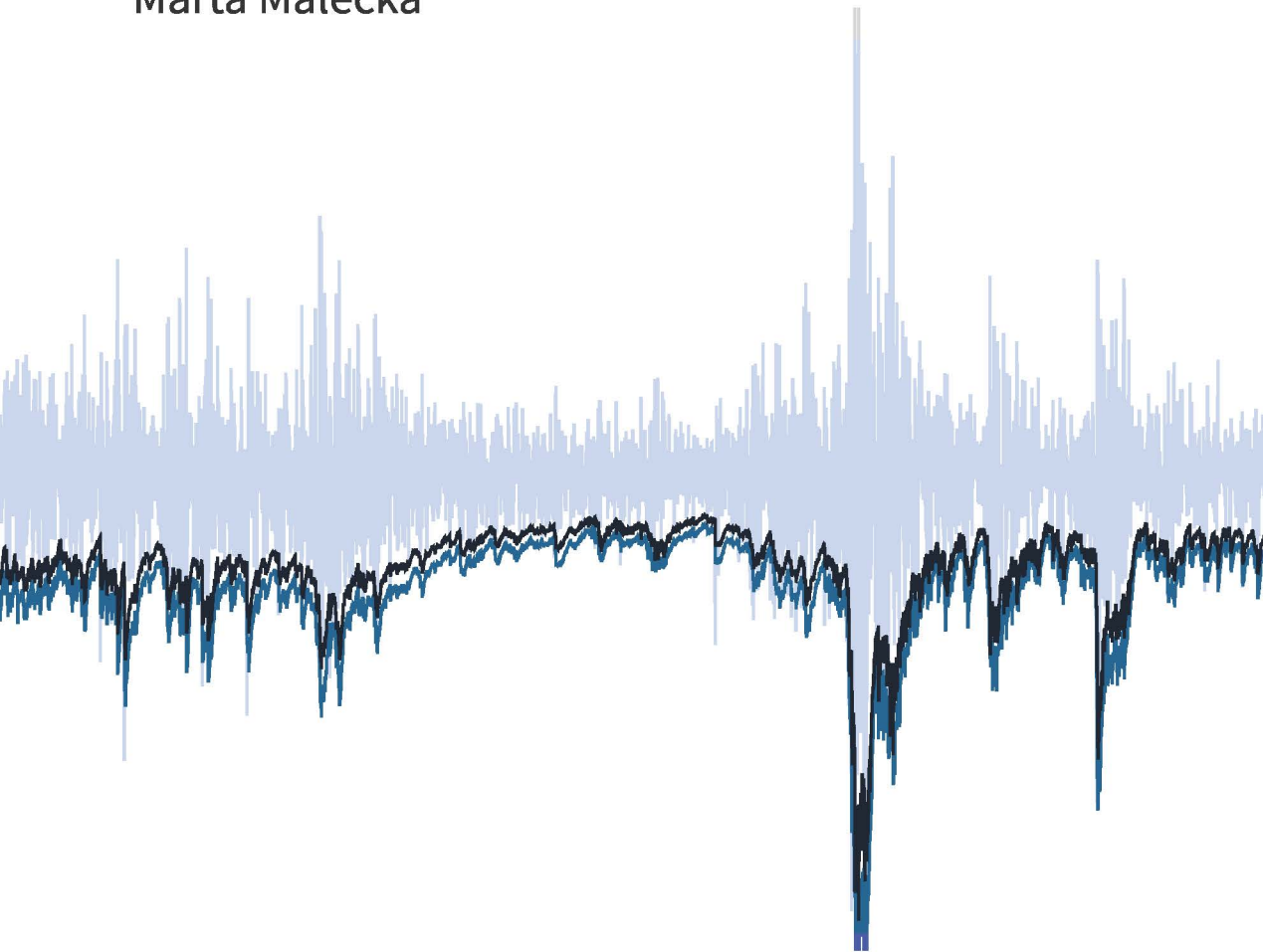


Ekonomia

Weryfikacja hipotez w ocenie ryzyka rynkowego

Marta Małecka



Weryfikacja hipotez w ocenie ryzyka rynkowego



WYDAWNICTWO
UNIwersytetu
ŁÓDZKIEGO

Ekonomia

Weryfikacja hipotez w ocenie ryzyka rynkowego

Marta Małecka



WYDAWNICTWO
UNIwersytetu
ŁÓDZKIEGO

ŁÓDŹ 2016

Marta Małecka – Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny
Katedra Metod Statystycznych, ul. Rewolucji 1905 r. nr 41, 90-214 Łódź

RECENZENT

Grażyna Dehnel

REDAKTOR INICJUJĄCY

Monika Borowczyk

KOREKTA TECHNICZNA

Elżbieta Rzymkowska

SKŁAD I ŁAMANIE

Marta Małecka

PROJEKT OKŁADKI

Stämpfli Polska Sp. z o.o.

Projekt został sfinansowany ze środków Narodowego Centrum Nauki
przyznanych na podstawie decyzji numer DEC-2013/11/N/HS4/03354.

© Copyright by Marta Małecka, Łódź 2016

© Copyright for this edition by Uniwersytet Łódzki, Łódź 2016

Wydrukowano z gotowych materiałów dostarczonych do Wydawnictwa UŁ

Wydane przez Wydawnictwo Uniwersytetu Łódzkiego
Wydanie I W.07800.16.0.M

Ark. druk. 11,25

ISBN 978-83-8088-536-3
e-ISBN 978-83-8088-537-0

Wydawnictwo Uniwersytetu Łódzkiego
90-131 Łódź, ul. Lindleya 8
www.wydawnictwo.uni.lodz.pl
e-mail: ksiegarnia@uni.lodz.pl
tel. (42) 665 58 63

<https://doi.org/10.18778/8088-536-3>

Spis treści

Wstęp	7
Rozdział 1	
Pojęcie i statystyczna ocena ryzyka rynkowego	13
1.1. Ryzyko rynkowe i jego rodzaje	14
1.1.1. <i>Wprowadzenie pojęcia ryzyka rynkowego</i>	14
1.1.2. <i>Rodzaje ryzyka rynkowego</i>	16
1.1.3. <i>Metody kwantyfikacji ryzyka</i>	19
1.2. Miara <i>VaR</i>	20
1.2.1. <i>Wprowadzenie pojęcia VaR</i>	20
1.2.2. <i>Kryteria oceny modeli VaR</i>	23
1.3. Koherentne miary ryzyka rynkowego	26
1.3.1. <i>Aksjomatyczna definicja ryzyka</i>	26
1.3.2. <i>Wprowadzenie pojęcia ES i innych miar koherentnych</i>	30
1.3.3. <i>Kryteria oceny modelu ES</i>	35
Rozdział 2	
Testy wartości zagrożonej (<i>VaR</i>) i oczekiwanego niedoboru (<i>ES</i>)	39
2.1. Testy <i>VaR</i> oparte na procesach Bernoulliego i Markowa	41
2.1.1. <i>Klasyczne testy bezwarunkowego rozkładu wyjątków VaR</i>	41
2.1.2. <i>Modyfikacje podejścia do testowania bezwarunkowego rozkładu wyjątków VaR</i>	43
2.1.3. <i>Testy warunkowego rozkładu wyjątków VaR</i>	45
2.2. Testy oparte na procesie odległości wyjątków <i>VaR</i>	53
2.2.1. <i>Testy braku pamięci</i>	53
2.2.2. <i>Testy parametrów regresji z rozkładem wykładniczym</i>	56
2.3. Testy bazujące na gęstości rozkładów	57
2.3.1. <i>Testy zgodności</i>	57
2.3.2. <i>Testy ilorazu wiarygodności</i>	59
2.3.3. <i>Testy gęstości spektralnej</i>	61

