \times [0, \hat{t}] \times \Theta$.

W pierwszym kroku transformacji proces Y otrzymuje się z procesu X poprzez przekształcenie Lampertiego (Lamperti 1964)⁸⁾:

$$Y = \gamma(X; t, \theta) = \int^X b(u, t; \theta)^{-1} du \quad (2.6)$$

Z przyjętych założeń wynika, że $b(\cdot, t; \theta) > 0$ na zbiorze D_X , zatem $\gamma(\cdot, t; \theta)$ jest funkcją rosnącą, a tym samym odwracalną⁹⁾. Niech $D_Y = (\underline{y}, \bar{y})$ oznacza zbiór wartości procesu Y , gdzie $\underline{y} = \lim_{x \rightarrow \underline{x}} \gamma(x, t; \theta)$ oraz $\bar{y} = \lim_{x \rightarrow \bar{x}} \gamma(x, t; \theta)$. Podobnie jak w przypadku procesu X przyjmuje się, że $D_Y = (-\infty, \infty)$ lub $D_Y = (0, \infty)$. Należy podkreślić, że warunki te spełnione są dla procesów wykorzystywanych do modelowania finansowych szeregów czasowych. Nowo otrzymany proces jest procesem Itô o stałej funkcji dyfuzji. Zgodnie ze wzorem Itô (por. twierdzenie 1.3) można ponadto pokazać, że jest on rozwiązaniem następującego układu równań:

$$\begin{cases} dY_t = a_Y(Y_t, t; \theta)dt + dW_t \\ Y_0 = y_0 \end{cases} \quad (2.7)$$

$$\text{gdzie } a_Y(y, t; \theta) = \frac{a(\gamma^{-1}(y, t; \theta), t; \theta)}{b(\gamma^{-1}(y, t; \theta), t; \theta)} + \frac{\partial \gamma}{\partial t}(\gamma^{-1}(y, t; \theta), t; \theta) - \frac{1}{2} \frac{\partial b}{\partial x}(\gamma^{-1}(y, t; \theta), t; \theta).$$

W drugim kroku transformacji definiuje się proces Z . Niech $t = s + \Delta$ oraz $Y_s = y_s$ wówczas:

$$Z_t = \frac{Y_t - y_s}{\sqrt{\Delta}}.$$

Gęstość przejścia p_X dla procesu wyjściowego X otrzymuje się na podstawie wzoru Jacobiego z gęstości p_Y dla procesu Y . Identyczna sytuacja ma miejsce w przypadku p_Y i gęstości przejścia p_Z procesu Z :

$$\begin{aligned} p_Y(y | y_s, s, \Delta; \theta) &= \Delta^{-1/2} p_Z(\Delta^{-1/2}(y - y_s) | y_s, s, \Delta; \theta) \\ p_X(x | x_s, s, \Delta; \theta) &= b(x, t; \theta)^{-1} p_Y(\gamma(x, t; \theta) | \gamma(x_s, s; \theta), s, \Delta; \theta) \end{aligned} \quad (2.8)$$

Aby sformułować kolejne założenia, definiuje się następujące pojęcia:

⁸⁾ $\int^X f(x)dx$ jest całką oznaczoną, której dolna granica całkowania jest arbitralnie dobranym punktem z dziedziny funkcji f .

⁹⁾ Własność ta odegra istotną rolę przy przekształceniu p_Y do p_X .

Definicja 2.2

Nieskończenie wiele razy różniczkowalna funkcja f ma wzrost co najwyżej wielomianowy w $-\infty (+\infty)$, jeśli istnieje liczba całkowita $p \geq 0$ taka, że odwzorowanie $|y|^{-p} |f(y)|$ jest ograniczone w pewnym sąsiedztwie $-\infty (+\infty)$. Pojęcie co najwyżej wielomianowego wzrostu w zerze oznacza ograniczoność funkcji $|y|^p |f(y)|$ w pewnym sąsiedztwie tego punktu. Jeśli $p = 1$, to mówi się, że f ma co najwyżej liniowy wzrost, natomiast dla $p = 2$ co najwyżej kwadratowy.

Założenie 2.3

Dla $(t, \theta) \in [0, \infty) \times \Theta$ funkcja a_Y wraz z pochodnymi względem y , t i θ ma w pobliżu granic dziedziny co najwyżej wzrost wielomianowy. Ponadto zachodzą następujące własności:

1. $\lim_{y \rightarrow \underline{y}^+} \lambda_Y(y, t; \theta) < \infty$ i $\lim_{y \rightarrow \bar{y}} \lambda_Y(y, t; \theta) < \infty$,

$$\text{gdzie } \lambda_Y(y, t; \theta) = -\frac{1}{2} \left[a_Y^2(y, t; \theta) + \frac{\partial}{\partial y} a_Y(y, t; \theta) \right] - \int_{y_s}^y \frac{\partial}{\partial t} a_Y(u, t; \theta) du.$$

2. Jeśli $\underline{y} = 0$, wówczas dla dowolnego $\hat{t} > 0$ istnieją stałe $\alpha > 1$ i $\kappa > 0$ lub $\alpha = 1$ i $\kappa \geq 1$ oraz $\varepsilon_0 > 0$ takie, że dla dowolnych $(y, t, \theta) \in (0, \varepsilon_0) \times [0, \hat{t}] \times \Theta$ prawdziwa jest nierówność: $a_Y(y, t; \theta) \geq \kappa y^{-\alpha}$.
3. Jeśli $\bar{y} = \infty$, wówczas istnieje stała $K > 0$ taka, że dla dowolnego $\hat{t} > 0$ można wskazać $\varepsilon_0 > 0$ takie, że prawdziwa jest nierówność $a_Y(y, t; \theta) \leq Ky$ dla dowolnych $(y, t, \theta) \in (\varepsilon_0, \infty) \times [0, \hat{t}] \times \Theta$.
Jeśli $\underline{y} = -\infty$, wówczas istnieje stała $K > 0$ taka, że dla dowolnego $\hat{t} > 0$ można wskazać $\varepsilon_0 < 0$ takie, że prawdziwa jest nierówność $a_Y(y, t; \theta) \geq Ky$ dla dowolnych $(y, t, \theta) \in (-\infty, \varepsilon_0] \times [0, \hat{t}] \times \Theta$.

Warto zwrócić uwagę, że założenia 2.1–2.3 nie nakładają na proces warunku stacjonarności. Ponadto na ich podstawie można udowodnić następujące twierdzenie mówiące o własnościach gęstości przejścia procesu Z :

Twierdzenie 2.3

Przy założeniach 2.1–2.3 dla dowolnego rozkładu Y_0 istnieje dokładnie jedno słabe rozwiązanie Y układu (2.7). Ponadto niech stała $L = 0$, jeśli $a_Y \leq 0$ w sąsiedztwie $+\infty$ i $a_Y \geq 0$ w sąsiedztwie dolnej granicy dziedziny (tj. $-\infty$ lub 0), natomiast w pozostałych przypadkach niech $L = K$, gdzie K jak w założeniu 2.3.

Wówczas dla dowolnych $\hat{t} > 0$ i $\hat{y} \in D_Y$ istnieją stałe $C_0 = C_0(\hat{y}, \hat{t})$ i $C_1 = C_1(\hat{y}, \hat{t})$ takie, że dla $(z, s, \theta, \Delta) \in D_Z \times [0, \hat{t}] \times \Theta \times (0, \hat{t} - s]$ zachodzą następujące nierówności:

$$0 \leq p_Z(z|\hat{y}, s, \Delta; \theta) \leq C_0 e^{-z^2(1-\Delta L)/2+C_1|z|},$$

$$\left| \frac{\partial}{\partial z} p_Z(z|\hat{y}, s, \Delta; \theta) \right| \leq C_0 e^{-z^2(1-\Delta L)/2+C_1|z|}.$$

Przy tym, jeśli $\underline{y} = 0$, to $\lim_{z \rightarrow 0} p_Z = 0$.

Na mocy twierdzenia 2.3, twierdzenia 2.2 i warunku $\Delta < 1/(2L)$ można wykazać zbieżność rozwinięcia Hermite'a $f_Z^{(J)}$ do gęstości p_Z , gdy $J \rightarrow \infty$, przy czym poprzez pojęcie rozwinięcia Hermite'a rozumiany jest ciąg funkcyjny zadany następującym wzorem:

$$f_Z^{(J)}(z|y_s, s, \Delta; \theta) = \phi(z) \sum_{j=0}^J \eta_Z^{(j)}(y_s, s, \Delta; \theta) H_j(z) \quad (2.9)$$

gdzie

$$\eta_Z^{(j)}(y_s, s, \Delta; \theta) = \frac{1}{j!} \int_{-\infty}^{\infty} H_j(z) p_Z(z|y_s, s, \Delta; \theta) dz.$$

Związki pomiędzy poszczególnymi aproksymacjami są analogiczne jak we wzorze (2.8):

$$f_Y^{(J)}(y|y_s, s, \Delta; \theta) = \Delta^{-1/2} f_Z^{(J)}(\Delta^{-1/2}(y - y_s)|y_s, s, \Delta; \theta) \quad (2.10)$$

$$f_X^{(J)}(x|x_s, s, \Delta; \theta) = b(x, t; \theta)^{-1} f_Y^{(J)}(\gamma(x, t; \theta)|\gamma(x_s, s; \theta), s, \Delta; \theta)$$

Poniższe twierdzenie precyzuje rodzaj zbieżności aproksymacji $f_X^{(J)}$ do prawdziwej, choć nieznannej, gęstości przejścia p_X :

Twierdzenie 2.4

Jeśli spełnione są założenia 2.1–2.3, to istnieje $\bar{\Delta} > 0$ takie, że dla każdego $(x, x_s, \Delta, s, \theta) \in D_X^2 \times (0, \bar{\Delta}) \times (0, \infty) \times \Theta$ zachodzi:

$$p_X^{(J)}(x|x_s, s, \Delta; \theta) \xrightarrow{J \rightarrow \infty} p_X(x|x_s, s, \Delta; \theta).$$

Zbieżność ta jest jednostajna ze względu na $\theta \in \Theta$, x_s należący do dowolnego zwartego podzbioru D_X oraz $s \in [0, \hat{t}]$ dla dowolnego $\hat{t} > 0$.

Ponadto, jeśli dla każdego $\xi > 0$ i $\hat{t} > 0$ istnieje stała c_ξ taka, że $b(x, t; \theta) \geq c_\xi > 0$ dla wszystkich $(x, t, \theta) \in [\xi, \infty) \times [0, \hat{t}] \times \Theta$, wówczas zbieżność jest jednostajna również względem x na dowolnym przedziale postaci $[\varepsilon, \infty)$, gdzie $\varepsilon > 0$.

Jeśli $a_Y \leq 0$ w sąsiedztwie $+\infty$ i $a_Y \geq 0$ w sąsiedztwie lewej granicy dziedziny (tj. $-\infty$ lub 0), wówczas $\bar{\Delta} = \infty$; w pozostałych przypadkach $\bar{\Delta} = 1/(2K)$ (gdzie K jak w założeniu 2.3).

Znaczenie twierdzenia 2.4 oraz pomysłu aproksymacji nieznanej gęstości przy wykorzystaniu wielomianów Hermite’a uzależnione jest od wielkości $\bar{\Delta}$. Przy stosownych założeniach na zbiór parametrów Θ okazuje się jednak, że dla większości najczęściej rozważanych w literaturze przedmiotu oraz w zastosowaniach modeli (w tym dla tych prezentowanych w monografii) $\bar{\Delta} = \infty$. Czyni to zaproponowaną aproksymację silnym narzędziem w rękach badacza. Jednocześnie należy podkreślić, że jest to ogromny krok naprzód w poszukiwaniu coraz dokładniejszych przybliżeń, wolnych od założenia, by czas pomiędzy obserwacjami był „odpowiednio bliski zeru”.

Istotnym mankamentem rozwinięcia Hermite’a jest brak w ogólnym przypadku analitycznych wzorów na współczynniki $\eta_Z^{(j)}$.

Można jednak zauważyć, że:

$$\begin{aligned}\eta_Z^{(j)}(y_s, s, \Delta; \theta) &= \frac{1}{j!} \int_{-\infty}^{\infty} H_j(z) p_Z(z | y_s, s, \Delta; \theta) dz = \\ &= \frac{1}{j!} \int_{-\infty}^{\infty} H_j(z) \Delta^{1/2} p_Y(\Delta^{1/2} z + y_s | y_s, s, \Delta; \theta) dz = \\ &= \frac{1}{j!} \int_{-\infty}^{\infty} H_j(\Delta^{-1/2}(y - y_s)) p_Y(y | y_s, s, \Delta; \theta) dy = \\ &= \frac{1}{j!} E \left[H_j(\Delta^{-1/2}(Y_{s+h} - Y_s)) | Y_s = y_s, s; \theta \right].\end{aligned}$$

Zatem współczynniki $\eta_Z^{(j)}$ są pewnego rodzaju warunkowymi momentami procesu Y i jako takie można je przybliżać na różne sposoby. W pracach Aït-Sahalia (2002) oraz Egorov, Li i Xu (2003) zaproponowano przybliżanie wspomnianych wartości poprzez stosowną wersję wzoru Taylora rzędu m z wykorzystaniem operatora infinitezymalnego. Wówczas aproksymacja gęstości p_Z dana jest analitycznym wzorem:

$$\begin{aligned}p_Z(z | y_s, s, \Delta; \theta) &\approx f_Z^{[m]}(z | y_s, s, \Delta; \theta) = \\ &= \phi(z) \sum_{k=0}^{2m} \eta_{Z,k}^{[m]}(y_s, s, \Delta; \theta) H_k(z)\end{aligned}\tag{2.11}$$

W cytowanych pracach sugerowano stosowanie wzoru Taylora rzędu $m = 3$ i wówczas współczynniki $\eta_{Z,k}^{[3]}$ mają postać¹⁰⁾:

$$\eta_{Z,0}^{[3]} = -1,$$

$$\begin{aligned} \eta_{Z,1}^{[3]} = & \Delta^{1/2} \zeta + \frac{1}{4} \Delta^{3/2} (2\zeta_{0,1} + 2\zeta_{\zeta 1,0} + \zeta_{2,0}) + \\ & + \frac{1}{24} \Delta^{5/2} (4\zeta_{0,2} + 4\zeta_{0,1}\zeta_{1,0} + 4\zeta_{\zeta 1,0}^2 + \\ & + 8\zeta_{\zeta 1,1} + 4\zeta_{\zeta 2,0}^2 + 6\zeta_{1,0}\zeta_{2,0} + 4\zeta_{2,1} + 4\zeta_{\zeta 3,0} + \zeta_{4,0}), \end{aligned}$$

$$\begin{aligned} \eta_{Z,2}^{[3]} = & -\frac{1}{2} \Delta (\zeta^2 + \zeta_{1,0}) - \frac{1}{12} \Delta^2 (6\zeta_{\zeta 0,1} + 6\zeta_{\zeta 1,0}^2 + 4\zeta_{1,0}^2 + 4\zeta_{1,1} + 7\zeta_{\zeta 2,0} + 2\zeta_{\zeta 3,0}) - \\ & - \frac{1}{96} \Delta^3 (12\zeta_{0,1}^2 + 16\zeta_{\zeta 0,2} + 40\zeta_{\zeta 0,1}\zeta_{1,0} + 28\zeta_{\zeta 1,0}^2 + 16\zeta_{1,0}^3 + 32\zeta_{\zeta 1,1}^2 + \\ & + 40\zeta_{1,0}\zeta_{1,1} + 12\zeta_{1,2} + 16\zeta_{\zeta 2,0}^3 + 24\zeta_{0,1}\zeta_{2,0} + 88\zeta_{\zeta 1,0}\zeta_{2,0} + 21\zeta_{\zeta 2,0}^2 + \\ & + 40\zeta_{\zeta 2,1} + 28\zeta_{\zeta 3,0}^2 + 32\zeta_{1,0}\zeta_{3,0} + 12\zeta_{3,1} + 16\zeta_{\zeta 4,0} + 3\zeta_{5,0}), \end{aligned}$$

$$\begin{aligned} \eta_{Z,3}^{[3]} = & \frac{1}{6} \Delta^{3/2} (\zeta^3 + 3\zeta_{\zeta 1,0} + \zeta_{2,0}) + \frac{1}{48} \Delta^{5/2} (12\zeta_{\zeta 0,1}^2 + 12\zeta_{\zeta 1,0}^3 + 12\zeta_{0,1}\zeta_{1,0} + \\ & + 28\zeta_{\zeta 1,0}^2 + 16\zeta_{\zeta 1,1} + 22\zeta_{\zeta 2,0}^2 + 24\zeta_{1,0}\zeta_{2,0} + 6\zeta_{2,1} + 14\zeta_{\zeta 3,0} + 3\zeta_{4,0}), \end{aligned}$$

$$\begin{aligned} \eta_{Z,4}^{[3]} = & -\frac{1}{24} \Delta^2 (\zeta^4 + 6\zeta_{\zeta 1,0}^2 + 3\zeta_{1,0}^2 + 4\zeta_{\zeta 2,0} + \zeta_{3,0}) - \frac{1}{240} \Delta^3 (20\zeta_{\zeta 0,1}^3 + 20\zeta_{\zeta 1,0}^4 + \\ & + 60\zeta_{\zeta 0,1}\zeta_{1,0} + 100\zeta_{\zeta 1,0}^2 + 40\zeta_{1,0}^3 + 40\zeta_{\zeta 1,1}^2 + 40\zeta_{1,0}\zeta_{1,1} + 50\zeta_{\zeta 2,0}^3 + \\ & + 20\zeta_{0,1}\zeta_{2,0} + 180\zeta_{\zeta 1,0}\zeta_{2,0} + 34\zeta_{\zeta 2,0}^2 + 30\zeta_{\zeta 2,1} + 50\zeta_{\zeta 3,0}^2 + 52\zeta_{1,0}\zeta_{3,0} + \\ & + 8\zeta_{3,1} + 23\zeta_{\zeta 4,0} + 4\zeta_{5,0}), \end{aligned}$$

¹⁰⁾ Wyznaczenie tych współczynników, podobnie jak i wielu prezentowanych w monografii analitycznych wzorów, wymaga żmudnych i często skomplikowanych obliczeń. Efektywne uproszczenie rachunków można uzyskać stosując gotowe pakiety komputerowe umożliwiające symboliczne obliczenia (np. program Mathematica).

$$\eta_{Z,5}^{[3]} = \frac{1}{120} \Delta^{5/2} \left(\varsigma^5 + 10\varsigma^3 \varsigma_{1,0} + 10\varsigma^2 \varsigma_{2,0} + 10\varsigma_{1,0} \varsigma_{2,0} + 15\varsigma \varsigma_{1,0}^2 + 5\varsigma \varsigma_{3,0} + \varsigma_{4,0} \right),$$

$$\eta_{Z,6}^{[3]} = -\frac{1}{720} \Delta^3 \left(\varsigma^6 + 15\varsigma^4 \varsigma_{1,0} + 20\varsigma^3 \varsigma_{2,0} + 10\varsigma_{2,0}^2 + 15\varsigma_{1,0}^3 + 15\varsigma_{1,0} \varsigma_{3,0} + \right. \\ \left. + 45\varsigma^2 \varsigma_{1,0}^2 + 15\varsigma^2 \varsigma_{3,0} + 60\varsigma \varsigma_{1,0} \varsigma_{2,0} + 6\varsigma \varsigma_{4,0} + \varsigma_{5,0} \right),$$

$$\eta_{Z,k}^{[3]} = 0 \text{ dla } k > 6, \text{ gdzie } \Delta = t - s, \varsigma = a_Y(s, y_s; \theta), \text{ ponadto dla dowolnych } i \text{ oraz } j \\ \text{zachodzi } \varsigma_{i,j} = \frac{\partial^{i+j} a_Y(s, y; \theta)}{\partial y^i \partial s^j} \Big|_{y=y_s}.$$

Warto zauważyć, że aproksymacja $f_X^{[m]}$, uzyskana za pomocą wzoru Jacobiego z aproksymacji dla procesu Z , może przyjmować wartości ujemne oraz może nie być funkcją unormowaną. Jest jednak bardzo wygodnym narzędziem aproksymacji ze względu na łatwość i szybkość jej wyliczenia.

B. Jensen i R. Poulsen w pracy z roku 1999 porównali dokładność wybranych metod aproksymacji gęstości przejścia. Rozważono modele jednorodne V, BS i CIR przy różnych układach parametrów i warunków początkowych.

Wzięto pod uwagę następujące sposoby aproksymacji:

- metodę drzewa dwumianowego (por. Nelson i Ramaswamy 1989);
- metodę opartą na $M \in \{0, 2, 4, 16, 32, 64\}$ zmiennych pośrednich (por. Pedersen 1995a i b);
- metodę opartą na schemacie Cranka–Nicolsona polegającą na numerycznym rozwiązaniu prospektywnego równania Kołmogorowa (por. Poulsen 1999a);
- metodę opartą na rozwinięciu hermitowskim dla $m \in \{1, 2, 3\}$.

Jako miarę dopasowania aproksymacji f^A do prawdziwej (znanej) gęstości przejścia p przyjęto:

$$\sup_i \left\{ \left| f^A(x_i) - p(x_i) \right| \right\} \text{ oraz } \sum_i \left| f^A(x_i) - p(x_i) \right| \Delta x \approx \int_{\mathbb{R}} \left| f^A(x) - p(x) \right| dx.$$

Okazało się wówczas, że metoda oparta na rozwinięciu hermitowskim jest w sposób wyraźny dokładniejsza od pozostałych. Ponadto ze względu na dokładność i czas obliczeń, metoda wielomianów Hermite’a okazała się również najlepsza. Następną w rankingu była metoda numerycznego rozwiązania prospektywnego równania Kołmogorowa, potem metoda drzewa dwumianowego, a na końcu metoda zmiennych pośrednich. Dla ostatniej ze wspomnianych metod zauważono, że wzrost liczby zmiennych pośrednich zwiększa istotnie czas obliczeń przy jednocześnie nieznacznej poprawie dokładności aproksymacji. Wyższość aproksymacji wielomianami Hermite’a wy-

kazali również A.V. Egorov, H. Li i Y. Xu, którzy w artykule z 2003 roku porównali dokładność przybliżenia gęstości przejścia dla wybranych niejednorodnych SRR, w tym modelu EOU.

Y. Aït-Sahalia (por. Aït-Sahalia 2002) zauważył, że w przypadku jednorodnych SRR, przy ustalonej liczbie K wyrazów rozwinięcia Hermite'a i pełnym rozwinięciu Taylora (tj. $m = \infty$), aproksymacja sformułowana dla procesu Y ma następującą postać:

$$\tilde{f}_Y^{(K)}(y | y_0; \Delta, \theta) = \Delta^{-1/2} \phi\left(\frac{y - y_0}{\Delta^{1/2}}\right) \exp\left(\int_{y_0}^y a_Y(w; \theta) dw\right) \sum_{k=0}^K c_k(y | y_0; \theta) \frac{\Delta^k}{k!} \quad (2.12)$$

gdzie $c_0(y | y_0; \theta) = 1$ dla $j \geq 1$ i $y \neq y_0$:

$$\begin{aligned} c_j(y | y_0; \theta) &= \\ &= j(y - y_0)^{-j} \int_{y_0}^y (w - y_0)^{j-1} \left\{ \lambda_Y(w) c_{j-1}(w | y_0; \theta) + \left(\partial^2 c_{j-1}(w | y_0; \theta) / \partial w^2 \right) / 2 \right\} dw \end{aligned} \quad (2.13)$$

natomiast dla $j \geq 1$ i $y = y_0$:

$$c_j(y_0 | y_0; \theta) = \lim_{\substack{y \rightarrow y_0 \\ y \neq y_0}} c_j(y | y_0; \theta).$$

gdzie λ_Y jak w założeniu 2.3.

W szczególności zachodzą następujące związki:

$$c_1(y | y_0; \theta) = \begin{cases} \frac{1}{y - y_0} \int_{y_0}^y \lambda_Y(w; \theta) dw & y \neq y_0 \\ \lambda_Y(y_0; \theta) & y = y_0 \end{cases},$$

$$\begin{aligned} c_2(y | y_0; \theta) &= \\ &= \begin{cases} \left[\frac{2}{(y - y_0)^2} \left[-\frac{1}{y - y_0} \int_{y_0}^y \lambda_Y(s) ds + \frac{1}{2} (\lambda_Y(y) + \lambda_Y(y_0)) \right] + \left(\frac{1}{y - y_0} \int_{y_0}^y \lambda_Y(w; \theta) dw \right)^2 \right] & y \neq y_0 \\ \lambda_Y(y_0; \theta)^2 + \frac{1}{6} \frac{\partial^2 \lambda_Y(y; \theta)}{\partial y^2} \Big|_{y=y_0} & y = y_0 \end{cases}. \end{aligned}$$

Y. Aït-Sahalia w artykule z 1999 roku zastosował aproksymację (2.12) dla $K = 1$ i $K = 2$ do oceny estymatora największej wiarygodności parametrów wybranych procesów Itô modelujących miesięczne stopy procentowe FED. Wyznaczone w ten sposób

estymatory okazały się bardzo bliskie rzeczywistym, tj. wyznaczonym z prawdziwej gęstości przejścia¹¹⁾. Ponadto powyższą aproksymację wykorzystał M. DiPietro (por. DiPietro 2001) do wnioskowania bayesowskiego opartego na wspomnianym powyżej szeregu czasowym. Zauważył on, że dla $K = 1$ aproksymacja może przyjmować wartości ujemne, natomiast wybór $K = 2$ gwarantuje nieujemność aproksymacji. M. DiPietro stosował unormowaną wersję wzoru (2.12) dla $K = 2$, tzn. aproksymacja została podzielona przez całkę po przestrzeni stanów lub jej numeryczną aproksymację.

Zaproponowana przez Y. Aït-Sahalię aproksymacja $\tilde{f}_Y^{(K)}$ gęstości przejścia nie zawsze całkuje się do jedności oraz może przyjmować wartości ujemne. W praktyce w miejsce ujemnych wartości aproksymacji gęstości przejścia przyjmuje się wartość zero. Odpowiada to pomnożeniu funkcji aproksymującej przez funkcję charakterystyczną zbioru, w którym przyjmuje ona wartości dodatnie. Należy podkreślić dużą łatwość i szybkość obliczeń numerycznych przy zastosowaniu powyższej aproksymacji – co jest przydatne w przeprowadzonym wnioskowaniu statystycznym.

Wśród poczynionych założeń prezentowanej teorii, wymagana jest znajomość jawnej postaci procesu Y , tj. możliwość wyliczenia całki we wzorze (2.6). W sytuacjach gdy założenie to nie jest spełnione, stosuje się odpowiednio zmodyfikowane podejście do wyżej zaproponowanego, niestety już nie tak efektywnie obliczeniowo (por. Aït-Sahalia 2003, 2004). Problem taki występuje na przykład dla równania, gdzie współczynnik dyfuzji dany jest wzorem:

$$b^2(x; \theta) = \theta_1 x^{-1} + \theta_2 + \theta_3 x + \theta_4 x^{\theta_5}.$$

Zagadnienie aproksymacji przy braku znajomości jawnej postaci procesu Y nie będzie rozważane w monografii.

Aproksymacja zadana wzorem (2.11) wykorzystywana będzie w niniejszej pracy do przybliżania wartości gęstości przejścia procesu EOU. Ma ona wówczas postać:

$$f_X^{(EOU)}(x | x_s, s, \Delta; \theta) = \frac{e^{-\sigma_1 t}}{\sigma \sqrt{\Delta}} \phi(z) \sum_{k=0}^{2m} \eta_{Z,k}^{[m]} \left(\frac{e^{-\sigma_1 s}}{\sigma} x_s, s, \Delta; \theta \right) H_k(z),$$

$$\text{gdzie } z = \frac{e^{-\sigma_1 t} x - e^{-\sigma_1 s} x_s}{\sigma \sqrt{\Delta}}.$$

¹¹⁾ Dla niektórych z rozważanych modeli istniała analityczna postać gęstości przejścia, która umożliwiła wyliczenie rzeczywistych wartości estymatorów największej wiarygodności, tym samym możliwe było wyznaczenie niedokładności aproksymacji estymatora.

Ponadto:

$$\varsigma = -(\beta + \sigma_1) y_s,$$

$$\varsigma_{1,0} = -(\beta + \sigma_1),$$

$$\varsigma_{i,j} = 0 \text{ dla } i > 1 \text{ i } j \geq 0.$$

Należy zwrócić uwagę na prostotę powyższej aproksymacji, która z punktu widzenia zastosowań jest dużym atutem tej metody.

2.5. Metoda rozwiązań równań Fokkera, Plancka i Kołmogorowa

Na początku tego podrozdziału warto wspomnieć o ważnym wyniku teoretycznym, zawartym w pracy Dacunha-Castelle i Florens-Zmirou z 1986 roku, gdzie przy pewnych założeniach wyznaczono wzór na gęstość przejścia procesu Itô spełniającego SRR o stałej funkcji dyfuzji. Istotnym mankamentem tego wyniku okazał się czynnik będący wartością oczekiwaną pewnej całki stochastycznej, której wartość nie jest znana.

We wspomnianej pracy skoncentrowano się na jednorodnych SRR o stałej funkcji zmienności, tj.:

$$dY_t = a_Y(Y_t; \theta)dt + dW_t \quad (2.14)$$

Okazuje się wówczas, że:

Twierdzenie 2.5 (por. Dacunha-Castelle i Florens-Zmirou 1986)

Jeśli istnieje stała $M > 0$ taka, że $|g(y)| \leq M |y|^2$ w sąsiedztwie $\pm\infty$ przy czym:

$$g(y) = -\frac{1}{2} \left(a_Y(y; \theta)^2 + \frac{da_Y(y; \theta)}{dy} \right),$$

to gęstość przejścia procesu Y ma następującą postać:

$$p(y | y_0; \Delta, \theta) = \frac{1}{\sqrt{2\pi\Delta}} \exp \left(-\frac{(y - y_0)^2}{2\Delta} + \int_{y_0}^y a(u; \theta) du \right) E[\Psi] \quad (2.15)$$

gdzie:

$$\begin{aligned}\psi &= \exp \left\{ \Delta \int_0^1 g \left(z_u(y_0, y) + \sqrt{\sigma^2 \Delta} B_u \right) du \right\}, \\ g(y) &= -\frac{1}{2} \left(a_Y(y; \theta)^2 + \frac{da_Y(y; \theta)}{dy} \right), \\ z_u(y_0, y) &= y_0(1-u) + yu, \\ B_t &= W_t - tW_1, \quad 0 \leq t \leq 1.^{12)}\end{aligned}$$

Przedstawiony wynik można rozszerzyć na przypadek dowolnego jednorodnego SRR, jeśli tylko można je przekształcić do postaci (2.14) (np.: poprzez przekształcenie Lampertiego – por. wzór (2.6)).

Jak już wspomniano, największym problemem z punktu widzenia zastosowań powyższego wzoru jest obliczenie $E[\Psi]$. W literaturze przedmiotu można znaleźć różne propozycje oszacowania tej wartości, na przykład poprzez stosowny szereg wartości oczekiwanych odpowiednich transformacji mostu Browna (por. Dacunha-Castelle i Florens-Zmirou 1986) lub oszacowanie jej za pomocą metod Monte Carlo (por. Nicolau 2002).

Ze względu na podobieństwo wzorów (2.12) i (2.15) ostatni czynnik wzoru (2.12) można potraktować jako alternatywne przybliżenie nieznannej wartości $E[\Psi]$. Mankament takiej aproksymacji jest przyjmowanie przez nią wartości ujemnych przy jednoczesnej nieujemności wartości oczekiwanej. Wady tej nie ma aproksymacja, zaproponowana w pracy Aït-Sahalia (2003, 2004), następującej postaci:

$$E[\Psi] \approx \exp \left\{ \sum_{k=1}^K C_Y^{(k)}(y | y_0; \theta) \frac{\Delta^k}{k!} \right\}.$$

Nieznane współczynniki $C_Y^{(k)}$ wyznacza się w sposób jednoznaczny z faktu, że aproksymacja gęstości przejścia postaci:

$$\begin{aligned}\hat{f}_Y^{(K)}(y | y_0; \Delta, \theta) &= \\ &= \Delta^{-1/2} \phi \left(\frac{y - y_0}{\Delta^{1/2}} \right) \exp \left(\int_{y_0}^y a_Y(w; \theta) dw \right) \exp \left\{ \sum_{k=1}^K C_Y^{(k)}(y | y_0; \theta) \frac{\Delta^k}{k!} \right\}\end{aligned}\quad (2.16)$$

¹²⁾ Proces stochastyczny B_t nazywany jest mostem Browna.

spełnia przybliżone równania Fokkera, Plancka i Kołmogorowa (por. wzory (1.5) i (1.6):

$$\left\{ \begin{aligned} \frac{\partial \hat{f}_Y^{(K)}(y|y_0; \Delta, \theta)}{\partial \Delta} + \frac{\partial}{\partial y} \{a_Y(y; \theta) \hat{f}_Y^{(K)}(y|y_0; \Delta, \theta)\} - \frac{1}{2} \frac{\partial^2 \hat{f}_Y^{(K)}(y|y_0; \Delta, \theta)}{\partial y^2} &= O(\Delta^K) \\ \frac{\partial \hat{f}_Y^{(K)}(y|y_0; \Delta, \theta)}{\partial \Delta} - a_Y(y_0; \theta) \frac{\partial \hat{f}_Y^{(K)}(y|y_0; \Delta, \theta)}{\partial y_0} - \frac{1}{2} \frac{\partial^2 \hat{f}_Y^{(K)}(y|y_0; \Delta, \theta)}{\partial y_0^2} &= O(\Delta^K) \end{aligned} \right. .$$

Można pokazać, że szukane współczynniki zadane są wzorem rekurencyjnym postaci:

$$C_Y^{(0)}(y|y_0; \theta) = \int_{y_0}^y a_Y(w; \theta) dw \quad (2.17)$$

$$C_Y^{(k)}(y|y_0; \theta) = k \int_0^1 G_Y^{(k)}(y_0 + u(y - y_0)|y_0; \theta) u^{k-1} du, \text{ dla } k > 0$$

gdzie

$$\begin{aligned} G_Y^{(1)}(y|y_0; \theta) &= \\ &= -\frac{\partial a_Y(y; \theta)}{\partial y} - a_Y(y; \theta) \frac{\partial C_Y^{(0)}(y|y_0; \theta)}{\partial y} + \frac{1}{2} \left[\frac{\partial^2 C_Y^{(0)}(y|y_0; \theta)}{\partial y^2} + \left(\frac{\partial C_Y^{(0)}(y|y_0; \theta)}{\partial y} \right)^2 \right] \end{aligned} \quad (2.18)$$

oraz dla $k \geq 2$

$$\begin{aligned} G_Y^{(k)}(y|y_0; \theta) &= \\ &= -a_Y(y; \theta) \frac{\partial C_Y^{(k-1)}(y|y_0; \theta)}{\partial y} + \frac{1}{2} \frac{\partial^2 C_Y^{(k-1)}(y|y_0; \theta)}{\partial y^2} + \\ &+ \frac{1}{2} \sum_{h=0}^{k-1} \binom{k-1}{h} \frac{\partial C_Y^{(h)}(y|y_0; \theta)}{\partial y} \frac{\partial C_Y^{(k-1-h)}(y|y_0; \theta)}{\partial y} . \end{aligned}$$

Ponadto istnieje związek pomiędzy współczynnikami c_k zdefiniowanymi w poprzednim podrozdziale (por. wzór (2.13)) i $C_Y^{(k)}$. Dla pierwszych dwóch wyrazów jest on następujący:

$$C_Y^{(1)}(y | y_0; \theta) = c_1(y | y_0; \theta) = \begin{cases} \frac{1}{y - y_0} \int_{y_0}^y \lambda_Y(w; \theta) dw & y \neq y_0 \\ \lambda_Y(y_0; \theta) & y = y_0 \end{cases},$$

$$C_Y^{(2)}(y | y_0; \theta) = c_2(y | y_0; \theta) - [C_Y^{(1)}(y | y_0; \theta)]^2 = \begin{cases} \frac{2}{(y - y_0)^2} \left[-\frac{1}{y - y_0} \int_{y_0}^y \lambda_Y(s; \theta) ds + \frac{1}{2} (\lambda_Y(y; \theta) + \lambda_Y(y_0; \theta)) \right] & y \neq y_0 \\ \frac{1}{6} \frac{\partial^2 \lambda_Y(y; \theta)}{\partial y^2} \Big|_{y=y_0} & y = y_0 \end{cases}.$$

Wspomniana zależność nie jest przypadkowa i wynika z równości pomiędzy sumą

$\sum_{k=0}^K c_k(y | y_0; \theta) \frac{\Delta^k}{k!}$ oraz rozwinięciem $\exp \left\{ \sum_{k=1}^K C_Y^{(k)}(y | y_0; \theta) \frac{\Delta^k}{k!} \right\}$ w szereg Taylora¹³⁾ z pominięciem wyrazów Δ^j dla $j > K$ (por. Aït-Sahalia 2003, 2004). Innymi słowy, można powiedzieć, że wspomniane współczynniki c_k i $C_Y^{(k)}$ są względem siebie wielomianami¹⁴⁾.

Aproksymacja $\hat{f}_Y^{(K)}$ jest nieujemna, lecz nie zawsze jest unormowana. Należy podkreślić łatwość implementacji powyższych wzorów oraz szybkość wyliczenia wartości omówionej aproksymacji w przypadku rozważanych w pracy modeli jednorodnych.

Aproksymacja zadana wzorem (2.16) przy $K = 2$ jest stosowana w niniejszej pracy do aproksymacji gęstości przejścia procesów jednorodnych. Wartości współczynników $C_Y^{(k)}$, dla $k \in \{0, 1, 2\}$, oraz postaci funkcji aproksymujących dla rozważanych modeli zestawiono poniżej.