2.4. Wielowymiarowe testy <i>VaR</i>	66
2.4.1. <i>Test Ljunga-Boxa dla wielu poziomów VaR</i>	66
2.4.2. <i>Propozycja zastosowania wielowymiarowego testu macierzy korelacji</i>	70
2.5. Testy <i>ES</i>	73
2.5.1. <i>Testy parametryczne</i>	73
2.5.2. <i>Testy nieparametryczne</i>	79
Rozdział 3	
Ocena własności statystycznych testów wartości zagrożonej (<i>VaR</i>) i oczekiwanego niedoboru (<i>ES</i>)	85
3.1. Projektowanie eksperymentów do oceny rozmiaru	87
3.1.1. <i>Eksperymenty oparte na procesie Bernoulliego</i>	87
3.1.2. <i>Zastosowanie metody Monte Carlo do uzyskania testu dokładnego</i>	88
3.2. Projektowanie eksperymentów do oceny mocy	90
3.2.1. <i>Eksperymenty wykorzystujące proces GARCH</i>	90
3.2.2. <i>Propozycja eksperymentu opartego na procesie BGAR</i>	93
3.2.3. <i>Propozycja eksperymentu opartego na procesie BGMA</i>	95
3.2.4. <i>Propozycja eksperymentu opartego na procesie Markowa</i>	95
3.3. Wyniki symulacyjnej oceny rozmiaru i mocy testów <i>VaR</i>	96
3.3.1. <i>Ocena rozmiaru testów VaR</i>	96
3.3.2. <i>Ocena mocy testów VaR</i>	110
3.3.3. <i>Wybór optymalnych testów VaR</i>	127
3.4. Wyniki symulacyjnej oceny rozmiaru i mocy testów <i>ES</i>	129
3.4.1. <i>Ocena rozmiaru testów ES</i>	129
3.4.2. <i>Ocena mocy testów ES</i>	132
3.4.3. <i>Wybór optymalnych testów ES</i>	136
Rozdział 4	
Weryfikacja modeli ryzyka na przykładzie szeregów empirycznych	139
4.1. Opis badania empirycznego	140
4.1.1. <i>Opisowa analiza szeregów czasowych</i>	140
4.1.2. <i>Zastosowane metody wnioskowania statystycznego</i>	147
4.2. Wyniki badania empirycznego dla rynku finansowego	153
4.2.1. <i>Ocena modeli VaR i ES dla indeksu WIG20</i>	153
4.2.2. <i>Ocena modeli VaR i ES dla indeksu S&P500</i>	158
4.3. Wyniki badania empirycznego dla rynku towarowego	163
4.3.1. <i>Ocena modeli VaR i ES dla indeksu Gold Bullion LBM</i>	163
4.3.2. <i>Ocena modeli VaR i ES dla indeksu S&P GSCI Wheat</i>	167
Podsumowanie	173

Wstęp

Rozwój współczesnej teorii ryzyka wiąże się z potrzebą kwantyfikacji i kontroli ryzyka instytucji w tym sensie, by zapewnić odpowiednie zabezpieczenie kapitałowe z tytułu poszczególnych rodzajów ryzyka oraz chronić przed utratą płynności. Z tego względu aktualne badania z zakresu zarządzania ryzykiem rynkowym koncentrują się na miarach zagrożenia, których konstrukcja ma na celu oszacowanie potencjalnej wielkości straty, związanej ze zmianami cen rynkowych. Dynamiczny rozwój statystycznych metod oceny ryzyka rynkowego, w szczególności metod bazujących na koncepcji pomiaru zagrożenia, nastąpił począwszy od lat dziewięćdziesiątych ubiegłego wieku. Punktem przełomowym było wprowadzenie miary *VaR* (wartość zagrożona, ang. *value-at-risk*) przez JP Morgan w 1994 roku [JP Morgan 1994]. W świetle praktyki rynkowej, na przestrzeni dwóch ostatnich dekad *VaR* stała się najpopularniejszą miarą ryzyka rynkowego, stanowiącą filar współczesnej koncepcji pomiaru ryzyka również w szerszym znaczeniu np. kredytowego lub operacyjnego. Modele *VaR* powszechnie używane są przez międzynarodowe instytucje finansowe oraz znajdują umocowanie w zaleceniach opracowywanych przez banki centralne oraz instytucje nadzorcze zajmujące się zarządzaniem ryzykiem [Basel Committee on Banking Supervision 1996, 2005a, 2005b].

Ze względu na rosnącą popularność, koncepcja *VaR* stała się impulsem do rozwoju nowych miar ryzyka, opartych na idei pomiaru zagrożenia. Począwszy od końca ubiegłego stulecia, stworzone zostały podwaliny aksjomatycznej teorii ryzyka i wprowadzona została definicja miary koherentnej. Za główne uzupełnienie *VaR* uznaje się koherentną miarę *ES* (oczekiwanego niedoboru, ang. *expected shortfall*), której celem jest informowanie o oczekiwanym rozmiarze straty w sytuacji przekroczenia *VaR*. Od 2012 roku *ES* stanowi podstawę systemów zarządzania ryzykiem w świetle zaleceń międzynarodowego nadzoru bankowego [Basel Committee on Banking Supervision 2012-2016].

Szybki rozwój metod związanych z estymacją modeli opartych na mierze VaR , powstanie modeli ES oraz ich coraz szersze zastosowanie wywołują potrzebę tworzenia nowych jak również usystematyzowania i oceny istniejących procedur weryfikacji modeli ryzyka. Różnorodność technik opartych na koncepcjach VaR i ES przesądza o konieczności stosowania uniwersalnych metod wnioskowania statystycznego, pozwalających na obiektywną ocenę konkurencyjnych modeli. Ze względu na brak obserwowalności realizacji VaR i ES ocena modeli ryzyka nie może opierać się typowych metodach, porównujących realizacje zmiennych ex-post z wartościami prognoz dokonywanych ex-ante. Konieczne jest konstruowanie testów statystycznych z wykorzystaniem procesów zawierających informację o nieobserwowalnej zmiennej, np. w oparciu o porównanie prognoz ryzyka i zrealizowanego ciągu wartości stopy zwrotu.

W rozwoju literatury poświęconej statystycznym metodom oceny modeli opartych na mierze VaR można wyróżnić trzy nurty badawcze. Historycznie najstarsze podejście, oparte na porównywaniu procesu przekroczeń VaR z procesem Bernoulliego, sprawdza hipotezę mówiącą, że udział przekroczeń w próbie powinien korespondować z założonym poziomem tolerancji [Kupiec 1995]. Jako rozwinięcie testu Kupca, zaproponowano metodę badania niezależności przekroczeń VaR , opartą na warunkowym rozkładzie prawdopodobieństwa, uwzględniającą systematyczny napływ nowych informacji w kolejnych okresach [Christoffersen 1998]. Rozszerzeniem tego nurtu badań były testy posiadające zdolność wykrywania bardziej złożonych form zależności niż autokorelacja pierwszego rzędu. W dynamicznym teście kwantylowym, do oceny autokorelacji między wyjątkami VaR , zastosowano liniowy model autoregresji [Engle i Manganelli 2004]. W celu sprawdzania procesu wyjątków VaR pod kątem autokorelacji dowolnego rzędu, zaproponowano wykorzystanie klasycznych testów opartych na statystyce Pearsona [Wang 2010], współczynnikach autokorelacji różnych rzędów [Berkowitz, Christoffersen i Pelletier 2011] oraz innowacyjnych w naukach ekonomicznych testów opartych na funkcji gęstości spektralnej [Berkowitz, Christoffersen i Pelletier 2011].

Inny kierunek badań nad weryfikacją modeli VaR obejmuje tzw. testy odległości między przekroczeniami. Koncepcja wykorzystania procesu odległości wywodzi się z testu $TUFF$ (ang. *time-until-first-failure*), w którym odwrotność czasu do pierwszego wyjątku VaR wykorzystana została jako oszacowanie prawdopodobieństwa sukcesu w procesie Bernoulliego [Kupiec 1995]. W uogólnionej wersji testu zaproponowano wykorzystanie wszystkich dostępnych odległości między wyjątkami [Haas 2001]. Transformacja procesu wyjątków w proces odległości umożliwiła badanie zależności między kolejnymi przekroczeniami VaR z wykorzystaniem własności rozkładów geo-

metrycznego i wykładniczego. W ramach tego nurtu badań zaproponowano podejście regresyjne ze składnikiem losowym o rozkładzie wykładniczym [Engle i Russel 1998] oraz test restrikcji parametru rodziny rozkładów, obejmujących rozkład wykładniczy [Christoffersen i Pelletier 2004].

Trzecia grupa testów *VaR* wiąże się z badaniem zgodności funkcji gęstości stopy zwrotu. W przeciwieństwie do testów weryfikujących modele *VaR* dla jednego poziomu tolerancji, przyjęto założenie, że empiryczne percentyle odpowiadające prognozom *VaR* na wszystkich poziomach tolerancji powinny być niezależne i równomiernie rozłożone, co umożliwiło wykorzystanie większej ilości informacji z próby [Crnkovic i Drachman 1995, 1996]. W późniejszej literaturze opracowano uzupełnienie formalnej procedury testowania zgodności rozkładów za pomocą analizy graficznej [Diebold, Gunther i Tay 1998]. Podejmując problem dużego rozmiaru próby, wymaganego do przeprowadzenia nieparametrycznych procedur testowania zgodności, zaproponowano transformację empirycznego szeregu czasowego, pozwalającą na skonstruowanie testu parametrycznego [Berkowitz 2001]. W odpowiedzi na zarzut jednakowego traktowania percentyli ze środka rozkładu jak i ogonów, powstały modyfikacje testów zgodności, pozwalające na wybór obszaru funkcji gęstości. Zaproponowano wielowymiarowe rozwinięcie testu niezależności *VaR*, oparte na współczynnikach autokorelacji, które umożliwiło jednoczesne testowanie modelu *VaR* dla wielu, dowolnie ustalonych poziomów tolerancji [Hurlin i Tokpavi 2007].

W związku z opracowaniem aksjomatycznej definicji ryzyka [Artzner et al. 1999] oraz wprowadzeniem koherentnej miary *ES* [Acerbi, Nardio i Sirtori 2001], powstały propozycje procedur, weryfikujących modele oparte na miarach koherentnych. Zaproponowano podejście oparte na tzw. resztach z procesu przekroczeń w połączeniu z wykorzystaniem metod bootstrapowych [McNeil i Frey 2000]. Poprzez zastosowanie uciętego rozkładu normalnego, wprowadzono test *ES* skonstruowany na zasadzie badania rozbieżności między ogonami rozkładów [Berkowitz 2001]. Omijając problem nieznaności parametrycznej postaci ogona stopy zwrotu, zaproponowano test *ES*, wykorzystujący technikę punktu siodłowego, w którym przybliżone *p*-wartości wyznaczone są w oparciu o rozwinięcie funkcji generującej momenty w szereg Taylora [Wong 2008]. W ślad za regresyjnymi testami *VaR*, opracowano statystykę testową dla modelu *ES*, opartą na badaniu istotności regresji liniowej [Christoffersen 2012]. Jednak ze względu na brak znajomości rozkładów statystyk opartych na skrajnych obserwacjach testowanie *ES* wciąż pozostaje tematem dyskusyjnym [np. Carver 2013, Carver 2014, Hull i White 2014].

Z tego względu, mimo powszechnej akceptacji tej miary, procedury weryfikacji *ES* nie zostały włączone do zaleceń nadzoru bankowego w ramach *Basel III* [Basel Committee on Banking Supervision 2016].

Celem pracy było porównanie własności statystycznych testów stosowanych do oceny modeli ryzyka rynkowego oraz przedstawienie propozycji nowych lub zmodyfikowanych testów. Zgodnie ze współczesną teorią pomiaru ryzyka badanie koncentrowało się na miarach zagrożenia *VaR* i *ES*. Punktem wyjścia do oceny własności statystycznych testów było sformułowanie uzasadnionych ekonomicznie kryteriów oceny modelu ryzyka. Określone zostały postulaty decydujące o uznaniu model za dobry. Kryteria oceny modelu przełożono następnie na język możliwych do testowania hipotez statystycznych. Analizę porównawczą poprzedzono prezentacją aktualnych badań teoretycznych związanych z określeniem miary ryzyka oraz kompleksowym przeglądem procedur testowych *VaR* i *ES*. Wskazano modyfikacje, dzięki którym można uzyskać większą dokładność lub poprawę mocy oraz przedstawiono nowe propozycje testowania modeli ryzyka. Porównanie teoretycznych własności testów zostało uzupełnione analizą porównawczą wyników badania empirycznego, dotyczącego indeksów rynków kapitałowych i towarowych.

Uwzględnienie aktualnych potrzeb rynkowych oraz najnowszych zaleceń międzynarodowego nadzoru bankowego wymaga testowania *VaR* na dwóch, niskich poziomach tolerancji. W pracy zaproponowano test, którego przewagą w stosunku do standardowych procedur jest określoność – zatem możliwość bezpośredniego stosowania – w przypadku szeregu, w którym nie zaobserwowano żadnego przekroczenia *VaR*. Ma to praktyczne znaczenie przy testowaniu *VaR* przy skrajnych poziomach tolerancji, zalecanych obecnie przez Komitet Bazylejski. Ponadto analizę testów *VaR* uzupełniono propozycją testu wielowymiarowego, co pozwoliło na jednoczesne testowanie modelu dla wielu poziomów tolerancji z zachowaniem pełnej kontroli błędu pierwszego rodzaju. Dzięki wielowymiarowemu podejściu wykorzystano większą ilość informacji z próby oraz zapewniono elastyczność w zarządzaniu ryzykiem, polegającą na możliwości dostosowania testowanego obszaru rozkładu do potrzeb analizy.