¹³⁾ $\exp\{z\} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots$

¹⁴⁾ $C_Y^{(k)}$ można wyrazić jako wielomiany współczynników c_k i na odwrót.

Model Vasička (V)

$$C_Y^{(0)}(y|y_0; \theta) = \int_{y_0}^y a_Y(w; \theta) dw = (y - y_0) \left[\frac{\alpha}{\sigma} - \frac{\beta}{2} (y + y_0) \right] \text{ dla } y \in D_Y,$$

$$C_Y^{(1)}(y|y_0; \theta) = \frac{\beta}{2} - \frac{\beta^2}{6\sigma^2} \left[\left(\frac{\alpha}{\beta} - \sigma y \right)^2 + \left(\frac{\alpha}{\beta} - \sigma y \right) \cdot \left(\frac{\alpha}{\beta} - \sigma y_0 \right) + \left(\frac{\alpha}{\beta} - \sigma y_0 \right)^2 \right] \text{ dla } y \in D_Y,$$

$$C_Y^{(2)}(y|y_0; \theta) = -\frac{\beta^2}{6} \text{ dla } y \in D_Y.$$

Dla $K = 2$ w celu unormowania aproksymacji gęstości przejścia $\hat{f}_Y^{(V)(2)}$ rozważanego modelu wystarczy zauważyć, że:

$$\int_{\underline{y}}^{\bar{y}} \hat{f}_Y^{(V)(2)}(y|y_0; \Delta, \theta) dy = \sqrt{1 + \beta\Delta + \frac{\beta^2\Delta^2}{3}} \exp\{-C - M\},$$

gdzie:

$$C = -\frac{y_0^2}{2\Delta} - \frac{\alpha}{\sigma} y_0 + \frac{\beta}{2} y_0^2 + \Delta \left[\frac{\beta}{2} - \frac{\alpha^2}{2\sigma^2} + \frac{\alpha\beta}{2\sigma} y_0 - \frac{\beta^2 y_0^2}{6} \right] - \frac{\beta^2}{12} \Delta^2,$$

$$M = \frac{\frac{1}{\Delta} y_0 + \frac{\alpha}{\sigma} + \left(\frac{\alpha\beta}{2\sigma} - \frac{\beta^2 y_0}{6} \right) \Delta}{2 \left(\frac{1}{\Delta} + \beta + \frac{\beta^2 \Delta}{3} \right)}.$$

Aproksymację gęstości przejścia $\hat{f}_X^{(V)(K)}$ dla procesu wyjściowego X dla $K \geq 1$ można otrzymać ze wzoru Jacobiego, tj.:

$$\hat{f}_X^{(V)(K)}(x|x_0; \Delta, \theta) = \frac{1}{\sigma} \hat{f}_Y^{(V)(K)}(x/\sigma | x_0/\sigma; \Delta, \theta).$$

Model Blacka i Scholesa (BS)

$$C_Y^{(0)}(y|y_0; \theta) = \int_{y_0}^y a_Y(w; \theta) dw = (y - y_0) \left(\frac{\alpha}{\sigma} - \frac{\sigma}{2} \right) \text{ dla } y \in D_Y,$$

$$C_Y^{(1)}(y|y_0; \theta) = -\frac{1}{2} \left(\frac{\alpha}{\sigma} - \frac{\sigma}{2} \right)^2 \quad \text{dla } y \in D_Y,$$

$$C_Y^{(k)}(y|y_0; \theta) = 0 \quad \text{dla } k \geq 2 \quad \text{i } y \in D_Y.$$

Na podstawie wzoru Jacobiego zachodzi równość:

$$\hat{f}_X^{(BS)(1)}(x|x_0; \Delta, \theta) = \frac{1}{\sigma x} \hat{f}_Y^{(BS)(1)}(\ln(x)/\sigma | \ln(x_0)/\sigma; \Delta, \theta).$$

Porównując powyższy wynik z rzeczywistą gęstością przejścia dla modelu Blacka i Scholesa (wzór (1.24)) otrzymujemy dla $K \geq 1$ następujący związek:

$$\hat{f}_X^{(BS)(K)}(x|x_0; \Delta, \theta) = \hat{f}_X^{(BS)(1)}(x|x_0; \Delta, \theta) = p_X^{(BS)}(x|x_0; \Delta, \theta).$$

Innymi słowy, zaproponowana aproksymacja już dla $K = 1$ jest równa prawdziwej gęstości przejścia!

Model Coxa, Ingersolla i Rossa (CIR)

$$C_Y^{(0)}(y|y_0; \theta) = \int_{y_0}^y a_Y(w; \theta) dw = \left(\frac{2\alpha}{\sigma^2} - \frac{1}{2} \right) \ln \left(\frac{y}{y_0} \right) - \frac{\beta}{4} (y^2 - y_0^2) \quad \text{dla } y \in D_Y,$$

$$C_Y^{(1)}(y|y_0; \theta) = -\frac{1}{2} \left(\frac{2\alpha}{\sigma^2} - \frac{1}{2} \right) \left(\frac{2\alpha}{\sigma^2} - \frac{3}{2} \right) \frac{1}{yy_0} - \frac{\beta^2}{24} (y^2 + yy_0 + y_0^2) + \frac{\beta\alpha}{\sigma^2} \quad \text{dla } y \in D_Y,$$

$$C_Y^{(2)}(y|y_0; \theta) = -\frac{1}{2} \left(\frac{2\alpha}{\sigma^2} - \frac{1}{2} \right) \left(\frac{2\alpha}{\sigma^2} - \frac{3}{2} \right) \frac{1}{y^2 y_0^2} - \frac{\beta^2}{24} \quad \text{dla } y \in D_Y.$$

Aproksymacja gęstości przejścia $\hat{f}_X^{(CIR)(K)}$ dla procesu wyjściowego X dla $K \geq 1$ wynosi:

$$\hat{f}_X^{(CIR)(K)}(x|x_0; \Delta, \theta) = \frac{1}{\sigma \sqrt{x}} \hat{f}_Y^{(CIR)(K)}(2\sqrt{x}/\sigma | 2\sqrt{x_0}/\sigma; \Delta, \theta).$$

Niestety nie istnieje analityczny wzór na $\int_{D_Y} \hat{f}_Y^{(CIR)(2)}(y|y_0; \Delta, \theta) dy$. Chcąc zatem unormować $\hat{f}_X^{(CIR)(2)}$, należy zastosować numeryczną aproksymację całki. W tym celu

w niniejszej pracy wykorzystana została metoda Simpsona. Funkcję $\hat{f}_X^{(CIR)(2)}$ można również wykorzystać do aproksymacji gęstości niecentralnego rozkładu χ^2 (por. wzór (1.29)).

Model Brennana i Schwarza (BSBS)

$$C_Y^{(0)}(y|y_0; \theta) = \int_{y_0}^y a_Y(w; \theta) dw = -\frac{\alpha}{\sigma^2} (e^{-\sigma y} - e^{-\sigma y_0}) - \left(\frac{\beta}{\sigma} + \frac{\sigma}{2} \right) (y - y_0) \text{ dla } y \in D_Y.$$

Dla $y \in D_Y$ i $y \neq y_0$:

$$C_Y^{(1)}(y|y_0; \theta) = \frac{\alpha^2}{4\sigma^3} \left(\frac{e^{-2\sigma y} - e^{-2\sigma y_0}}{y - y_0} \right) - \left(\frac{\alpha\beta}{\sigma^3} + \frac{\alpha}{\sigma} \right) \frac{(e^{-\sigma y} - e^{-\sigma y_0})}{y - y_0} - \frac{1}{2} \left(\frac{\beta}{\sigma} + \frac{1}{2}\sigma \right)^2,$$

$$\begin{aligned} C_Y^{(2)}(y|y_0; \theta) = & \frac{2}{(y - y_0)^2} \left[-\frac{\alpha^2}{4\sigma^3} \left(\frac{e^{-2\sigma y} - e^{-2\sigma y_0}}{y - y_0} \right) + \right. \\ & + \left(\frac{\alpha\beta}{\sigma^3} + \frac{\alpha}{\sigma} \right) \left(\frac{e^{-\sigma y} - e^{-\sigma y_0}}{y - y_0} \right) - \frac{\alpha^2}{4\sigma^2} (e^{-2\sigma y} + e^{-2\sigma y_0}) + \\ & \left. + \left(\frac{\alpha\beta}{2\sigma^2} + \frac{\alpha}{4} \right) (e^{-\sigma y} + e^{-\sigma y_0}) + \frac{\alpha}{4} (e^{-\sigma y} + e^{-\sigma y_0}) \right], \end{aligned}$$

zaś dla $y \in D_Y$ i $y = y_0$:

$$C_Y^{(1)}(y_0|y_0; \theta) = -\frac{\alpha^2}{2\sigma^2} e^{-2\sigma y_0} + \frac{\alpha\beta}{\sigma^2} e^{-\sigma y_0} - \frac{1}{2} \left(\frac{\beta}{\sigma} + \frac{1}{2}\sigma \right)^2,$$

$$C_Y^{(2)}(y_0|y_0; \theta) = -\frac{\alpha^2}{3} e^{-2\sigma y_0} + \frac{\alpha}{6} (\sigma^2 + \beta) e^{-\sigma y_0}.$$

Aproksymacja gęstości przejścia dla procesu wyjściowego X dla $K \geq 1$ wynosi:

$$\hat{f}_X^{(BSBS)(K)}(x|x_0; \Delta, \theta) = \frac{1}{\sigma x} \hat{f}_Y^{(BSBS)(K)}(\ln(x)/\sigma | \ln(x_0)/\sigma; \Delta, \theta).$$

W niniejszej pracy w celu unormowania funkcji $\hat{f}_X^{(BSBS)(2)}$ zastosowano metody numeryczne – algorytm Simpsona.

Należy zwrócić uwagę na prostotę powyższej aproksymacji, która z punktu widzenia zastosowań jest jej dużą zaletą. Prezentowaną metodę można rozszerzyć na przypadek nieautonomiczny – wymaga to jednak rozważenia dwuwymiarowego SRR. Zagadnienie to nie będzie rozważane w monografii.

2.6. Uwagi końcowe

W rozdziale drugim zaprezentowane zostały trzy podejścia do aproksymacji gęstości przejścia. Rozważono podejścia opierające się na:

- 1) schematach dyskretyzacji,
- 2) rozwinięciu funkcji w szereg wielomianów Hermite’a,
- 3) przybliżonych rozwiązaniach równań Fokkera, Plancka i Kołmogorowa (FPK).

Metody oparte na schematach dyskretyzacji oraz ich uogólnienie w postaci metody zmiennych pośrednich zostały opisane ze względu na naturalny związek pomiędzy modelami ciągłymi i ich dyskretnymi odpowiednikami, a także ze względu na ich wciąż dużą popularność. Pozostałe dwie metody opierają się na stosunkowo nowych pomysłach. Autorowi nie jest znane porównanie metody wielomianów Hermite’a oraz metody rozwiązań równań FPK, również prezentowana monografia nie miała tego na celu – metody te zostały bowiem zastosowane do odmiennych procesów dyfuzji (por. rozdz. 4 i 5).

Przykład empiryczny ilustrujący zastosowanie metody zmiennych pośrednich przedstawiony zostanie w rozdziale trzecim. Rozważanymi szeregami czasowymi będą trajektorie wygenerowane zgodnie z logarytmem procesu pierwiastka kwadratowego. W rozdziałach czwartym i piątym zaprezentowane zostaną wyniki opracowanego w niniejszej pracy wnioskowania bayesowskiego, wykorzystującego metody wielomianów Hermite’a i rozwiązań równań FPK. Przedmiotem modelowania będzie amerykańska stopa procentowa FED i szereg WIG20. Metoda rozwiązań równań FPK zastosowana zostanie dla modeli autonomicznych, natomiast metoda wielomianów Hermite’a wykorzystana zostanie w przypadku modelu nieautonomicznego EOU. Metodę rozwiązań równań FPK można zastosować również dla przypadku nieautonomicznego. Wymaga ona jednak rozpatrzenia dwuwymiarowego SRR, co wykracza poza zakres rozważań podjętych w monografii.

3. Wnioskowanie bayesowskie dla procesów Itô

3.1. Uwagi wstępne

W rozdziale zdefiniowana zostanie ogólna postać bayesowskiego modelu statystycznego oraz przedstawione zostaną zasady wnioskowania bayesowskiego, w tym: estymacja nieznanymi parametrów, predykcja realizacji badanego zjawiska oraz porównanie konkurencyjnych modeli. W rozdziale podane zostaną definicje modeli bayesowskich zadanych przez wybrane stochastyczne równania różniczkowe. Ponieważ metody numeryczne są często jedynymi narzędziami obliczania charakterystyk wielowymiarowych rozkładów a posteriori i rozkładów predyktywnych, poświęcona jest im znaczna część tego rozdziału. Ponadto omówione zostanie zagadnienie estymacji parametrów wybranego procesu Itô przy zastosowaniu metody zmiennych pośrednich. Celem tego eksperymentu jest ukazanie nieefektywności tej metody, co uzasadnia pominięcie w kolejnych badaniach empirycznych idei zmiennych pośrednich (por. rozdz. 4 i 5).

Wybór podejścia bayesowskiego jako narzędzia rozwiązywania stawianych w pracy problemów podyktowany był jego następującymi własnościami:

- poprzez rozkład a priori istnieje możliwość włączenia do wnioskowania w sposób jawny i sformalizowany wstępnej wiedzy o parametrach SRR posiadanej przed wglądem w dane;
- w sposób czysto probabilistyczny i wolny od asymptotycznych aproksymacji rozkład a posteriori dostarcza informacji o nieznanymi parametrach, czyniąc problem „dostatecznie licznej próby” nieistotnym w ujęciu bayesowskim;
- wynikiem wnioskowania są pełne rozkłady nieznanymi parametrów modeli oraz wektora prognozowanego, a nie jedynie oceny punktowe (i ew. błędy szacunku) nieznanymi wielkości;
- istnieje możliwość łączenia (uśredniania) wiedzy płynącej z różnych modeli.

Informacje dotyczące teoretycznych aspektów wnioskowania oraz metod numerycznych opisanych w niniejszym rozdziale zaczerpnięto w dużym stopniu z następujących opracowań: Zellner (1971), Box i Tiao (1973), Osiewalski (1991, 2001), Bernardo i Smith (2002), O’Hagan (1994), Szreder (1994), Gamerman (1997), Męczarski (1998), Krzykowski (2000), Pipień (2000) oraz Pajor (2003).

3.2. Ogólne zasady wnioskowania bayesowskiego

Metody bayesowskie uznawane są obecnie za nieklasyczne metody statystyczne. Ich odmienność w stosunku do metod klasycznych (tzw. metod teorio-próbkowych) jest wynikiem przyjęcia innych interpretacji prawdopodobieństwa. W podejściu klasycznym prawdopodobieństwo interpretowane jest jako idealizacja częstości, podczas gdy w ujęciu bayesowskim prawdopodobieństwo jest miarą stopnia przekonania (*degree of belief*) badacza co do prawdziwości zajścia danego zdarzenia. Konsekwencją tego jest uznanie parametrów klasycznego modelu za zmienne losowe oraz formalne uwzględnienie w procesie wnioskowania informacji spoza próby, które zostają wyrażone za pomocą rozkładu a priori dla parametrów modelu. Pełny układ aksjomatów gwarantujących istnienie i jednoznaczność miary stopnia przekonania zaprezentowany został m.in. w pracy DeGroot (1981).

W celu zdefiniowania modelu bayesowskiego wprowadzone zostaną dodatkowe oznaczenia. Niech wektor $x \in X \subset \mathbb{R}^T$ oznacza realizację badanego zjawiska empirycznego (obserwowaną w doświadczeniu). Natomiast przez $\tilde{x} \in \tilde{X} \subset \mathbb{R}^S$ oznaczono prognozowaną realizację tego zjawiska. Model statystyczny definiuje trójka $(X \times \tilde{X}, \Sigma, \mathbb{P})$, gdzie Σ jest σ -algebrą określoną na przestrzeni probabilistycznej $X \times \tilde{X}$, natomiast $\mathbb{P}$ jest rodziną rozkładów prawdopodobieństwa określonych na Σ . Dla potrzeb prezentowanej monografii przyjęto, że $\mathbb{P}$ jest parametryczną rodziną absolutnie ciągłych (względem miary Lebesgue'a) rozkładów prawdopodobieństwa zadanych poprzez gęstości p , tzn.:

$$\mathbb{P} = \{p(\cdot | \theta) : X \times \tilde{X} \rightarrow \mathbb{R}^+ \cup \{0\} \mid \theta \in \Theta \subset \mathbb{R}^d\}.$$

W podejściu klasycznym (niebayesowskim) rodzina $\mathbb{P}$ definiuje parametryczny model statystyczny (tzw. model próbkowy), w którego ramach rozważa się np.: zagadnienie estymacji wektora θ oraz predykcji wektora $\tilde{x}$. Na gruncie bayesowskim rodzina $\mathbb{P}$ jednak nie wystarcza do jednoznacznego zdefiniowania modelu statystycznego (tzw. bayesowskiego modelu statystycznego). Do pełnego jego określenia potrzebne jest włączenie informacji o estymowanych parametrach modelu, jaką ma badacz przed wglądem w dane, poprzez określenie miary probabilistycznej na zbiorze Θ za pomocą gęstości $p(\theta)$ – czyli tzw. rozkładu a priori. Statystyczny model bayesowski definiuje się wówczas poprzez określenie łącznej gęstości dla wektora obserwacji i wartości prognozowanych $(x, \tilde{x})$ oraz wektora parametrów θ :

$$p(x, \tilde{x}, \theta) = p(x, \tilde{x} | \theta) p(\theta).$$

Wyróżnić można dwa podstawowe kryteria doboru rozkładu a priori. W pierwszym z nich rozkład dobiera się tak, by odzwierciedlał całą wstępną wiedzę badacza (przed wglądem w dane) – mówi się wówczas o informacyjnym rozkładzie a priori.

W drugim kryterium rozważa się tzw. nieinformacyjne rozkłady a priori nazywane również referencyjnymi (*reference priors*) lub wzorcowymi. Idea wyboru nieinformacyjnego rozkładu ma pozwolić, aby dane „mówiły” same za siebie. Rozkłady takie zazwyczaj nie są miarami probabilistycznymi, a wyłącznie miarami σ -skończonymi (zwłaszcza gdy przestrzeń parametrów jest nieograniczona). W literaturze przedmiotu nazywa się je rozkładami niewłaściwymi (*improper distributions*), ponieważ miara całej przestrzeni liczona względem nich jest nieskończona. Mankamentem struktury a priori opartej na sformalizowanych rozkładach wzorcowych jest możliwość zastosowania jej wyłącznie w przypadku prostego modelu statystycznego. Dla bardziej skomplikowanych modeli przyjmuje się rozkłady a priori, które jedynie w sensie intuicyjnym są nieinformacyjne. W dalszej części pracy rozważane są tylko miary probabilistyczne (rozkłady „właściwe”).

Estymacja parametrów modelu bayesowskiego

Istota bayesowskiej estymacji parametrów polega na wyznaczeniu rozkładu warunkowego wektora θ przy zadanym wektorze obserwacji x , czyli wyznaczeniu rozkładu a posteriori o gęstości zadanej zgodnie ze wzorem Bayesa:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)},$$

gdzie $p(x|\theta)$ jest funkcją wiarygodności (inaczej gęstością próbkową¹⁾) (*likelihood function*), natomiast $p(x) = \int_{\Theta} p(x|\theta)p(\theta)d\theta$ jest brzegową gęstością wektora obserwacji x (*marginal data density*).

Wnioskowanie o każdej ze współrzędnych wektora parametrów $\theta = (\theta_1, \dots, \theta_d) \in \Theta$, $\Theta = \Theta_1 \times \dots \times \Theta_d \subseteq \mathbb{R}^d$ odbywa się na podstawie brzegowych rozkładów a posteriori, czyli gęstości postaci:

$$p(\theta_i|x) = \int_{\Theta_d} \dots \int_{\Theta_{i+1}} \int_{\Theta_{i-1}} \dots \int_{\Theta_1} p(\theta|x) d\theta_1 \dots d\theta_{i-1} d\theta_{i+1} \dots d\theta_d.$$

Wartość oczekiwana pewnej funkcji $h(\cdot)$ zmiennej losowej θ (na przykład wyliczanej w celu wyznaczenia momentów) zadana jest następującym wzorem (wprost z definicji):

$$E(h(\theta)|x) = \int_{\Theta} h(\theta)p(\theta|x)d\theta = \frac{\int_{\Theta} h(\theta)p(x|\theta)p(\theta)d\theta}{p(x)}.$$

¹⁾ Ścisłej funkcja wiarygodności jest odwzorowaniem zmiennej θ przy ustalonym x , tzn. $p(x|\cdot)$, natomiast gęstość próbkowa jest odwzorowaniem zmiennej x przy ustalonym θ , tzn. $p(\cdot|\theta)$.

Prognoza bayesowska

Celem prognozowania jest m.in. sprawdzenie własności predyktywnych modelu, tzn. zbadanie, z jaką niepewnością *ex ante* i dokładnością *ex post* jest on w stanie przewidzieć wartości przyszłe. Bayesowska prognoza opiera się na wyznaczeniu gęstości rozkładu warunkowego dla wektora prognozowanego $\tilde{x}$ przy znanym wektorze obserwacji x , tzn. gęstości predyktywnej $p(\tilde{x}|x)$. Dla gęstości tej spełnione są wynikające z podstawowych zasad rachunku prawdopodobieństwa następujące równości:

$$\begin{aligned} p(\tilde{x}|x) &= \frac{p(x, \tilde{x})}{p(x)} = \frac{\int_{\Theta} p(x, \tilde{x}, \theta) d\theta}{p(x)} = \\ &= \int_{\Theta} p(\tilde{x}|x, \theta) \frac{p(x, \theta)}{p(x)} d\theta = \int_{\Theta} p(\tilde{x}|x, \theta) p(\theta|x) d\theta. \end{aligned}$$

Innymi słowy, funkcja $p(\tilde{x}|x)$ jest uśrednieniem (po całej przestrzeni parametrów) tzw. próbkowej gęstości predyktywnej $p(\tilde{x}|x, \theta)$ (*sampling predictive density*), przy czym funkcją wagową jest gęstość rozkładu a posteriori $p(\theta|x)$. Warto podkreślić, że zdefiniowane powyżej podejście do problemu prognozy jest odmienne w porównaniu z ujęciem, w którym rozkład predyktywny otrzymuje się z próbkowej gęstości predyktywnej po uprzednim oszacowaniu nieznanego wektora parametrów.

Reasumując powyższe rozważania, można stwierdzić, że istotą wnioskowania bayesowskiego jest wyznaczenie odpowiedniego rozkładu warunkowego dla badanej funkcji parametrów θ oraz przyszłych realizacji $\tilde{x}$ przy zadanym wektorze obserwacji x . Jednak w praktyce nie zawsze można wyznaczyć analityczną postać gęstości rozkładu warunkowego, a jeśli nawet można to zrobić, to gęstość ta ma często na tyle skomplikowaną postać, że skutecznie utrudnia estymację i prognozę. W takich sytuacjach stosowane są metody numeryczne, które zaprezentowano w dalszej części niniejszego rozdziału.

Chcąc określić oceny punktowe nieznanymi wielkościami, np. w celu porównania z ocenami uzyskanymi za pomocą klasycznych metod wnioskowania, wyznacza się najczęściej wartości oczekiwane lub mediany a posteriori. Należy jednak podkreślić, że sama ocena (prognoza) punktowa niesie znacznie mniej informacji niż cały rozkład a posteriori (predyktywny), prowadząc w specyficznych przypadkach (np.: dla rozkładów silnie asymetrycznych lub wielomodalnych) do niewłaściwych wniosków (por. Box i Tiao 1973).

3.3. Bayesowskie porównanie konkurencyjnych modeli

Na gruncie bayesowskim możliwe jest w pełni sformalizowane porównanie konkurencyjnych modeli opisujących badane zjawisko. Miarą porównania modeli są wiel-

kości prawdopodobieństw a posteriori każdego z nich. Model o najwyższym prawdopodobieństwie może być uznany za najlepszy i niekiedy jedynie dla niego przedstawia się i szczegółowo interpretuje wyniki bayesowskiej estymacji i prognozy.

W celu zaprezentowania idei porównywania modeli rozważono m konkurencyjnych bayesowskich specyfikacji:

$$M_i: \quad p_i(x, \theta_{(i)}) = p_i(x | \theta_{(i)}) p_i(\theta_{(i)}), \quad i = 1, \dots, m,$$

gdzie $p_i(x | \theta_{(i)})$, $p_i(\theta_{(i)})$ i $\theta_{(i)} = (\omega, \ell_{(i)}) \in \Theta_{(i)} = \Omega \times L_{(i)}$ są odpowiednio funkcją wiarygodności, gęstością a priori i wektorem parametrów modelu M_i . Przy czym wektor $\theta_{(i)}$ składa się z części grupującej parametry wspólne ω dla każdego z rozważanych modeli oraz parametry swoiste $\ell_{(i)}$ dla i -tego modelu. Przy założeniu, że rozpatrywane modele wzajemnie się wykluczają (są niezagnieżdżone) i tworzą pełny zbiór modeli opisujących badane zjawisko, prawdopodobieństwa a priori każdego z nich $p(M_1), \dots, p(M_m)$ sumują się do jedności: $p(M_1) + \dots + p(M_m) = 1$. Ponadto na mocy wzoru Bayesa prawdopodobieństwo a posteriori modelu wynosi:

$$p(M_i | x) = \frac{p(M_i) p(x | M_i)}{\sum_{k=1}^m p(M_k) p(x | M_k)}, \quad i = 1, \dots, m \quad (3.1)$$

gdzie $p(x | M_i) = p_i(x) = \int_{\Theta_{(i)}} p_i(x | \theta_{(i)}) p_i(\theta_{(i)}) d\theta_{(i)}$ jest brzegową gęstością wektora obserwacji dla modelu M_i . Funkcja ta jest średnią ważoną po przestrzeni parametrów funkcji wiarygodności względem gęstości a priori jako funkcji wagowej.

W literaturze przedmiotu zaleca się przyjęcie równych prawdopodobieństw $p(M_i)$ dla $i = 1, \dots, m$ lub dobór ich (zgodnie z tzw. zasadą brzytwy Ockhama) w taki sposób, by modele prostsze, tzn. o mniejszej liczbie parametrów, uczynić bardziej prawdopodobnymi (por. Jeffreys 1967; Osiewalski i Steel 1993). Jednym ze sposobów jest przyjęcie²⁾:

$$p(M_i) \propto 2^{-l_i} \quad (3.2)$$

gdzie l_i oznacza wymiar przestrzeni parametrów swoistych $l_{(i)}$. Ponadto, jeśli jakiś model ze względów merytorycznych zasługuje na uznanie go za bardziej prawdopodobny niż inne, wówczas przyznaje się mu duże prawdopodobieństwo a priori, np.: $\frac{1}{2}$, nato-

²⁾ Symbol $\propto$ oznacza równość z dokładnością do dodatniej stałej multiplikatywnej, tzn. $(f \propto g \Leftrightarrow \exists A > 0 : f = Ag)$.

miast pozostałe modele otrzymują równe prawdopodobieństwa lub przyjmuje się dla nich strukturę a priori zgodną ze wzorem (3.2).

Prawdopodobieństwo a posteriori modelu w widoczny sposób zależy od rozkładu a priori parametrów $p_i(\theta_{(i)})$ oraz prawdopodobieństw a priori rozważanych modeli $p(M_i)$. Wpływ prawdopodobieństw a priori modeli można wyeliminować, jeśli modele będą porównywane parami. W tym celu definiuje się tzw. iloraz szans a posteriori (*posterior odds ratio*):

$$\frac{p(M_i | x)}{p(M_j | x)} = \frac{p(M_i)}{p(M_j)} \cdot \frac{p(x | M_i)}{p(x | M_j)},$$

tzn. iloczyn ilorazu szans a priori (*prior odds ratio*) $\frac{p(M_i)}{p(M_j)}$ oraz czynnika Bayesa (*Bayes factor*) $B_{ij} = \frac{p_i(x)}{p_j(x)} = \frac{p(x | M_i)}{p(x | M_j)}$. Ten ostatni czynnik służy za miarę porów-

nawczą modeli, dostarczając informacji wnoszonych przez obserwacje badanego zjawiska o wartościowaniu modelu M_i względem M_j . Przy czym jeśli prawdopodobieństwa a priori modeli są równe, wówczas czynnik ten równa się ilorazowi szans a posteriori. Niestety B_{ij} nie jest wolny od założeń a priori dla parametrów. Warto zauważyć, że jeśli wysokie wartości funkcji wiarygodności nie towarzyszą dużemu prawdopodobieństwu a priori na tych samych obszarach przestrzeni parametrów, wówczas gęstość wektora obserwacji przyjmuje małe wartości (por. Osiewalski i Steel 1993). W przypadku niewłaściwych rozkładów a priori porównanie modeli nie jest już takie proste. Związane jest to z możliwą niejednoznacznością zarówno prawdopodobieństw a posteriori modeli, jak i czynników Bayesa. Dlatego w tej pracy zastosowane zostaną wyłącznie właściwe (choć rozproszone) rozkłady a priori parametrów modeli.

Warto nadmienić, że chcąc wnioskować o parametrach wspólnych w rozważanych modeli oraz chcąc dokonać bayesowskiej prognozy wektora przyszłych obserwacji $\tilde{x}$, można uniknąć wyboru „najlepszego” modelu, stosując tzw. metodę bayesowskiego łączenia wiedzy (*Bayesian pooling approach*). Metoda ta polega na uśrednieniu indywidualnych gęstości a posteriori parametrów (gęstości predyktywnych przyszłych obserwacji) z wagami równymi prawdopodobieństwom a posteriori modeli. Wówczas:

$$p(\omega | x) = \sum_{i=1}^m p(M_i | x) p_i(\omega | x),$$

gdzie $p_i(\omega | x) = \int_{L_i} p_i(\omega | x, \ell_i) p_i(\ell_i | x) d\ell_i$ jest gęstością a posteriori parametrów wspólnych, wyznaczoną na podstawie i -tego modelu. Natomiast w przypadku bayesowskiej prognozy:

$$p(\tilde{x}|x) = \sum_{i=1}^m p(M_i|x) p_i(\tilde{x}|x),$$

gdzie $p_i(\tilde{x}|x) = \int_{\Theta_{(i)}} p_i(\tilde{x}|x, \theta_{(i)}) p_i(\theta_{(i)}|x) d\theta_{(i)}$ jest gęstością predyktywną dla i -tego modelu.

Prezentowane powyżej podejście do zagadnienia estymacji i prognozy uwzględnia dwa rodzaje niepewności, tj.:

- 1) niepewność co do wartości nieznanymi parametrów rozważanych modeli lub nie-obszerowanej realizacji zjawiska,
- 2) niepewność co do prawidłowej specyfikacji w pewnej skończonej klasie modeli (uwzględniona poprzez łączenie – uśrednienie – wiedzy dostarczanej przez rozważane modele).

3.4. Estymacja, predykcja oraz porównanie modeli zadanych przez wybrane SRR

W niniejszym podrozdziale, na podstawie rozważanych w pracy SRR, zdefiniowanych zostanie pięć modeli bayesowskich: V, BS, CIR, EOU i BSBS. Oprócz określenia struktury a priori zaprezentowane zostaną wzory na gęstości a posteriori, gęstości predykcyjne oraz prawdopodobieństwa a posteriori konkurencyjnych specyfikacji bayesowskich.

Rozważane w pracy modele zadane są przez układy stochastycznych równań różniczkowych, które w ogólnym przypadku mają następującą postać:

$$\begin{cases} dX_t = a(t, X_t; \theta) dt + b(t, X_t; \theta) dW_t \\ X_{t_0} = x_0 \end{cases}.$$

Niech $(x_0, x_1, \dots, x_T)'$ oznacza wektor kolejnych obserwacji, a $\tilde{x} = (x_{T+1}, \dots, x_{T+S})'$ ich dalszą, prognozowaną realizację. Dodatkowo niech $x = (x_1, \dots, x_T)'$. Zakłada się jednocześnie, że wektor $(x_0, x', \tilde{x})'$ jest dyskretną realizacją silnego rozwiązania powyższego układu równań, czyli ciągu zmiennych losowych $\{X_t\}_{t=0}^{T+S}$. Wnioskowanie bayesowskie przeprowadzane w niniejszej pracy będzie warunkowane względem x_0 , tj. wartości początkowej szeregu czasowego (czyli warunku początkowego SRR).

W ogólnym przypadku gęstość próbkową (inaczej funkcję wiarygodności) możemy zapisać w następującej postaci:

$$p(x|x_0, \theta) = p(x_1, \dots, x_T | x_0, \theta).$$

Ponieważ rozważane w pracy silne rozwiązania SRR są procesami Markowa, zatem prawdziwa jest tożsamość (por. wzór (1.2)):

$$p(x | x_0, \theta) = p(x_1, \dots, x_T | x_0, \theta) = \prod_{i=1}^T p(x_i | x_{i-1}; t_{i-1}, t_i - t_{i-1}; \theta) \quad (3.3)$$

Funkcja wiarygodności jest zatem iloczynem gęstości przejścia. W monografii funkcja $p(x_i | x_{i-1}; t_{i-1}, t_i - t_{i-1}; \theta)$ będzie rozumiana jako prawdziwa gęstość przejścia lub jej aproksymacja (por. rozdz. 2). Z własności Markowa zachodzą ponadto następujące związki:

$$p(\tilde{x} | x_0, x, \theta) = \prod_{i=T+1}^{T+S} p(x_i | x_{i-1}; t_{i-1}, t_i - t_{i-1}; \theta) \quad (3.4)$$

$$p(x, \tilde{x} | x_0, \theta) = \prod_{i=1}^{T+S} p(x_i | x_{i-1}; t_{i-1}, t_i - t_{i-1}; \theta) \quad (3.5)$$

Aby uzupełnić definicje modeli bayesowskich, określone zostaną poniżej rozkłady a priori parametrów.

Model Vasička (V)

Niech proces X_t będzie silnym rozwiązaniem SRR definiującego model V (por. wzór (1.20)). W tej sytuacji przedmiotem estymacji jest wektor parametrów modelu postaci $\theta_{(1)} = (\alpha, \beta, \sigma^2)$, gdzie $\alpha, \beta, \sigma^2 > 0$. Przyjęto, że łączna gęstość a priori wektora $\theta_{(1)}$ jest iloczynem gęstości rozkładów a priori jego współrzędnych (innymi słowy przyjmujemy, że parametry są a priori niezależne):

$$p_1(\theta_{(1)}) = p(\alpha) p(\beta) p(\sigma^2),$$

gdzie:

$$p(\alpha) = \frac{1}{\Phi\left(\frac{m_\alpha}{\sqrt{w_\alpha}}\right) \sqrt{2\pi w_\alpha}} \exp\left\{-\frac{(\alpha - m_\alpha)^2}{2w_\alpha}\right\} I_{(0,\infty)}(\alpha) \propto \phi(\alpha | m_\alpha, w_\alpha) I_{(0,\infty)}(\alpha),$$

$$m_\alpha = 0,01, w_\alpha = 10^7,$$

$$p(\beta) = \frac{1}{\Phi\left(\frac{m_\beta}{\sqrt{w_\beta}}\right) \sqrt{2\pi w_\beta}} \exp\left\{-\frac{(\beta - m_\beta)^2}{2w_\beta}\right\} I_{(0,\infty)}(\beta) \propto \phi(\beta | m_\beta, w_\beta) I_{(0,\infty)}(\beta),$$

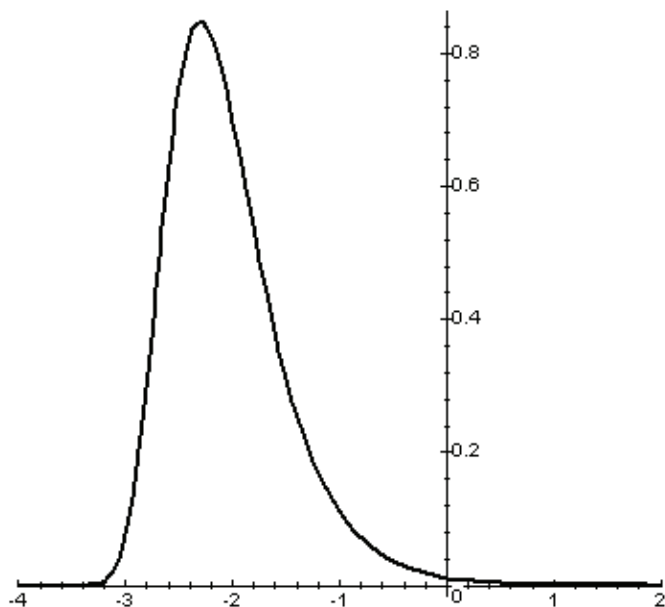
$$m_\beta = 0,01, w_\beta = 10^7,$$

przy czym funkcja $\phi(\cdot|m, w)$ oznacza gęstość rozkładu normalnego o średniej m i wariancji w , $\Phi(\cdot)$ jest dystrybuantą standardowego rozkładu normalnego, a odwzorowanie I_A jest funkcją charakterystyczną zbioru A . Dla parametru dyfuzji przyjęto natomiast:

$$p(\sigma^2) = p_{IG}(\sigma^2 | a, b) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2} \right)^{a+1} e^{-\frac{b}{\sigma^2}} I_{(0, \infty)}(\sigma^2), \quad a=1, \quad b=0,005,$$

gdzie $p_{IG}(\cdot|a, b)$ jest gęstością odwróconego rozkładu gamma ze średnią $b/(a-1)$, gdy $a > 1$, oraz wariancją $b^2/((a-1)^2(a-2))$, gdy $a > 2$, zaś w pozostałych odpowiednio przypadkach wartości te nie istnieją. Wobec tego dla przyjętych wielkości a i b wartość oczekiwana i wariancja rozkładu a priori parametru σ^2 nie istnieje.