W porównaniu z weryfikacją modeli *VaR*, testowanie modeli *ES* jest zadaniem bardziej wymagającym z punktu widzenia teorii statystyki, ze względu na nieznaną odpowiednich rozkładów statystyk testowych z próby. W tym kontekście przedstawiono możliwość uzupełnienia testów korzystających z rozkładów asymptotycznych testem bootstrapowym oraz testem, w którym przybliżone *p*-wartości wyznaczane są w oparciu o rozwinięcie funkcji generującej momenty w szereg Taylora. Zaproponowano przeprowadzenie standaryza-

cji zmiennej objaśnianej w teście regresyjnym *ES*, co umożliwiło uzyskanie przybliżonej stacjonarności składnika losowego.

Kluczową częścią pracy było przeprowadzenie oceny i porównania rozważanych testów pod kątem rozmiaru i mocy. Główne problemy metodologiczne wiązały się z nieznaną dokładnością rozkładów statystyk testowych oraz projektowaniem odpowiednich eksperymentów symulacyjnych. Przy ocenie mocy, ze względu na brak analitycznej postaci rozkładu statystyk testowych dla prób skończonych, wykorzystano technikę zrandomizowanego testu Monte Carlo, gwarantującą zachowanie założonego rozmiaru. Dzięki temu osiągnięto porównywalność ocen mocy rozważanych testów. W obszernym badaniu symulacyjnym, poza typowymi eksperymentami opartymi na procesie *GARCH*, zaproponowano wykorzystanie różnych schematów generowania danych cechujących się określoną strukturą autokorelacji. Propozycje eksperymentów zostały dobrane w taki sposób by reprezentowały zjawisko gromadzenia wariancji, uznawane za główny problem związany z kontrolą ryzyka rynkowego. W tym celu wykorzystano procesy *BGAR*, *BGMA* i proces Markowa. Praca stanowi zatem rozwinięcie dotychczasowych analiz mocy testów *VaR* i *ES*, obejmujące większą klasę procesów ekonomicznych niż wcześniejsze badania, odnoszące się głównie do danych finansowych.

W wyniku analizy rozmiaru i mocy, wyłoniono grupy testów gwarantujące najwyższy stopień kontroli nad błędem pierwszego rodzaju oraz cechujące się wyższą niż w innych grupach skutecznością wykrywania autokorelacji przekroczeń *VaR* oraz błędnych modeli *ES*. Wnioski dotyczące oceny proponowanych testów statystycznych zostały wykorzystane następnie w części empirycznej, stanowiącej czwarty rozdział pracy. Wykorzystując cztery rzeczywiste szeregi czasowe, pochodzące z rynku finansowego i towarowego, przedstawiono przykład zastosowania wybranych procedur testowych do oceny konkurencyjnych modeli ryzyka.

Praca składa się z czterech rozdziałów. Pierwszy rozdział poświęcony został wprowadzeniu pojęć wykorzystywanych w dalszej części pracy, omówieniu aksjomatycznej definicji ryzyka oraz sformułowaniu ekonomicznych uzasadnionych kryteriów oceny modeli ryzyka rynkowego. W drugim rozdziale zebrano dotychczasowe wnioski dotyczące metod weryfikacji hipotez dotyczących modeli *VaR* i *ES*, wskazano możliwe modyfikacje oraz uzupełniono o nowe propozycje. Rozdział trzeci poświęcony został analizie porównawczej rozmiaru i mocy rozważanych testów statystycznych. Czwarty rozdział pracy objął badanie empiryczne, wykorzystujące wnioski z wcześniejszych jej części oraz stanowiące ilustrację zastosowania omówionych wcześniej metod wnioskowania statystycznego.

Rozdział 1

Pojęcie i statystyczna ocena ryzyka rynkowego

Metody związane z oceną ryzyka rynkowego zajmują obszerne miejsce w literaturze z zakresu ekonomii. Szeroka dyskusja odnosząca się do jego pomiaru oraz kontroli wywołana została przez zmiany światowej gospodarki, które, począwszy od lat 70-tych XX wieku, spowodowały skokowy wzrost zmienności cen na rynku walutowym, kapitałowym i towarowym. Wzrost ryzyka rynkowego spowodował, że wcześniej wykorzystywane metody jego analizy stały się niewystarczające. Zrodziło to potrzebę opracowania całych systemów analizy i zarządzania ryzykiem, a co za tym idzie również odpowiednich miar, które pozwalałyby na ocenę ryzyka pozycji zajmowanych w różnych instrumentach finansowych a nawet różnych rodzajów ryzyka. Od pierwszej połowy lat 90-tych, za podstawową miarę ryzyka rynkowego uznawana jest miara *VaR*, definiowana jako kwantyl rozkładu straty lub stopy zwrotu. W ślad za rozwojem aksjomatycznej teorii ryzyka na przełomie XX i XXI wieku, powstały nowe propozycje miar, nazwanych miarami koherentnymi. Za główne uzupełnienie miary *VaR* uznaje się miarę koherentną *ES*, informującą o oczekiwanym rozmiarze straty w sytuacji ekstremalnych wahań na rynku.

W niniejszym rozdziale wprowadzono pojęcie ryzyka rynkowego, przedstawiono jego rodzaje oraz przytoczono podstawowe fakty dotyczące ewolucji tego pojęcia. W dalszej części rozdziału zaprezentowano definicje miar *VaR*, *ES* oraz innych miar powstałych na gruncie aksjomatycznej teorii ryzyka. Wprowadzono proces wyjątków *VaR*, stanowiący podstawę oceny modeli ry-

zyka. Zagadnienia związane z przedstawieniem powyższych definicji zakończono sformułowaniem ekonomicznie uzasadnionych kryteriów oceny modeli *VaR* i *ES*. Kryteria te przełożono na język hipotez statystycznych, co stanowiło punkt wyjścia do rozważań związanych z testowaniem modeli ryzyka, zawartych w kolejnych rozdziałach pracy.

1.1. Ryzyko rynkowe i jego rodzaje

1.1.1. Wprowadzenie pojęcia ryzyka rynkowego

Ryzyko rynkowe określa się w literaturze jako ryzyko wynikające ze zmian cen rynkowych, przy czym cena rynkowa rozumiana jest jako jedna z czterech kategorii: cena towaru, stopa procentowa, kurs walutowy lub kurs akcji. Wzrost zainteresowania ryzykiem rynkowym w naukach ekonomicznych nastąpił w latach 70-tych ubiegłego wieku, wraz z szokowymi zmianami ówczesnej gospodarki jak m.in. krach systemu z Bretton Woods, pierwszy kryzys naftowy oraz deregulacja stóp procentowych w Stanach Zjednoczonych i innych krajach.

Ustalone w 1944 roku w Bretton Woods zasady określały sposób funkcjonowania międzynarodowego systemu walutowego do 1971 roku. System ten był oparty na obowiązku utrzymywania przez państwa stałych kursów walutowych, które posiadały parytet określony w dolarach amerykańskich lub złocie. Zawieszenie w 1971 roku przez Rezerwę Federalną wymienialności dolarów na złoto spowodowało destrukcję zasad z Bretton Woods. Stałe kursy walutowe zostały zastąpione systemem kursów płynnych, co doprowadziło do gwałtownego wzrostu ryzyka walutowego.

Drugim znaczącym czynnikiem wzrostu ryzyka rynkowego na rynku światowym była rezygnacja przez Stany Zjednoczone w latach 70-tych ubiegłego stulecia z odgórnego limitowania wysokości oprocentowania oferowanego przez banki, ze względu na niekorzystny wpływ takiej regulacji na konkurencyjność sektora bankowego. Dało to początek konkurencji cenowej między bankami oraz znacznym wahaniom stopy procentowej, uzależnionym od zmian stopy dyskontowej oferowanej przez Rezerwę Federalną. Efektem był wzrost ryzyka związanego ze stopą procentową.

Kryzys naftowy roku 1973 był kolejnym czynnikiem wzrostu ryzyka rynkowego. Konflikty na Bliskim Wschodzie doprowadziły do zakazu eksportu ropy naftowej przez kraje OPEC do Stanów Zjednoczonych oraz do czterokrotnego wzrostu cen ropy naftowej w ciągu kilku miesięcy. Konsekwencją

tych wydarzeń, jak również drugiego kryzysu naftowego z lat 1978-80, były duże wahania cen na rynkach towarowych.

Rezultatem zmian na rynku światowym, które doprowadziły do nasilenia ryzyka rynkowego, był trend do zaostrzania wymogów nadzoru bankowego. Ewolucja ryzyka rynkowego była związana z przemianami w sektorze bankowym, gdzie ostrzejsze wymogi dotyczące proporcji kapitałów własnych do obcych wymusiły ograniczenie akcji depozytowo-kredytowej. Zrodziło to większe oczekiwania dotyczące zysku przy tej samej skali prowadzonej działalności. Efektem zmian w światowym systemie gospodarczo-finansowym było przekształcenie głównego profilu działalności bankowej z działalności kredytowej w kierunku działalności handlowej, polegającej na nabywaniu instrumentów finansowych w celu osiągnięcia zysków spowodowanych zmianami cen tych instrumentów.

Zmienność warunków gospodarowania, będąca skutkiem przemian lat 70-tych, stanowiła bodziec dla rynku finansowego do rozwijania oferty produktowej nakierowanej na hedging bądź spekulację na ryzyku rynkowym. Dalsza ewolucja ryzyka rynkowego związana była z rozwojem rynków instrumentów pochodnych. Rozwój oferty instrumentów finansowych, których pierwotnym celem było ograniczanie ryzyka, umożliwił podejmowanie ryzyka w celach spekulacyjnych.

Definicje ryzyka rynkowego występujące w literaturze w sposób jednoznaczny określają jego źródło jako zmiany cen rynkowych, natomiast nie zawsze są zgodne co do zakresu pojęciowego tej kategorii. Definicja Komitetu Bazylejskiego¹ mówi, że ryzyko rynkowe jest to „ryzyko strat na pozycjach bilansowych i pozabilansowych wynikających ze zmian cen rynkowych” [Basel Committee on Banking Supervision 1996, 2005a, 2005b]. Według innej klasyfikacji rodzajów ryzyka, występującej w literaturze, określane jest jako „ryzyko straty w wyniku zmiany wartości aktywów będących przedmiotem obrotu i znajdujących się w posiadaniu przedsiębiorstwa” [Tarczyński i Mojsiewicz 2001, s. 20]. W terminologii bankowej „ryzyko ceny jest ryzykiem wystąpienia niekorzystnych zmian ceny rynkowej w czasie, gdy bank zajmuje spekulacyjną lub związaną z obsługą klienta pozycję netto w towarach, stopach procentowych czy też zmienności implikowanej w opcjach” [Riehl 2001, s. 42].

Pomiędzy pierwszą a drugą definicją występuje różnica dotycząca przedmiotu, który narażony jest na ryzyko rynkowe, ujmowanego jako pozycje

¹ Międzynarodowy organ działający przy Banku Rozliczeń Międzynarodowych, zajmujący się opracowywaniem standardów nadzoru bankowego.

bilansowe i pozabilansowe lub aktywa przedsiębiorstwa. Komitet Bazylejski określa więc przedmiot ryzyka bardziej ogólnie. Stwierdzenie, że na ryzyko rynkowe narażone są aktywa wyjmuje spod rozważań pasywa oraz pozycje pozabilansowe. Ponadto według drugiej definicji przedmiotem rozważań są jedynie aktywa przeznaczone do obrotu oraz znajdujące się w posiadaniu przedsiębiorstwa, czyli tzw. portfel handlowy. Portfel handlowy obejmuje pozycje przeznaczone do odsprzedaży, otwierane z zamiarem uzyskania krótkoterminowych korzyści płynących z aktualnych lub spodziewanych różnic pomiędzy cenami kupna i sprzedaży. Nieuwzględniony w definicji portfel lokacyjny tworzą pozycje otwierane z zamiarem przetrzymania do terminu płatności oraz czerpania do tego czasu korzyści w postaci odsetek. Głównymi pozycjami na kontach lokacyjnych są niezbywalne instrumenty oparte na stopie procentowej (kredyty i depozyty). Dotychczasowe regulacje Komitetu Bazylejskiego nakładające obciążenia kapitałowe z tytułu podejmowanego ryzyka rynkowego odnosiły się wyłącznie do portfela handlowego, jednak definicja Komitetu traktuje ryzyko to w sposób szeroki, z uwzględnieniem portfela lokacyjnego.

W trzeciej przytoczonej definicji znajduje się odwołanie do przyczyn ryzyka rynkowego oraz podkreślone zostają pozycje netto w towarach, stopach procentowych czy też zmienności implikowanej w opcjach jako warunek konieczny wystąpienia tego ryzyka. Ryzyko rynkowe nazwane zostaje ryzykiem ceny, co odzwierciedla fakt, że wszystkie źródła tego ryzyka – stopy procentowe, kursy akcji, kursy walut, ceny towarów – są cenami rynkowymi.

1.1.2. Rodzaje ryzyka rynkowego

Odnosząc się do źródeł ryzyka rynkowego, wyróżnia się cztery jego podkategorie związane z różnymi rodzajami cen [por. Jajuga, Jajuga 2011, s. 180]:

- ryzyko stopy procentowej,
- ryzyko cen akcji,
- ryzyko walutowe,
- ryzyko cen towarów.

Instrumentami służącymi do otwarcia pozycji w stopie procentowej są, według Komitetu Bazylejskiego, „wszystkie instrumenty dłużne o stałej i zmiennej stopie procentowej oraz instrumenty, które zachowują się podobnie do nich.” Można wyróżnić trzy grupy takich instrumentów [Basel Committee on Banking Supervision 2005a]:

- niezbywalne aktywa/pasywa – depozyty lub kredyty,
- zbywalne aktywa/pasywa – dłużne papiery wartościowe,
- instrumenty pochodne oparte na stopie procentowej czyli:
 - walutowe lub procentowe transakcje forward,
 - walutowe lub procentowe transakcje futures,
 - swapy,
 - opcje na jeden z powyższych instrumentów.