Zdefiniowane powyżej rozkłady a priori odzwierciedlają subiektywnie nikłą wstępną wiedzę (przed wglądem w dane) na temat parametrów modelu. Jest to oczywiste w przypadku uciętych rozkładów normalnych dla α i β , ponieważ wariancje w_α i w_β są bardzo duże. Natomiast dość duże rozproszenie rozkładu a priori parametru σ^2 ilustruje przedstawiona na rysunku 3.1 funkcja gęstości indukowanego rozkładu a priori dla $\log_{10}(\sigma^2)$. Dopuszcza on szeroki (praktycznie wystarczający) zakres rzędów wielkości parametru σ^2 .



Rys. 3.1. Wykres gęstości a priori zmiennej $\log_{10}(\sigma^2)$

Źródło: Opracowanie własne

Gęstość a posteriori wektora parametrów, gęstość predyktywna oraz brzegowa gęstość wektora obserwacji dla modelu V mają odpowiednio postać:

$$p_1(\theta_{(1)} | x_0, x) \propto \prod_{i=1}^T p_1(x_i | x_{i-1}, \theta_{(1)}) \cdot \exp \left\{ -\frac{(\alpha - m_\alpha)^2}{2w_\alpha} - \frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\beta) I_{(0,\infty)}(\sigma^2) \quad (3.7)$$

$$p_1(\tilde{x} | x_0, x) \propto \int_{\Theta_{(1)}} \prod_{i=1}^{T+S} p_1(x_i | x_{i-1}, \theta_{(1)}) \cdot \exp \left\{ -\frac{(\alpha - m_\alpha)^2}{2w_\alpha} - \frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\beta) I_{(0,\infty)}(\sigma^2) d\theta_{(1)} \quad (3.8)$$

$$p_1(x | x_0, M_1) \propto \int_{\Theta_{(1)}} \prod_{i=1}^T p_1(x_i | x_{i-1}, \theta_{(1)}) \cdot \exp \left\{ -\frac{(\alpha - m_\alpha)^2}{2w_\alpha} - \frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\beta) I_{(0,\infty)}(\sigma^2) d\theta_{(1)} \quad (3.9)$$

Model Blacka i Scholesa (BS)

Niech proces X_t będzie geometrycznym ruchem Browna, tj. silnym rozwiązaniem SRR definiującego model BS (por. wzór (1.23)). Estymowany wektor parametrów modelu jest postaci $\theta_{(2)} = (\beta, \sigma^2)$, gdzie $\beta \in \mathbb{R}$, $\sigma^2 > 0$. Przyjęto, że łączna gęstość a priori wektora $\theta_{(2)}$ jest iloczynem gęstości rozkładów a priori jego współrzędnych:

$$p_2(\theta_{(2)}) = p(\beta)p(\sigma^2).$$

Wybór rozkładów a priori odzwierciedla subiektywnie nikłą wstępną wiedzę na temat parametrów modelu:

$$p(\beta) = \phi(\beta | m_\beta, w_\beta) = \frac{1}{\sqrt{2\pi w_\beta}} \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} \right\}, \quad m_\beta = 0,01, w_\beta = 10^7,$$

$$p(\sigma^2) = p_{IG}(\sigma^2 | a, b) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2} \right)^{a+1} e^{-\frac{b}{\sigma^2}} I_{(0,\infty)}(\sigma^2), \quad a=1, b=0,005.$$

Ostatecznie można zapisać, że gęstość a posteriori wektora parametrów, gęstość predyktywna oraz brzegowa gęstość wektora obserwacji dla modelu BS mają postać, odpowiednio:

$$p_2(\theta_{(2)} | x_0, x) \propto \prod_{t=1}^T p_2(x_t | x_{t-1}, \theta_{(2)}) \cdot \exp \left\{ -\frac{(\beta - m_\alpha)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\sigma^2) \quad (3.10)$$

$$p_2(\tilde{x} | x_0, x) \propto \int_{\Theta_{(2)}} \prod_{t=1}^{T+S} p_2(x_t | x_{t-1}, \theta_{(2)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\sigma^2) d\theta_{(2)} \quad (3.11)$$

$$p_2(x | x_0, M_2) \propto \int_{\Theta_{(2)}} \prod_{t=1}^T p_2(x_t | x_{t-1}, \theta_{(2)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\sigma^2) d\theta_{(2)} \quad (3.12)$$

Model Coxa, Ingersolla i Rossa (CIR)

Niech X_t będzie procesem pierwiastka kwadratowego, tj. silnym rozwiązaniem SRR definiującego model CIR (por. wzór (1.26)). W tej sytuacji przedmiotem bayesowskiej estymacji jest wektor parametrów modelu postaci $\theta_{(3)} = (\alpha, \beta, \sigma^2)$, gdzie $\alpha, \beta, \sigma^2 > 0$ i $\alpha > \frac{3}{4}\sigma^2$. Przyjęto, że łączna gęstość a priori wektora $\theta_{(3)}$ zadana jest w następujący sposób:

$$p_3(\theta_{(3)}) = p(\beta)p(\alpha, \sigma^2),$$

gdzie:

$$p(\beta) = \frac{1}{\Phi\left(\frac{m_\beta}{\sqrt{w_\beta}}\right)\sqrt{2\pi w_\beta}} \exp\left\{-\frac{(\beta - m_\beta)^2}{2w_\beta}\right\} I_{(0,\infty)}(\beta) \propto \phi(\beta | m_\beta, w_\beta) I_{(0,\infty)}(\beta),$$

$$m_\beta = 0,01, w_\beta = 10^7,$$

$$p(\alpha, \sigma^2) \propto \exp\left\{-\frac{(\alpha - m_\alpha)^2}{2w_\alpha}\right\} \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2}\right)^{a+1} e^{-\frac{b}{\sigma^2}} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\sigma^2) I_{\left(\alpha > \frac{3}{4}\sigma^2\right)}(\alpha, \sigma^2)$$

dla $m_\alpha = 0,01, w_\alpha = 10^7$ $a = 1$ i $b = 0,005$.

Zdefiniowane powyżej rozkłady a priori są rozkładami prawdopodobieństwa odzwierciedlającymi subiektywnie nikłą wstępną wiedzę na temat parametrów modelu. Ostatecznie można zapisać, że gęstość a posteriori dla wektora parametrów, gęstość predykcyjna oraz brzegowa gęstość wektora obserwacji w przypadku modelu CIR wynoszą odpowiednio:

$$p_3(\theta_{(3)} | x_0, x) \propto \prod_{t=1}^T p_3(x_t | x_{t-1}, \theta_{(3)}) \cdot \exp\left\{-\frac{(\alpha - m_\alpha)^2}{2w_\alpha} - \frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2}\right\} \cdot \left(\frac{1}{\sigma^2}\right)^{a+1} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\beta) I_{(0,\infty)}(\sigma^2) I_{\left(\alpha > \frac{3\sigma^2}{4}\right)}(\alpha, \sigma^2) \quad (3.13)$$

$$p_3(\tilde{x} | x_0, x) \propto \int \prod_{t=1}^{T+S} p_3(x_t | x_{t-1}, \theta_{(3)}) \cdot \exp \left\{ -\frac{(\alpha - m_\alpha)^2}{2w_\alpha} - \frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\beta) I_{(0,\infty)}(\sigma^2) I_{\left(\alpha > \frac{3\sigma^2}{4}\right)}(\alpha, \sigma^2) d\theta_{(3)} \quad (3.14)$$

$$p_3(x | x_0, M_3) \propto \int \prod_{t=1}^T p_3(x_t | x_{t-1}, \theta_{(3)}) \cdot \exp \left\{ -\frac{(\alpha - m_\alpha)^2}{2w_\alpha} - \frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,\infty)}(\alpha) I_{(0,\infty)}(\beta) I_{(0,\infty)}(\sigma^2) I_{\left(\alpha > \frac{3\sigma^2}{4}\right)}(\alpha, \sigma^2) d\theta_{(3)} \quad (3.15)$$

Uogólniony model Ornsteina i Uhlenbecka (EOU)

Niech X_t będzie uogólnionym procesem Ornsteina i Uhlenbecka, tj. silnym rozwiązaniem SRR definiującego model EOU (por. wzór (1.31)). Estymowany wektor parametrów modelu jest postaci $\theta_{(4)} = (\beta, \sigma_1, \sigma^2)$, gdzie $\beta \in \mathbb{R}$, $\sigma_1 \in [-10, 10]$, $\sigma^2 > 0$ i $\beta + \sigma_1 \geq -1$. Przyjęto, że łączna gęstość a priori wektora $\theta_{(4)}$ zadana jest w następujący sposób³⁾:

$$p_4(\theta_{(4)}) = p(\beta, \sigma_1) p(\sigma^2),$$

gdzie:

$$p(\beta, \sigma_1) \propto \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} \right\} I_{[-10, 10]}(\sigma_1) I_{(\beta + \sigma_1 \geq -1)}(\beta, \sigma_1),$$

$$p(\sigma^2) = p_{IG}(\sigma^2 | a, b) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2} \right)^{a+1} e^{-\frac{b}{\sigma^2}} I_{(0,\infty)}(\sigma^2)$$

dla $m_\beta = 0,01$, $w_\beta = 10^7$ $a = 1$ i $b = 0,005$.

³⁾ W przypadku modelu EOU do aproksymacji gęstości przejścia wykorzystana jest metoda wielomianów Hermite'a. Jednym z jej założeń jest zwartość przestrzeni parametrów. W przypadku tego modelu wystarcza zwartość zbioru parametrów (β, σ_1) (por. Egorov, Li i Xu 2003). Dodatkowe zawężenie zbioru wartości β do maszynowego zakresu liczb jest restrykcją gwarantującą zwartość. Dla przejrzystości warunek ten zostanie jednak pominięty w kolejnych wzorach.

Rozkłady a priori są rozkładami prawdopodobieństwa odzwierciedlającymi subiektywnie nikłą wstępną wiedzę na temat parametrów modelu. Gęstość a posteriori dla wektora parametrów, gęstość predyktywna oraz brzegowa gęstość obserwacji w przypadku modelu EOU zadane są następującymi wzorami:

$$p_4(\theta_{(4)} | x_0, x) \propto \prod_{t=1}^T p_4(x_t | x_{t-1}, \theta_{(4)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\}. \quad (3.16)$$

$$\cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{[-10,10]}(\sigma_1) I_{(\beta+\sigma_1 \geq -1)}(\beta, \sigma_1) I_{(0,\infty)}(\sigma^2)$$

$$p_4(\tilde{x} | x_0, x) \propto \int_{\Theta_{(4)}} \prod_{t=1}^{T+S} p_4(x_t | x_{t-1}, \theta_{(4)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\}. \quad (3.17)$$

$$\cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{[-10,10]}(\sigma_1) I_{(\beta+\sigma_1 \geq -1)}(\beta, \sigma_1) I_{(0,\infty)}(\sigma^2) d\theta_{(4)}$$

$$p_4(x | x_0, M_4) \propto \int_{\Theta_{(4)}} \prod_{t=1}^T p_4(x_t | x_{t-1}, \theta_{(4)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\}. \quad (3.18)$$

$$\cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{[-10,10]}(\sigma_1) I_{(\beta+\sigma_1 \geq -1)}(\beta, \sigma_1) I_{(0,\infty)}(\sigma^2) d\theta_{(4)}$$

Model Brennana i Schwartza (BSBS)

Niech proces X_t będzie silnym rozwiązaniem SRR definiującego model BSBS (por. wzór (1.34)). Estymacją bayesowską objęty jest wektor parametrów modelu postaci $\theta_{(5)} = (\alpha, \beta, \sigma^2)$, gdzie $\alpha \in (0, 1)$, $\beta \in \mathbb{R}$, $\sigma^2 > 0$.

Łączną gęstość a priori wektora $\theta_{(5)}$ określono jako iloczyn gęstości rozkładów a priori jego współrzędnych:

$$p_5(\theta_{(5)}) = p(\alpha) p(\beta) p(\sigma^2),$$

gdzie rozkłady a priori odzwierciedlają subiektywnie nikłą wiedzę wstępną o parametrach modelu:

$$p(\alpha) = I_{(0,1)}(\alpha),$$

$$p(\beta) = \phi(\beta | m_\beta, w_\beta) = \frac{1}{\sqrt{2\pi w_\beta}} \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} \right\}, \quad m_\beta = 0,01, w_\beta = 10^7,$$

$$p(\sigma^2) = p_{IG}(\sigma^2 | a, b) = \frac{b^a}{\Gamma(a)} \left(\frac{1}{\sigma^2} \right)^{a+1} e^{-\frac{b}{\sigma^2}} I_{(0,\infty)}(\sigma^2), \quad a = 1, b = 0,005.$$

Gęstość a posteriori wektora parametrów, gęstość predyktywna oraz brzegowa gęstość obserwacji dla modelu BSBS wynoszą odpowiednio:

$$p_5(\theta_{(5)} | x_0, x) \propto \prod_{i=1}^T p_5(x_i | x_{i-1}, \theta_{(5)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\}. \quad (3.19)$$

$$\cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,1)}(\alpha) I_{(0,\infty)}(\sigma^2)$$

$$p_5(\tilde{x} | x_0, x) \propto \int_{\Theta_{(5)}} \prod_{i=1}^{T+S} p_5(x_i | x_{i-1}, \theta_{(5)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\}. \quad (3.20)$$

$$\cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,1)}(\alpha) I_{(0,\infty)}(\sigma^2) d\theta_{(5)}$$

$$p_5(x | x_0, M_5) \propto \int_{\Theta_{(5)}} \prod_{i=1}^T p_5(x_i | x_{i-1}, \theta_{(5)}) \cdot \exp \left\{ -\frac{(\beta - m_\beta)^2}{2w_\beta} - \frac{b}{\sigma^2} \right\} \quad (3.21)$$

$$\cdot \left(\frac{1}{\sigma^2} \right)^{a+1} I_{(0,1)}(\alpha) I_{(0,\infty)}(\sigma^2) d\theta_{(5)}$$

Z braku przesłanek do faworyzowania któregośkolwiek z przedstawionych powyżej modeli wydaje się właściwe uczynienie *a priori* każdego z nich równie prawdopodobnym, tj.:

$$p(M_1) = \dots = p(M_5) = 1/5,$$

gdzie $M_1, \dots, M_5$ odpowiadają rozważanym w pracy modelom.

Na mocy wzoru Bayesa prawdopodobieństwo a posteriori i -tego modelu wynosi:

$$p(M_i | x_0, x) = \frac{p(M_i) p(x | x_0, M_i)}{\sum_{k=1}^5 p(M_k) p(x | x_0, M_k)} = \frac{p(x | x_0, M_i)}{\sum_{k=1}^5 p(x | x_0, M_k)}, \quad i = 1, \dots, 5 \quad (3.6)$$

natomiast iloraz szans a posteriori równa się czynnikowi Bayesa:

$$\frac{p(M_i | x_0, x)}{p(M_j | x_0, x)} = \frac{p(M_i)}{p(M_j)} \cdot \frac{p(x | x_0, M_i)}{p(x | x_0, M_j)} = \frac{p(x | x_0, M_i)}{p(x | x_0, M_j)} = B_{ij}, \quad i, j = 1, \dots, 5.$$

Nieodłącznym zagadnieniem towarzyszącym wnioskowaniu bayesowskiemu jest analiza wrażliwości jego wyników na dobór rozkładu a priori nieznanymi parametrów. Aby zbadać wpływ rozkładów a priori na uzyskane wyniki, w pracy wykorzystano m.in. rozkład (lub jego odpowiednie ucięcie) zadany poprzez gęstość:

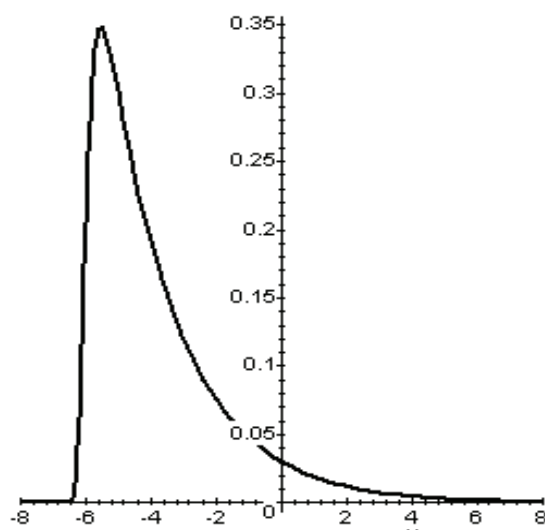
$$\phi(\cdot | m_\beta, w_\beta), \quad m_\beta = 0,01, w_\beta = 10^2$$

w przypadku parametrów, dla których przyjęto *a priori* rozkład normalny (lub jego odpowiednie ucięcie). Natomiast dla parametru σ^2 analizowano rozkłady o gęstości:

$$p_{IG}(\sigma^2 | a, b),$$

gdzie $(a, b) \in \{(3, 3/50), (0,1, 10^{-12})\}$.

Nieinformacyjny charakter odwróconego rozkładu gamma o parametrach $a = 0,1$ i $b = 10^{-12}$ (w porównaniu z rozkładem zilustrowanym na rys. 3.1) ilustruje funkcja gęstości indukowanego rozkładu a priori dla $\log_{10}(\sigma^2)$, obrazującego szanse różnych rzędów wielkości dla σ^2 (por. rys. 3.2).



Rys. 3.2. Wykres gęstości a priori zmiennej $\log_{10}(\sigma^2)$ dla parametrów $a = 0,1$, $b = 10^{-12}$
 Źródło: Opracowanie własne

3.5. Metody numeryczne wykorzystywane we wnioskowaniu bayesowskim

W niniejszym podrozdziale opisane zostaną metody numeryczne wykorzystywane w dalszej części pracy do przeprowadzania wnioskowania bayesowskiego, przy czym ograniczono się jedynie do opisu podstawowych własności oraz metod oceny zbieżności wybranych technik numerycznych. Przedstawione zostaną niektóre metody całkowania numerycznego służące do oszacowania charakterystyk rozpatrywanych rozkładów, a także podstawowe metody generowania prób z rozkładów, dla których losowania bezpośrednie są niemożliwe do wykonania.

Ważną klasę metod numerycznych stanowią metody Monte Carlo (MC), których idea oparta jest na generowaniu realizacji zmiennych losowych w celu oszacowania parametrów rozkładów tych zmiennych (por. Jermakow 1976). Rok 1733 uważa się za datę narodzin pomysłu wykorzystywania zjawisk losowych w zagadnieniach obliczeniowych. W roku tym G-L. Buffon⁴⁾ postawił problem aproksymacji liczby π poprzez oszacowanie prawdopodobieństwa przecięcia się prostej i rzuconej w sposób losowy igły (długości $2l$) na płaszczyznę poprzecinaną równoległymi prostymi (odległymi od

⁴⁾ Znany również pod nazwiskiem Georges-Louis Leclerc (por. Wiedza i Życie 10/2001).

siebie o $2a$, przy czym $l < a$). Następnie w roku 1777 podał rozwiązanie postawionego przez siebie problemu (por. Buffon 1777⁵⁾). Wiele lat później, gdyż dopiero w 1949 roku, w pracy Metropolis i Ulam (1949) po raz pierwszy zaprezentowano w sposób ścisły podstawy matematyczne metod Monte Carlo. Ze względu na wciąż rosnące możliwości komputerów, metody MC stały się bardzo popularnym i wygodnym narzędziem obliczeniowym. Ich rozwój przyczynił się do powstania różnorodnych technik numerycznych. Wśród nich można wymienić metody Monte Carlo z funkcją ważności skupiające wokół siebie algorytmy, w których losowanie odbywa się w sposób niezależny z tego samego rozkładu (próba prosta). Jeśli natomiast procedura generująca opiera się na jądrze przejścia, a jej wynikiem jest łańcuch Markowa, wówczas mówi się o markowskich metodach Monte Carlo, do których należy omówiony poniżej algorytm Metropolis i Hastingsa.

W przypadku wygenerowania próby prostej zbieżność do obliczanych charakterystyk rozkładu (ściślej odpowiednich wartości oczekiwanych) zapewnia prawo wielkich liczb, natomiast w przypadku próby będącej łańcuchem Markowa zbieżność gwarantują twierdzenia ergodyczne (por. np.: Tierney 1994).

W celu zaprezentowania wybranych metod numerycznych rozważono rozkład prawdopodobieństwa Π zadany przez gęstość:

$$p_{\Pi}(x) = f(x)/K \text{ dla } x \in \mathbb{R}^d \quad (3.22)$$

gdzie f nie musi być unormowana – jej stała normująca wynosi K (równa całce z f). Funkcja f nazywana jest jądrem rozkładu Π . Należy dodać, że prezentowane poniżej algorytmy numeryczne wykorzystuje się w sytuacji, gdy generowanie wprost z rozkładu Π jest trudne lub wręcz niemożliwe.

Metoda eliminacji (A-R)

Metoda eliminacji (*Acceptance-Rejection Sampling*, A-R), nazywana również procedurą A-R, umożliwia generowanie pseudolosowej próby z absolutnie ciągłego rozkładu (por. Zieliński 1970; Chib i Greenberg 1995).

Jeśli istnieje pewien rozkład prawdopodobieństwa wyznaczony przez gęstość s , z którego da się w prosty sposób generować próbę losową, oraz przy pewnej stałej C zachodzi nierówność $f(x) \leq Cs(x)$ dla wszystkich $x \in \mathbb{R}^d$, wówczas algorytm 3.1 umożliwia losowanie z rozkładu Π .

⁵⁾ Por. <http://mathworld.wolfram.com/BufconsNeedleProblem.html>. Prawdopodobieństwo to wynosi $2l(\pi a)$ (por. Plucińska i Pluciński 1970). G-L Buffon dokonywał oceny prawdopodobieństwa, rzucając za siebie na poliniowaną podłogę wysuszone bułki bagietki. W 1901 roku Włoch M. Lazzarini po 3408 rzutach „igłą Buffona” wyznaczył liczbę π z dokładnością do 6 miejsc po przecinku (por. Wiedza i Życie 10/2001).

Algorytm 3.3.1 (A-R)

1. Wylosowanie y z rozkładu zadanego przez gęstość s .
2. Wylosowanie u z rozkładu jednostajnego określonego na przedziale $(0, 1)$.
- 3a. Jeśli $u \leq \frac{f(y)}{Cs(y)}$, wówczas y traktowany jest jako wartość wygenerowana z rozkładu o gęstości p_{Π} .
- 3b. Jeśli $u > \frac{f(y)}{Cs(y)}$, wówczas powtarzane są kroki algorytmu (powrót do kroku 1).

Efektywność algorytmu uzależniona jest od wielkości stałej C , gdyż oczekiwana liczba przebiegów algorytmu, aż do uzyskania wartości wygenerowanej z rozkładu Π , wynosi $\frac{C}{K}$. Zatem, jak łatwo zauważyć, optymalizacja metody polega na przyjęciu

$C = \sup \left\{ \frac{f(x)}{s(x)}, x \in \text{supp}(s) \right\}$, gdzie $\text{supp}(\cdot)$ oznacza nośnik funkcji⁶⁾. Jednak wybór takiej wartości nie eliminuje pustych przebiegów algorytmu, tzn. braku wartości wygenerowanej z rozkładu Π po wykonaniu kroków 1–3 algorytmu.

Warto podkreślić, że procedura A-R umożliwia generowanie próby prostej (ściślej pseudolosowej). Ponadto metoda ta nie wymaga znajomości stałej normującej K .

Metoda Monte Carlo z funkcją ważności (MC-IS)

Metoda Monte Carlo z funkcją ważności (*Monte Carlo with Importance Sampling*, MC-IS) (por. Hammersley i Handscomb 1964) jest metodą całkowania numerycznego. We wnioskowaniu statystycznym ważna jest wiedza o momentach badanej zmiennej losowej X , jednak często nie wystarcza informacja o samym ich istnieniu, a potrzebna jest znajomość konkretnej wartości. Zagadnienie to związane jest z obliczaniem całek postaci:

$$I = E(h(X)) = \int_{\mathbb{R}^d} h(x) p_{\Pi}(x) dx.$$

W sytuacjach gdy wspomnianych całek nie da się obliczyć analitycznie, metoda MC-IS umożliwia numeryczne oszacowanie ich wartości. Wielkość I można zapisać w sposób równoważny, jako:

$$I = \int_{\text{supp}(s)} \frac{h(x) p_{\Pi}(x)}{s(x)} s(x) dx = \int_{\text{supp}(s)} \frac{h(x) p_{\Pi}(x)}{s(x)} dS(x),$$

⁶⁾ Nośnik funkcji (ozn. $\text{supp}(\cdot)$) jest zbiorem wszystkich elementów dziedziny, dla których funkcja przyjmuje wartości niezerowe.

gdzie S jest pewnym rozkładem prawdopodobieństwa zadany przez gęstość s (względem miary Lebesgue'a), przy czym s nazywa się funkcją ważności (stąd pochodzi nazwa metody). Zakłada się ponadto, że $\text{supp}(p_\Pi) \subset \text{supp}(s)$. Na mocy wzoru (3.22) można ostatecznie zapisać:

$$\begin{aligned} I &= \int_{\text{supp}(s)} \frac{h(x) p_\Pi(x)}{s(x)} dS(x) = \frac{1}{K} \int_{\text{supp}(s)} \frac{h(x) f(x)}{s(x)} dS(x) = \\ &= \frac{\int_{\text{supp}(s)} \frac{h(x) f(x)}{s(x)} dS(x)}{\int_{\text{supp}(s)} \frac{f(x)}{s(x)} dS(x)}. \end{aligned}$$

Zatem problem wyliczenia całki sprowadzony został do oszacowania ilorazu dwóch wartości oczekiwanych. Estymator $\hat{I}_N$ nieznanej wartości I definiowany jest w następujący sposób:

$$\hat{I}_N = \frac{\frac{1}{N} \sum_{i=1}^N h(x^{(i)}) f(x^{(i)}) / s(x^{(i)})}{\frac{1}{N} \sum_{i=1}^N f(x^{(i)}) / s(x^{(i)})},$$

gdzie zbiór $\{x^{(1)}, \dots, x^{(N)}\}$ oznacza próbę losową z rozkładu S .

Jeśli poprzez $w(\cdot)$ oznaczona zostanie funkcja wagowa, tzn. funkcja postaci $w(x) = f(x)/s(x)$, wówczas:

$$\hat{I}_N = \frac{\sum_{i=1}^N h(x^{(i)}) w(x^{(i)})}{\sum_{i=1}^N w(x^{(i)})}.$$

Nietrudno zauważyć, że estymator stałej normującej K równy jest średniej arytmetycznej wag:

$$\frac{1}{N} \sum_{i=1}^N w(x^{(i)}).$$

Z mocnego prawa wielkich liczb wynika zgodność estymatora $\hat{I}_N$.

Ponadto, jeśli zachodzą warunki:

$$E(w(X)) < \infty \quad \text{ i } \quad E(h^2(X)w(X)) < \infty,$$

wówczas rozkład zmiennej losowej $\sqrt{N}(\hat{I}_N - I)$ jest asymptotycznie normalny, co jest wnioskiem z centralnego twierdzenia granicznego (por. Geweke 1989; Piepiń i Osiewalski 1999). Warto zauważyć, że wyliczenie wartości estymatora nie wymaga znajomości stałej normującej K . Powyższe własności uzasadniają stosowanie metody MC-IS, jednak nie informują o sposobie jej zbieżności. Ponadto efektywność estymacji uzależniona jest od wyboru funkcji ważności s . Niestety literatura przedmiotu dostarcza jedynie wskazówek, a nie jednoznacznych zasad wyboru funkcji s . Znane są natomiast kryteria oceny jakości funkcji s (tzn. sposoby oceny efektywności numerycznej metody MC-IS przy zadanej funkcji ważności), wśród których należy wymienić opisane poniżej współczynniki względnej efektywności numerycznej i zmienności wagowej.

Współczynnik RNE względnej efektywności numerycznej (*relative numerical efficiency*), a dokładniej jego estymator $\hat{RNE}$, definiowany jest w następujący sposób (por. Geweke 1989):

$$\hat{RNE} = \frac{1}{N\hat{\phi}_N^2} \left[\frac{\sum_{i=1}^N h^2(x^{(i)}) w(x^{(i)})}{\sum_{i=1}^N w(x^{(i)})} - \hat{I}_N^2 \right],$$

gdzie $\hat{\phi}_N^2 = \frac{\sum_{i=1}^N (h(x^{(i)}) - \hat{I}_N)^2 w^2(x^{(i)})}{\left(\sum_{i=1}^N w(x^{(i)}) \right)^2}$ nazywany jest standardowym błędem numerycznym metody Monte Carlo (*Numerical Standard Error*, NSE).

Interpretacja wartości $\hat{RNE}$ jest następująca:

- jeśli $\hat{RNE}$ jest bliski jedności, to funkcję ważności s można uznać za nieznacznie różniącą się od gęstości rozkładu Π ;
- jeśli $\hat{RNE}$ przyjmuje małe, bliskie zeru wartości, wówczas funkcję s należy zastąpić inną gęstością, która uczyniłaby algorytm bardziej efektywnym;
- jeśli wartość $\hat{RNE}$ jest powyżej jedności, wówczas w celu zwiększenia efektywności metody sugeruje się, aby funkcję ważności podzielić przez funkcję h (wymagane jest w takim przypadku, by h była funkcja dodatnią).

Współczynnik zmienności funkcji wagowej γ_N zadany jest następującym wzorem:

$$\gamma_N = \frac{\sqrt{\sum_{i=1}^N \frac{w^2(x^{(i)})}{N} - \left(\sum_{i=1}^N \frac{w(x^{(i)})}{N} \right)^2}}{\frac{1}{N} \sum_{i=1}^N w(x^{(i)})}.$$

Interpretacja wartości γ_N jest następująca:

- jeśli funkcja s jest proporcjonalna do gęstości p_Π , to $\gamma_N = 0$;
- jeśli wraz ze wzrostem N rosną w sposób nieograniczony wartości współczynnika γ_N (w postaci nagłych skoków), oznacza to, że wybrana funkcja ważności jest niewłaściwa;
- jeśli wartości współczynnika γ_N stabilizują się (wraz ze wzrostem N), a obserwowane skoki wartości są niewielkie, wówczas akceptuje się wstępnie funkcję ważności.

Metoda Metropolisa i Hastingsa (M-H)

Metoda Metropolisa i Hastingsa, nazywana również algorytmem M-H, umożliwia wygenerowanie skorelowanej próby z rozkładu absolutnie ciągłego Π . Należy ona do klasy metod Monte Carlo łańcuchów Markowa (*Markov Chain Monte Carlo Method*, MCMC). W wyniku jej zastosowania otrzymywana jest realizacja jednorodnego procesu Markowa, którego rozkładem stacjonarnym (inaczej niezmienniczym lub granicznym) (*stationary, invariant, equilibrium distribution*) jest rozkład Π . Aby zdefiniować wspomniany proces, wprowadzone zostanie pojęcie gęstości generującej q i prawdopodobieństwa przejścia α .

Odwzorowanie $q(x, \cdot)$ oznacza gęstość rozkładu prawdopodobieństwa, z którego generowana jest propozycja kolejnego stanu łańcucha Markowa y (*candidate-generating density*), przy założeniu, że x jest jego stanem obecnym. W literaturze nie istnieją jednoznaczne definicje funkcji generujących, a wyłącznie sugestie ich wyboru, który uzależniony jest od konkretnego zagadnienia.

Prawdopodobieństwo przejścia α (inaczej funkcję akceptacji) określa się w następujący sposób:

$$\alpha(x, y) = \begin{cases} \min \left\{ \frac{f(y)q(y, x)}{f(x)q(x, y)}, 1 \right\} & \text{dla } f(x)q(x, y) > 0 \\ 1 & \text{dla } f(x)q(x, y) = 0 \end{cases}.$$

Łańcuch Markowa definiowany jest poprzez rozkład prawdopodobieństwa (tzw. rozkład przejścia) określony na σ -algebrze zbiorów borelowskich $B(\mathbb{R}^d)$ w następujący sposób:

$$P_{MH}(x, A) = \int_A q(x, y) \alpha(x, y) dy + \left[1 - \int_{\mathbb{R}^d} q(x, z) \alpha(x, z) dz \right] I_A(x)$$

dla $A \in B(\mathbb{R}^d)$.

Miara ta określa prawdopodobieństwo przejścia procesu ze stanu x do któregoś z dowolnych stanów należących do zbioru A . Ponadto rozkład warunkowy n -tego stanu łańcucha określony jest następującą zależnością:

$$P(x^{(n)} \in A | x^{(1)} = x) = P^{(n)}(x, A) = \int_{\mathbb{R}^d} P_{MH}(y, A) P^{(n-1)}(x, dy) \text{ dla } n > 1,$$

gdzie $P^{(1)}(x, A) = P_{MH}(x, A)$, zaś x jest arbitralnie dobranym stanem początkowym. Powyższe definicje zapewniają asymptotyczną stacjonarność procesu Markowa, a jego gęstość graniczna wynosi Π . Ostatecznie po odrzuceniu początkowych stanów łańcucha (tzw. cykli spalonych) można uznać pozostałe realizacje procesu za skorelowaną próbę z rozkładu stacjonarnego Π .

Algorytm M-H polega na wygenerowaniu łańcucha Markowa indeksowanego poprzez kolejne liczby naturalne $j = 1, \dots, S, S+1, \dots, S+N, \dots$, przy czym pierwsze S stanów łańcucha traktowanych jest jako cykle spalone i odrzucane. Poniżej zapisano kolejne kroki algorytmu.

Algorytm 3.2 (M-H)

Określenie w sposób arbitralny stanu początkowego łańcucha $x^{(1)}$.

Dla $j = 1, \dots, S, S+1, \dots, S+N-1$:

1. Wylosowanie propozycji kolejnego stanu y z rozkładu o gęstości $q(x^{(j)}, \cdot)$.
2. Wylosowanie wartości u z rozkładu jednostajnego na przedziale $(0, 1)$.
- 3a. Jeśli $u \leq \alpha(x^{(j)}, y)$, wówczas $x^{(j+1)} = y$.
- 3b. Jeśli $u > \alpha(x^{(j)}, y)$, wówczas $x^{(j+1)} = x^{(j)}$.

W literaturze przedmiotu sugeruje się przyjęcie $x^{(1)}$ równego wartości modalnej rozkładu Π . Dowolność w określeniu stanu początkowego, czy wręcz nałożenie na niego dowolnego rozkładu prawdopodobieństwa, nie wpływa na fakt zbieżności do asymptotycznego rozkładu Π (znaczenie stanu początkowego maleje wraz ze wzrostem liczby iteracji). Ponadto na wybór funkcji generującej mają wpływ: łatwość losowania

wania z rozkładu wyznaczonego przez tę gęstość, procent akceptacji algorytmu (wyrażający stosunek liczby przejść do nowego stanu łańcucha do liczby wszystkich wygenerowanych stanów) oraz szybkość zbieżności algorytmu.

Warto zauważyć, że algorytm nie wymaga znajomości stałej normującej K . Ponadto, jeśli funkcja generująca jest symetryczna, tzn.: $q(x, y) = q(y, x)$, wówczas:

$$\min \left\{ \frac{f(y)q(y, x)}{f(x)q(x, y)}, 1 \right\} = \min \left\{ \frac{f(y)}{f(x)}, 1 \right\}.$$

Zatem bezpośrednio z postaci prawdopodobieństwa przejścia wynika, że jeśli $f(y) \geq f(x)$, to propozycja następnego stanu łańcucha y jest akceptowana, w przeciwnym razie wartość ta jest przyjmowana z prawdopodobieństwem $\frac{f(y)}{f(x)}$ (por. Metropolis *et al.* 1953).

Należy nadmienić, że jeśli funkcja generująca nie zależy od poprzedniego stanu łańcucha, tzn. $q(x, y) = q(y)$, wówczas algorytm M-H nazywa się niezależnym (*independent version of the M-H algorithm*). W przeciwnym przypadku mamy do czynienia z sekwencyjnym algorytmem M-H (*sequential version of the M-H algorithm*) (por. Liu 2001; Osiewalski i Pipień 2004 a).

Algorytm M-H umożliwia losowanie z wielowymiarowych rozkładów prawdopodobieństwa, jednak nie wymaga generowania całego wektora w pojedynczym kroku metody. Możliwe jest bowiem zastosowanie tzw. blokowej wersji algorytmu M-H (por. Hastings 1970), której istota polega na podziale generowanego wektora na bloki i generowanie każdego z nich oddzielnie. Wiele możliwości podziału na bloki oraz odmienne sposoby generowania pociągnęły za sobą definicje różnych metod będących wciąż szczególnymi przypadkami algorytmu M-H. W literaturze przedmiotu zaleca się podział wektora w taki sposób, by każdy blok zawierał skorelowane ze sobą współrzędne⁷⁾. Okazuje się, że możliwe jest generowanie kolejnych bloków w sposób sekwencyjny tak, by losowanie bloku warunkowane było względem najnowszych stanów pozostałych współrzędnych wektora. Jednym ze szczególnych przypadków blokowego algorytmu M-H jest tzw. próbnik Gibbsa, inaczej algorytm Gibbsa (*Gibbs sampler*; *Gibbs algorithm*). Jego istotą jest generowanie każdego bloku wprost ze stosownego rozkładu warunkowego Π (a nie z arbitralnie wybranego rozkładu generującego). W konsekwencji prawdopodobieństwo przejścia dla tego algorytmu równe jest jedności, a procent akceptacji wynosi 100%. Wspomniane wartości świadczą o wysokiej efektywności próbnika Gibbsa, co w dużej mierze przyczyniło się do jego popularności.

⁷⁾ Przykłady empiryczne wskazują, że taki podział na bloki zmniejsza wpływ współzależności współrzędnych generowanego wektora, przyspieszając tym samym zbieżność łańcucha Markowa (por. Gamerman 1997).

Poniżej przedstawiono algorytm Gibbsa służący do generowania wektora zmiennej losowej x , przy czym wektor został podzielony na r bloków (niekoniecznie tej samej długości), tzn. $x = (x_1, \dots, x_r)$.

Algorytm 3.3 (próbnik Gibbsa)

Określenie stanu początkowego łańcucha $x^{(1)}$.

Dla $j = 1, \dots, S, S+1, \dots, S+N-1$ generowanie realizacji $x^{(j+1)}$ odbywa się w następujących krokach:

1. $x_1^{(j+1)} \sim p_{\Pi} \left(x_1 \mid x_2 = x_2^{(j)}, \dots, x_r = x_r^{(j)} \right),$
2. $x_2^{(j+1)} \sim p_{\Pi} \left(x_2 \mid x_1 = x_1^{(j+1)}, x_2 = x_2^{(j)}, \dots, x_r = x_r^{(j)} \right),$
- $\vdots$
- $r.$ $x_r^{(j+1)} \sim p_{\Pi} \left(x_r \mid x_1 = x_1^{(j+1)}, x_2 = x_2^{(j+1)}, \dots, x_{r-1} = x_{r-1}^{(j+1)} \right).$

W sytuacjach gdy nie jest możliwe zastosowanie algorytmu Gibbsa, tj. wówczas, gdy generowanie niektórych bloków bezpośrednio z rozkładów warunkowych nie jest możliwe, należy dla nich zastosować metodę A-R lub podstawową wersję algorytmu M-H, natomiast dla pozostałych bloków przeprowadzić bezpośrednie losowanie. Taka metoda losowania nazywana jest algorytmem hybrydowym (*hybrid algorithm*), tzn. zagnieżdżeniem podstawowego algorytmu M-H w jego wersji blokowej. Algorytm ten nazywany jest często próbnikiem Gibbsa z zagnieżdżonym w nim algorytmem M-H (*Metropolis–Hastings within Gibbs sampling*).

Wybór wielkości bloku może być ustalany również w sposób losowy przy każdym przebiegu algorytmu. Wówczas technika generowania nazywana jest algorytmem M-H o losowej wielkości bloku (*random block size M-H algorithm*) (por. Shephard i Pitt 1997). Przykłady empiryczne świadczą o przyspieszeniu zbieżności łańcucha Markowa po zastosowaniu wspomnianego algorytmu, co związane jest ze zmniejszeniem autokorelacji łańcucha.

Możliwe jest również uzyskanie próby prostej przy wykorzystaniu algorytmu M-H. W tym celu generuje się n niezależnych łańcuchów Markowa. Po osiągnięciu zbieżności procesów do rozkładu granicznego określa się n -elementową próbę prostą jako zbiór powstały poprzez wybór po jednym elemencie z każdego łańcucha.

Jedną z metod oceny zbieżności procesu Markowa jest badanie **statystyki CUM-SUM** (por. Yu i Mykland 1994; Osiewalski i Pipień 2002)⁸⁾.

⁸⁾ Zagadnienie zbieżności procesu Markowa omawiane jest ponadto w pracy Cowles i Carlin (1996).

W ogólnej postaci zadana jest ona wzorem:

$$CS_t = \frac{\left(\frac{1}{t} \sum_{j=S+1}^{S+t} h(x^{(j)}) - \bar{h} \right)}{\sigma} \quad \text{dla } t = 1, \dots, N,$$

gdzie $\bar{h}$ i σ są średnią i odchyleniem standardowym z próby $\{h(x^{(S+1)}), \dots, h(x^{(S+N)})\}$, czyli z realizacji łańcucha Markowa po odrzuceniu cykli spalonych. Przy czym $h(\cdot)$ jest pewną funkcją generowanego wektora x . Najczęściej przyjmuje się, że $h(x) = x$, co odpowiada badaniu zbieżności średnich procesu Markowa do wartości oczekiwanej rozkładu granicznego Π . W literaturze przedmiotu sugeruje się wybór liczby cykli spalonych S tak, by wykres statystyki CUMSUM „zbiegał gładko” do zera, pozostając od pewnego momentu (tj. od jednej z początkowych wartości t) w przedziale $(-\tau, \tau)$ dla ustalonej małej wartości τ . W praktyce przyjmuje się, że $\tau \in \{0,05, 0,1\}$.

Bayesowskie testowanie konkurencyjnych modeli, przeprowadzone bezpośrednio z porównania prawdopodobieństw a posteriori modeli albo poprzez porównanie czynników Bayesa, wymaga wyliczenia wartości brzegowych gęstości wektora obserwacji x (por. wzór (3.1)). Analityczne wyliczenie tych wartości jest często niemożliwe ze względu na skomplikowaną postać funkcji podcałkowych. Dlatego też w literaturze przedmiotu zaleca się stosowanie różnych metod aproksymujących prawdziwe wartości stałych normujących (por. np.: Newton i Raftery 1994; Polson i Roberts 1994; Kass i Raftery 1995). W pracy zastosowano aproksymację stałych normujących zaproponowaną przez Newtona i Raftery’ego.

Estymator Newtona i Raftery’ego (N-R)

Mając próbkę $\{\theta^{(k)}\}_{k=1}^N$ pochodzącą z rozkładu a posteriori, brzegową gęstość wektora obserwacji x można aproksymować poprzez średnią harmoniczną wartości funkcji wiarygodności:

$$p(x|M) \approx \left[\frac{1}{N} \sum_{k=1}^N \frac{1}{p(x|\theta^{(k)})} \right]^{-1}.$$

Należy podkreślić, że $p(x|\theta^{(k)})$ jest wartością gęstości, a nie jej jądra. Powyższy estymator Newtona i Raftery’ego (w skrócie N-R) jest zgodny, ale pozbawiony skończonej asymptotycznej wariancji, co w konsekwencji prowadzić może do jego dużej niestabilności. Ponadto jest on bardzo wrażliwy na małe wartości funkcji wiarygodności, co pociąga za sobą trudności numeryczne przy jego obliczaniu. Pomimo wspomnianych wad estymator N-R jest stosowany w praktyce przede wszystkim z powodu

swojej prostej formy oraz ze względu na fakt, że przy dużym N zachowuje się wystarczająco stabilnie, by można go było wykorzystać do porównania konkurencyjnych modeli.

W niniejszej pracy w podrozdziale 3.6 wykorzystano algorytm hybrydowy, a dokładniej losową wersję blokową algorytmu M-H, gdzie współrzędne wektora losowano za pomocą sekwencyjnego algorytmu M-H i procedury A-R. Za funkcję generującą przyjęto gęstość wielowymiarowego rozkładu normalnego, a do oceny zbieżności łańcucha Markowa wykorzystano statystykę CUMSUM. Natomiast wyniki prezentowane w rozdziałach 4 i 5 uzyskano, wykorzystując sekwencyjną wersję M-H. Gęstość wielowymiarowego rozkładu t -Studenta z trzema stopniami swobody przyjęto jako funkcję generującą, przy czym wartość oczekiwana i wariancja tego rozkładu oszacowane zostały za pomocą metody MC-IS. Funkcję ważności zdefiniowano również jako gęstość wielowymiarowego rozkładu t -Studenta o trzech stopniach swobody, przy czym za wartość oczekiwaną przyjęto punkt (lub jego numeryczną aproksymację) z przestrzeni parametrów, dla którego jądro gęstości a posteriori przyjmuje wartość największą, a za macierz kowariancji przyjęto macierz jednostkową. Jako wartość oczekiwaną funkcji generującej przyjmowano kolejne akceptowane stany łańcucha Markowa. Pierwszy stan łańcucha przyjmowano na poziomie wartości oczekiwanej a posteriori (uzyskanej za pomocą metody MC-IS). Ponadto w rozdziałach 4 i 5 dokonano porównania siły wyjaśniającej rozważanych modeli, wykorzystując estymator N-R, a do oceny zbieżności łańcucha Markowa wykorzystano statystykę CUMSUM. Przyjęte założenia są zgodne z zaleceniami podawanymi w literaturze przedmiotu.

3.6. Bayesowska estymacja parametrów przy zastosowaniu metody zmiennych pośrednich

W niniejszym podrozdziale zaprezentowane zostaną zasady bayesowskiej estymacji parametrów procesów Itô przy zastosowaniu metody zmiennych pośrednich do aproksymacji gęstości przejścia (por. podrozdz. 2.3). Aproksymację opartą na idei wygenerowania (odtworzenia) brakujących danych pomiędzy każdą parą obserwacji wykorzystano w zagadnieniach bayesowskiej estymacji parametrów wybranych SRR w następujących pracach: Eraker (1998), Jones (1998), Elerian (1998, 1999), Elerian, Chib i Shephard (2001), DiPietro (2001), Roberts i Stramer (2001) oraz Kostrzewski (2004).

Oprócz prezentacji samej metody, poniżej przedstawione zostaną wyniki estymacji parametrów wybranego procesu Itô uzyskane na podstawie danych symulowanych, będących dyskretnymi obserwacjami trajektorii logarytmu procesu pierwiastka kwadratowego. Problemy numeryczne rozwiązane zostały na podstawie pracy Elerian, Chib i Shephard (2001) za pomocą metod Monte Carlo łańcuchów Markowa. Ponadto przyjęto notację taką jak w podrozdziale 2.3.

Z własności Markowa silnych rozwiązań SRR gęstość a posteriori wektora parametrów można zapisać w następującej postaci:

$$\begin{aligned} p(\theta | x_T, \dots, x_1; x_0, \Delta) &\propto p(x_T, \dots, x_1 | x_0, \Delta, \theta) p(\theta) = \\ &= \prod_{t=0}^{T-1} p_X(x_{t+1} | x_t, t, \Delta, \theta) p(\theta) \approx \dots, \end{aligned}$$

a poprzez zastąpienie nieznanej gęstości p_X funkcją $p_X^{(M)}$ (por. wzory (2.2) i (2.3)) uzyskano aproksymację gęstości a posteriori:

$$\dots \approx \prod_{t=0}^{T-1} p_X^{(M)}(x_{t+1} | x_t, t, \Delta, \theta) p(\theta).$$

W ogólnym przypadku nie jest znana jawna postać funkcji $p_X^{(M)}(\cdot | x, t, \Delta, \theta)$. Przybliżanie jej wartości zgodnie ze wzorem (2.4) jest związane z dużym obciążeniem obliczeniowym algorytmów numerycznych towarzyszących wnioskowaniu bayesowskiemu. Rozwiązaniem tego problemu jest rozważenie łącznej gęstości a posteriori dla parametru θ i wektora danych pośrednich. Z własności Markowa można napisać:

$$\begin{aligned} p_Y^{(M)}(\theta, y^* | y) &\propto p_Y^{(M)}(\hat{y}_{\setminus 0} | y_0, \theta) p(\theta) = \\ &= \prod_{t=0}^{T-1} \prod_{j=0}^M p_Y(y_{t,j+1}^* | y_{t,j}^*, \tau_{t,j}^*, \Delta^*; \theta) p(\theta). \end{aligned}$$

Ostatecznie rozkład a posteriori dla parametrów θ można aproksymować w następujący sposób:

$$\begin{aligned} p_X(\theta | x_T, \dots, x_1, x_0) &\approx p_X^{(M)}(\theta | x_T, \dots, x_1, x_0) = \\ &= p_Y^{(M)}(\theta | y) = \int_{\mathbb{R}^{MT}} p_Y^{(M)}(\theta, y^* | y) dy^*. \end{aligned}$$

Do wnioskowania o charakterystykach brzegowych aproksymacji rozkładu a posteriori stosuje się metody MCMC (por. podrozdz. 3.5). Metody te generują łańcuch Markowa $\{\theta^j, y^{*j}\}$, który po odrzuceniu cykli spalonych jest skorelowaną próbą z rozkładu $p_Y^{(M)}(\theta, y^* | y)$. W ten sposób wygenerowane wartości θ^j są realizacjami

brzegowego rozkładu $p_Y^{(M)}(\theta | y)$. Łańcuch Markowa $\{\theta^j, y^{*j}\}$, $j=1, \dots, N$, otrzymuje się stosując próbnik Gibbsa, generując kolejno y^{*j} , a następnie θ^j . Dane pośrednie y^* są losowane z rozkładu warunkowego $y^* | y, \theta$ za pomocą algorytmu Metropolis-a i Hastingsa o losowej wielkości bloku. Losowa wielkość bloku ustalana jest z rozkładu Poissona (Po) z parametrem intensywności μ .

Schemat generowania łańcucha Markowa $\{\theta^j, y^{*j}\}_{j=1 \dots N}$

Dla $j = 1$ arbitralny wybór wartości początkowych $y^{*(1)}, \theta^{(1)}$.

Dla $j > 1$ dla każdego $t \in \{0, 1, \dots, T-1\}$:

1. $k = 0$.
2. Generowanie $m-1 \sim Po(\mu)$ oraz zastąpienie k przez $k+m$. Jeśli $k > M$, to należy przyjąć $k = M$.
3. Wygenerowanie punktów pośrednich z rozkładu warunkowego:

$$y_{t(k,m)}^{*j} | y_{t,k-1}^{*j}, y_{t,k+m}^{*j-1}, \theta^{j-1}$$

(porównaj poniżej: Generowanie punktów pośrednich). Jeśli $k < M$, to należy powrócić do punktu 2.

Wygenerowanie wektora parametrów z rozkładu warunkowego $\theta^j | y^{*j}, y$ (porównaj: Generowanie parametrów dyfuzji).

Generowanie punktów pośrednich

Przy założonym schemacie dyskretyzacji Eulera oraz na podstawie własności Markowa, gęstość, z której generuje się blok punktów pośrednich długości m , ma postać:

$$p(y_{t(k,m)}^* | y_{t,k-1}^*, y_{t,k+m}^*, \theta)$$

$$\propto \prod_{j=k-1}^{k+m-1} \varphi \left\{ y_{t,j+1}^* | y_{t,j}^* + a(\tau_{t,j}^*, y_{t,j}^*, \theta) \Delta, b^2(\tau_{t,j}^*, y_{t,j}^*, \theta) \Delta \right\},$$

przy czym $y_{t(k,m)}^*$ generuje się przy użyciu algorytmu M-H. Wymaga to określenia gęstości generującej z której będą generowane proponowane stany łańcucha Markowa, oraz określenia funkcji akceptacji $\alpha(y_{t(k,m)}^*, w | y_{t,k-1}^*, y_{t+m}^*, \theta)$. Na jej podstawie stwierdza się, czy zaproponowaną wartość można uznać za kolejny wyraz łańcucha. Zgodnie z sugestiami występującymi w literaturze przedmiotu, proponowaną gęstość q zdefiniowano jako gęstość wielowymiarowego rozkładu normalnego z wartością oczekiwaną równą modzie rozkładu zadanego przez gęstość $p(\cdot | y_{t,k-1}^*, y_{t+m}^*, \theta)$ oraz macierzy kowariancji równej odwrotności macierzy drugich pochodnych cząstkowych

funkcji $\ln(1/p(\cdot | y_{t,k-1}^*, y_{t+m}^*, \theta))$ obliczonych w punkcie równym wspomnianej modzie⁹⁾.

W celu wyznaczenia charakterystyk funkcji q przyjęto następujące oznaczenia:

- $w = (w_1, \dots, w_m)$ – propozycja realizacji wektora punktów pośrednich
- $y_{t(k,m)}^* = (y_{t,k}^*, y_{t,k+1}^*, \dots, y_{t,k+m-1}^*) = (w_1, \dots, w_m) = w$;
- $w_0 = y_{t,k-1}^*, w_{m+1} = y_{t,k+m}^*$;
- dla $j = -1, 0, \dots, m-1$ zdefiniowano:

$$a_j = a(w_{j+1}, \tau_{t,k+j}^*, \theta), \quad g_j = \left\{ b^2(w_{j+1}, \tau_{t,k+j}^*, \theta) \right\}^{-1},$$

$$d_j = w_{j+2} - (w_{j+1} + a_j \Delta), \quad c_j = 1 + a'_j \Delta;$$

$$- \quad u = \frac{\partial \ln p(w | w_0, w_{m+1}, \theta)}{\partial w} = (u_j)_{j=1, \dots, m},$$

$$- \quad V = - \frac{\partial^2 \ln p(w | w_0, w_{m+1}, \theta)}{\partial w \partial w'} = [V_{ij}]_{i,j=1, \dots, m}.$$

Wówczas można pokazać, że zachodzą następujące związki:

$$u_j = -\frac{1}{\Delta} \left(g_{j-2} d_{j-2} - g_{j-1} d_{j-1} c_{j-1} + \frac{1}{2} g'_{j-1} d_{j-1}^2 - \frac{\Delta g'_{j-1}}{2 g_{j-1}} \right),$$

$$V_{jj} = \frac{1}{\Delta} \left(g_{j-2} + g_{j-1} c_{j-1}^2 - g_{j-1} d_{j-1} a''_{j-1} \Delta - 2 g'_{j-1} d_{j-1} c_{j-1} + \frac{1}{2} g''_{j-1} d_{j-1}^2 - \frac{(g_{j-1} g''_{j-1} - (g'_{j-1})^2) \Delta}{2 g_{j-1}^2} \right),$$

$$V_{ij} = V_{ji} = \begin{cases} \frac{1}{\Delta} (g'_{j-1} d_{j-1} - g_{j-1} c_{j-1}), & i = j+1, \\ 0, & i > j+1. \end{cases}$$

⁹⁾ Innymi słowy, gęstość q równa się przekształceniu poprzez funkcję $(\exp(\cdot))$ rozwinięcia funkcji $\ln(p(\cdot | y_{t,k-1}^*, y_{t+m}^*, \theta))$ w szereg Taylora drugiego rzędu w otoczeniu mody rozkładu o gęstości $p(\cdot | y_{t,k-1}^*, y_{t+m}^*, \theta)$.

Ostatecznie:

$$q(y_{t(k,m)}^* | y_{t,k-1}^*, y_{t+m}^*, \theta) \sim N(\hat{w} + V^{-1}u, V^{-1}),$$

gdzie $\hat{w}$ oznacza wspomnianą wartość mody.

Natomiast funkcja akceptacji definiowana jest w następujący sposób:

$$\begin{aligned} \alpha(y_{t(k,m)}^{*(j-1)}, w | y_{t,k-1}^{*j}, y_{t+m}^{*(j-1)}, \theta) = \\ = \min \left\{ 1, \frac{p(w | y_{t,k-1}^*, y_{t,k+m}^*, \theta) q(y_{t(k,m)}^{*(j-1)} | y_{t,k-1}^*, y_{t,k+m}^*, \theta)}{p(y_{t(k,m)}^{*(j-1)} | y_{t,k-1}^*, y_{t,k+m}^*, \theta) q(w | y_{t,k-1}^*, y_{t,k+m}^*, \theta)} \right\}. \end{aligned}$$

Generowanie parametrów procesu dyfuzji

Niech wstępną wiedzę o wektorze parametrów wyraża iloczyn gęstości a priori każdego z parametrów $p(\theta) = p(\theta_1) \cdot \dots \cdot p(\theta_k)$. Realizacje wektora $\theta = (\theta_1, \dots, \theta_k)$ otrzymuje się z rozkładów brzegowych a posteriori:

$$\begin{aligned} p(\theta_i | \hat{y}, \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k) &\propto p_Y^{(M)}(\hat{y}_{\setminus 0} | y_0, \theta) p(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k) = \\ &= p_Y^{(M)}(\hat{y}_{\setminus 0} | y_0, \theta) p(\theta_i). \end{aligned}$$

W celu poprawienia efektywności generowania wektora θ jego składowe dzielone są na bloki i kolejno losowane. Sposób losowania uzależniony jest od konkretnego modelu i wyboru specyfikacji rozkładów a priori.

Przykład zastosowania metody

W celu ilustracji metody rozważono proces stochastyczny X , będący rozwiązaniem nieliniowego układu równań:

$$\begin{cases} dX_t = \left[\frac{\alpha}{\exp(X_t)} - \beta - \frac{\sigma^2}{2 \exp(X_t)} \right] dt + \frac{\sigma}{\sqrt{\exp(X_t)}} dW \\ X_0 = x_0 \end{cases}.$$

Interesujący jest związek wspomnianego procesu z procesem pierwiastka kwadratowego – dla procesu Z zdefiniowanego poprzez tożsamość $Z = \exp(X)$ można wykazać (ze wzoru Itô), że spełnia on następujący nieliniowy układ równań:

$$\begin{cases} dZ_t = (\alpha - \beta Z_t)dt + \sigma \sqrt{Z_t} dW \\ Z_0 = z_0 \end{cases}.$$

Zatem Z jest procesem pierwiastka kwadratowego, a powyższy układ wyznacza model CIR. Warto zauważyć, że funkcje dryfu i dyfuzji SRR definiujących procesy Z i X znacznie różnią się od siebie. Ponadto proces X przyjmuje wartości z całego zbioru liczb rzeczywistych, zaś proces Z wyłącznie ze zbioru liczb nieujemnych. Własność ta jest istotna, ze względu na fakt stosowania schematów dyskretyzacji przyjmujących wartości z całego zbioru liczb rzeczywistych do aproksymacji procesów Itô. Z tego właśnie powodu zagadnienie estymacji parametrów modelu CIR (opartej na schematach dyskretyzacji) w praktyce sprowadza się do estymacji parametrów procesu X .

Przyjęto, że estymowany wektor parametrów procesu ma postać $\theta = (\alpha, \beta, \sigma^2)$, gdzie $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$, $\sigma^2 > 0$. Do aproksymacji gęstości przejścia nowego procesu stosowana jest metoda zmiennych pośrednich. Niech Y oznacza proces uzyskany poprzez zastosowanie dyskretyzacji Eulera (por. wzór (1.16)) dla procesu X . Z własności Markowa zachodzi związek:

$$p(\hat{y}_0 | y_0; \Delta, \theta) = \prod_{t=1}^n p_Y(\hat{y}_t | \hat{y}_{t-1}; \Delta, \theta).$$

Gęstość przejścia dla procesu Y jest gęstością rozkładu normalnego (por. wzór (2.1)):

$$p_Y(\cdot | y; \Delta, \theta) = \varphi\left(\cdot | y + \left[\frac{\alpha}{\exp(y)} - \beta - \frac{\sigma^2}{2\exp(y)}\right]\Delta; \frac{\sigma^2}{\exp(X_t)}\Delta\right).$$

Zatem aproksymacja logarytmu gęstości próbkowej wynosi:

$$\begin{aligned} \ln(p(\hat{y}_0 | y_0; \Delta, \alpha, \beta, \sigma^2)) = & -\left(\frac{n-1}{2}\right)\ln(2\pi\Delta) - \left(\frac{n-1}{2}\right)\ln \sigma^2 + \frac{1}{2}\sum_{t=1}^n \hat{y}_{t-1} - \\ & - \frac{1}{2\sigma^2\Delta}\sum_{t=1}^n \left[\sqrt{\exp(\hat{y}_{t-1})} \left(\tilde{y}_t + \frac{\sigma^2\Delta}{2\exp(\hat{y}_{t-1})} - \hat{y}_{t-1} - \frac{\alpha\Delta}{\exp(\hat{y}_{t-1})} + \beta\Delta \right) \right]^2. \end{aligned}$$

Przez ψ oznaczono wektor parametrów $\psi = (\alpha, \beta)'$ oraz przyjęto założenie, że rozkładem a priori dla tego wektora jest dwuwymiarowy rozkład normalny:

$$p(\psi) \propto \exp\left(-\frac{1}{2}(\psi - \gamma_0)' \Gamma_0 (\psi - \gamma_0)\right), \text{ gdzie } \gamma_0 = (0, 1, 0, 1)', \Gamma_0 = \begin{bmatrix} 10^{-7} & 0 \\ 0 & 10^{-7} \end{bmatrix}.$$

Charakterystyki tego rozkładu odzwierciedlają dużą niepewność a priori o parametrach (czego wyrazem są ich duże wariancje) oraz brak zależności a priori pomiędzy α i β .

Ponieważ:

$$p(\psi | \hat{y}, \sigma^2) \propto p(\hat{y}_0 | y_0; \Delta, \psi, \sigma^2) p(\psi | \sigma^2) = p(\hat{y}_0 | y_0; \Delta, \psi, \sigma^2) p(\psi),$$

zatem można napisać:

$$\ln(p(\psi | \hat{y}, \sigma^2)) = c - \frac{1}{2\sigma^2 \Delta} (v - K\psi\Delta)' (v - K\psi\Delta) - \frac{1}{2} (\psi - \gamma_0)' \Gamma_0 (\psi - \gamma_0),$$

gdzie c jest pewną stałą, natomiast:

$$K = \begin{bmatrix} \exp(-0,5\hat{y}_0) & -\exp(0,5\hat{y}_0) \\ \vdots & \vdots \\ \exp(-0,5\hat{y}_n) & -\exp(0,5\hat{y}_n) \end{bmatrix}, \quad v = \begin{bmatrix} \exp(0,5\hat{y}_0) \left(\hat{y}_1 + \frac{\sigma^2 \Delta}{2\exp(\hat{y}_0)} - \hat{y}_0 \right) \\ \vdots \\ \exp(0,5\hat{y}_n) \left(\hat{y}_n + \frac{\sigma^2 \Delta}{2\exp(\hat{y}_n)} - \hat{y}_{n-1} \right) \end{bmatrix}.$$

Ostatecznie:

$$p(\psi | \hat{y}, \sigma^2) \sim N(V^{-1}u, V^{-1}), \text{ gdzie } u = \frac{K'v}{\sigma^2} + \Gamma_0 \gamma_0, V = \frac{K'K\Delta}{\sigma^2} + \Gamma_0.$$

Podążając za wskazówkami występującymi w literaturze przedmiotu, założono, że rozkładem a priori dla σ^2 jest odwrócony rozkład gamma $IG(a, b)$. W niniejszej części przyjęto rozkład informacyjny o wartości oczekiwanej i odchyleniu standardowym równym 0,03, tzn. $IG\left(3, \frac{3}{50}\right)$.

Gęstość tego rozkładu ma postać:

$$p(\sigma^2 | \psi) \propto (\sigma^2)^{-(\frac{p}{2}+1)} \exp\left(-\frac{S_0 p}{2\sigma^2}\right),$$

gdzie $p = 6, S_0 = \frac{1}{50}$.

Ponieważ:

$$p(\sigma^2 | \hat{y}, \psi) \propto p(\hat{y}_0 | y_0; \Delta, \psi, \sigma^2) p(\sigma^2 | \psi) = p(\hat{y}_0 | y_0; \Delta, \psi, \sigma^2) p(\sigma^2),$$

zatem:

$$\ln p(\sigma^2 | \hat{y}; \alpha, \beta) = c' - \xi_1 \ln(\sigma^2) - \xi_2 \sigma^{-2} - S_1 \sigma^2,$$

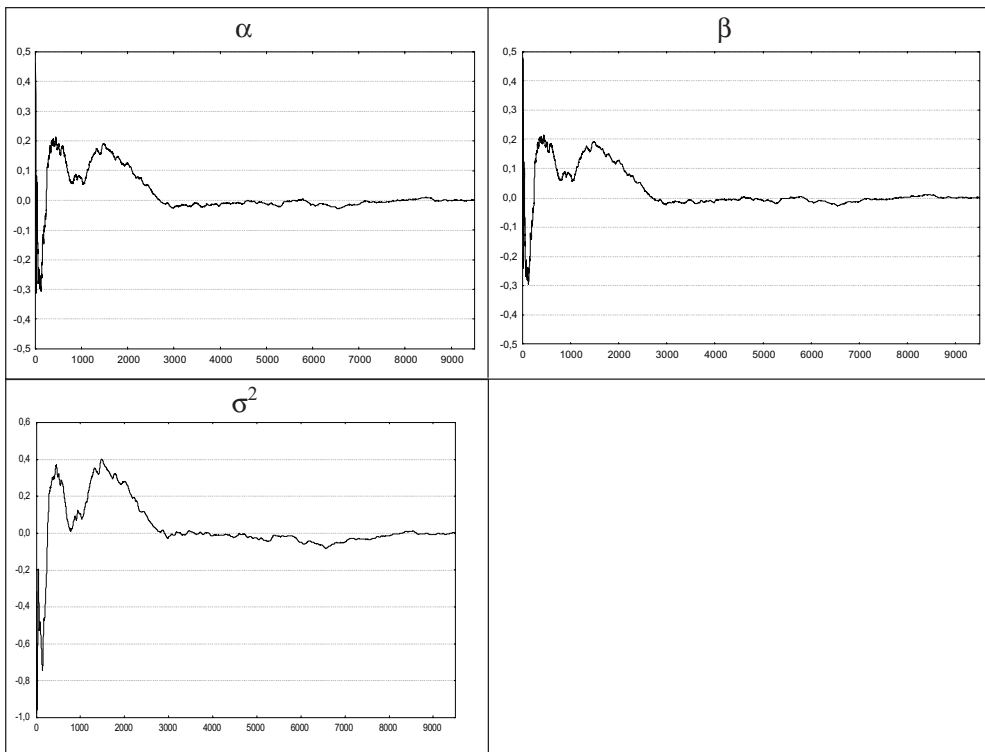
gdzie:

$$\xi_1 = \left(\frac{n+p+2}{2} \right), \quad \xi_2 = \frac{1}{2\Delta} \left(\sum_{t=1}^n \exp(\hat{y}_{t-1}) d_{t-1}^2 + S_0 p \Delta \right),$$

$$S_1 = \frac{\Delta}{8} \sum_{t=1}^n \exp(2\hat{y}_{t-1}), \quad d_{t-1} = \hat{y}_t - \hat{y}_{t-1} - \frac{\alpha \Delta}{\exp(\hat{y}_{t-1})} + \beta \Delta.$$

Losowania wartości z rozkładu o gęstości $p(\sigma^2 | \hat{y}, \psi)$ dokonano za pomocą metody A-R (por. podrozdz. 3.5), generując z rozkładu $IG(\xi_1 - p'(\hat{\lambda}) - 1, \xi_2)$, gdzie $p(\lambda) = -S_1 \exp(\lambda)$ oraz $\hat{\lambda}$ jest pewną ustaloną stałą. Prawdopodobieństwo akceptacji jest równe: $\exp(p(\lambda) - p(\hat{\lambda}) - p'(\hat{\lambda})(\lambda - \hat{\lambda}))$, gdzie $\lambda = \ln(\sigma^2)$. Stałą $\hat{\lambda}$ dobiera się tak, by prawdopodobieństwo akceptacji było jak największe. Okazuje się, że przy stałej $\hat{\lambda}$ równej modzie rozkładu zadanego przez gęstość $p(\sigma^2 | \hat{y}; \psi)$, prawdopodobieństwo to jest największe (por. Elerian 1999).

Aby zilustrować działanie metody zmiennych pośrednich jako narzędzia aproksymacji gęstości przejścia, wygenerowano dziesięć trajektorii procesu X (por. Posten 1989; Fraser, Wong i Wu 1998), o wielkości 500 elementów każda, przy ustalonych parametrach i zadanym warunku początkowym: $\alpha = 0,5, \beta = 0,2, \sigma^2 = 0,05, x_0 = 0$. Celem eksperymentu jest oszacowanie wielkości α, β i σ^2 przy zadanym wektorze obserwacji. Wielkość Δ jest interpretowana jako czas pomiędzy obserwowanymi (w tej sytuacji wygenerowanymi) punktami; M oznacza liczbę punktów pośrednich. Na potrzeby eksperymentu numerycznego przyjęto $\Delta = 1$. Parametr intensywności μ dla rozkładu Poissona ustalono w sposób arbitralny jako $\mu = 1$ (dla $M = 1$ oraz $M = 3$); $\mu = 3$ (dla $M = 5$ oraz $M = 10$); $\mu = 5$ dla ($M = 15$ oraz $M = 20$). Porównanie wyników dla różnej liczby iteracji oraz badanie wartości statystyki CUMSUM wykazały zbieżność łańcucha Markowa dla każdej rozważanej liczby punktów pośrednich. Wykresy statystyki dla wybranej trajektorii przedstawiono na rysunku 3.3, dla pozostałych ścieżek wykresy wyglądają podobnie i z tego względu nie są prezentowane w pracy. Wyniki ostateczne zostały osiągnięte po: 13 000 iteracji (dla $M = 1$ oraz $M = 3$); 40 000 iteracji (dla $M = 5, M = 10$ oraz $M = 15$); 51 000 iteracji (dla $M = 20$). Dla każdej wielkości M początkowych 1000 iteracji zostało odrzuconych jako cykle spalone. W tabeli 3.1 zaprezentowano uśrednione oszacowania wartości oczekiwanych i odchyłeń standardowych a posteriori estymowanych parametrów. Wielkości te zostały otrzymane poprzez uśrednienie wyników a posteriori dla każdej z dziesięciu trajektorii przy odpowiedniej liczbie punktów pośrednich.



Rys. 3.3. Wykresy statystyki CUMSUM dla wybranej trajektorii procesu
 Źródło: Opracowanie własne

Tabela 3.1

Uśrednione po 10 trajektoriach wartości oczekiwane (i odchylenia standardowe) a posteriori parametrów

M	α	β	σ^2
1	0,492 (0,068)	0,197 (0,028)	0,046 (0,0029)
3	0,506 (0,071)	0,2031 (0,029)	0,048 (0,0032)
5	0,508 (0,072)	0,2036 (0,03)	0,049 (0,0033)
10	0,515 (0,075)	0,206 (0,031)	0,05 (0,0034)
15	0,517 (0,075)	0,207 (0,031)	0,05 (0,0035)
20	0,517 (0,075)	0,207 (0,031)	0,05 (0,0035)

Źródło: Opracowanie własne

Przedstawione wyniki estymacji parametrów nie zależały od zadanych wartości początkowych. Zależały natomiast od wygenerowanej ścieżki (por. tab. 3.2) i stąd pomysł zbadania uśrednionych wyników. W miarę wzrostu M rosną prezentowane w tabeli 3.1 odchylenia standardowe. Jest to wynik włączenia do wnioskowania coraz większej liczby danych pośrednich, które są jedynie aproksymacją rzeczywistych danych. Zbieżność uśrednionych wartości oczekiwanych a posteriori dla parametru σ^2 do rzeczywistej jego wartości 0,05 uzyskano już przy $M = 10$. Natomiast dla $M = 15$ zaobserwowano zbieżność oszacowań parametrów α i β . Należy zatem zauważyć, że właściwy dobór wielkości M zależy nie tylko od konkretnej specyfikacji modelu, ale również od szacowanych parametrów. Ponadto zauważono znaczną poprawę uzyskiwanych wyników podczas zwiększenia dokładności aproksymacji mody rozkładu zadanego przez $p(\cdot | y_{t,k-1}^*, y_{t+m}^*, \theta)$ wykorzystywanej w algorytmie M-H. Miało to jednak duży wpływ na wydłużanie się czasu obliczeń. Procent akceptacji algorytmu M-H wynosił w przybliżeniu 90%. Fakt ten świadczy o właściwym dobraniu funkcji generującej.

Tabela 3.2

Wartości oczekiwane (i odchylenia standardowe) a posteriori parametrów
– uzyskane dla dwóch wybranych trajektorii

M	α		β		σ^2	
1	0,492 (0,068)	0,484 (0,067)	0,195 (0,028)	0,197 (0,028)	0,048 (0,0031)	0,047 (0,0030)
3	0,492 (0,069)	0,484 (0,069)	0,207 (0,030)	0,195 (0,029)	0,047 (0,0031)	0,050 (0,0034)
5	0,493 (0,070)	0,485 (0,070)	0,207 (0,030)	0,195 (0,029)	0,047 (0,0031)	0,051 (0,0034)
10	0,499 (0,072)	0,492 (0,072)	0,210 (0,031)	0,198 (0,030)	0,048 (0,0033)	0,052 (0,0036)
15	0,502 (0,073)	0,494 (0,073)	0,211 (0,031)	0,199 (0,030)	0,049 (0,0034)	0,053 (0,0037)
20	0,502 (0,073)	0,494 (0,073)	0,211 (0,031)	0,199 (0,030)	0,049 (0,0034)	0,053 (0,0036)

Źródło: Opracowanie własne

W niniejszym podrozdziale dokonano eksperymentu numerycznego dla $\Delta = 1$ w odróżnieniu od pracy Elerian, Chib i Shephard (2001), gdzie autorzy zaprezentowali oszacowania parametrów, przyjmując $\Delta = 5$. Prezentowana metoda danych pośrednich tworzy pomost pomiędzy procesami ciągłymi i ich dyskretnymi odpowiednikami. Z teoretycznego punktu widzenia wydaje się silnym narzędziem w rękach badacza. Jednakże ceną płaconą przy jej zastosowaniu jest ogromny wysiłek numeryczny. Obliczenia, których rezultaty przedstawiono powyżej, przeprowadzone na komputerze osobistym, zajęły kilkanaście dni. Duży czas obliczeń był bezpośrednią przyczyną nieprzeprowadzenia badania wrażliwości rozkładów a posteriori na dobór rozkładu a priori.