Przez ryzyko cen akcji rozumie się ryzyko wynikające ze zmian ich kursów. Pojęcie to odnosi się także do innych instrumentów kapitałowych, które są podobne do akcji w tym sensie, że ich ceny kształtują się w oparciu o te same mechanizmy. Pojęcie ryzyka rynkowego obejmuje wyłącznie ryzyko systematyczne, które wynika z ogólnej tendencji spadkowej lub wzrostowej na rynku instrumentów kapitałowych, nie dotyczy natomiast ryzyka emitenta, inaczej kredytowego lub specyficznego, które polega na tym, że cena akcji może spaść na skutek czynników związanych z danym emitentem (np. zaostrożenie konkurencji w branży, wadliwe zarządzanie i inne). W przypadku ryzyka specyficznego potencjalne zagrożenie dotyczy tylko określonego typu akcji i można się go pozbyć przez dobór wiarygodnych emitentów oraz dywersyfikację portfela – jest to przedmiot badań teorii portfela. Ryzyko rynkowe powstaje poprzez otwarcie pozycji w instrumentach, które wykazują zachowanie rynkowe podobne do akcji, do których w szczególności zalicza się [Basel Committee on Banking Supervision 2005a]:

- akcje zwykłe bez względu na prawo głosu,
- zamienne papiery wartościowe, które zachowują się podobnie jak akcje,
- zobowiązania do nabycia lub sprzedaży akcyjnych papierów wartościowych,
- instrumenty pochodne oparte na akcjach:
 - kontrakty forward na akcje lub indeksy,
 - kontrakty futures na akcje lub indeksy,
 - opcje na akcje lub indeksy.

Ryzyko walutowe odnosi się do zmian wartości otwartych pozycji przedsiębiorstwa, zachodzących na skutek wahań kursów walut. Do otwierania pozycji walutowych może służyć każdy rodzaj aktywów, pasywów czy też pozycji pozabilansowych, który denominowany jest w walucie obcej [Basel Committee on Banking Supervision 2005a]:

- pozycje spot bilansu (wszystkie pozycje aktywów oraz wszystkie pozycje pasywów denominowane w walucie obcej),
- pozycje terminowe (terminowe transakcje związane z walutami np. długo-terminowe kontrakty na nabycie środków trwałych denominowane w walucie obcej, walutowe kontrakty forward i futures),
- transakcje, które nie zostały jeszcze zafakturowane,
- przyszłe transakcje, które pojawią się jeśli dojdzie do zaplanowanej działalności,
- przyszłe transakcje, które mogą się pojawić w określonych warunkach w przyszłości (np. wynikające z wygrania przetargu lub udzielenia gwarancji).

Towar jest czwartą kategorią, która może służyć podejmowaniu ryzyka rynkowego. Według definicji Komitetu Bazylejskiego: towar jest to „fizyczny produkt, który jest w obrocie lub nadaje się do niego na rynku wtórnym, np. produkty rolne, surowce mineralne (włącznie z ropą naftową) oraz metale szlachetne” [Basel Committee on Banking Supervision 2005a]. Ryzyko cen dla przedsiębiorstw handlowych wynika ze zmian cen spot towarów oraz z zawierania terminowych kontraktów na towary. Na rynkach towarowych zwykle istnieją pewne konwencje regulujące obrót. Towary nabywane są na zasadzie kontraktów bieżących, gdzie stosowane są aktualne ceny rynkowe lub na rynkach, gdzie typowe są kontrakty długoterminowe, np. jednoroczne, a przy ich podpisywaniu cena sprzedaży nie jest znana lecz ustalany jest sposób jej kalkulacji w przyszłości. Ceną referencyjną może być przykładowo średnia cena z danych notowań oficjalnych z określonego okresu (np. na rynku miedzi). W przypadku podmiotów handlowych podejmowanie ryzyka rynkowego wynika z głównego nurtu prowadzonej działalności, natomiast celowe operacje spekulacyjne na cenach towarów prowadzone są przez banki oraz inne instytucje finansowe.

Niektórzy autorzy jako piątą kategorię ryzyka rynkowego wymieniają ryzyko zmienności implikowanej w opcjach, co wiąże się ze specyfiką gry rynkowej podejmowanej za pomocą tego instrumentu. Wyróżnia się dwa rodzaje opcyjnych strategii spekulacyjnych:

- spekulacja na cenach instrumentów bazowych,
- spekulacja na zmienności implikowanej.

Spekulacja na cenach instrumentów bazowych może służyć do podejmowania każdego z podstawowych rodzajów ryzyka rynkowego – ryzyka stóp

procentowych, akcji, walut lub cen towarów. Spekulacja na zmienności implikowanej polega na podejmowaniu specyficznego rodzaju ryzyka związanego ze zmiennością cen. Ponieważ ryzyko zmienności cen nie klasyfikuje się do żadnej z powyższych kategorii ryzyka, proponuje się uznanie zmienności za piątą kategorię cen rynkowych [Riehl 2001, s. 72].

Zmienność cen instrumentu bazowego w okresie od nabycia opcji do momentu jej wygaśnięcia jest parametrem mającym wpływ na ustalanie ceny opcji. Spekulacja na zmienności implikowanej opiera się na utrzymywaniu pozycji delta-neutralnej, czyli zabezpieczeniu się przed niepożądanymi ruchami cen instrumentów bazowych i nie zakłada, że spekulant jest w stanie przewidzieć przyszłe kształtowanie się cen instrumentu bazowego. Jeśli zmienność faktyczna od momentu zakupu do wygaśnięcia opcji jest równa zmienności założonej przy ustalaniu ceny, wówczas premia otrzymana przez sprzedawcę dokładnie pokrywa koszt zabezpieczania się przed niepożądanymi zmianami cen instrumentu bazowego. Z punktu widzenia nabywcy opcji, przychody otrzymywane dzięki utrzymywaniu pozycji delta-neutralnej równoważą koszt w postaci zapłaconej premii. Gra na zmienności implikowanej polega zatem na tym, że spekulant przewiduje, że faktyczna zmienność będzie inna niż ta założona przy ustalaniu premii. Ponieważ parametrem prognozowanym jest zmienność instrumentu bazowego, przejmuje ona rolę ceny, której zmiany są przedmiotem spekulacji w tradycyjnych transakcjach.

1.1.3. *Metody kwantyfikacji ryzyka*

Wyróżnia się trzy grupy miar ryzyka: miary wrażliwości, zmienności i zagrożenia [por. Jajuga, Jajuga 2011, s. 182–192]. Miary wrażliwości kwantyfikują wpływ pewnych czynników na inne. Do miar z tej grupy zalicza się współczynniki równań regresji budowanych w celu określenia korelacji między zmiennymi, co ma zastosowanie w analizie ryzyka w momencie podejmowania decyzji inwestycyjnych. Przy kwantyfikacji ryzyka inwestycji już dokonanych, miary wrażliwości są wykorzystywane do ustalenia wpływu jednostkowej zmiany ceny rynkowej na wartość pozycji. Ich kalkulacja bazuje na wolumenie pozycji oraz wycenie instrumentów finansowych i jest szczególnie potrzebna w przypadku nieliniowych modeli wyceny. Miary wrażliwości, objaśniając wpływ jednostkowych zmian rynkowych, nie wnoszą informacji o rynku, są zatem traktowane jako narzędzie pomocnicze w zarządzaniu ryzykiem rynkowym.