Mogłoby się wydawać, że dobór szybciej zbieżnego schematu dyskretyzacji niż specyfikacja Eulera może przyspieszyć obliczenia metody zmiennych pośrednich (por. Elerian 1998). Coraz bardziej wyrafinowane sposoby dyskretyzacji są dokładniejsze, jednak wymagają uwzględnienia np. większej liczby wyrazów rozwinięcia Taylora, a tym samym zwiększają liczbę wykonywanych obliczeń i ich skalę trudności. Z praktycznego punktu widzenia, metoda zmiennych pośrednich nie zdaje egzaminu, gdyż jest zbyt wolna. Natomiast jej udoskonalenia nie napawają większą nadzieją na przyspieszenie uzyskiwania wyników. Z tych powodów zaniechano stosowania metody w dalszych badaniach, których rezultaty prezentowane są w kolejnych rozdziałach monografii. Wydaje się, że podejście do zagadnienia aproksymacji gęstości przejścia poprzez zastosowanie wielomianów Hermite'a i metody opartej na rozwiązaniach równań FPK jest z punktu widzenia zastosowań bardziej korzystne. Zagadnieniom tym poświęcone będą następne dwa rozdziały pracy.

3.7. Uwagi końcowe

W niniejszym rozdziale zaprezentowano podstawy wnioskowania bayesowskiego. W szczególności omówiono zagadnienia: estymacji nieznanymi parametrów modeli, prognozy przyszłych wartości badanego szeregu czasowego oraz porównania konkurencyjnych specyfikacji. Ponadto zdefiniowanych zostało pięć modeli bayesowskich, na których koncentruje się wysiłek badawczy podjęty w niniejszej pracy. Duża część rozdziału poświęcona została opisowi wybranych metod numerycznych, bez których przeprowadzone w pracy wnioskowania bayesowskie byłyby niemożliwe do wykonania. W ostatniej części rozdziału przedstawiono wyniki bayesowskiej estymacji parametrów logarytmu procesu pierwiastka kwadratowego przy założeniu aproksymacji gęstości przejścia za pomocą metody zmiennych pośrednich. Wnioskowanie oparte zostało na danych symulowanych – trajektoriach procesu pierwiastka kwadratowego. Na podstawie przeprowadzonego doświadczenia wykazano małą efektywność tej metody.

4. Modelowanie stopy procentowej FED

4.1. Uwagi wstępne

W niniejszym rozdziale prezentowane są wyniki wnioskowania bayesowskiego przeprowadzonego na podstawie szeregu miesięcznych stóp procentowych FED. Szereg ten wykorzystany został między innymi w pracach Ait-Sahalia (1999) i DiPietro (2001) poświęconych estymacji parametrów wybranych procesów Itô. Warto zaznaczyć, że w pierwszej z cytowanych prac wykorzystano metodę największej wiarygodności, natomiast w drugiej zastosowano wnioskowanie bayesowskie. W obydwu pracach rozważono wyłącznie autonomiczne SRR, stosując wzór (2.12) do aproksymacji gęstości przejścia.

Obecnie zaprezentowane zostaną wyniki estymacji parametrów wybranych procesów Itô. Ponadto przedstawione zostaną rezultaty porównania siły wyjaśniającej konkurencyjnych modeli oraz prognozy dalszego poziomu stopy FED. Należy w tym miejscu wspomnieć, że M. DiPietro w rozprawie doktorskiej pod tytułem *Bayesian Inference for Discretely Sampled Diffusion Processes with Financial Applications* (por. DiPietro 2001) nie rozważał zagadnienia bayesowskiego porównania konkurencyjnych modeli oraz problemu bayesowskiej prognozy. Ponadto w niniejszej pracy aproksymację gęstości przejścia oparto na pewnego rodzaju uogólnieniu wzoru stosowanego w pracy DiPietro (2001), co było konieczne ze względu na rozpatrywanie modelu niejednorodnego EOU. Również metoda rozwiązań równań FPK wykorzystywana tutaj dla autonomicznych SRR nie była przedmiotem rozważań pracy DiPietro.

Rozważania rozdziału ograniczono do następujących modeli: V, BS, CIR, BSBS i EOU. Ponadto prezentowane poniżej wyniki otrzymano przy założeniu prawdziwej gęstości przejścia oraz jej aproksymacji. Dla modelu BS aproksymacja pokrywa się z prawdziwą funkcją, stąd ograniczenie prezentowanych wyników do przypadku rzeczywistego. Pomimo znajomości jawnej postaci gęstości przejścia w modelu CIR nie można wyliczyć jej bez zastosowania metod aproksymacji. W monografii zastosowano algorytm przybliżania wartości gęstości niecentralnego rozkładu χ^2 podany w pracy Abramowitz i Stegun (1968) – okazał się on zbieżny na całym zbiorze dopuszczalnych

wartości parametrów modelu. W ten sposób uzyskane rezultaty zostały nazwane prawdziwymi dla odróżnienia od wyników opartych na aproksymacji gęstości przejścia wyznaczonej na podstawie metody rozwiązań równań FPK. Dla modelu BSBS nieznana jest postać gęstości przejścia, stąd ograniczono wnioskowanie wyłącznie do przypadku aproksymacji.

Prezentowane rezultaty empiryczne uzyskano za pomocą programów napisanych przez autora w środowisku GAUSS 3.2. Algorytmy numeryczne zostały oparte na metodach Monte Carlo (MC) oraz łańcuchach Markowa MCMC. Zastosowano technikę MC z funkcją ważności (*Monte Carlo – Importance Sampling*, MC-IS) oraz algorytm Metropolisa i Hastingsa (M-H), a w szczególności próbnik Gibbsa i wersję sekwencyjną M-H (*sequential Metropolis Hastings algorithm*) (por. podrozdz. 3.5). Podążając za literaturą przedmiotu, za funkcję ważności (wykorzystaną w technikach MC-IS) i funkcję generującą (wykorzystaną w algorytmach M-H) przyjęto wielowymiarowy rozkład *t*-Studenta z trzema stopniami swobody, dla którego średnia i macierz kowariancji zostały ustalone po wstępnych przebiegach algorytmu MC-IS.

Długość wygenerowanego łańcucha Markowa, w tym również liczba cykli spalonych, ustalona była oddzielnie dla każdego modelu, a zależała od szybkości zbieżności algorytmu. Liczba cykli spalonych wahała się od 10 000, w przypadku modelu EOU (w wersji aproksymowanej), aż do 60 000, w przypadku modelu BSBS. Procent akceptacji algorytmu M-H wynosił od 45% (w przypadku modelu V) aż do 65% (w przypadku modelu BS). Stabilne zachowanie się współczynnika zmienności wagowej techniki MC-IS oraz wysoki procent akceptacji algorytmu M-H przemawiał za tym, że wybór rozkładu *t*-Studenta jako funkcji ważności i funkcji generującej był właściwy. Wysoki procent akceptacji pociągał za sobą szybką zbieżność algorytmów MCMC. Do oceny zbieżności algorytmu M-H wykorzystano statystykę CUMSUM (por. podrozdz. 3.5). Natomiast estymator średniej harmonicznej Newtona i Raftery’ego zastosowano do aproksymacji wartości $p(x | x_0; M_i)$ (por. podrozdz. 3.5), wykorzystywanej w celu porównania mocy wyjaśniającej rozważanych modeli. Należy dodać, że uzyskane wyniki empiryczne nie zależą od przyjętych wartości początkowych, ponadto nie zaobserwowano istotnej wrażliwości wyników na dobór rozkładów a priori.

Ponieważ przedmiotem analizy jest szereg danych miesięcznych, zatem parametr Δ odpowiadający odległości pomiędzy kolejnymi chwilami występowania danych jest wartością stałą, równą $\frac{1}{12}$ ($\Delta = 1$ odpowiada danym rocznym).

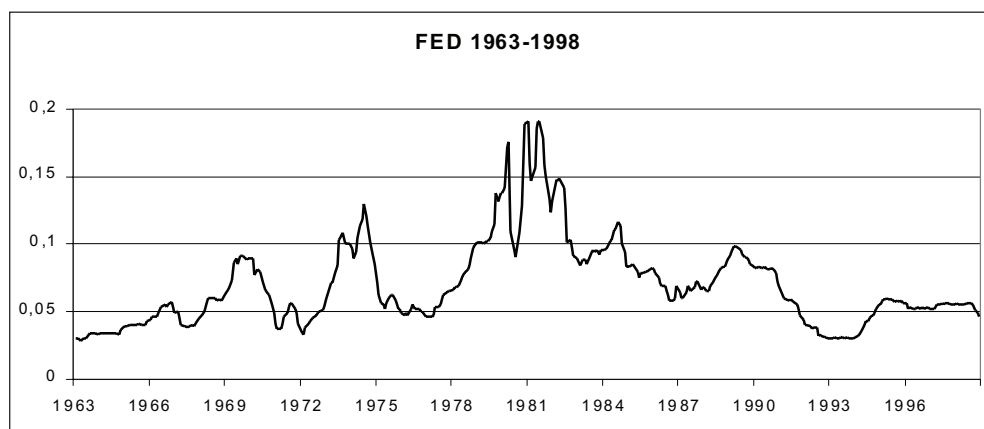
4.2. Prezentacja danych

W Stanach Zjednoczonych Ameryki Północnej prawo wymaga od banków i innych instytucji depozytowych gromadzenia stosownego kapitału (rezerwy obowiązkowej) na kontach tzw. Systemu Rezerwy Federalnej USA nazywanego również Amery-

kańskim Bankiem Centralnym. Nadwyżka ponad wymagany kapitał może być pożyczana innym instytucjom. Najwięcej obserwuje się transakcji jednodniowych O/N (*overnight*) względem dziennej stopy procentowej FED (*Federal Funds rate*). Jest to konsekwencja codziennego ustalania wymaganego kapitału (*minimum average balance*).

Stopa FED odgrywa istotną rolę w kształtowaniu innych amerykańskich stóp procentowych, zwłaszcza krótkoterminowych. Rozpatrywany w pracy szereg czasowy miesięcznych stóp procentowych FED (*monthly Federal Funds rates*) jest średnią ważoną dziennych wielkości. Dodatkowe informacje dotyczące tych danych można znaleźć między innymi w pracach Stigum (1990) i Baka (2001).

Szereg czasowy będący przedmiotem zainteresowania składa się z 432 obserwacji obejmujących okres od stycznia 1963 roku do grudnia 1998 roku¹⁾ (por. rys. 4.1).



Rys. 4.1. Miesięczne stopy procentowe FED
w okresie od stycznia 1963 roku do grudnia 1998 roku

Źródło: Opracowanie własne

Pierwsza obserwacja została wykorzystana jako warunek początkowy x_0 , a za moment jej obserwacji przyjęto $t_0 = 0$. Podstawowe statystyki opisowe modelowanych obserwacji oraz ich histogram przedstawiono w tabeli 4.1 i na rysunku 4.2. Współczynnik skośności wskazuje na asymetrię prawostronną rozkładu. Widoczne są liczne wzrosty, po których następują wyraźne spadki poziomu badanej stopy. Dane oscylują względem średniej w odległości do 3,8 odchylenia standardowego. Ponadto odległość pomiędzy najmniejszą i największą wartością wynosi pięć odchylenia standardowych.

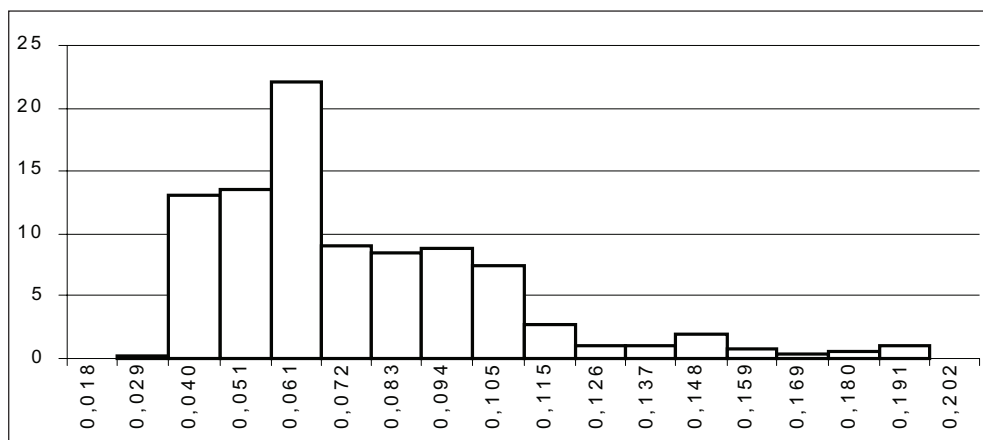
¹⁾ Dane pochodzą z internetowej strony www.federalreserve.gov/releases.

Tabela 4.1

Podstawowe statystyki opisowe dla szeregu czasowego FED

Średnia	Mediana	Odchylenie standardowe	Wariancja	Kurtoza ²⁾	Skośność	Minimum	Maksimum
0,070	0,060	0,032	0,001	5,130	1,360	0,029	0,191

Źródło: Opracowanie własne

**Rys. 4.2.** Histogram szeregu FED w okresie od stycznia 1963 roku do grudnia 1998 roku

Źródło: Opracowanie własne

4.3. Wyniki estymacji nieznanymi parametrów modeli

Wnioskowanie o nieznanymi wielkościach parametrów oparte zostało na metodzie wielomianów Hermite'a przy $m = 3$ (por. wzór (2.11)) w przypadku modelu EOU, natomiast dla modeli autonomicznych zastosowano metodę rozwiązań Fokkera, Plancka i Kołmogorowa z $K = 2$ (por. wzór (2.16)). Założenia te są zgodne z sugestiami występującymi w literaturze przedmiotu (por. Aït-Sahalia 2002; Egorov, Li i Xu 2003).

W tabeli 4.2 przedstawiono podstawowe charakterystyki rozkładów a posteriori parametrów modeli. Widoczna jest zgodność oszacowań parametrów na poziomie wartości oczekiwanej i odchylenia standardowego a posteriori w przypadku wnioskowania bayesowskiego opartego zarówno na aproksymowanej, jak i prawdziwej gęstości przejścia. Należy zwrócić uwagę na identyczną postać funkcji dyfuzji oraz podobne oszacowanie wartości średnich i odchyłeń standardowych parametru zmienności w przypadku modelu BS i BSBS.

²⁾ Współczynnik ekscesu (spłaszczenia) wyniósł 2,130.

Tabela 4.2

Wartości oczekiwane (i odchylenia standardowe) a posteriori parametrów rozpatrywanych modeli przy założeniu aproksymowanej i prawdziwej gęstości przejścia

Model	α	β	σ_1	σ^2
V Aproksymacja gęstości	0,01989 (0,00883)	0,27727 (0,11535)	-----	0,00053 (0,00004)
V Prawdziwa gęstość	0,01991 (0,00896)	0,27718 (0,11756)	-----	0,00053 (0,00004)
BS Prawdziwa gęstość	-----	0,03720 (0,03658)	-----	0,04812 (0,00329)
CIR Aproksymacja gęstości	0,01645 (0,00616)	0,22915 (0,09751)	-----	0,00448 (0,00031)
CIR Prawdziwa gęstość	0,01649 (0,00610)	0,22928 (0,09680)	-----	0,00449 (0,00031)
BSBS Aproksymacja gęstości	0,01150 (0,00499)	0,15974 (0,09296)	-----	0,04851 (0,00333)
EOU Aproksymacja gęstości	-----	0,05031 (0,05054)	-0,03575 (0,00704)	0,00182 (0,00049)
EOU Prawdziwa gęstość	-----	0,05501 (0,04413)	-0,03568 (0,00704)	0,00182 (0,00048)

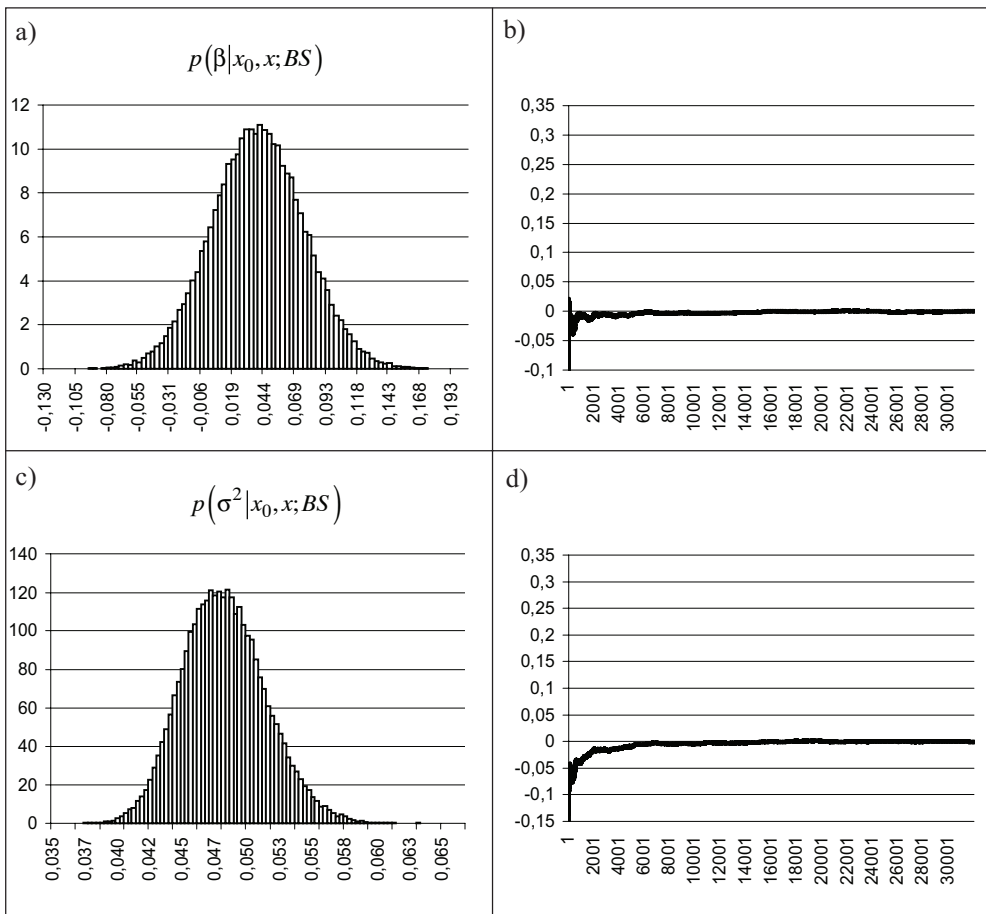
Źródło: Opracowanie własne

W przypadku modeli V i EOU wartość procesu nie ma wpływu na funkcję zmienności. W modelach tych widoczna jest również zgodność oszacowania charakterystyk parametru σ^2 . Zgodność w oszacowaniu parametrów modelu CIR w przypadku prawdziwej i aproksymowanej gęstości przejścia przemawia za stosowaniem metody rozwiązań równań FPK jako alternatywnej techniki oszacowania wartości gęstości niecentralnego rozkładu χ^2 .

Zbadano również wpływ unormowania aproksymacji gęstości przejścia na wyniki oszacowania parametrów. Stałe normujące wyliczono analitycznie lub oszacowano za pomocą metod numerycznych. Nie zaobserwowano jednak istotnych zmian wyników – z tego też powodu ich zestawienia nie przedstawiono w pracy.

Prezentowany eksperyment potwierdza sugestię o tym, że niewielkie wartości parametrów m i K wystarczają, by otrzymać zadowalające przybliżenie rozkładu a posteriori parametrów SRR.

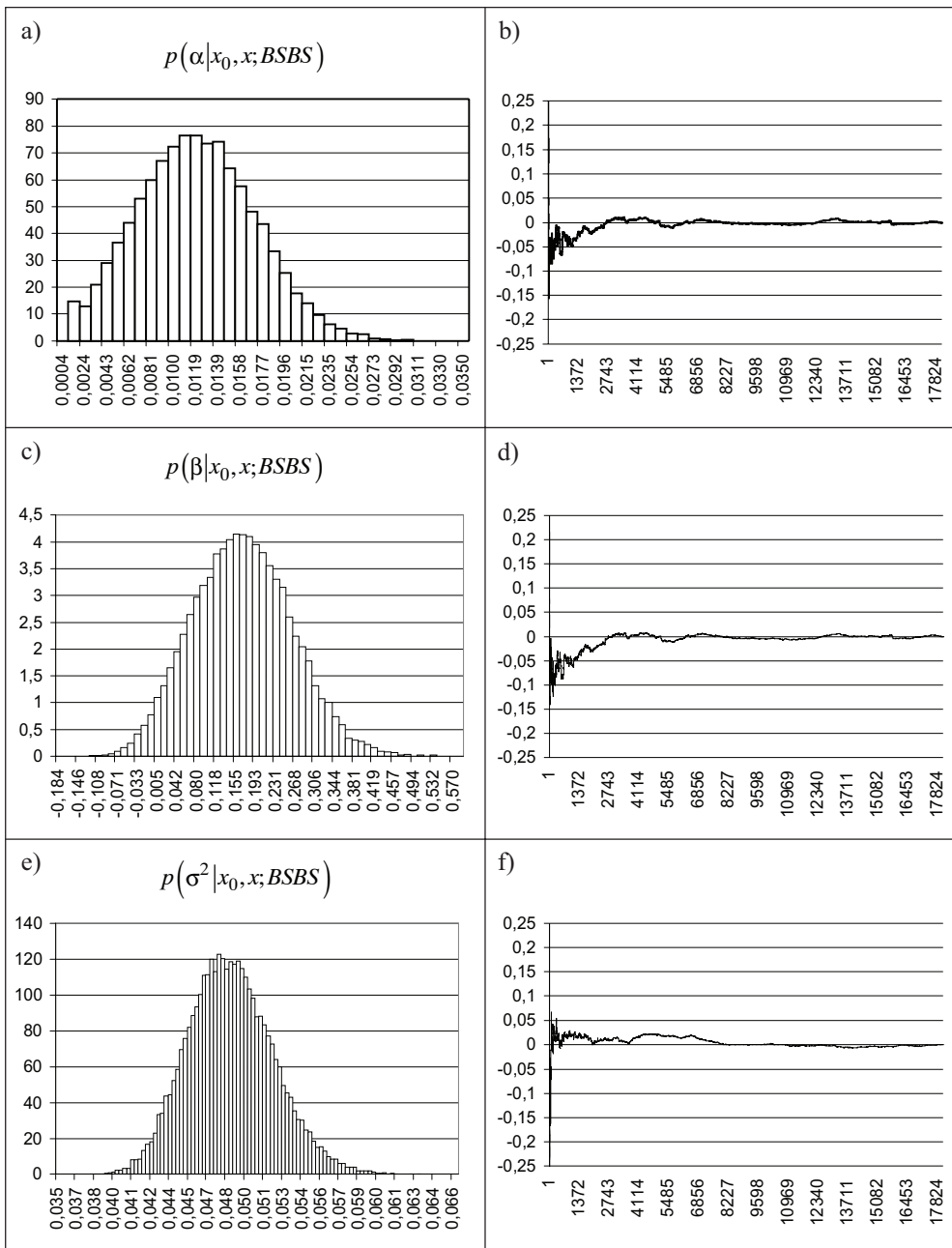
Na rysunkach 4.3 i 4.4 przedstawiono brzegowe rozkłady parametrów i wykresy statystyki CUMSUM dla modeli BS i BSBS. Wykresy skonstruowane zostały po uprzednim odrzuceniu cykli spalonych. Widoczna jest szybka zbieżność łańcucha Markowa do rozkładu stacjonarnego, tj. rozkładu a posteriori, oraz rozproszenie tego rozkładu. W przypadku innych modeli wykresy statystyki CUMSUM są podobne, tzn. równie szybko pozostają w przedziale $(-0,05; 0,05)$.



Rys. 4.3. Histogramy brzegowych rozkładów a posteriori parametrów (a, c) i wykresy (co 100.) wartości statystyki CUMSUM (b, d) po odrzuceniu cykli spalonych dla modelu BS.

Kolejne wiersze odpowiadają parametrom β i σ^2

Źródło: Opracowanie własne



Rys. 4.4. Histogramy brzegowych rozkładów a posteriori parametrów (a, c, e) i wykresy (co 100.) wartości statystyki CUSUM (b, d, f) po odrzuceniu cykli spalonych dla modelu BSBS.

Kolejne wiersze odpowiadają parametrom α , β i σ^2

Źródło: Opracowanie własne

Modelami rozważanymi wspólnie w pracach Aït-Sahalia (1999), DiPietro (2001) oraz prezentowanej monografii są V i CIR. Zatem jedynie w przypadku tych modeli możliwe jest porównanie uzyskanych wyników estymacji parametrów. W tabeli 4.3 przedstawiono oszacowania nieznanymi parametrów poprzez: podanie wartości estymatora największej wiarygodności, wyznaczonych w pracy Aït-Sahalia (1999), zaprezentowanie wartości oczekiwanych i median rozkładów a posteriori oraz błędów standardowych³⁾ uzyskanych w pracy DiPietro (2001), dodatkowo zestawiono wartości oczekiwane, odchylenia standardowe i mediany rozkładów a posteriori uzyskane za pomocą metod omówionych w monografii. Ponieważ w pracach Aït-Sahalia (1999) i DiPietro (2001) wektor nieznanymi parametrów był postaci $\theta = (\mu, \beta, \sigma)$, gdzie $\mu = \frac{\alpha}{\beta}$ oraz $\sigma = \sqrt{\sigma^2}$, zatem dla niego prezentowane są wyniki estymacji. Warto zauważyć, że wartości estymatorów największej wiarygodności znajdują się w bliskim sąsiedztwie median rozkładu a posteriori obliczonych w pracy DiPietro (2001) oraz wartości oczekiwanych i median a posteriori uzyskanych w niniejszej monografii. Wartości oczekiwane a posteriori parametrów dryfu, podane za pracą DiPietro (2001), są istotnie różne w porównaniu z pozostałymi charakterystykami przedstawionymi w tabeli 4.3. Jest to przyczyną rozważania odmiennych zbiorów parametrów dryfu – w pracy DiPietro (2001) przyjęto warunek $0 \leq \mu = \frac{\alpha}{\beta} \leq 1$, podczas gdy w pozostałych pracach założono $\mu = \frac{\alpha}{\beta} > 0$. Należy zwrócić uwagę na dużą zgodność punktowego oszacowania wartości parametru funkcji zmienności σ . Ponadto można zauważyć duże odchylenie standardowe rozkładu a posteriori wyznaczonego w monografii na podstawie dokładnej (nieasymptotycznej) analizy bayesowskiej i bardzo długich przebiegów MCMC. Wydaje się, że błędy standardowe podane przez DiPietro mogą być znacząco niedoszacowane.

W tabeli 4.4 zestawiono oceny wartości oczekiwanych z długiego horyzontu czasu dla tych rozpatrywanych w pracach procesów, dla których wystąpiło zjawisko powracania do średniej (por. podrozdz. 1.5)⁴⁾. Przy czym:

- w przypadku pracy Aït-Sahalia (1999) założono, że parametry równają się wartościom estymatorów największej wiarygodności;
- w przypadku pracy DiPietro (2001) przyjęto, że wartości parametrów to oszacowania median brzegowych rozkładów a posteriori;
- w przypadku niniejszej pracy przyjęto, że wartości oczekiwane z długiego horyzontu czasu równają się medianom rozkładów a posteriori zmiennej α/β .

³⁾ Błąd standardowy (*standard error*; *s.e.*) wyznaczony został na podstawie 10 łańcuchów Markowa o długości 10 000 każdy (por. DiPietro 2001).

⁴⁾ Ponieważ modele BS, BSBS i EOU nie były rozpatrywane w cytowanych pracach, stąd brak w zestawieniu stosownych wartości.

Tabela 4.3

Wartości estymatora największej wiarygodności, błędy standardowe oraz wartości oczekiwane, odchylenia standardowe i mediany rozkładów a posteriori parametrów dla modeli V i CIR

Model	Parametr	Aït-Sahalia (1999)			DiPietro (2001)			Niniejsza praca		
		Estymator największej wiarygodności	Wartość oczekiwana	Błąd standardowy	Mediana	Wartość oczekiwana	Odchylenie standardowe	Mediana		
V Aproksymacja gęstości	α/β	0,0717	0,099	0,02	0,074	0,079	0,33	0,072		
	β	0,261	0,192	0,007	0,185	0,277	0,115	0,277		
	σ	0,02237	0,0224	0,00001	0,0224	0,0229	0,00079	0,0229		
V Prawdziwa gęstość	α/β	0,0717	0,105	0,004	0,075	0,076	0,116	0,072		
	β	0,261	0,19	0,003	0,183	0,277	0,117	0,276		
	σ	0,02237	0,0224	0,000008	0,0224	0,0229	0,000792	0,0229		
CIR Aproksymacja gęstości	α/β	0,0742	0,132	0,04	0,079	0,080	0,09	0,072		
	β	0,189	0,157	0,01	0,148	0,229	0,10	0,227		
	σ	0,06658	0,0668	0,00005	0,0667	0,0670	0,00232	0,0669		
CIR Prawdziwa gęstość	α/β	0,0721	0,116	0,02	0,078	0,081	0,09	0,072		
	β	0,219	0,161	0,009	0,152	0,229	0,097	0,225		
	σ	0,06665	0,0668	0,00005	0,0667	0,0669	0,0023	0,0669		

Źródło: Opracowano na podstawie wyników własnych oraz rezultatów prac Aït-Sahalia (1999) i DiPietro (2001)

Tabela 4.4

Zestawienie wartości oczekiwanych z długiego horyzontu czasu

Model	Aït-Sahalia (1999)	DiPietro (2001)	Niniejsza praca
V Aproksymacja gęstości	0,072	0,074	0,072
V Prawdziwa gęstość	0,072	0,075	0,072
CIR Aproksymacja gęstości	0,074	0,079	0,072
CIR Prawdziwa gęstość	0,072	0,078	0,072

Źródło: Opracowanie na podstawie wyników własnych oraz prac Aït-Sahalia (1999) i DiPietro (2001)

Widoczna jest zgodność w oszacowaniu wartości oczekiwanych z długiego horyzontu czasu dla rozpatrywanych podejść. Na mocy przeprowadzonych w monografii badań wynika, że z prawdopodobieństwem a posteriori, równym 0,96, proces określony poprzez model BSBS ma własność powracania do średniej. Przy czym mediana indukowanego rozkładu a posteriori wartości oczekiwanej z długiego horyzontu czasu wynosi 0,071. Ponadto okazuje się, że z prawdopodobieństwem 0,85 geometryczny proces Browna (BS) nie ma własności powracania do średniej. Natomiast w przypadku modelu EOU, zarówno dla aproksymowanej, jak i prawdziwej gęstości przejścia, uogólniony proces Ornsteina i Uhlenbecka nie ma, z prawdopodobieństwem a posteriori równym 1, własności powracania do średniej. Warte podkreślenia jest identyczność uzyskanych w pracy ocen badanej charakterystyki dla modeli V, CIR i BSBS. Ponadto można zauważyć, że wartość średnia z próby 0,070 (por. tab. 4.1), traktowana w praktyce jako przybliżenie długoterminowego średniego poziomu badanego zjawiska, jest zbliżona do wartości oczekiwanych w długim horyzoncie czasu dla modeli V, CIR oraz BSBS.

4.4. Wyniki porównania modeli

W celu porównania siły wyjaśniającej konkurencyjnych specyfikacji procesów Itô założono identyczne prawdopodobieństwa a priori i obliczono prawdopodobieństwa a posteriori każdego z nich. W konsekwencji żaden z modeli nie był *a priori* bardziej preferowany. Przyjęte założenie jest uzasadnione małą liczbą parametrów rozważanych specyfikacji. Opisany w podrozdziale 3.5 estymator średniej harmonicznej Newtona i Raftery'ego (N-R) został wykorzystany do oszacowania brzegowej gęstości wektora obserwacji $p(x|M_i; x_0)$ (por. podrozdz. 3.3).

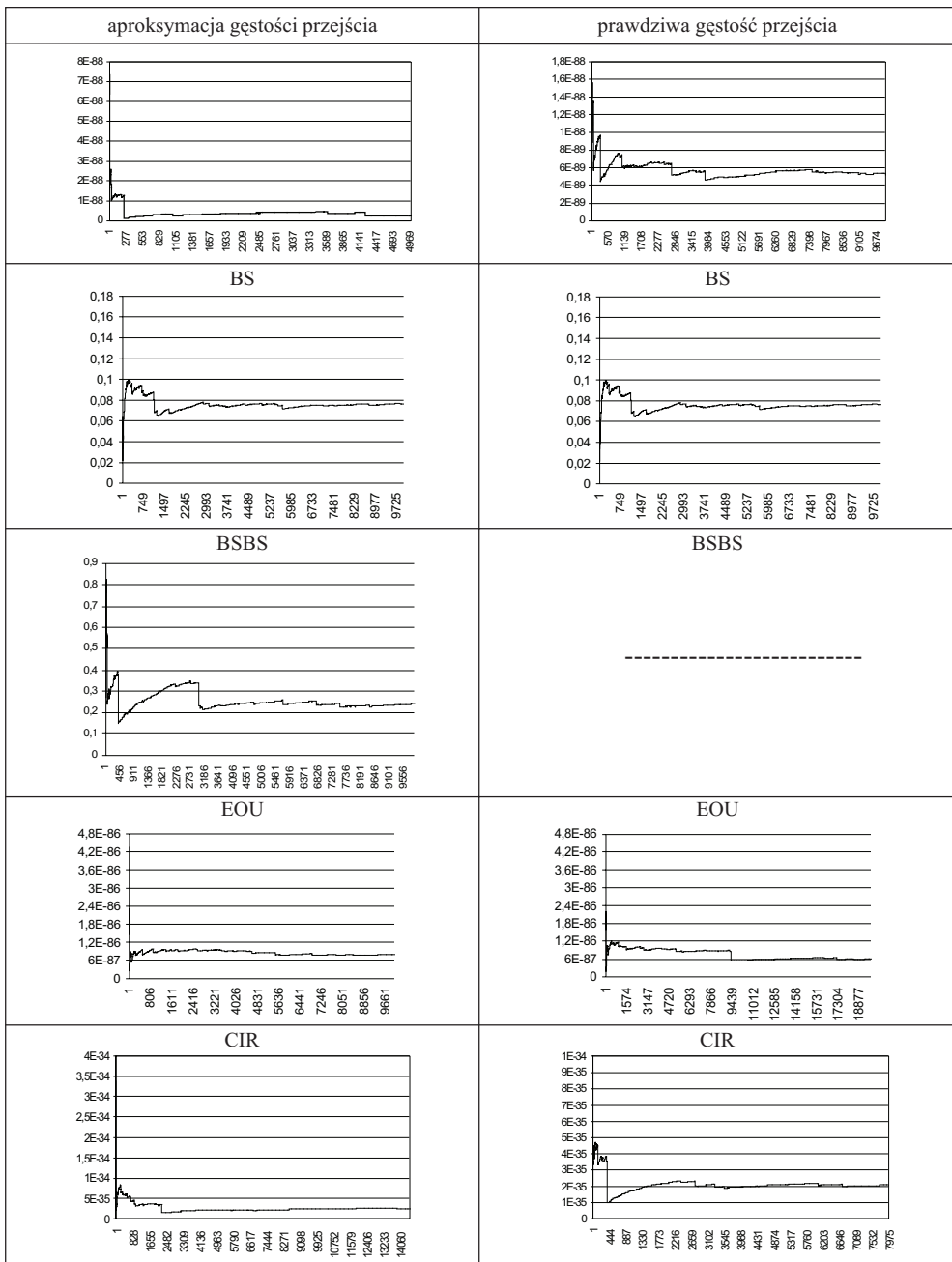
Tabela 4.5 przedstawia ranking konkurencyjnych modeli. W kolumnie drugiej (aproksymowana gęstość przejścia) zestawione zostały wartości brzegowych gęstości wektora obserwacji oraz prawdopodobieństwa a posteriori modeli wyliczone przy założeniu aproksymowanych gęstości przejścia, rezultaty kolumny trzeciej (prawdziwa gęstość przejścia) oparte zostały na prawdziwych gęstościach przejścia (w sytuacjach, w których taka analiza była możliwa). Natomiast ostatnią kolumnę tabeli (wyniki ostateczne) skonstruowano na podstawie prawdziwych gęstości w sytuacjach, kiedy było to możliwe, zaś w pozostałych przypadkach wykorzystano aproksymację funkcji. Model BS uzyskał najwyższe, bliskie jeden, prawdopodobieństwo a posteriori w zestawieniu wyników otrzymanych na bazie prawdziwych gęstości (tj. w sytuacji, gdy modelu BSBS nie brano pod uwagę). Sytuacja zmieniła się, gdy w rozważaniach uwzględniono aproksymację gęstości przejścia – wówczas model BSBS otrzymał niecałe 76% masy prawdopodobieństwa. Jednak model BS i w tym zestawieniu odgrywa istotną rolę, uzyskując ponad 24% masy prawdopodobieństwa. Modele V, CIR i EOU z powodu niskich prawdopodobieństw a posteriori uznane zostały w klasie rozpatrywanych modeli za nieadekwatne narzędzia opisu dynamiki rozważanego szeregu stóp procentowych. Z tego też powodu pominięto je w dalszych rozważaniach. Warto zwrócić uwagę na zbieżność oszacowań wartości brzegowych gęstości obserwacji w przypadku prawdziwej i aproksymowanej gęstości przejścia. Powyższe porównanie przeprowadzono przy założeniu równych prawdopodobieństw a priori modeli. Zmieniając strukturę a priori tak, by była zgodna z zasadą brzytwy Ockhama, tzn. przyznając modelowi BS prawdopodobieństwo a priori równe 1/3, a pozostałym czterem modelom po 1/6, ranking modeli nie zmienił się. Masy prawdopodobieństw skupione wówczas wokół modeli BS i BSBS wynosiły (w przybliżeniu) odpowiednio 0,39 i 0,61.