W przeciwieństwie do miar wrażliwości, miary zmienności dostarczają informacje o warunkach rynkowych. Są często wykorzystywane przy ocenie ren-

towności inwestycji. Analizuje się wówczas przewidywane stopy zwrotu wraz z ryzykiem kalkulowanym za pomocą wybranej miary zmienności. W tym zastosowaniu zaletą miar zmienności jest zależność wyłącznie od danych płynących z rynku i brak powiązania z wielkością pozycji. To sprawia, że można je wykorzystywać do porównywania ryzyka związanego z konkurencyjnymi projektami inwestycyjnymi. Rozważając pomiar ryzyka rynkowego inwestycji już dokonanych, wskaźniki zmienności nie informują o wielkości strat, ponieważ nie jest brany pod uwagę wolumen posiadanych pozycji.

Trzecią grupą miar ryzyka są miary zagrożenia. Celem kwantyfikacji zagrożenia jest oszacowanie potencjalnej wielkości straty, którą może przynieść dana inwestycja. Wyznaczenie potencjalnej straty wymaga informacji o wolumenie pozycji i o zmienności rynkowej, zatem zarówno miary wrażliwości jak i zmienności wykorzystywane są w procedurze kalkulacji ryzyka rozumianego jako zagrożenie. Zarządzanie ryzykiem oparte na pomiarze zagrożenia jest coraz szerzej wdrażane przez instytucje finansowe. Miary ryzyka należące do grupy miar zagrożenia bazują na modelu *VaR* (wartość zagrożona, ang. *value-at-risk*) opracowanym w 1994 roku przez JP Morgan [JP Morgan 1994]. W świetle praktyki rynkowej, na przestrzeni dwóch ostatnich dekad *VaR* stała się najpopularniejszą miarą ryzyka rynkowego, stanowiącą filar współczesnej koncepcji pomiaru ryzyka również w szerszym sensie np. kredytowego lub operacyjnego [Mentel 2011, s. 47]. W szczególności podejście to jest stosowane w przypadku inwestycji krótkoterminowych i tzw. ryzyka „overnight”. Jako główne cechy decydujące o popularności miar zagrożenia, podkreśla się bezpośrednie wskazanie rozmiaru potencjalnych strat z określonym prawdopodobieństwem oraz możliwość szacowania ryzyka dla całych portfeli z uwzględnieniem korelacji cen [Jajuga 2000a, 2000b].

1.2. Miara *VaR*

1.2.1. Wprowadzenie pojęcia *VaR*

VaR należy do grupy tzw. kwantylowych miar ryzyka, których definicje oparte są na pojęciach dolnego i górnego kwantyla zmiennej losowej. Rozważmy przestrzeń probabilistyczną (Ω, \mathcal{F}, P) oraz zmienną $X : \Omega \rightarrow \mathbb{R}$.

Definicja 1.1. Dolnym i górnym kwantylem rzędu $p \in (0, 1)$ zmiennej losowej X , oznaczanymi odpowiednio $q_p(X)$ i $q^p(X)$, nazywamy

$$q_p(X) = \inf\{x \in \mathbb{R} : P(X \leq x) \geq p\},$$

$$q^p(X) = \inf\{x \in \mathbb{R} : P(X \leq x) > p\}$$

[Embrechts, Kluppelberg and Mikosch 1997].

Z określenia kwantyla dolnego i górnego wynika, że jeżeli dystrybuanta F_X zmiennej losowej X jest ciągłą i rosnącą funkcją x , to kwantyle te są sobie równe. Jeżeli X ma rozkład dyskretny, to dla $p \notin F_X(\mathbb{R}) \cap (0, 1)$ kwantyl górny i dolny są sobie równe, natomiast dla $p \in F_X(\mathbb{R}) \cap (0, 1)$ kwantyl górny jest większy niż dolny.

Definicja 1.2 [Artzner 1999]. Wartością narażoną na ryzyko (VaR) na poziomie tolerancji $p \in (0, 1)$ zmiennej losowej X nazywamy²

$$VaR_p(X) = -q^p(X). \quad (1.1)$$

Znak minus przed wyrażeniem po prawej stronie wzoru (1.1) wynika z konwencji przedstawiania tej miary jako wartości dodatniej. Ze związku między kwantylem dolnym i górnym $-q^p(X) = q_{1-p}(-X)$ wynika, że równoważnie VaR można definiować w oparciu o pojęcie kwantyla dolnego:

$$VaR_p(X) = q_{1-p}(-X). \quad (1.2)$$

Wykorzystując prawdopodobieństwo zdarzenia przeciwnego otrzymuje się równoważną definicję:

$$VaR_p(X) = \inf\{x \in \mathbb{R} : P(-X > x) \leq p\}. \quad (1.3)$$

Dla zmiennych losowych ciągłych o ściśle rosnącej dystrybuancie VaR jest kwantylem rozkładu

$$VaR_p(X) = F_X^{-1}(p) \quad (1.4)$$

[por. np. Rockafellar, Uryasev 2000].

Załóżmy, że Ω reprezentuje zbiór wszystkich możliwych stanów końcowych rynku. Niech \mathcal{F}_t będzie filtracją generowaną przez podzbiory Ω w czasie³ t , $t = 1, \dots, T$. Niech $\mathbf{P}_t = (P_{t_1}, P_{t_2}, \dots, P_{t_k})$ będzie wektorem losowym przekształcającym Ω w zbiór wartości czynników ryzyka. Oznaczmy przez

² Definicja VaR jako kwantyla górnego rozkładu zmiennej losowej pochodzi od Artznera [1999], natomiast koncepcja pomiaru ryzyka za pomocą VaR opracowana została pierwotnie w ramach strategii banku JP Morgan [1994].

³ W ogólności w niniejszej pracy zakłada się, że rozważane procesy rozciągają się w nieskończoność w czasie dyskretnym t.j. $t \in \mathbb{Z}$, jednak wywód prowadzony jest w odniesieniu do obserwowanego okresu $t = 1, \dots, T$.

$v : \mathbb{R}^k \rightarrow \mathbb{R}$ funkcję wypłaty określoną na przestrzeni wartości czynników ryzyka, V_t – wartość funkcji wypłaty w momencie t : $V_t = v(\mathbf{P}_t)$, $t = 1, \dots, T$. Niech $g : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$ będzie funkcją zysku określoną jako:

$$g(\mathbf{P}_{t-1}, \mathbf{R}_t) = V_t - V_{t-1} = v(\mathbf{P}_t) - v(\mathbf{P}_{t-1}) \quad (1.5)$$

gdzie $\mathbf{P}_t = \mathbf{P}_{t-1} \cdot \exp(\mathbf{R}_t)$, $\mathbf{R}_t = (\ln \frac{P_{t,1}}{P_{t-1,1}}, \ln \frac{P_{t,2}}{P_{t-1,2}}, \dots, \ln \frac{P_{t,k}}{P_{t-1,k}})$, $t = 2, \dots, T$. Oznaczmy $G_t = g(\mathbf{P}_{t-1}, \mathbf{R}_t)$.