Tabela 4.5

Wartości brzegowej gęstości wektora obserwacji i prawdopodobieństwa a posteriori modeli

Model	Aproksymowana gęstość przejścia		Prawdziwa gęstość przejścia	Wyniki ostateczne	
	$\log_{10}(p(x x_0, M))$	$P(M x_0, x)$	$\log_{10}(p(x x_0, M))$	$\log_{10}(p(x x_0, M))$	$P(M x_0, x)$
V	677,96	3,38427E-88	678,26	678,26	6,77414E-88
BS	765,43	0,240170127	765,43	765,43	0,240170127
CIR	731,85	6,31336E-35	731,86	731,86	6,44147E-35
BSBS	765,93	0,759829873	-----	765,93	0,759829873
EOU	681,04	9,69142E-86	681,15	681,15	1,64592E-85

Źródło: Opracowanie własne



Rys. 4.5. Wykresy statystyk N-R (pomnożone przez $0,0166^{431}$) aproksymujących wartości $p(x|M_i; x_0)$. Na osi odciętych zaznaczono (co 500.) liczbę wykonanych iteracji po odrzuceniu cykli spalonych

Źródło: Opracowanie własne

Rysunek 4.5 przedstawia wykresy statystyk Newtona i Raftery’ego aproksymujące wielkości $p(x|M_i; x_0)$, które wykorzystano do oszacowania prawdopodobieństw a posteriori modeli. Wykresy pokazują zbieżność wspomnianych statystyk, wskazując jednocześnie wyższość modeli BS i BSBS nad pozostałymi specyfikacjami równań Itô. Wnioskowanie takie jest uprawnione, gdyż porównanie wielkości prezentowanych na rysunkach odpowiada określeniu wartości czynników Bayesa, co przy założeniu równych prawdopodobieństw a priori modeli prowadzi do porównania prawdopodobieństw a posteriori modeli.

Na koniec warto zauważyć, że modele V, BSBS, CIR i EOU mają identyczną postać funkcji dryfu, ale różne funkcje dyfuzji. Natomiast najbardziej prawdopodobne a posteriori modele BS i BSBS mają różne funkcje dryfu, ale takie same funkcje dyfuzji. Można zatem przypuszczać, że postać funkcji dyfuzji miała decydujący wpływ na ranking modeli.

4.5. Prognoza poziomu stopy procentowej FED

Niniejszy podrozdział ten przedstawia wyniki bayesowskiej prognozy nieuwzględnionych w estymacji („przyszłych”) wartości szeregu FED. Horyzont prognozy wynosi jeden, dwa i trzy kroki naprzód. Odpowiada to prognozie na jeden, dwa oraz trzy miesiące. Prognoza przeprowadzona została w dwóch wariantach. W pierwszym z nich predyktywne rozkłady wyznaczone zostały na podstawie modelu BS, gdyż okazał się on najlepszy w zestawieniu opartym na prawdziwych gęstościach przejścia. W wariancie drugim prognoza wyznaczona została poprzez zastosowanie metody bayesowskiego łączenia wiedzy (por. podrozdz. 3.3) zawężonej do modeli BS i BSBS. Podejście to jest uzasadnione, gdyż modele te skupiają niemalże całe prawdopodobieństwo a posteriori w zestawieniu opartym na prawdziwych i aproksymowanych gęstościach przejścia. Uśrednienie gęstości predyktywnych wykonane zostało przy założeniu, że wagi są równe prawdopodobieństwom a posteriori modeli, tj. $p(BS|x; x_0) = 0,240170127$, $p(BSBS|x; x_0) = 0,759829873$ (por. tab. 4.5).

W obydwu wariantach (tj. wnioskowania opartego wyłącznie na modelu BS oraz poprzez zastosowanie bayesowskiego łączenia wiedzy dla modeli BS i BSBS) zaobserwowane wartości stopy FED znalazły się w przedziałach wysokiego prawdopodobieństwa a posteriori – pomiędzy kwantylami rzędu 0,25 i 0,75 w bliskim sąsiedztwie median. Należy zwrócić również uwagę na mniejsze rozproszenie rozkładów (mierzone rozstępem wartości kwantyli) w przypadku bayesowskiego łączenia wiedzy dla modeli BS i BSBS w porównaniu z wnioskowaniem opartym wyłącznie na modelu BS. Formalne uwzględnienie niepewności co do właściwej specyfikacji modelu prowadzi do bardziej skupionego ostatecznego rozkładu predyktywnego, gdyż rozkład predyktyw-

ny w bayesowskim modelu BSBS jest podobnie zlokalizowany i bardziej skupiony niż w przypadku BS. Ponadto w obu przypadkach widoczny jest wzrost rozproszenia rozkładów predyktywnych wraz z wydłużaniem się horyzontu prognozy. Do identycznych wniosków prowadzi obserwacja histogramów rozkładów predyktywnych.

Tabela 4.6

Kwantyle predyktywnych rozkładów a posteriori poziomu stopy FED
skonstruowanych na podstawie modelu BS

Kwantyl rzędu	x_{T+1}	x_{T+2}	x_{T+3}
0,05	0,04220	0,04046	0,03918
0,25	0,04489	0,04415	0,04360
0,5	0,04685	0,04690	0,04695
0,75	0,04890	0,04983	0,05056
0,95	0,05201	0,05437	0,05626
Zaobserwowane wartości	0,0463	0,0476	0,0481

Źródło: Opracowanie własne

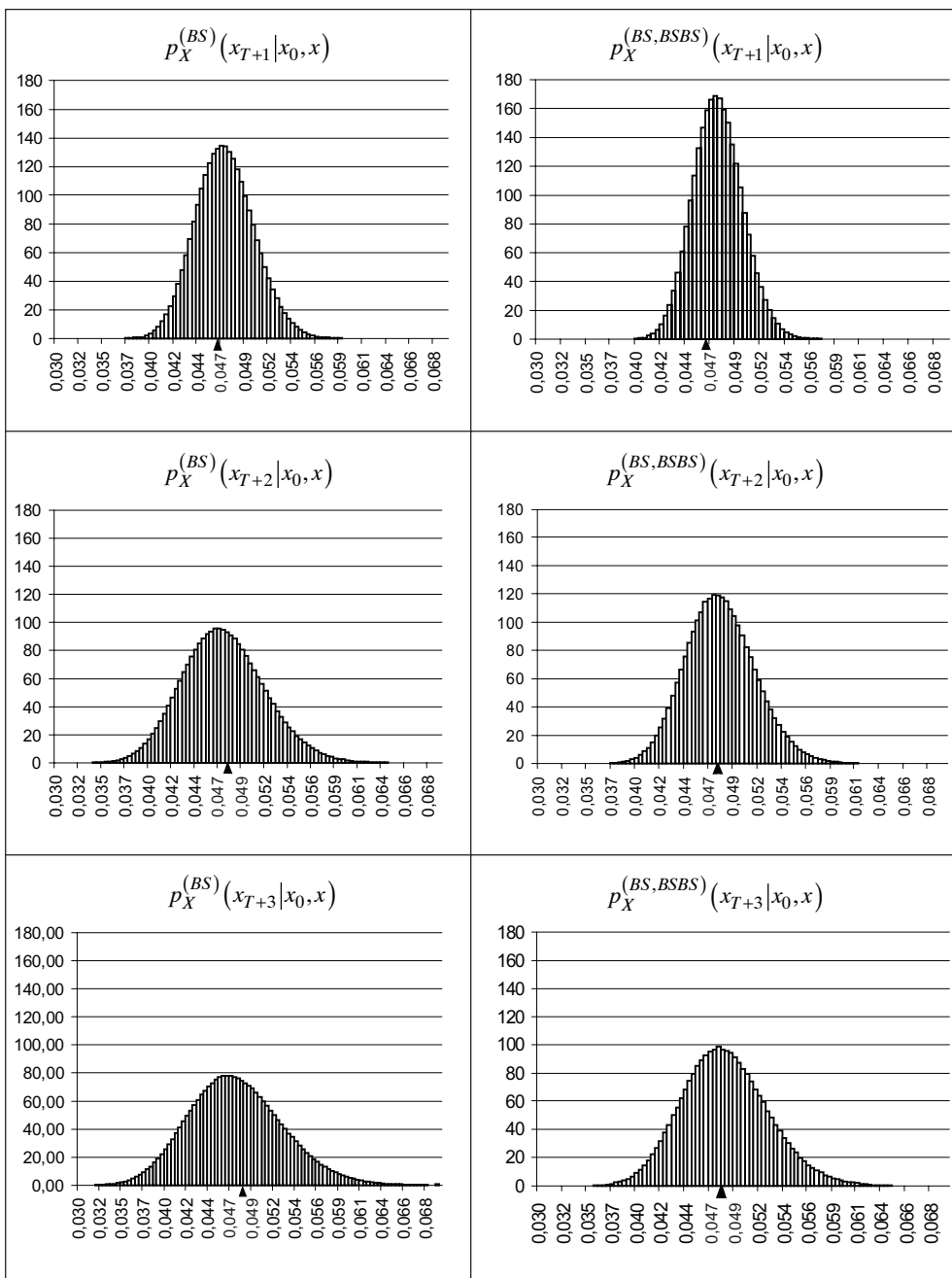
Tabela 4.7

Kwantyle predyktywnych rozkładów a posteriori poziomu stopy FED
skonstruowanych za pomocą metody bayesowskiego łączenia wiedzy
zastosowanej do modeli BS i BSBS

Kwantyl rzędu	x_{T+1}	x_{T+2}	x_{T+3}
0,05	0,04344	0,04235	0,04161
0,25	0,04559	0,04532	0,04519
0,5	0,04716	0,04752	0,04787
0,75	0,04879	0,04984	0,05074
0,95	0,05126	0,05342	0,05520
Zaobserwowane wartości	0,0463	0,0476	0,0481

Źródło: Opracowanie własne

Oszacowanie kwantyli rozkładów predyktywnych wyznaczonych w oparciu o model BS oraz modeli BS i BSBS przedstawiono w tabelach 4.6 i 4.7. Natomiast wykresy gęstości predyktywnych zestawiono na rysunku 4.6.



Rys. 4.6. Histogramy predyktywnych rozkładów poziomu stopy procentowej FED oraz rzeczywiste wartości oznaczone za pomocą trójkątów

Źródło: Opracowanie własne

4.6. Uwagi końcowe

W niniejszym rozdziale zaprezentowane zostały wyniki wnioskowania bayesowskiego przeprowadzonego na podstawie szeregu czasowego miesięcznej stopy procentowej FED.

Zgodność w oszacowaniu nieznanych parametrów w przypadku prawdziwej i aproksymowanej gęstości przejścia jest empirycznym dowodem zbieżności prezentowanych metod. Ponadto należy zwrócić uwagę na małą liczbę kroków w prezentowanych aproksymacjach, co w konsekwencji zwiększa szybkość algorytmów numerycznych, czyniąc przedstawiane metody efektywnym narzędziem w ręku badacza.

Interesującym wnioskiem rozdziału czwartego jest możliwość modelowania w sposób zadowalający dynamiki szeregu FED za pomocą prostych, gdyż zadanych poprzez liniowe stochastyczne równania różniczkowe, procesów dyfuzji. W szczególności mowa tutaj o powszechnie krytykowanym modelu Blacka i Scholesa. W praktyce do modelowania struktury czasowej stóp procentowych stosuje się między innymi modele CIR i V, które w rozważanym eksperymencie zostały odrzucone przez dane. Nieoczekiwanie model BS, który wykorzystywany jest do opisu wartości akcji, okazał się dobrym narzędziem modelowania rozważanego szeregu stopy procentowej FED. Nieco lepszym modelem okazał się model BSBS będący uogólnieniem specyfikacji BS poprzez stosowną zmianę funkcji dryfu. Warto przypomnieć, że dla modelu BSBS nieznana jest gęstość przejścia, co jest główną przyczyną trudności jego badania i zastosowania. Opisane w monografii metody aproksymacji umożliwiły przeprowadzenie wnioskowania bayesowskiego na podstawie modelu BSBS. Ponadto wyniki wnioskowania wykazały jego przydatność w opisie dynamiki miesięcznej stopy procentowej FED.

5. Modelowanie indeksu WIG20

5.1. Uwagi wstępne

Niniejsza, ostatnia część monografii poświęcona jest ilustracji empirycznej zastosowania procesów Itô w modelowaniu polskich finansowych szeregów czasowych. Rozpatrywany jest szereg dziennych notowań z zamknięcia sesji indeksu WIG20. Wnioskowanie bayesowskie wykorzystane zostało do oszacowania nieznanych wartości parametrów rozpatrywanych procesów, następnie dokonano porównania rozważanych modeli oraz przeprowadzono bayesowską prognozę i wycenę europejskiej opcji kupna na indeks WIG20. Ponadto porównano wyniki bayesowskiej prognozy poziomu indeksu WIG20 i wyceny europejskich opcji kupna na ten indeks z wykorzystaniem procesu EOU oraz dyskretnego procesu wariancji stochastycznej FCSV (*Fat-Tailed and Correlated Stochastic Variance*).

Prezentowane w niniejszym rozdziale rezultaty empiryczne uzyskane zostały za pomocą programów napisanych przez autora w środowisku GAUSS 3.2. Algorytmy numeryczne zostały oparte na metodach Monte Carlo (MC) oraz łańcuchach Markowa. Zastosowano technikę MC z funkcją ważności oraz algorytm Metropolisa i Hastingsa, a w szczególności próbnik Gibbsa i wersję sekwencyjną M-H (por. podrozdz. 3.5). Podobnie jak w poprzednim rozdziale, za funkcję ważności (wykorzystaną w technikach MC-IS) i funkcję generującą (wykorzystaną w algorytmach M-H) przyjęto wielowymiarowy rozkład t -Studenta z trzema stopniami swobody, dla którego średnia i macierz kowariancji zostały ustalone po wstępnych przebiegach algorytmu MC-IS.

Długość wygenerowanego łańcucha Markowa, w tym również liczba cykli spalonych, ustalona była oddzielnie dla każdego modelu, a zależała od szybkości zbieżności algorytmu. Liczba cykli spalonych wahała się od 8000 w przypadku modelu EOU (z prawdziwą gęstością przejścia) aż do 100 000 w przypadku modelu BSBS. Procent akceptacji algorytmu M-H był większy niż 37%, z wyjątkiem modelu BSBS, gdzie wyniósł on 13%. Stabilne zachowanie się współczynnika zmienności wagowej techniki MC-IS oraz dość wysoki procent akceptacji algorytmu M-H przemawia za tym, że wybór rozkładu t -Studenta był właściwy. Wysoki procent akceptacji pociągał za sobą

szybka zbieżność algorytmów MCMC¹⁾. Do oceny zbieżności algorytmu M-H wykorzystano statystykę CUMSUM (por. podrozdz. 3.5). Natomiast estymator średniej harmonicznej Newtona i Raftery'ego (por. podrozdz. 3.5) został wykorzystany do aproksymacji wartości $p(x|x_0, M_i)$ w celu porównania mocy wyjaśniającej rozważanych modeli. Należy dodać, że uzyskane wyniki empiryczne nie zależą od przyjętych wartości początkowych, ponadto nie zaobserwowano istotnej wrażliwości wyników na dobór rozkładów a priori.

Naturalną konsekwencją rozpatrywania w pracy procesów definiowanych w przedziale czasu, a nie w chwilach dyskretnych, jest „urealnienie” czasu. Za rok przyjęto 365 dni kalendarzowych, co jest odróżnieniem od 251 dni sesyjnych przyjmowanych w przypadku procesów dyskretnych. Przy wyznaczaniu wielkości Δ pod uwagę brano rzeczywistą liczbę dni, włączając w to na przykład dni wolne od pracy. W konsekwencji Δ jest wartością zmieniającą się w czasie, przyjmuje wartość $1/365$ lub jej wielokrotność. Założenie to przyjmowane jest również przez Giełdę Papierów Wartościowych w Warszawie (GPW), czego wyrazem jest sposób wyceny opcji na indeks WIG20, a dokładnie wyznaczanie kursu odniesienia odbywające się przy założeniu, że szereg czasowy wartości indeksu opisywany jest przez proces ciągły spełniający równanie modelu BS.

Pomimo znajomości gęstości przejścia w modelu CIR nie można wyliczyć jej wartości bez zastosowania metod aproksymacji. Z powodu problemów numerycznych, algorytm przybliżania wartości gęstości niecentralnego rozkładu χ^2 , wykorzystany w poprzednim rozdziale, okazał się tutaj niemożliwy do zastosowania na całym zbiorze dopuszczalnych wartości parametrów. Ostatecznie przedstawienie rezultatów odpowiadających modelowi CIR ograniczono do przypadku rozważanej w pracy aproksymacji z wykorzystaniem metody rozwiązań równań FPK.

Ze względu na problemy numeryczne badano przeskalowany szereg wartości indeksu WIG20, tzn. założono, że szereg czasowy podzielony przez arbitralnie wybrany skalar równy 2000 jest realizacją procesu Itô. Dla przeskalowanego szeregu przyjęto przedstawioną w podrozdziale 3.4 strukturę a priori. Uzyskane wyniki empiryczne przekształcono tak, by odpowiadały oryginalnemu, nieprzeskalowanemu szeregowi wartości WIG20 i w takiej formie je przedstawiono.

5.2. Prezentacja danych

Indeks WIG20 notowany jest przez Giełdę Papierów Wartościowych w Warszawie od 16 kwietnia 1994 roku. Jest obliczany na podstawie wartości portfela akcji 20 największych i najbardziej płynnych spółek z podstawowego rynku akcji. WIG20 jest indeksem typu cenowego, co oznacza, że przy jego obliczaniu bierze się pod uwagę jedynie ceny zawartych w nim transakcji, a nie uwzględnia się dochodów z akcji (dywi-

¹⁾ Czas oczekiwania na wygenerowanie aż 10 000 000 realizacji łańcucha Markowa na komputerze osobistym o wysokiej mocy obliczeniowej wahał się od 2 godzin w przypadku modelu BS do 19 godzin w przypadku modelu EOU dla aproksymowanej gęstości przejścia.

dend, praw poboru). W indeksie WIG20 nie mogą uczestniczyć fundusze inwestycyjne oraz więcej niż pięć spółek z jednego sektora giełdowego. Jego wartość podaje się poprzez liczbę punktów. Indeks WIG20 jest instrumentem bazowym dla kontraktów terminowych i opcji (por. *Rocznik...* 2005)²⁾.



Rys. 5.1. Dzielne notowania indeksu WIG20 (02.01.2001–31.12.2004)

Źródło: Opracowanie własne

Rozważany szereg czasowy składa się z 1005 obserwacji obejmujących okres od dnia 2 stycznia 2001 roku do dnia 31 grudnia 2004 roku (por. rys. 5.1). Pierwsza obserwacja została wykorzystana jako warunek początkowy x_0 , a za moment jej obserwacji przyjęto $t_0 = 0$. Podstawowe statystyki opisowe indeksu WIG20 oraz histogram przedstawiono w tabeli 5.1 i na rysunku 5.2. Analizowany szereg charakteryzuje się dużą liczbą spadków i wzrostów. Współczynnik skośności wskazuje na niewielką asymetrię prawostronną rozkładu; wartość kurtozy ujawnia niską koncentrację rozkładu badanego szeregu wokół wartości średniej. Dane oscylują względem średniej w odległości do 2,14 odchylenia standardowego. Ponadto odległość pomiędzy najmniejszą i największą wartością wynosi 3,8 odchylen standardowych.

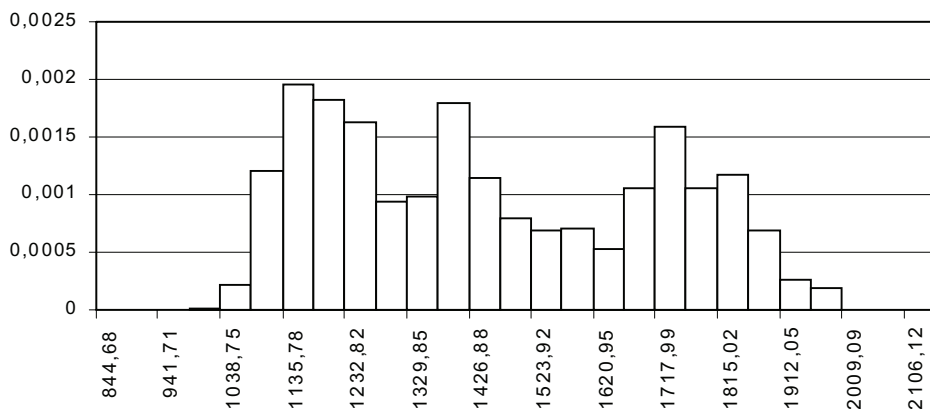
Tabela 5.1

Podstawowe statystyki opisowe dla szeregu danych WIG20

Średnia	Mediana	Odchylenie standardowe	Wariancja	Kurtoza	Skośność	Minimum	Maksimum
1414,46	1374,44	255,28	65167,98	1,76	0,27	990,23	1960,60

Źródło: Opracowanie własne

²⁾ Więcej informacji można znaleźć na stronie internetowej GPW – www.gpw.com.pl.



Rys. 5.2. Histogram rozważanego szeregu WIG20

Źródło: Opracowanie własne

5.3. Estymacja nieznanych parametrów

Wyniki prezentowanego wnioskowania oparte zostały na metodzie wielomianów Hermite'a z $m = 3$ (por. wzór (2.11)) zastosowanej dla modelu EOU, oraz metodzie rozwiązań równań FPK przy $K = 2$ (por. wzór (2.16)) w przypadku modeli autonomicznych. Założenia te są zgodne z sugestiami występującymi w literaturze przedmiotu (por. Aït-Sahalia 2002; Egorov, Li i Xu 2003).

Tabela 5.2 przedstawia wartości oczekiwane i odchylenia standardowe a posteriori parametrów rozpatrywanych modeli. Widoczna jest zgodność oszacowań parametrów w przypadku wnioskowania bayesowskiego opartego na aproksymowanej, jak i prawdziwej gęstości przejścia, potwierdzając zarazem, że małe wartości K oraz m są wystarczające. Porównanie oszacowań parametrów w przypadku modelu BS jest niepotrzebne, gdyż aproksymacja gęstości przejścia uzyskana za pomocą metody rozwiązań równań FPK pokrywa się z prawdziwą funkcją. Natomiast w przypadku modelu BSBS nie jest znana prawdziwa gęstość przejścia, zatem porównania nie można dokonać. Brak również w zestawieniu rezultatów osiągniętych przy założeniu prawdziwej gęstości w przypadku modelu CIR. Spowodowane jest to koniecznością zastosowania metod aproksymacji jej wartości na przyjętym dopuszczalnym zbiorze parametrów.

Parametr α w każdym z modeli, w którym występuje, ma wartość oczekiwaną a posteriori powyżej tysiąca z dość dużym odchyleniem standardowym. W przypadku modelu CIR wartość parametru β jest nieco ponad jeden, natomiast w pozostałych modelach nie przekracza jedności. Wartość oczekiwana a posteriori parametru σ_1 w przypadku modelu EOU jest ujemna, co powoduje, że czynnik $\exp(\sigma_1 t)$ jest mniejszy od jedności i zmniejsza tym samym, wraz ze wzrostem czasu, wartości funkcji dy-

fuzji³⁾. Można zauważyć, że w przypadku modeli V i EOU, gdzie wartość procesu w chwili t nie wpływa na funkcję zmienności, wartość oczekiwana i odchylenie standardowe a posteriori parametru σ^2 są bardzo duże. Dla procesu pierwiastka kwadratowego (CIR), gdzie w funkcji zmienności występuje pierwiastek z wartości procesu w chwili t , charakterystyki wspomnianego parametru są znacznie mniejsze i wynoszą odpowiednio ok. 93 i 4,16. W przypadku modeli BS i BSBS, tj. modeli, w których na funkcję zmienności ma wpływ wartość procesu w chwili t , wartość oczekiwana a posteriori tegoż parametru jest niewielka (ok. 0,069). Warto zauważyć, że w obydwu modelach postać funkcji zmienności jest taka sama. Należy również nadmienić, że wartość oczekiwana oraz mediana a posteriori parametru zmienności σ dla modelu BS wynoszą w przybliżeniu 0,262 – jednocześnie Krajowy Depozyt Papierów Wartościowych (KDPW) oszacował ten parametr na dzień 31 grudnia 2004 roku⁴⁾ na poziomie 0,192.

Tabela 5.2

Wartości oczekiwane (odchylenia standardowe) a posteriori parametrów rozważanych modeli dla prawdziwej i aproksymowanej gęstości przejścia

Model	α	β	σ_1	σ^2
V Aproksymacja gęstości	1348,4133 (803,48466)	0,9227 (0,55884)	-----	129868 (5820)
V Prawdziwa gęstość	1350,9942 (806,99854)	0,9241 (0,56121)	-----	129864 (6116)
BS	-----	0,0557 (0,13061)	-----	0,0687 (0,00307)
CIR Aproksymacja gęstości	1458,0426 (803,06854)	1,0054 (0,57880)	-----	93,2074 (4,15714)
BSBS Aproksymacja gęstości	1071,4118 (536,02668)	0,7261 (0,41385)	-----	0,0688 (0,00308)
EOU Aproksymacja gęstości	-----	-0,0615 (0,12083)	-0,0832 (0,01847)	177495 (15384)
EOU Prawdziwa gęstość	-----	-0,0620 (0,12198)	-0,0832 (0,01849)	177408 (15412)

Źródło: Opracowanie własne

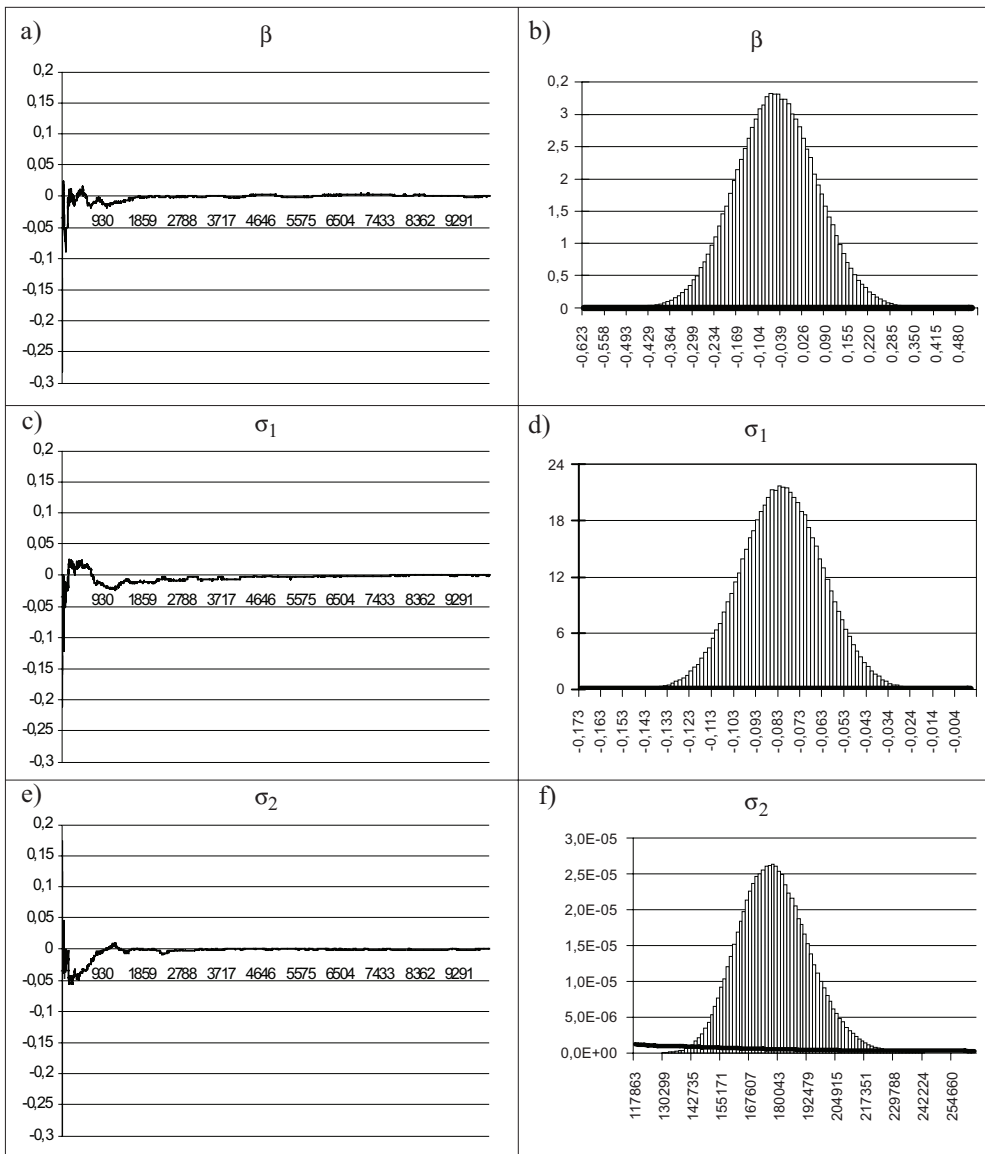
³⁾ Spadek wartości funkcji dyfuzji (zmienności) wraz ze wzrostem czasu nie musi pociągać za sobą spadku wariancji warunkowej procesu (tj. przy ustalonej poprzedniej obserwacji). Przyjmując za oszacowania punktowe parametrów modelu EOU prezentowane wartości oczekiwane a posteriori, okazuje się, że wspomniana wariancja rośnie do nieskończoności wraz ze wzrostem czasu. Własność ta nie jest wyłącznie cechą modelu EOU, posiadają ją także między innymi procesy dane następującymi równaniami: $dX_t = e^{-t} dW_t$, $dX_t = (1/\sqrt{t+1}) dW_t$.

⁴⁾ Porównaj: Informacja o wysokości depozytów zabezpieczających na 31 grudnia 2004 r. na podstawie komunikatu NR 393/DZ/2004 z dnia 31 grudnia 2004 r.

Wielkość tę Giełda Papierów Wartościowych wykorzystuje do wyliczenia kursu odniesienia zgodnego z modelem BS dla opcji na WIG20. Czytelnika mogą zastanawiać wartości niektórych wyestymowanych wielkości zawartych w tabeli 5.2, zwłaszcza pod kątem nałożonych restrykcji a priori. Dla przykładu wartość oczekiwana a posteriori parametru α w przypadku modelu BSBS jest większa niż 1071, podczas gdy rozkładem a priori dla α (por. podrozdz. 3.4) jest rozkład jednostajny w przedziale (0, 1). Wyjaśnieniem tej pozornej niezgodności jest transformacja wartości indeksu WIG20 do wartości indeksu podzielonych przez arbitralnie dobrany skalar 2000. Transformacja ta ma wpływ na wielkość parametrów. W szczególności dla modelu BSBS opisującego szereg „1/2000·WIG20” parametr α odpowiada zmiennej 2000α , gdy za pomocą BSBS wyjaśniana jest dynamika szeregu czasowego WIG20 (por. twierdzenie 1.3).

Na rysunkach 5.3a, c i e przedstawiono wykresy statystyki CUMSUM po cyklach spalonych dla parametrów β , σ_1 , σ^2 modelu EOU (dla prawdziwej gęstości przejścia). Wykresy wskazują na szybką zbieżność łańcucha Markowa do rozkładu stacjonarnego, tj. rozkładu a posteriori. Dla pozostałych modeli wykresy statystyki są podobne, tzn. równie szybko pozostają w przedziale $(-0,05; 0,05)$ – z tego też powodu nie są prezentowane w pracy. Na rysunkach 5.3b, d i f przedstawione zostały brzegowe rozkłady a priori (ściślej fragmenty gęstości) oraz a posteriori rozważanego wektora parametrów modelu EOU dla prawdziwej gęstości przejścia. Wykresy uwiadcniają brak asymetrii rozkładów oraz duże rozproszenie rozkładu parametru σ^2 . Ponadto ukazują one istotny wpływ danych na wyniki estymacji. Do podobnego wniosku prowadzi porównanie wartości oczekiwanych (oraz odchyłeń standardowych) a posteriori (por. tab. 5.2) oraz a priori. Dla parametru β omawiane charakterystyki a priori wynoszą 2522,5 (1906,44), a dla σ_1 są to 0,0084083 (5,7735). Ponieważ dla parametru σ^2 nie istnieją omawiane momenty a priori warto rozważyć zmienną losową σ^{-2} – w przypadku a posteriori wynoszą one 0,00000569 (0,00000049), natomiast dla rozkładu a priori równają się 0,00005 (0,00005). W każdym przypadku brzegowe rozkłady a posteriori parametrów modelu są bardziej skupione wokół średnich niż odpowiednie brzegowe rozkłady a priori (odchylenia standardowe a posteriori są niższe od odchyłeń standardowych a priori).

Rozważane w pracy modele V i CIR określają procesy stochastyczne o własności powracania do średniej równej α/β (por. podrozdz. 1.5). Mediana rozkładu a posteriori zmiennej losowej α/β w przypadku modelu V dla prawdziwej gęstości przejścia wynosi 1462,06, a dla aproksymowanej gęstości przejścia 1461,90. Dla modelu CIR mediana rozkładu a posteriori zmiennej losowej α/β równa się 1451,28. Ponadto proces określony poprzez model BSBS z prawdopodobieństwem równym ponad 0,96 ma własność powracania do średniej, przy czym mediana rozkładu indukowanego a posteriori tej charakterystyki wynosi 1464,56. Warto podkreślić podobieństwo oszacowania badanej charakterystyki za pomocą odmiennych specyfikacji bayesowskich. W przypadku modelu EOU, dla prawdziwej i aproksymowanej gęstości przejścia, z prawdopodobieństwem a posteriori równym jeden, proces nie ma własności powracania do średniej – jest to spodziewany rezultat, bowiem w przeciwnym razie wartość oczekiwana z długiego horyzontu czasu wynosiłaby zero, tzn. wartość trudną do zaakceptowania w rozważanym badaniu empirycznym.



Rys. 5.3. Wartości statystyki CUMSUM po cyklach spalonych dla wektora parametrów (β , σ_1 , σ^2) modelu EOU (na osi odciętych zaznaczono liczbę wykonanych iteracji podzieloną przez 100) (a, c, e); brzegowe rozkłady a priori (pogrubiona linia ciągła⁵⁾) oraz a posteriori (belki) rozważanego wektora parametrów (b, d, f)

Źródło: Opracowanie własne

⁵⁾ Ze względu na duże rozproszenie rozkładów a priori linie przedstawiające fragmenty ich gęstości niemal pokrywają się z osią odciętych.

Dla modelu BS proces ma, z prawdopodobieństwem równym 0,33, własność powracania do średniej równej zero. Jednak jak zostanie wykazane w następnym podrozdziale, model ten jest niewłaściwym narzędziem opisu dynamiki szeregu czasowego WIG20. Warto w tym momencie przypomnieć o wartości średniej z próby, która wynosi 1414,56 (por. tab. 5.1). Wartość tę w praktyce często przyjmuje się za przybliżenie średniej szeregu czasowego w długim horyzoncie czasu. Jest ona niższa niż wyżej przedstawione wielkości.

5.4. Porównanie modeli

Rozważając różne specyfikacje SRR opisujące to samo zjawisko, wydaje się naturalne zadanie pytania, która z nich najdokładniej (najwierniej) oddaje rzeczywistość. Wyborowi najlepszego spośród ujętych w niniejszej pracy modeli towarzyszy ranking wskazujący również na najgorsze specyfikacje (tj. najmniej prawdopodobne), które badacz może potraktować jako nieadekwatne.

W ujęciu bayesowskim, porównanie konkurencyjnych modeli odbywa się na podstawie prawdopodobieństw a posteriori rozważanych specyfikacji (por. wzór (3.1)).

Tabela 5.3

Brzegowe gęstości obserwacji i prawdopodobieństwa a posteriori modeli

Model	Aproksymowana gęstość przejścia		Prawdziwa gęstość przejścia	Wyniki ostateczne	
	$\log_{10}(p(x x_0,M))$	$P(M x_0,x)$	$\log_{10}(p(x x_0,M))$	$\log_{10}(p(x x_0,M))$	$P(M x_0,x)$
V	-1952,95	8,36110E-05	-1952,97	-1952,97	7,48733E-05
BS	-1958,87	1,01970E-10	-1958,87	-1958,87	9,53223E-11
CIR	-1952,79	1,22284E-04	-----	-1952,79	1,14312E-04
BSBS	-1958,52	2,25282E-10	-----	-1958,52	2,10594E-10
EOU	-1948,88	0,9997941	-1948,85	-1948,85	0,9998108

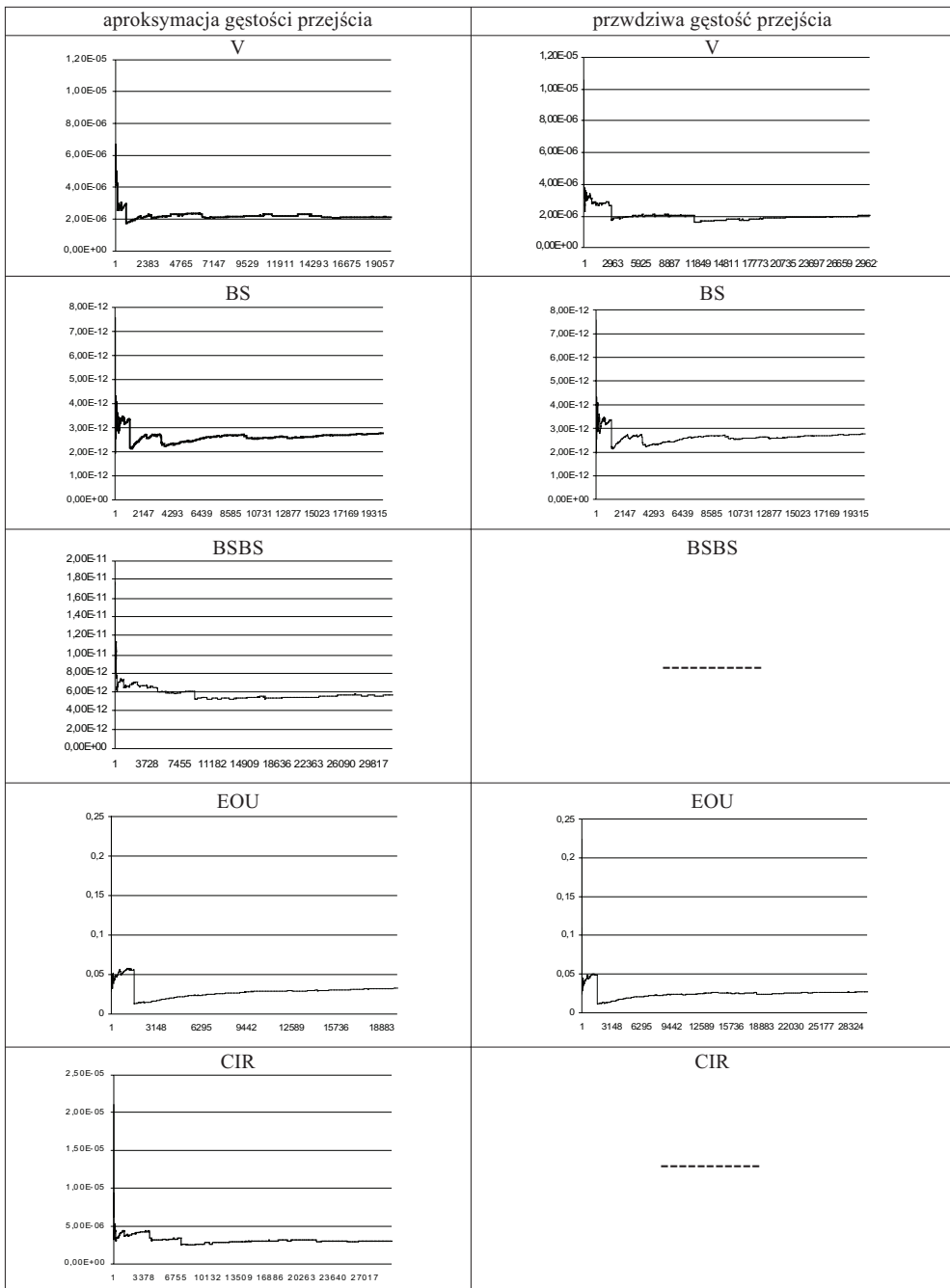
Źródło: Opracowanie własne

Tabela 5.3 przedstawia ranking konkurencyjnych modeli. W kolumnie drugiej (aproksymowana gęstość przejścia) zestawione zostały brzegowe gęstości obserwacji oraz prawdopodobieństwa a posteriori modeli wyliczone przy założeniu aproksymowa-

nych gęstości przejścia. Rezultaty kolumny trzeciej (prawdziwa gęstość przejścia) oparte zostały na prawdziwych gęstościach przejścia (w sytuacjach, w których taka analiza była możliwa). Natomiast ostatnią kolumnę tabeli (wyniki ostateczne) skonstruowano, opierając się na prawdziwych gęstościach w sytuacjach, kiedy było to możliwe, zaś w pozostałych przypadkach wykorzystano aproksymację funkcji.

Wyniki w każdym z zestawień przemawiają w sposób jednoznaczny za modelem EOU, który skupia ponad 99% masy prawdopodobieństwa. Warto zwrócić uwagę na zbieżność oszacowań wartości brzegowych gęstości obserwacji w przypadku prawdziwej i aproksymowanej gęstości przejścia. Należy zauważyć, że stochastyczne równania różniczkowe definiujące modele EOU, CIR, V i BSBS mają identyczną postać funkcji dryfu, natomiast różne funkcje dyfuzji. Postać funkcji zmienności miała zatem decydujący wpływ na wybór najlepszego modelu. W rankingu „wygrał” model, dla którego wartość procesu w chwili t nie wpływa na funkcję dyfuzji, natomiast ma na nią wpływ zmieniający się czas. Jest to o tyle ciekawe, że model ten należy do klasy nieautonomicznych stochastycznych równań różniczkowych, relatywnie rzadko spotykanych w rozważaniach teoretycznych i zastosowaniach. Wynik ten może być impulsem do częstszego stosowania modeli, w których funkcje dryfu i dyfuzji zależą nie tylko od samego procesu i jego parametrów, ale również od czasu. Kolejnymi w rankingu modeli okazały się modele CIR i V. Jednak nie stanowią one żadnej konkurencji dla modelu EOU w modelowaniu indeksu WIG20. Ostatnim w rankingu jest model BS, z niemalże zerowym prawdopodobieństwem a posteriori. Pomimo częstego stosowania (np.: GPW wyliczając kurs odniesienia dla opcji na WIG20, zakłada, że model ten opisuje dynamikę szeregu czasowego WIG20), został on w sposób jednoznaczny odrzucony na podstawie przeprowadzonej analizy. Powyższe porównanie przeprowadzono przy założeniu równych prawdopodobieństw a priori modeli. Zmieniając strukturę a priori tak, by była zgodna z zasadą brzytwy Ockhama, tzn. przyznając modelowi BS prawdopodobieństwo a priori równe $1/3$, a pozostałym czterem modelom po $1/6$, ranking modeli nie zmienił się, natomiast prawdopodobieństwa a posteriori modeli zmieniły się w sposób nieznaczny (np. $P(BS|x_0, x) \approx 1,5E-10$). Modele V, BS, CIR i BSBS z powodu niskich prawdopodobieństw a posteriori uznane zostały, w klasie rozważanych modeli, za nieadekwatne narzędzia opisu dynamiki rozważanego szeregu WIG20; pominęto je w dalszych rozważaniach.

Rysunek 5.4 przedstawia wykresy wartości statystyki Newtona i Raftery’ego aproksymujących wielkości $p(x|M_i; x_0)$, którą wykorzystano do oszacowania prawdopodobieństw a posteriori modeli. Wykresy pokazują zbieżność tej statystyki dla każdego z modeli, wykazując jednocześnie wyższość modelu EOU nad pozostałymi specyfikacjami. Wnioskowanie takie jest uprawnione, gdyż porównanie wielkości prezentowanych na rysunkach odpowiada określeniu wartości czynników Bayesa, co przy założeniu równych prawdopodobieństw a priori modeli prowadzi do porównania prawdopodobieństw a posteriori modeli.



Rys. 5.4. Wartości statystyk N-R (pomnożone przez 87^{1004}) aproksymujących wartości $p(x|M_i;x_0)$. Na osi odciętych zaznaczono (co 500.) wartość statystyki

Źródło: Opracowanie własne

5.5. Prognoza poziomu indeksu WIG20

Na mocy wyników prezentowanych w podrozdziale 5.4 uzasadnione jest przyjęcie modelu EOU jako najlepiej opisującego dynamikę indeksu WIG20. W niniejszej części pracy ograniczono wyznaczenie rozkładów predyktywnych przyszłych poziomów rozważanego indeksu do modelu EOU z prawdziwą gęstością przejścia. Prognozą objęte zostały dwa momenty czasu: 18 marca 2005 roku i 17 czerwca 2005 roku. Warto w tym miejscu przypomnieć, że ostatnia z obserwacji, na podstawie których dokonano prognozy, pojawiła się dnia 31 grudnia 2004 roku (tj. w momencie T). W konsekwencji horyzont prognozy wyniósł 77 i 168 dni kalendarzowych. Zatem czas S , jaki obejmuje prognoza – wyrażony w skali roku – wynosi 77/365 lub 168/365. Wybrane daty pokryły się z momentami wygaśnięcia europejskich opcji kupna oznaczonymi symbolami OW20C5180, OW20C5210, OW20F5180 i OW20F5210, których wycena na podstawie predyktywnych rozkładów a posteriori stanowić będzie temat następnego podrozdziału. W tabeli 5.4 zostały przedstawione kwantyle rozkładów predyktywnych poziomu indeksu WIG20. Wartości rzeczywiste indeksu znajdują się między kwantylami 0,25 i 0,75 – w bardzo bliskim sąsiedztwie median. Widoczna jest duża rozpiętość kwantyli wzrastająca wraz z wydłużaniem się horyzontu prognozy – przemawia ona za dużą niepewnością prognozy. W konsekwencji prognozy punktowe (np. na poziomie mediany) obarczone są dużym błędem ex ante, znacznie przewyższającym błąd liczony ex post.

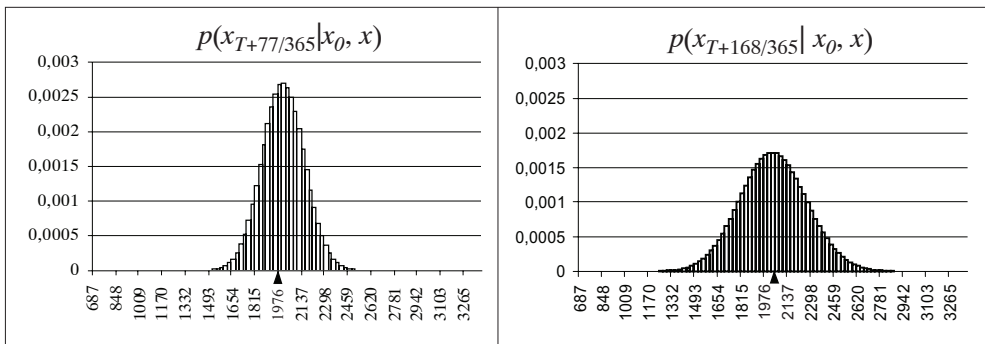
Tabela 5.4

Kwantyle predyktywnych rozkładów poziomów indeksu WIG20 otrzymane na podstawie modelu EOU

Kwantyl rzędu	$x_{T+77/365}$	$x_{T+168/365}$
0,05	1745,13	1641,52
0,25	1887,44	1861,87
0,5	1986,55	2017,82
0,75	2086,15	2176,19
0,95	2230,47	2409,20
Różnica międzykwantylowa	198,71	314,32
Rzeczywiste wartości	1970,17	2029,99

Źródło: Opracowanie własne

Rysunek 5.5 przedstawia histogramy predyktywnych rozkładów a posteriori. Rzeczywiste wartości indeksu (oznaczone na wykresach za pomocą trójkątów) znajdują się w przedziałach dużego prawdopodobieństwa. Widoczny jest wzrost rozproszenia rozkładu wraz ze zwiększaniem się horyzontu prognozy.



Rys. 5.5. Gęstości predyktywne rozkładów poziomu indeksu WIG20 oraz rzeczywiste wartości indeksu oznaczone za pomocą trójkątów

Źródło: Opracowanie własne

Należy przypomnieć o podstawowej wadzie modelu EOU w zagadnieniach modelowania nieujemnych szeregów czasowych, a mianowicie proces będący rozwiązaniem SRR definiującego EOU z niezerowym prawdopodobieństwem przyjmuje wartości ujemne (por. podrozdz. 1.5). Warto więc zbadać, jaka jest wartość takiego prawdopodobieństwa. W tym celu przyjęto, że wartości nieznanymi parametrów modelu EOU równają się ich wartościom oczekiwany a posteriori (por. tab. 5.2). Wówczas prawdopodobieństwo przyjęcia przez proces wartości ujemnych w chwilach czasu 77/365 i 168/365 wynosi odpowiednio 3,8291431E-108 i 3,7883303E-052. Przyjęcie przez proces wartości ujemnej jest zatem wysoce nieprawdopodobne. Z punktu widzenia praktycznego można nawet powiedzieć niemożliwe. Ponadto dla rozkładu predyktywnego prawdopodobieństwo przyjęcia przez proces wartości dodatnich dla momentów czasu objętych prognozą jest bardzo bliskie jedności.

5.6. Wycena europejskich opcji kupna

Interesującym i ważnym problemem poruszonym we współczesnych zagadnieniach finansowych jest wycena instrumentów pochodnych. Często w literaturze i praktyce zakłada się, że obserwowany szereg czasowy jest realizacją pewnego procesu stochastycznego, a w szczególności procesu Itô. W kontekście takiego podejścia w podrozdziale tym zostaną wycenione europejskie opcje kupna (*European call options*), dla których instrumentem podstawowym (bazowym) jest indeks WIG20.

Europejska opcja kupna jest kontraktem dającym jej nabywcy prawo do zakupu umówionej wcześniej liczby jednostek instrumentu podstawowego w określonym momencie czasu $T + S$ (*maturity, expiration date*) po ustalonej z góry cenie K (*strike, exercise price*). Posiadacz opcji ma prawo, ale nie obowiązek, jej wykonania – w konsekwencji maksymalna strata nabywcy opcji jest z góry znana i równa zapłaconej w chwili zakupu cenie, czyli tzw. premii (*premium*). Wartość opcji w chwili wygaśnięcia określa zmienna losowa (nazywana funkcją wypłaty) $(x_{T+S} - K)^+ = \max\{x_{T+S} - K, 0\}$, gdzie

x_{T+S} jest wartością instrumentu bazowego w chwili wygaśnięcia opcji. Opcje wykorzystywane są zarówno do zabezpieczenia inwestycji finansowych, jak i do celów spekulacyjnych.

W literaturze przedmiotu znane są różne podejścia do zagadnienia wyceny opcji. Najbardziej znane jest tzw. podejście martyngałowe, inaczej zwane wyceną arbitrażową (*arbitrage pricing*). Jego istota polega na przyjęciu założeń gwarantujących istnienie portfela oraz odpowiadającej mu strategii replikującej instrument pochodny, tzn. opiera się na założeniu rynku zupełnego⁶⁾ (*complete market*). Wówczas za cenę opcji przyjmuje się wartość wspomnianego portfela w chwili wyceny. W takiej sytuacji można pokazać, że cena europejskiej opcji kupna w chwili T wynosi $E_{P^*} \left(e^{-rS} (x_{T+S} - K)^+ | \mathfrak{F}_T \right)$ (*risk-neutral valuation formula*), gdzie P^* jest miarą martyngałową. Należy w tym miejscu wspomnieć o pracy F. Blacka i M. Scholesa opublikowanej w 1973 roku, którzy przy dość silnych założeniach wyznaczyli cenę europejskiej opcji na akcję nie wypłacającą dywidendy.

Przyjęli oni następujące warunki:

- dynamika instrumentu bazowego opisywana jest poprzez geometryczny ruch Browna (por. wzór (1.23)) przy czym parametry β i σ nie zmieniają się;
- nie ma kosztów transakcji oraz podatków;
- nie istnieje możliwość arbitrażu;
- papiery wartościowe są doskonale podzielne, a ich obrót odbywa się w sposób ciągły;
- możliwe jest pożyczanie oraz inwestowanie środków według tej samej wolnej od ryzyka stopy procentowej r ; stopa ta jest stała w okresie trwania opcji.

Na mocy powyższych założeń pokazano, że:

$$E_{P^*} \left(e^{-rS} (x_{T+S} - K)^+ | \mathfrak{F}_T \right) = x_T \Phi(d_1) - K e^{-rS} \Phi(-d_2) \quad (5.1)$$

gdzie:

$$d_1 = \frac{\ln \left(\frac{x_T}{K \exp(-rS)} \right) + \frac{1}{2} \sigma^2 S}{\sigma \sqrt{S}},$$

$$d_2 = d_1 - \sigma \sqrt{S}, \Phi(\cdot) - \text{dystrybuenta standardowego rozkładu normalnego.}$$

Należy zauważyć, że do zastosowania wzoru (5.1) nie potrzeba znajomości oczekiwanej stopy zwrotu z instrumentu bazowego, tj. parametru dryfu β . Ponadto wzór ten jest stosowany między innymi przez GPW do teoretycznej wyceny europejskiej opcji kupna

⁶⁾ Rynek zupełny, to taki, dla którego każdy instrument pochodny jest replikowany przez stosowną strategię samofinansującą się. Instrument finansowy jest replikowany, jeśli istnieje portfel oraz odpowiednia strategia finansowa, względem której wartość portfela jest identyczna z wartością instrumentu pochodnego w chwili jego wygaśnięcia. Dla strategii samofinansującej się zmiana wartości portfela wynika wyłącznie ze zmian wartości jego składowych.

na indeks WIG20. Uzyskana w ten sposób cena nazywana jest kursem odniesienia. Szczegóły dotyczące arbitrażowej wyceny opcji oraz rozważań na temat osłabienia założeń wyceny w przypadku modelu Blacka i Scholesa można znaleźć między innymi w pracach: Black i Scholes (1972, 1973), Hull (1993, 1998), Lamberton i Lapeyre (1996), Kostrzewski (2000, 2001a, b) oraz Jakubowski, Palczewski, Rutkowski i Stettner (2003).

W prezentowanej monografii do zagadnienia wyceny opcji zastosowano podejście bayesowskie. Jego istota polega na wyznaczeniu rozkładu zmiennej losowej $e^{-rS}(x_{T+S} - K)^+$ z rozkładu predyktywnego dla x_{T+S} . Za ocenę punktową wartości opcji przyjmuje się często wartość oczekiwaną lub medianę tego rozkładu. Bayesowska wycena opcji po raz pierwszy została przedstawiona w pracy Bauwens i Lubrano (1997), a rozwijana w pracach: Bauwens i Lubrano (1998, 2002), Osiewalski i Pipień (2003, 2004b), Pajor (2003a, b), Pipień i Pajor (2005) oraz Pipień (2003), gdzie wykorzystywano procesy dyskretne typu GARCH lub SV.

Przedmiotem zainteresowania tej części pracy są europejskie opcje kupna z datą wygaśnięcia na dzień 18 marca 2005 roku (z kursem wykonania 1800 i 2100 punktów) oraz na dzień 17 czerwca 2005 roku (z kursem wykonania 1800 i 2100 punktów). Opcje te są notowane na GPW, mają odpowiednio oznaczenia: OW20C5180, OW20C5210, OW20F5180 i OW20F5210. Wartość instrumentu bazowego (indeks WIG20) określa się poprzez liczbę punktów, natomiast opcję rozlicza się gotówkowo, zakładając, że każdy punkt wart jest 10 złotych⁷⁾. Celem niniejszego podrozdziału jest wyznaczenie rozkładu predyktywnego zmiennej losowej opisującej wartość opcji na dzień 31 grudnia 2004 roku, przy założeniu włączenia do analizy również danych z tego dnia. Chcąc porównać uzyskane wyniki (w postaci na przykład wartości oczekiwanych lub median rozkładu predyktywnego) z rzeczywistymi cenami podawanymi przez GPW, należy potraktować je jako propozycje cen otwarcia na najbliższy dzień sesyjny, tj. 3 stycznia 2005 roku. W dniu wyceny wartość indeksu na zamknięciu sesji wynosiła 1960,57 punktów, zatem pierwsza i trzecia z rozważanych opcji są opcjami w cenie (*in the money*), a pozostałe są tzw. opcjami nie będącymi w cenie (*out of the money*).

Podobnie jak w podrozdziale 5.5, rozważania ograniczono do przypadku modelu EOU z prawdziwą gęstością przejścia. Czas S do wygaśnięcia opcji został przyjęty na poziomie: 77/365 oraz 168/365. Przyjęto model kapitalizacji ciągłej względem stopy procentowej $r = 6,5\%$, zgodnie z założeniami GPW i KDPW⁸⁾ dla rozważanego momentu czasu. Zmienna losowa opisująca wartość opcji w chwili jej wystawienia, tzn. w momencie T (tzw. zdyskontowana funkcja wypłaty), której wygaśnięcie nastąpi po czasie S , jest postaci $C_{T+S} = e^{-rS}(x_{T+S} - K)^+$. Należy zauważyć, że dla rozważanej sytuacji zmienna C_{T+S} ma rozkład dyskretno-ciągły ze względu na fakt niezerowego prawdopodobieństwa zdarzenia $C_{T+S} = 0$. Tabela 5.5 przedstawia kwantyle rozważanych rozkładów, wartości oczekiwane, kursy odniesienia podane przez GPW, ceny opcji

⁷⁾ Szczegółowe informacje dotyczące opcji na indeks WIG20 można znaleźć na stronach internetowych GPW – www.gpw.com.pl.

⁸⁾ Porównaj: Informacja KDPW o wysokości depozytów zabezpieczających na 31 grudnia 2004 r. na podstawie komunikatu NR 393/DZ/2004 z dnia 31 grudnia 2004 r.

(kursy otwarcia) odnotowane dnia 3 stycznia 2005 roku oraz rzeczywiste (inaczej: zrealizowane) wartości zdyskontowanej funkcji wypłaty⁹⁾ C_{T+S} . W przypadku opcji *in the money*, ceny opcji i zrealizowane wartości wypłaty C_{T+S} znajdują się w sąsiedztwie mediany pomiędzy kwantylami rzędu 0,5 i 0,75. W tym samym przedziale znajduje się również cena opcji *out of the money*, o terminie wykonania $S = 168/365$. Natomiast inna sytuacja jest w przypadku opcji o terminie wykonania $S = 77/365$, której cena znajduje się między kwantylami rzędu 0,75 i 0,95. Dla opcji *out of the money* mediany równają się rzeczywistym wartościom zmiennej C_{T+S} . Wartość oczekiwana w trzech przypadkach wyceny na cztery jest bliższa zaobserwowanej cenie opcji w porównaniu z kursem odniesienia podanym przez GPW. Ponadto w przypadku opcji w cenie wartości oczekiwane oraz kursy odniesienia są podobne do siebie. Duża rozpiętość wartości kwantyli świadczy o znacznym rozproszeniu badanych rozkładów, co w konsekwencji oznacza dużą niepewność wyceny opcji, dla której prognozy punktowe, np. na poziomie mediany lub wartości oczekiwanej, obarczone są dużym błędem *ex ante*.

Tabela 5.5

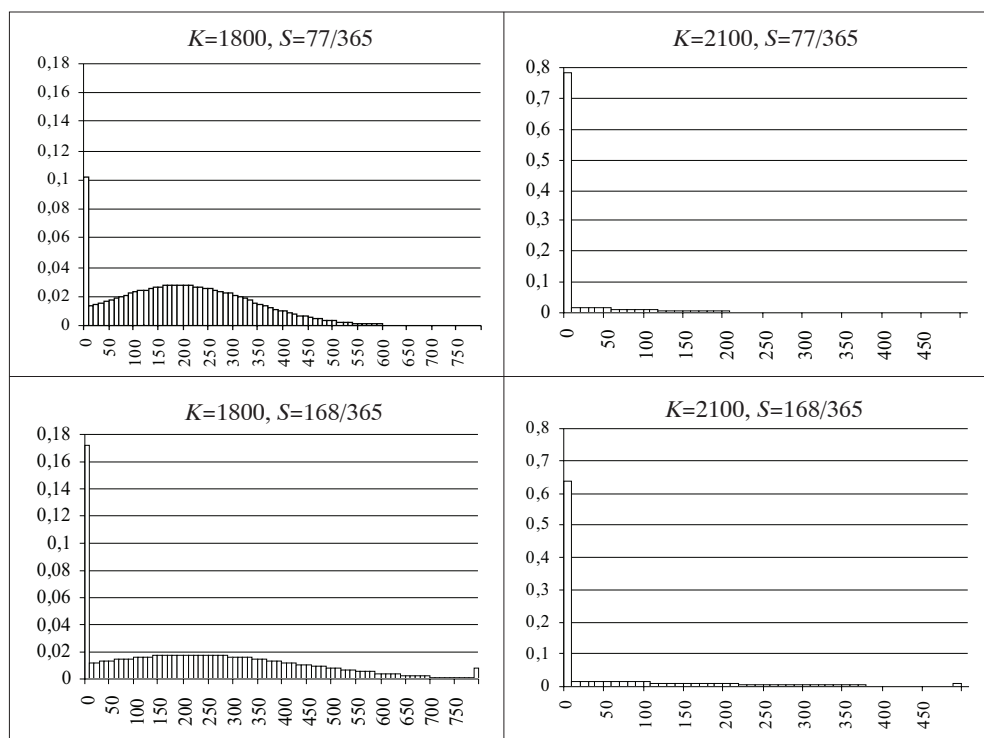
Kwantyle rozkładów predyktywnych zdyskontowanej wypłaty, ich wartości oczekiwane, kurs odniesienia podany przez GPW (wyznaczony z modelu BS przy parametrze zmienności podanym przez KDPW na 31 grudnia 2004 roku), rzeczywiste ceny opcji z otwarcia sesji z 3 stycznia 2005 roku oraz rzeczywiste wartości zdyskontowanej funkcji wypłaty C_{T+S} . Wszystkie wartości podane zostały w punktach

Kwantyl rzędu	$C_{T+77/365}$ 1800	$C_{T+77/365}$ 2100	$C_{T+168/365}$ 1800	$C_{T+168/365}$ 2100
0,05	0	0	0	0
0,25	86,25	0	60,04	0
0,5	184,01	0	211,39	0
0,75	282,25	0	365,10	73,95
0,95	424,61	128,70	591,25	300,09
Wartość oczekiwana	191,47	18,68	234,34	57,33
Kurs odniesienia	195,85	28,85	238,45	69,60
Cena opcji (kurs otwarcia)	210,15	15,00	228,00	37,90
Rzeczywista wartość C_{T+S}	172,62	0	217,40	0

Źródło: Opracowanie własne

⁹⁾ Wielkości te wyliczane są zgodnie ze wzorem na C_{T+S} , przy czym zamiast x_{T+S} podstawia się podany przez GPW kurs rozliczeniowy dla danej opcji. W praktyce kurs rozliczeniowy wyliczany jest jako średnia arytmetyczna ze wszystkich wartości indeksu WIG20 z ostatniej godziny notowań ciągłych oraz wartości tego indeksu ustalonej na zamknięciu sesji z dnia wygaśnięcia opcji.

Ceny opcji równe medianom predyktownych rozkładów zdyskontowanej wypłaty są niższe od cen podanych przez GPW oraz kursu odniesienia. Natomiast są „bliższe” rzeczywistym wielkościom wypłaty w chwili T . Należy w tym momencie podkreślić, że przy wycenie opcji nie zostały uwzględnione koszty dodatkowe, w szczególności koszty transakcji. Duże rozproszenie rozkładów predyktownych zdyskontowanej funkcji wypłaty jest również zauważalne na wykresach histogramów przedstawionych na rysunku 5.6. Ponieważ rozkład C_{T+S} jest rozkładem mieszanym (dyskretno-ciągłym), stąd pierwsza belka histogramu przedstawia masę prawdopodobieństwa skupioną w zerze. Wysokość belki jest równa prawdopodobieństwu niewykonania opcji. Można zauważyć, że im dłuższy czas do terminu wygaśnięcia opcji, tym grubsze prawe ogony rozkładu zdyskontowanej wypłaty. Należy zwrócić uwagę na znaczną koncentrację masy prawdopodobieństwa wokół zera w przypadku opcji *out of the money*. Własność ta odpowiada niskiemu prawdopodobieństwu ich wykonania.



Rys. 5.6. Histogramy rozkładów predyktownych zdyskontowanej wypłaty C_{T+S}

Źródło: Opracowanie własne

Relatywnie niskie oceny wartości opcji *out of the money* OW20C5210 i OW20F5210 mają swoje odzwierciedlenie w dużym prawdopodobieństwie ich niewykonania, odpowiednio ponad 0,77 i 0,63 (por. tab. 5.6). Dla opcji OW20C5180

i OW20F5180 oszacowane prawdopodobieństwa są o wiele mniejsze i wynoszą odpowiednio niewiele ponad 0,1 i 0,17. Kurs rozliczeniowy dla rozważanych instrumentów pochodnych wyniósł: 1975 (dla $S = 77/365$) i 2024 (dla $S = 168/365$). Opcji z kursem wykonania 2100 punktów (tj. OW20C5210 i OW20F5210) nie warto było realizować. Powyższa uwaga jest zgodna z prezentowanym oszacowaniem prawdopodobieństw ich niewykonania.

Tabela 5.6

Prawdopodobieństwa niewykonania opcji kupna na indeks WIG20 oraz informacja o ich realizacji na podstawie wartości opcji w momencie wygaśnięcia

	$C_{T+77/365}$ 1800	$C_{T+77/365}$ 2100	$C_{T+168/365}$ 1800	$C_{T+168/365}$ 2100
$\Pr\{C_{T+s}=0 \mid x, x_0\}$	0,1019	0,7788	0,1728	0,6371
Rzeczywista wielkość wypłaty w chwili $T+S$	175	0	175	0
Wykonano opcję	TAK	NIE	TAK	NIE

Źródło: Opracowanie własne

Warto na koniec tego rozdziału wspomnieć o dokonanym w pracy Kostrzewski i Pajor (2006) porównaniu wyników wnioskowania bayesowskiego dla rozważanego szeregu czasowego WIG20 przy zastosowaniu modelu EOU oraz dyskretnego procesu zmienności (wariancji) stochastycznej z korelacją i grubymi ogonami (w skrócie FCSV) (por. Pajor 2003). W przypadku modelu FCSV zakłada się, że obserwowany szereg czasowy jest realizacją pewnego procesu dyskretnego. Natomiast w modelu EOU zakłada się, że ceny instrumentu finansowego kształtowane są w sposób ciągły, a ich wartości obserwowane są w dyskretnych (niekoniecznie równych) odstępach czasu. Przy jednakowych prawdopodobieństwach a priori obu modeli, model FCSV okazał się istotnie bardziej prawdopodobny a posteriori niż model EOU (por. Kostrzewski i Pajor 2005). Przedmiotem porównania były wyniki prognozy dalszego poziomu indeksu WIG20 oraz wyceny europejskich opcji kupna. Daty objęte prognozą oraz charakterystyki opcji były identyczne jak w niniejszej monografii. Zaobserwowano zgodność median rozkładów predyktywnych oraz rozkładów zdyskontowanej funkcji wypłaty dla obydwu modeli. Rozproszenie rozkładów, mierzone rozstępem międzykwartylowym było duże, jednak istotnie większe w przypadku modelu FCSV. W modelu tym rozkłady posiadały znacznie grubsze ogony. Większe rozproszenie rozkładów w przypadku procesów SV uczyniło bardziej prawdopodobnym pojawianie się obserwacji nietypowych i to właśnie było powodem, dla którego model FCSV był bardziej prawdopodobny a posteriori w porównaniu z modelem EOU. Znaczna rozpiętość kwantyli

przemawia za dużą niepewnością prognozy i wyceny opcji, która wzrasta wraz z wydłużaniem się jej horyzontu. W konsekwencji prognozy punktowe (na poziomie mediany) dla obydwu z rozważanych modeli obarczone są dużym błędem *ex ante*. Ponadto oszacowania prawdopodobieństw niewykonania rozważanych europejskich opcji kupna w obu modelach były bardzo podobne.

5.7. Uwagi końcowe

W ostatnim rozdziale monografii zaprezentowane zostały wyniki wnioskowania bayesowskiego przeprowadzonego na rzeczywistym szeregu czasowym notowanym przez Giełdę Papierów Wartościowych w Warszawie. Przedmiotem analizy był szereg dziennych wartości indeksu WIG20 z zamknięcia sesji giełdowych. Dokonano bayesowskiego porównania konkurencyjnych specyfikacji SRR. Najwyższe prawdopodobieństwo *a posteriori* (równe prawie 1) uzyskał model EOU. W rozdziale zmierzono się również z problemem wyceny czterech opcji kupna na ww. indeks. Wyniki, choć zadowalające, obarczone są sporą niepewnością *ex ante* mierzoną rozpiętością kwantyli rozkładów predyktywnych zdyskontowanej wypłaty.

W rozdziale wspomniano o wynikach prognozy i wyceny opcji dokonanej przy zastosowaniu modelu EOU i FCSV. Zaskakująca okazała się zgodność uzyskanych w obydwu modelach wielkości median rozkładów predyktywnych poziomu indeksu WIG20 oraz rozkładów zdyskontowanej wypłaty. Ponadto oszacowania prawdopodobieństw niewykonania rozważanych europejskich opcji kupna w obu modelach okazały się bardzo podobne. Wyniki uzyskane na podstawie modeli FCSV i EOU różnią się niepewnością (mierzoną rozpiętością kwantyli) co do przyszłej wartości indeksu oraz wielkości zdyskontowanej wypłaty, a także grubością ogonów rozkładów predyktywnych. Zaobserwowano większe rozproszenie oraz znacznie grubsze ogony tych rozkładów w dyskretnym modelu zmienności stochastycznej. Rozkłady predyktywne zdyskontowanej funkcji wypłaty są jednak w obu modelach na tyle rozproszone, że przyjęcie za cenę opcji jakiegokolwiek punktowego oszacowania jest obarczone dużym błędem *ex ante*. Powyższa uwaga potwierdza nietrywialność problemu precyzyjnej wyceny instrumentów pochodnych.

Zagadnienie wyznaczenia analitycznego wzoru na europejską opcję, gdy dynamika instrumentu bazowego opisywana jest poprzez model EOU, jest tematem dalszych badań autora.

Zewsząd płynąca krytyka zasadności stosowania najpopularniejszego modelu zadanego równaniem Itô, tj. modelu Blacka i Scholesa, znalazła swoje potwierdzenie w niniejszym badaniu empirycznym. Wyniki przeprowadzonej analizy wykazały całkowitą nieprzydatność modelu w opisie dynamiki indeksu WIG20.

Ciągła potrzeba poszukiwania coraz to bardziej wyrafinowanych narzędzi modelowania zjawisk zachodzących w świecie finansów, przysłania i czyni niepopularnym stosowanie stosunkowo dobrze teoretycznie poznanych procesów Itô. Jednak wyniki niniejszego rozdziału wydają się zachęcać do ich uwzględnienia przy podejmowaniu decyzji finansowych. Na uwagę zasługują zwłaszcza procesy niejednorodne, które, być może ze względu na skalę trudności ich modelowania, nie znalazły dotychczas wystarczającego zainteresowania w literaturze przedmiotu i praktyce.

Zakończenie

W monografii podjęto badania nad opracowaniem i zastosowaniem metod bayesowskich w modelowaniu finansowych szeregów czasowych za pomocą procesów dyfuzji. Procesy te wykorzystywane są m.in. w nowoczesnej teorii finansów, a w szczególności stosowane są w zagadnieniach związanych z konstrukcją optymalnych strategii inwestycyjnych, modelowaniem struktury czasowej instrumentów bazowych oraz do wyceny instrumentów pochodnych. Wzrastająca rola procesów dyfuzji Itô w zastosowaniach finansowych zmusza do podjęcia tematu estymacji parametrów funkcji dryfu i funkcji zmienności. W ogólnym przypadku, brak jawnego przepisu na funkcję wiarygodności stanowi główną trudność w przeprowadzaniu wnioskowania statystycznego. Pomimo to, że problem estymacji parametrów procesów dyfuzji nie jest nowym zagadnieniem, wciąż cieszy się dużym zainteresowaniem, przynoszącym w ostatnich kilku latach ciekawe wyniki teoretyczne. Widoczne w światowej literaturze zainteresowanie tą tematyką może świadczyć o nietrywialności rozważanego w pracy problemu.

W pierwszym rozdziale monografii przedstawiono definicje i podstawowe własności rozważanych w pracy procesów stochastycznych. Omówiono procesy dyfuzji, procesy Itô oraz procesy dyfuzji Itô. W rozdziale tym wspomniano o numerycznych aproksymacjach rozwiązań SRR. Rozważono zagadnienie schematów dyskretyzacji i ich zastosowania w aproksymacji trajektorii oraz charakterystyk rozkładów procesów. W szczególności omówiono schemat Eulera i Milsteina. Dodatkowo przedstawiono pojęcie linii kwantylowych, jako metody aproksymacji i graficznej ilustracji rozwiązań SRR. Ponadto omówiono zadane poprzez stochastyczne równania różniczkowe modele matematyczne: Vasička (V), Blacka i Scholesa (BS), Coxa, Ingersolla i Rossa (CIR), Brennana i Schwartza (BSBS) oraz uogólniony model Ornsteina i Uhlenbecka (EOU). Dla każdego z nich, z wyjątkiem BSBS, znana jest jawna postać gęstości przejścia rozwiązania. Taki wybór modeli umożliwił porównanie wyników wnioskowania bayesowskiego uzyskanych za pomocą aproksymowanej i prawdziwej funkcji wiarygodności.

U podstaw prezentowanego w monografii wnioskowania statystycznego leży zagadnienie wyznaczenia funkcji wiarygodności, a ściślej jej aproksymacji. Problem ten, z własności silnych rozwiązań równań Itô, sprowadza się do zagadnienia aproksymacji gęstości przejścia.

Opierając się na najnowszych wynikach teoretycznych literatury przedmiotu, w rozdziale drugim rozważono trzy metody:

- 1) zmiennych pośrednich,
- 2) wielomianów Hermite’a,
- 3) rozwiązań równań Fokkera, Plancka i Kołmogorowa.

Celem pracy nie było porównanie wspomnianych metod, a wyłącznie opracowanie na ich podstawie narzędzi wnioskowania statystycznego.

W rozdziale trzecim przedstawiono ogólne zasady wnioskowania bayesowskiego. Zdefiniowano przyjętą w monografii strukturę rozkładu a priori oraz określono rozkłady a posteriori i rozkłady predyktywne dla omawianych modeli. Poruszono temat porównania konkurencyjnych modeli. Znaczną część rozdziału poświęcono metodom numerycznym będącym często jedynymi narzędziami umożliwiającymi obliczenie charakterystyk wielowymiarowych rozkładów a posteriori i predyktywnych. W szczególności omówiono metody: eliminacji, Monte Carlo z funkcją ważności, Metropolis’a i Hastingsa (a w szczególności próbnik Gibbsa) oraz przedstawiono podstawowe własności statystyki CUMSUM i estymatora Newtona i Raftery’ego. Ponadto omówiono zagadnienie estymacji parametrów za pomocą metody zmiennych pośrednich. W szczególności przedstawiono wyniki bayesowskiej estymacji parametrów opartej na wspomnianej metodzie, gdzie przedmiotem badania były trajektorie logarytmu procesu pierwiastka kwadratowego. Wnioskiem z tego eksperymentu było stwierdzenie nieefektywności metody zmiennych pośrednich uzasadniające pominięcie jej w kolejnych badaniach empirycznych.

W rozdziale czwartym zaprezentowano wyniki bayesowskiego wnioskowania przeprowadzonego, opierając się na szeregu miesięcznych stóp procentowych Amerykańskiego Banku Centralnego (FED). W szczególności omówiono zagadnienie estymacji parametrów, porównania siły wyjaśniającej konkurencyjnych modeli oraz przedstawiono wyniki krótkoterminowej prognozy dalszego poziomu stopy FED. Warte podkreślenia jest w szczególności trafność uzyskanej prognozy, dla której przyszłe wartości stopy FED znalazły się w bardzo bliskim sąsiedztwie median rozkładów predyktywnych – pomiędzy kwantylami rzędu 0,25 i 0,75. Prezentowane wyniki otrzymano, wykorzystując prawdziwą gęstość przejścia oraz jej aproksymację. Zgodność w oszacowaniu nieznanymi parametrów w obu tych przypadkach jest dodatkowym empirycznym dowodem zbieżności prezentowanych metod. W rozdziale zwrócono uwagę na szybkość algorytmów numerycznych, świadczącą o dużej efektywności omawianych w pracy metod.

Interesującym wnioskiem rozdziału czwartego jest możliwość modelowania w sposób zadowalający dynamiki szeregu FED za pomocą prostych procesów Itô, zadanych poprzez liniowe stochastyczne równania różniczkowe. W szczególności mowa tutaj o powszechnie krytykowanym modelu Blacka i Scholesa. W praktyce do modelowania struktury czasowej stóp procentowych stosuje się między innymi modele CIR i V, które w rozważanym eksperymencie zostały odrzucone przez dane. Nieoczekiwanie model

BS, który wykorzystywany jest zwykle do opisu wartości akcji, okazał się dobrym narzędziem modelowania rozważanego szeregu stopy procentowej FED.

Ostatni, piąty rozdział monografii, poświęcony został ilustracji empirycznej zastosowania procesów dyfuzji w modelowaniu polskich danych finansowych. Przedmiotem zainteresowania był szereg dziennych wartości indeksu WIG20 z zamknięcia sesji, notowanych przez Giełdę Papierów Wartościowych w Warszawie. Wnioskowanie bayesowskie zastosowane zostało do oszacowania nieznanych wartości parametrów rozpatrywanych procesów, dokonano porównania rozważanych modeli oraz przeprowadzono bayesowską prognozę i wycenę europejskiej opcji kupna na indeks WIG20. Przyszłe wartości indeksu WIG20 znalazły się w bardzo bliskim sąsiedztwie median rozkładów predyktywnych – pomiędzy kwantylami rzędu 0,25 i 0,75. Podobnie rzeczywiste wartości wypłaty rozpatrywanych opcji znalazły się w bliskim sąsiedztwie median rozkładów predyktywnych zdyskontowanej funkcji wypłaty – pomiędzy kwantylami rzędu 0,25 i 0,75. Ponadto wartość oczekiwana rozkładu predyktywnego zdyskontowanej funkcji wypłaty w trzech przypadkach wyceny na cztery okazała się bliższa zaobserwowanej cenie opcji w porównaniu z kursem odniesienia podanym przez GPW.

Naturalną konsekwencją rozpatrywania w pracy procesów zdefiniowanych w przedziale czasu, a nie w chwilach dyskretnych, jest jego „urealnienie”. W szczególności wyznaczając czas pomiędzy występowaniem kolejnych wartości szeregu czasowego brano pod uwagę rzeczywistą liczbę dni, włączając w to np. dni wolne od pracy.

Krytyka zasadności stosowania modelu Blacka i Scholesa znalazła swoje potwierdzenie w badaniu empirycznym rozdziału piątego. Wyniki przeprowadzonej analizy wykazały całkowitą nieprzydatność tego modelu w opisie dynamiki indeksu WIG20. W rozważanej klasie konkurencyjnych specyfikacji bayesowskich model EOU najwierniej opisuje dynamikę szeregu WIG20. Wadą rozwiązania modelu EOU jest możliwość przyjmowania przez nie wartości ujemnych. Przyjmując za oceny punktowe parametrów ich wartości oczekiwane a posteriori, wykazano jednak, że prawdopodobieństwo takiego zdarzenia (dla horyzontu rozważanej prognozy) jest bardzo bliskie zeru. Ponadto dla rozkładu predyktywnego prawdopodobieństwo przyjęcia przez proces wartości dodatnich jest bliskie jedności.

Wyniki badań empirycznych wskazały na to, że o wyborze najlepszego modelu decydujący wpływ miała postać funkcji zmienności. Modele o takich samych funkcjach dryfu, ale odmiennych funkcjach dyfuzji, uzyskiwały wyraźnie różne prawdopodobieństwa a posteriori.

Ponadto w rozdziale piątym wspomniano o porównaniu wyników prognozy i wyceny opcji dokonanej przy zastosowaniu modelu EOU oraz dyskretnego procesu zmienności stochastycznej z korelacją i grubymi ogonami (FCSV). Zaskakująca okazała się zgodność uzyskanych w obydwu modelach wielkości median rozkładów predyktywnych poziomu indeksu WIG20 oraz rozkładów zdyskontowanej wypłaty. Ponadto oszacowania prawdopodobieństw niewykonania rozważanych europejskich opcji kupna w obu modelach okazały się bardzo podobne. Wyniki uzyskane na podstawie modeli EOU i FCSV różnią się niepewnością (mierzoną rozpiętością kwantyli) co do przyszłej

wartości indeksu oraz wielkości zdyskontowanej wypłaty. Zaobserwowano większe rozproszenie oraz znacznie grubsze ogony tych rozkładów dla drugiego ze wspomnianych modeli, świadczące o mniejszej precyzji prognozy i wyceny opartej na modelu SV. Rozkłady predyktywne zdyskontowanej funkcji wypłaty były w obu modelach jednak na tyle rozproszone, że przyjęcie za cenę opcji jakiegokolwiek punktowego oszacowania jest obarczone dużym błędem *ex ante*. Powyższa uwaga potwierdza nietrywialność problemu precyzyjnej wyceny instrumentów pochodnych.

Przedstawione w pracy rezultaty empiryczne uzyskano za pomocą algorytmów numerycznych opartych na metodach Monte Carlo oraz łańcuchach Markowa. Zastosowano technikę MC z funkcją ważności oraz algorytm Metropolisa i Hastingsa, a w szczególności próbnik Gibbsa, wersję sekwencyjną M-H i losową wersję blokową algorytmu M-H. Ponadto wykorzystano metodę eliminacji. Przy czym za funkcję ważności (wykorzystaną w technikach MC-IS) i funkcję generującą (wykorzystaną w algorytmach M-H) przyjęto wielowymiarowy rozkład *t*-Studenta z trzema stopniami swobody, dla którego średnia i macierz kowariancji zostały ustalone po wstępnych przebiegach algorytmu MC-IS. Wysoki procent akceptacji pociągał za sobą szybką zbieżność algorytmów MCMC. Do oceny zbieżności algorytmu M-H wykorzystano statystykę CUMSUM. Natomiast estymator średniej harmonicznej Newtona i Raftery'ego został wykorzystany do porównania mocy wyjaśniającej rozważanych modeli.

Rosnąca potrzeba poszukiwania coraz to bardziej wyrafinowanych narzędzi modelowania zjawisk finansowych przysłania i czyni niepopularnym stosowanie stosunkowo dobrze teoretycznie poznanych procesów dyfuzji Itô. Jednak wyniki prezentowanej monografii przemawiają za uwzględnieniem ich przy podejmowaniu decyzji finansowych. Na uwagę zasługują zwłaszcza procesy niejednorodne, które, być może ze względu na skalę trudności ich zastosowania, nie znalazły dotychczas wystarczającego zainteresowania w literaturze przedmiotu i praktyce.

Warto podkreślić, że modele V, CIR i BSBS są najczęściej stosowane do opisu stóp procentowych, natomiast za pomocą modelu BS wyjaśniana jest dynamika cen akcji. Z punktu widzenia teorii nie ma jednak przeciwwskazań, by modele te stosować do opisu innych zjawisk – co zresztą jest powszechną praktyką w szeroko rozumianej literaturze z zakresu ekonomii. Rozważane modele: V, BS, CIR, BSBS i EOU wykorzystano w monografii do opisu zarówno dynamiki szeregu czasowego stopy procentowej, jak i wartości indeksu giełdowego. Interesującym wynikiem przeprowadzonych badań empirycznych jest wysokie prawdopodobieństwo *a posteriori* przyznane modelowi BS w przypadku analizy stopy procentowej FED.

Stosowane w części empirycznej pracy metody wielomianów Hermite'a oraz rozwiązań równań Fokkera, Plancka i Kołmogorowa są wygodnym narzędziem przybliżania funkcji wiarygodności. Należy podkreślić szybkość obliczeniową algorytmów numerycznych opartych na wspomnianych metodach.

Na podstawie przeprowadzonych w niniejszej pracy badań, można uznać podejście bayesowskie za wygodne i efektywne narzędzie wnioskowania o nieznanach własnościach badanych szeregów czasowych. W wyniku jego zastosowania badacz uży-

skuje wolne od asymptotycznych aproksymacji rozkłady nieznanych parametrów modeli, wektora prognozowanego, zdyskontowanej funkcji wypłaty opcji czy też jakiegokolwiek innej badanej charakterystyki – a nie jedynie oceny punktowe nieznanych wielkości, jak ma to miejsce w ujęciu teorio-próbkowym. Dodatkowo poprzez określenie rozkładów a priori istnieje możliwość włączenia do wnioskowania w sposób jawny i sformalizowany wstępnej wiedzy posiadanej przed wglądem w dane. Ponadto ujęcie bayesowskie poprzez możliwość łączenia wiedzy płynącej z konkurencyjnych modeli daje sposobność przeprowadzenia wnioskowania statystycznego bez konieczności ograniczania rozważań tylko do jednego modelu.

W monografii ograniczono rozważania do jednowymiarowych procesów dyfuzji, pozostawiając analizę wielowymiarowego przypadku dalszej pracy autora. Ponadto w pracy skoncentrowano się na SRR z zaburzeniem gaussowskim, pozostawiając tematykę SRR z zaburzeniami α -stabilnymi (*SDE with α -stable measure*) czy też z zaburzeniami o rozkładzie Poissona (*jump-diffusion SDE*) dalszemu zakresowi badań. Opracowanie i zastosowanie bayesowskiego wnioskowania dla jednowymiarowych SRR z zaburzeniem gaussowskim było samo w sobie dużym wyzwaniem, a jednocześnie wspaniałą przygodą, której owocem jest niniejsza monografia.

Literatura

- Abramowitz M., Stegun I.A. 1968, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York.
- Aït-Sahalia Y. 1999, *Transition Densities for Interest Rate and other Nonlinear Diffusions*, Journal of Finance, 54, 1361–1395.
- Aït-Sahalia Y. 2002, *Maximum-Likelihood Estimation of Discretely-Sampled Diffusions: A Closed-Form Approximation Approach*, Econometrica, 70, 223–262.
- Aït-Sahalia Y. 2003, 2004, *Closed-Form Likelihood Expansions for Multivariate Diffusions*, Working Paper of Princeton University.
- Arnold L. 1974, *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons, Inc.
- Bachelier L. 1900, *Théorie de la spéculation*, Annales Scientifiques de l'Ecole Normale Supérieure, III-17, 21–86.
- Bailey N.T.J. 1990, *The Elements of Stochastic Processes with Applications to the Natural Sciences*, John Wiley & Sons, Inc.
- Baka W. 2001, *Bankowość Centralna*, OW Zarządzanie i Finanse, Warszawa.
- Bally V., Talay D. 1996, *The Law of the Euler Scheme for Stochastic Differential Equations*, Monte Carlo Methods and Applications 2, 93–128.
- Barndorff-Nielsen O.E., Sørensen M. 1994, *A Review of Some Aspects of Asymptotic Likelihood Theory for Stochastic Processes*, International Statistical Review, 62, 133–165.
- Bauwens L., Lubrano M. 1997, *Bayesian Option Pricing Using Asymmetric GARCH*, CORE Discussion Paper, No. 9759 (Université Catholique de Louvain, Louvain-la-Neuve).
- Bauwens L., Lubrano M. 1998, *Bayesian Inference on GARCH Models Using the Gibbs Sampler*, Econometrics Journal, 1, C23–C46.
- Bauwens L., Lubrano M. 2002, *Bayesian Option Pricing Using Asymmetric GARCH Models*, Journal of Empirical Finance, 9, 321–342.
- Bernardo J.M., Smith A.F.M. 2002, *Bayesian Theory*, John Wiley & Sons, Inc., New York.
- Bernstein S.N. 1934, *Principes de la théorie des équations différentielles stochastiques*, Trudy Fiz.-Mat. in-ta im. Steklova, 5, 95–124.

- Billingsley P. 1961a, *Statistical Methods in Markov Chains*, AMS, vol. 32, 12–40.
- Billingsley P. 1961b, *Statistical Inference for Markov Processes*, Chicago University Press, Chicago.
- Bingham N.H., Kiesel R. 2004, *Risk-Neutral Valuation. Pricing and Hedging of Financial Derivatives*, Springer-Verlag, Londyn – Berlin – Heidelberg.
- Black F., Scholes M. 1972, *The Valuation of Option Contracts and a Test of Market Efficiency*, Journal of Finance, 27, 399–418.
- Black F., Scholes M. 1973, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, vol. 81, 637–654.
- Box G.E.P., Tiao G.C. 1973, *Bayesian Inference in Statistical Analysis*, Addison-Wesley Publishing Company, London.
- Brandt W, Santa-Clara P. 2002, *Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets*, Journal of Financial Economics, 63, 161–210.
- Brennan M., Schwartz E. 1979, *A Continuous Time Approach to the Pricing of Bonds*, Journal of Banking and Finance, 3, 133–155.
- Brennan M., Schwartz E. 1980, *Analyzing Convertible Bonds*, Journal of Financial and Quantitative Analysis, 15, 907–929.
- Brigo D., Mercurio F. 2001, *Interest Rate Models Theory and Practice*, Springer-Verlag, Berlin – Heidelberg – New York.
- Brown R. 1827, *A Brief Account of Microscopical Observations Made in the Months of June, July, and August 1827, on the Particles Contained in the Pollen of Plants; and on the General Existence of Active Molecules in Organic and Inorganic Bodies*, Philosophical Magazine N. S., 4 (1928), 161–173.
- Brzeźniak Z., Zastawniak T. 2002, *Basic Stochastic Processes: A Course Through Exercises*, Springer-Verlag, London.
- Buffon G. 1733, Editor's note concerning a lecture given 1733 by Mr. Le Clerc de Buffon to the Royal Academy of Sciences in Paris, Histoire de l'Acad. Roy. des Sci., 43–45.
- Buffon G. 1777, *Essai d'arithmétique morale*, Histoire naturelle, générale et particulière, Supplément 4, 46–123.
- Butler C. 2001, *Tajniki Value at Risk*, Wydawnictwo K. E. Liber.
- Chib S., Greenberg E. 1995, *Understanding the Metropolis-Hastings Algorithm*, The American Statistician, 49, 327–335.
- Clark J.M.C., Cameron R.J. 1980, *The Maximum Rate of Convergence of Discrete Approximations for Stochastic Differential Equations*, Springer Lecture Notes in Control and Inform. Sc. Vol., 25, 162–171.
- Clements D.J., Anderson B.D.O. 1973, *Well Behaved Ito Equations with Simulations that Always Misbehave*. IEEE Trans. Autom. Control, AC-18, 676–677.
- Cowles M.K., Carlin B.P. 1996, *Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review*, Journal of the American Statistical Association, 91, 883–904.

- Cox J. C., Ingersoll J. E., Ross S.A. 1985, *A Theory of the Term Structure of Interest Rates*, *Econometrica*, 53, 384–407.
- Cox J.C. 1996, *The Constant Elasticity of Variance Option Pricing Model*, *The Journal of Portfolio Management*, Special issue.
- Cramer H. 1925, *On Some Classes of Series Used in Mathematical Statistics*, *Proceedings of the Sixth Scandinavian Mathematical Congress*, 399–425.
- Cramer H. 1957, *Mathematical Methods of Statistics*, Princeton University Press.
- Crawford G., Sen B. 1998, *Instrumenty pochodne. Narzędzie podejmowania decyzji finansowych*, Wydawnictwo K.E. Liber.
- Cyganowski S., Kloeden P., Ombach J. 2001, *From Elementary Probability to Stochastic Differential Equations with MAPLE*. Springer-Verlag, Berlin.
- Dacunha-Castelle D., Florens-Zmirou D. 1986, *Estimation of the Coefficients of Diffusion from Discrete Observations*, *Stochastics*, 19, 263–284.
- DeGroot M.H. 1981, *Optymalne decyzje statystyczne*, PWN, Warszawa.
- DiPietro M. 2001, *Bayesian Inference for Discretely Sampled Diffusion Processes with Financial Applications*, PhD Thesis, Department of Statistics, Carnegie-Mellon University.
- Doob J.L. 1953, *Stochastic Processes*, John Wiley and Sons, Inc., New York – London.
- Egorov A.V., Li H., Xu Y. 2003, *Maximum Likelihood Estimation of Time Inhomogeneous Diffusions*, *Journal of Econometrics*, 114, 107–139.
- Einstein A. 1905, *On the Movement of Small Particles Suspended in a Stationary Liquid Demanded by the Molecular-Kinetic Theory of Heat*, *Ann. Physik*, 17.
- Einstein A. 1906, *Zur Theorie der Brownschen Bewegung*, *Ann. Phys.*, IV 19, 371.
- Elerian O. 1998, *A Note on the Existence of a Closed-form Conditional Transition Density for the Milstein Scheme*, *Economics Discussion Paper*, 1998-W18, Nuffield College, Oxford.
- Elerian O. 1999, *Simulation Estimation of Continuous Time Series Models with Applications to Finance*, unpublished PhD Thesis, Nuffield College, Oxford.
- Elerian O., Chib S., Shephard N. (2001), *Likelihood Inference for Discretely Observed Nonlinear Diffusions*, *Econometrica*, 69, 959–993.
- Eraker B. 1998, *Markov Chain Monte Carlo Analysis of Diffusion Models with Application to Finance*, HAE Thesis, Norwegian School of Economics and Business Administration.
- Feller W. 1936, *Zur Theorie der Stochastischen Prozesse (Existenz und Eindeutigkeitsätze)*, *Math. Ann.*, 113, 113–160.
- Feller W. 1952, *The Parabolic Differential Equations and the Associated Semigroups of Transformations*, *Ann. Math.*, 55, 468–519.
- Feller W. 1978, *Wstęp do rachunku prawdopodobieństwa*, PWN, Warszawa.
- Filipović D. 2001, *Consistency Problems for Heath-Jarrow-Morton Interest Rate Models*, Springer-Verlag, Berlin – Heidelberg.

- Fisz M. 1967, *Rachunek prawdopodobieństwa i statystyka matematyczna*, PWN, Warszawa.
- Florens-Zmirou D. 1989, *Approximate Discrete-Time Schemes for Statistics of Diffusion Processes*, Statistics, 20, 547–557.
- Fraser D.A.S., Wong A.C.M., Wu J. 1998, *An Approximation for the Noncentral Chi-Squared Distribution*, Communications in Statistics: Simulation and Computation, 27, 257–287.
- Friedman A. 1975, *Stochastic Differential Equations and Applications*, Vol. 1, New York: Academic Press.
- Gajek L., Kałuszka M. 1996, *Wnioskowanie statystyczne*, WNT, Warszawa.
- Gamerman D. 1997, *Markov Chain Monte Carlo (Stochastic Simulation for Bayesian Inference)*, Chapman and Hall, London.
- Geweke J. 1989, *Bayesian Inference in Econometric Models Using Monte Carlo Integration*, Econometrica, 57, 1317–1339.
- Gichman I.I., Skorochod A.V. 1968, *Wstęp do teorii procesów stochastycznych*, PWN, Warszawa.
- Girsanov I.V. 1962, *An Example of Non-Uniqueness of the Solution to the Stochastic Differential Equation of K. Ito*, Theory Probab. Appl., 7, 325–331.
- Gutiérrez R., Nafidi A., Gutiérrez Sánchez R. 2004, *Inference in the Stochastic Gompertz Diffusion Model with Continuous Sampling*, Monografías del Seminario Matemático García de Galdeano, 31, 347–353.
- Hammersley J.M., Handscomb D.C. 1964, *Monte Carlo Methods*, Methuen, London.
- Hastings W.K. 1970, *Monte Carlo Sampling Methods Using Markov Chains and Their Applications*, Biometrika, 57, 97–109.
- Higham D. J. 2001, *An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations*, SIAM Review, vol. 43, no. 3, 525–546.
- Hull J.C. 1993, *Options, Futures, and other Derivatives*. New Jersey, Prentice-Hall.
- Hull J.C. 1998, *Kontrakty terminowe i opcje: wprowadzenie*, WIG-Press, Warszawa
- Ikeda N., Watanabe S. 1981, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Amsterdam, New York.
- Instrumenty pochodne: symposium matematyki finansowej, Uniwersytet Jagielloński, 10–12 kwietnia 1997*, zespół red. Kobak P., Życzkowski K., Capinski M., TAIWPN Universitas, Krakow.
- Itô K. 1942, *Differential Equations Determining Markov Processes* (in Japanese), Zenkoku Shijō Sūgaku Danwakai 1077, 1352–1400.
- Itô K. 1944, *Stochastic Integral*, Proc. Imp. Acad. Tokyo, 20, 519–524.
- Itô K. 1946, *On A Stochastic Integral Equation*, Proc. Imp. Acad. Tokyo 22, 32–35.
- Itô K. 1951, *On Stochastic Differential Equations*, Mem. Amer. Math. Soc., 4, 1–51.
- Jajuga K., Jajuga T. 2000, *Inwestycje*, PWN, Warszawa.
- Jakubowski J., Palczewski A., Rutkowski M., Stettner Ł. 2003, *Matematyka finansowa. Instrumenty pochodne*. WNT, Warszawa.

- Janicki A. 1996, *Numerical and Statistical Approximation of Stochastic Differential Equations with Non-Gaussian Measures*, Monograph No. 1, H. Steinhaus Center for Stochastic Methods in Science and Technology, Wrocław.
- Janicki A., Izydorczyk A. 2001, *Komputerowe metody w modelowaniu stochastycznym*, WNT, Warszawa.
- Janicki A., Weron A. 1994, *Simulation and Chaotic Behavior of α -Stable Stochastic Processes*, Marcel Dekker, New York.
- Jeffreys H. 1967, *Theory of Probability*, Oxford University Press, London.
- Jensen B., Poulsen R. 1999, *A Comparison of Approximation Techniques for Transition Densities of Diffusion Processes*, Working Paper, No. 38, University of Aarhus.
- Jensen J.L. 1986, *Nogle Asymptotiske Resultater*, Working Paper, University of Aarhus.
- Jermakow S.M. 1976, *Metoda Monte Carlo i zagadnienia pokrewne*, PWN, Warszawa.
- Jones C.S. 1998, *Bayesian Estimation of Continuous-Time Finance Models*, Working Paper, Simon School of Business, University of Rochester.
- Karatzas I., Shreve S.E. 1988, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York.
- Kass R.E., Raftery A.E. 1995, *Bayes Factors*, Journal of the American Statistical Association, vol. 90.
- Kim S., Shephard N., Chib S. 1998, *Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models*, Review of Economic Studies, 65, 361–393.
- Kloeden P.E., Platen E. 1992, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin – Heidelberg – New York.
- Komori Y., Saito Y., Mitsui T. 1994, *Some Issues in Discrete Approximate Solution for Stochastic Differential Equations*, Computers Math. Applic. vol. 28, No. 10–12, 269–278.
- Korbicz J., Zacharewicz Zgurowski M. 1991, *Estymacja i sterowanie stochastyczne układami o parametrach rozłożonych*, PWN, Warszawa.
- Kostrzewski M. 2000, *Hedging a koszty transakcji w wycenie europejskiej opcji call na akcję nie wypłacającą dywidendy*, praca magisterska obroniona 4 lipca 2000 roku w Instytucie Matematyki Uniwersytetu Jagiellońskiego w Krakowie.
- Kostrzewski M. 2001a, *Delta Hedging and Transaction Costs in Pricing European Call Options on a Non-Dividend-Paying Stock*, Modelling Economies in Transition. Proceedings of the Fifth AMFET Conference, Ed. Władysław Welfe (Łódź 2001), 146–153.
- Kostrzewski M. 2001b, *Wpływ kosztów transakcji w zabezpieczeniu europejskiej opcji call*, Rynek Terminowy, 14, 108–112.
- Kostrzewski M. 2004, *Bayesowska estymacja parametrów dyskretnie obserwowalnych procesów dyfuzji (na przykładzie modelu CIR)*, Przegląd Statystyczny, 3, 129–139.
- Kostrzewski M. 2006a, *Bayesian Inference on Discretely Sampled Itô Processes*, 4th Annual International Conference on Forecasting Financial Markets and Economic Decision-making (Łódź, May 2005), Acta Universitatis Lodzensis. Folia Oeconomica (w druku).

- Kostrzewski M. 2006b, *Bayesowska estymacja, prognoza i porównanie procesów Itô modelujących finansowe szeregi czasowe*, Metody ilościowe w naukach ekonomicznych. Szóste warsztaty doktorskie z zakresu Ekonometrii i Statystyki (Zakopane, czerwiec 2005), red. A. Welfe, Wydawnictwo SGH, Warszawa (w druku).
- Kostrzewski M., Pajor A. 2005, *Bayesowska wycena opcji na indeks WIG20: procesy Itô a dyskretne procesy SV*. Folia Oeconomica Cracoviensia, vol. 46 (w druku)
- Kowalenko I.N., Kuzniecowa N.J., Szurienkow W.M. 1989, *Procesy stochastyczne. Poradnik*, PWN, Warszawa.
- Krzykowski G. 2000, *Analiza wyników badań marketingowych. Metody symulacyjne*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
- Lamberton D., Lapeyre B. 1996, *Introduction to Stochastic Calculus Applied to Finance*, Chapman & Hall.
- Lamperti J. 1964, *A Simple Construction of Certain Diffusion Processes*, J. Math. Kyoto Univ., 4, 161–170.
- Lévy P. 1937, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, Paris.
- Lévy P. 1939, *Sur certains processus stochastiques homogènes*, Compositio Math. 7, 283–339.
- Lévy P. 1948, *Processus Stochastiques et Mouvement Brownien*, Gauthier-Villars, Paris.
- Lipcer R.S., Szirajew A.N. 1981, *Statystyka procesów stochastycznych*, PWN, Warszawa.
- Liu J.S. 2001, *Monte Carlo Strategies in Scientific Computing*, Springer-Verlag, Berlin.
- Lo A. 1988, *Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data*, Econometric Theory, 4, 231–247.
- Markow A.A. (Марков, А.А.) 1906, *Распределение закона больших чисел на величины, зависящие друг от друга*, Изв. физ. матем. о-ва при Казанском ун-те, (2) 15, 135–156.
- Maruyama G. 1955, *Continuous Markov Processes and Stochastic Equations*, Rend. Circolo Math. Palermo, 4, 48–90.
- Męczarski M. 1998, *Problemy odporności w bayesowskiej analizie statystycznej*, Monografie i Opracowania 446, Szkoła Główna Handlowa, Warszawa.
- Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E. 1953, *Equations of State Calculations by Fast Computing Machines*, Journal of Chemical Physics, 21, 1087–1092.
- Metropolis N., Ulam S. 1949, *The Monte Carlo Method*, J. Amer. Stat. Assoc., 44, 247, 335–341.
- Milstein G.N. 1974, *Approximate Integration of Stochastic Differential Equations*, Theory of Probability and Its Applications, 19, 557–562.
- Milstein G.N. 1978, *A Method of Second-Order Accuracy Integration of Stochastic Differential Equations*. Theory of Probability and Its Applications, 23, 396–401.

- Musiela M., Rutkowski M. 1998, *Martingale Methods in Financial Modelling*, Springer-Verlag, Berlin.
- Neftci S.N. 2000, *An Introduction to the Mathematics of Financial Derivatives*, Academic Press, San Diego.
- Nelson D., Ramaswamy K. 1989, *Simple Binomial Processes as Diffusion Approximations in Financial Models*, Review of Financial Studies, 3, 393–430.
- Newton M.A., Raftery A.E. 1994, *Approximate Bayesian Inference by the Weighted Likelihood Bootstrap (with Discussion)*, Journal of the Royal Statistical Society B, vol. 56, No. 1, 3–48.
- Nicolau J. 2002, *A New Technique for Simulating the Likelihood of Stochastic Differential Equations*, Econometrics Journal, vol. 5, 91–103.
- Nicolau J. 2004, *A Method for Simulating Non-Linear Stochastic Differential Equations in R*, Working Paper, Department of Economics, Technical University of Lisbon.
- Øksendal B. 2000, *Stochastic Differential Equation: An Introduction with Applications*, Springer-Verlag, Berlin.
- O'Hagan A. 1994, *Bayesian Inference*, Halsted Press, New York.
- Osborne M.F.M. 1959, *Brownian Motion in the Stock Market*, Operations Research, 7, 145–173.
- Osiewalski J. 1991, *Bayesowska estymacja i predykcja dla jednorównaniowych modeli ekonometrycznych*, Monografie Akademii Ekonomicznej, nr 100, Kraków.
- Osiewalski J. 2001, *Ekonometria bayesowska w zastosowaniach*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Osiewalski J., Pipień M. 2002, *Multivariate t-GARCH Models: Bayesian Analysis for Exchange Rates*, Modelling Economies in Transition. Proceedings of the Sixth AMFET Conference, ed. Welfe W., Łódź 2002, 151–167.
- Osiewalski J., Pipień M. 2003, *Bayesian Analysis and Option Pricing in Univariate GARCH Models with Asymmetries and GARCH-In-Mean Effects*, Przegląd Statystyczny, 50, 5–29.
- Osiewalski J., Pipień M. 2004 a, *Bayesian Comparison of Bivariate ARCH-type Models for the Main Exchange Rates in Poland*, Journal of Econometrics, 123, 371–391.
- Osiewalski J., Pipień M. 2004 b, *Bayesian Pricing of an European Call Option Using a GARCH Model with Asymmetries*, Acta Universitatis Lodziensis – Folia Oeconomica 177, 219–238.
- Osiewalski J., Steel M.F.J. 1993, *A Bayesian Perspective on Model Selection* (maszynopis), opublikowano w j. hiszpańskim: *Una perspectiva bayesiana en seleccion de modelos*, Cuadernos Economicos, no. 55/3.
- Pajor A. 2003a, *Bayesowska wycena europejskiej opcji kupna na podstawie modelu CSV*, Prace Naukowe Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, nr 1006, 173–185.
- Pajor A. 2003b, *Procesy zmienności stochastycznej SV w bayesowskiej analizie finansowych szeregów czasowych*, Monografie: Prace Doktorskie, nr 2, Wydawnictwo AE w Krakowie, Kraków.

- Pedersen A.R. 1995a, *A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations*, Scand. J. Statist., 22, 55–71.
- Pedersen A.R. 1995b, *Consistency and Asymptotic Normality of an Approximate Maximum Likelihood Estimator for Discretely Observed Diffusion Processes*, Bernoulli, 1 (3), 257–279.
- Pipień M. 2000, *Bayesowska analiza finansowych szeregów czasowych z wykorzystaniem procesów GARCH*, maszynopis rozprawy doktorskiej, Katedra Ekonometrii Akademii Ekonomicznej w Krakowie.
- Pipień M. 2003, *Zastosowanie wnioskowania bayesowskiego do wyceny opcji*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie, nr 628.
- Pipień M., Osiewalski J. 1999, *Monte Carlo z funkcją ważności w bayesowskiej analizie procesów GARCH* w: Dynamiczne modele ekonometryczne, Materiały zgłoszone na VI Ogólnopolskie Seminarium Naukowe, 7–9 września 1999, Uniwersytet Mikołaja Kopernika, Toruń.
- Pipień M., Pajor A. 2005, *Bayesowska analiza europejskiej opcji kupna i strategii delta neutralnej z wykorzystaniem procesów GARCH i CSV*, Przegląd Statystyczny, 52, 3, 37–63.
- Platen E. 1999, *An Introduction to Numerical Methods for Stochastic Differential Equations*, Acta Numerica, 197–246.
- Polson N.G., Roberts G.O. 1994, *Bayes Factors for Discrete Observations from Diffusion Processes*, Biometrika, 81, 1, 11–26.
- Posten H.O. 1989, *An Effective Algorithm for the Noncentral Chi-Squared Distribution Function*, American Statistical Association, vol. 43, No. 4, 261–269.
- Poulsen R. 1999a, *Approximate Maximum Likelihood Estimation of Discretely Observed Diffusion Processes*, Working Paper, University of Aarhus.
- Poulsen R. 1999b, *Excursion on Main Street – Excursions in Fixed Income Modelling*, PhD Thesis, University of Aarhus.
- Prohorov Yu.V., Rozanov Yu.A. 1969, *Probability Theory*, Springer-Verlag, Berlin – Heidelberg – New York.
- Protter P. 1990, *Stochastic Integration and Differential Equations*, Springer-Verlag, New York.
- Roberts G.O., Stramer O. 2001, *On Inference for Partially Observed Nonlinear Diffusion Models Using the Metropolis-Hastings Algorithm*, Biometrika, vol. 88, No. 3, 603–621.
- Rocznik Giełdowy 2005. *Dane statystyczne za rok 2004*. 2005, Giełda Papierów Wartościowych, Warszawa.
- Rogers L.C.G., Williams D. 1987, *Diffusions, Markov Processes and Martingales, Volume 2: Itô Calculus*, John Wiley & Sons.
- Rozanov J.A. 1974, *Wstęp do teorii procesów stochastycznych*, PWN, Warszawa.
- Saito Y., Mitsui T. 1996, *Stability Analysis of Numerical Schemes for Stochastic Differential Equations*, SIAM J Numer. Anal., vol. 88, No. 6, 2254–2267.

- Samuelson P.A. 1965, *Rational theory of warrant pricing*, *Industrial Management Rev.*, 6, 13–31.
- Schuss Z. 1989, *Teoria i zastosowania stochastycznych równań różniczkowych*, PWN, Warszawa.
- Shephard N., Pitt M. K. (1997), *Likelihood Analysis of Non-Gaussian Measurement Time Series*, *Biometrika*, 84, 653–667.
- Skorochod A.V. 1961, *On the Existence and Uniqueness of Solutions of Stochastic Differential Equations*, *Sib. Mat. Žurn.*, 2, 129–137 (w jęz. ros.).
- Smoluchowski M. 1906, *Zarys kinetycznej teorii ruchów Browna i roztworów mętnych*, *Rozprawy wydziału matematyczno-przyrodniczego Akademii Umiejętności w Krakowie* T XLVI, serja A, 257–281.
- Sobczyk K. 1996, *Stochastyczne równania różniczkowe*, PWN, Warszawa.
- Sørensen H. 2002, *Parametric Inference for Diffusion Processes Observed at Discrete Points in Time: A Survey*, Working Paper, University of Copenhagen.
- Stigum M. 1990, *The Money Market*, Dow Jones-Irwin.
- Stratonovich R.L. 1965, *Conditional Markov Processes and Their Application to the Theory of Optimal Control*, Elsevier, New York.
- Stratonovich R.L. 1966, *A New Representation of Stochastic Integrals and Equations*, *SIAM J. Control*, 4, 362–371.
- Stroock D.W., Varadhan S.R.S. 1979, *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin.
- Szreder M. 1994, *Informacje a priori w klasycznej i bayesowskiej estymacji modeli regresji*, Wydawnictwo Uniwersytetu Gdańskiego, Gdańsk.
- Talay D. 1984, *Efficient Numerical Schemes for the Approximation of Expectations of Functionals of the Solution of an SDE and Applications*, in *Filtering and Control of Random Processes*, Springer Lecture Notes in Control and Inform. Sci., vol. 61, 294–313.
- Talay D., Tubaro L. 1990, *Expansion of the Global Error for Numerical Schemes Solving Stochastic Differential Equations*, *Stochastic Analysis and Applications*, 8, 94–120.
- Tierney L. 1994, *Markov Chains for Exploring Posterior Distributions (with Discussion)*, *Annals of Statistics*, 22.
- Vasiček O.A. 1977, *An Equilibrium Characterization of the Term Structure*, *Journal of Financial Economics*, 5, 177–188.
- Wagner W., Platen E. 1978, *Approximation of Itô Integral Equations*. Preprint ZIMM, Akad. der Wiss. der DDR, Berlin.
- Weisstein E.W. 2003, *CRC Concise Encyclopedia of Mathematics*, Chapman & Hall/CRC.
- Wentzell A.D. 1980, *Wykłady z teorii procesów stochastycznych*, PWN, Warszawa.
- Weron A., Weron R. 1999, *Inżynieria finansowa*, WNT, Warszawa.
- Wiener N. 1923, *Differential Space*, *J. Math. Phys.*, 2, 131–174.

- Wilmott P. 1998, *Derivatives: The Theory and Practice of Financial Engineering*, John Wiley & Sons, Chichester.
- Wright D.J. 1974, *The Digital Simulation of Stochastic Differential Equations*, IEEE Trans. Autom. Control, AC-19, 75–76.
- Yu B., Mykland P. 1994, *Looking at Markov Samplers Through CUMSUM Paths Plots: A Simple Diagnostic Idea*, Technical Report 413, Department of Statistics, University of Carolina at Berkeley.
- Zellner A. 1971, *An Introduction to Bayesian Inference in Econometrics*, J. Wiley, New York.