Jeden ze sposobów wykorzystania miary VaR w analizach empirycznych sprowadza się do wyznaczenia jej wartości dla zmiennej losowej reprezentującej zysk, czyli

$$VaR_p(G_t) = -q^p(G_t), \quad (1.6)$$

$t = 2, \dots, T$. Interpretując wzór (1.6), VaR można określić jako maksymalną kwotę, narażoną na stratę w wyniku inwestycji w portfel o określonym horyzoncie czasowym i przy założonym poziomie tolerancji [por. np. Best 2000, s. 23]. Podobne definicje podaje wiele pozycji literatury [Jajuga 2000a, 2000b, Jorion 2007, s. 106].

Z punktu widzenia interpretacji VaR jako straty ponoszonej z zadaniem prawdopodobieństwem, wygodne jest przedstawianie jej w terminach funkcji straty. Wprowadźmy funkcję straty $L_t = -G_t$. Wtedy

$$VaR_p(G_t) = VaR_t(-L_t) = q_{1-p}(L_t) \quad (1.7)$$

lub równoważnie

$$VaR_p(G_t) = \inf\{x \in \mathbb{R} : P(L_t > x) \leq p\} \quad (1.8)$$

$t = 2, \dots, T$, [por. np. Bałamut 2002].

Wykorzystując równoważność wyrażeń $P(G_t \leq x) > p \Leftrightarrow P(V_t - V_{t-1} \leq x) \Leftrightarrow P(V_t \leq x + V_{t-1})$ otrzymuje się VaR w terminach wartości portfela:

$$VaR_p(G_t) = VaR_p(V_t - V_{t-1}) = V_{t-1} - q^p(V_t), \quad (1.9)$$

$t = 2, \dots, T$.

Przekształcenie wzoru (1.9) pozwala na określenie VaR w terminach logarytmicznej stopy zwrotu z portfela R_t , za pomocą formuły

$$VaR_p(G_t) = VaR_p(V_{t-1}e^{R_t} - V_{t-1}) = V_{t-1}(1 - e^{q^p(R_t)}), \quad (1.10)$$

$t = 2, \dots, T$.

Posługując się zwykłą stopą zwrotu \tilde{R}_t można wzór (1.10) zapisać w postaci

$$VaR_p(G_t) = VaR_p(\tilde{R}_t V_{t-1}) = -q^p(\tilde{R}_t) V_{t-1} \quad (1.11)$$

$t = 2, \dots, T$.

Wprowadzenie do definicji funkcji zysku G_t wartości oczekiwanej portfela EV_t zamiast wartości zrealizowanej w okresie poprzednim V_{t-1} prowadzi do określenia względnej wartości VaR [por. np. Osińska 2006, s. 106-108]. W myśl tej koncepcji zakłada się, że zysk jest różnicą między aktualną wartością portfela a wartością oczekiwaną, $G_t^* = V_t - EV_t$. Wtedy wzory (1.9), (1.10) i (1.11) przyjmują odpowiednio postaci:

$$VaR_p(G_t^*) = EV_t - q^p(V_t), \quad (1.12)$$

$$VaR_p(G_t^*) = V_{t-1}(e^{\mu_t} - e^{q^p(R_t)}), \quad (1.13)$$

$$VaR_p(G_t^*) = V_{t-1}(E\tilde{R}_t - q^p(\tilde{R}_t)), \quad (1.14)$$

gdzie μ_t oznacza wartość oczekiwaną zmiennej R_t , $t = 2, \dots, T$.

Innym podejściem, stosowanym w analizach empirycznych zamiast wprowadzania funkcji zysku, jest wyznaczenie miary VaR bezpośrednio dla zmiennej reprezentującej stopę zwrotu R_t :

$$VaR_p(R_t) = -q^p(R_t) \quad (1.15)$$

[por. np. Wong 2008]. Ujęcie dane wzorem (1.15) będzie stosowane w dalszej części pracy.

Ogólny charakter przytoczonych powyżej definicji wynika z braku założeń o typie rozkładu prawdopodobieństwa opisującego niepewność co do rozważanej możliwej straty. Każde praktyczne wykorzystanie wartości zagrożonej wymaga określenia typu rozkładu lub szacowania jego parametrów. Przyjmowanie (w sposób jawny bądź nie) różnych założeń o rozkładzie straty jest powodem występowania w literaturze różnych metod estymacji VaR [Pipień 2006, s. 137].

1.2.2. Kryteria oceny modeli VaR

Ze względu na brak obserwowalności realizacji VaR , ocena modeli VaR oraz innych modeli ryzyka, bazujących na koncepcji wartości zagrożonej, oparta jest na porównywaniu prognoz ryzyka oraz zrealizowanego ciągu wartości stopy zwrotu. Rozważa się proces informujący o tym, czy strata z portfela

przekroczyła poziom VaR . W tym celu wprowadza się zmienną indykatorkową, która określana jest jako⁴

$$I_t = \mathbf{1}_{\{-R_t > VaR_p(R_t)\}}, \quad (1.16)$$

$t = 1, \dots, T$. Proces (1.16) określany jest jako proces przekroczeń lub proces wyjątków VaR .

Statystyczna weryfikacja modeli bazujących na mierze VaR może odbywać się poprzez testowanie warunkowego oraz bezwarunkowego rozkładu przekroczeń. Przyjmuje się założenie, że prognozy wartości zagrożonej powinny spełniać dwa postulaty. Pierwszy z nich odnosi się do rozkładu bezwarunkowego zmiennej losowej I_t i mówi, że prawdopodobieństwo przekroczenia prognozy wartości narażonej na ryzyko powinno być równe założonemu poziomowi tolerancji p , czyli $P(I_t = 1) = p$ (tzw. własność bezwarunkowego pokrycia, ang. unconditional coverage property). Testowanie prawdopodobieństwa przekroczenia VaR oparte jest na sprawdzeniu relacji liczby wyjątków do całkowitego czasu trwania obserwacji. W sytuacji gdy straty z portfela przekraczające prognozę VaR zdarzają się częściej niż wskazuje odsetek czasu wyznaczony przez prawdopodobieństwo p , można podejrzewać, że zastosowany model systematycznie niedoszacowuje ryzyka portfela. Sytuacja przeciwna, kiedy straty występują w mniejszej proporcji niż założony poziom tolerancji, sygnalizuje zbyt konserwatywny model.

Własność bezwarunkowego pokrycia odnosi się do częstotliwości występowania wyjątków VaR , natomiast drugi postulat, zwany własnością warunkowego pokrycia lub niezależności (ang. independence property), ocenia sposób rozłożenia wyjątków w próbie. Występowanie nierównomiernego rozkładu może sygnalizować istnienie zależności między kolejnymi przekroczeniami prognozowanego zagrożenia. Drugie kryterium odnosi się zatem do rozkładu warunkowego względem zbioru informacji o przeszłości posiadanych w danym okresie. Zakłada się, że w dobrym modelu pary zmiennych, będące składowymi procesu wyjątków (I_{t+g}, I_{t+h}) , $g, h > 0$, powinny być niezależne. Warunek ten implikuje, że historia procesu $(I_t)_{t=1}^T$ poprzedzająca okres $t \in \{2, \dots, T\}$, czyli $(I_1, \dots, I_{t-2}, I_{t-1})$ nie powinna zawierać żadnej informacji o tym, czy w okresie tym nastąpi przekroczenie VaR zapisywane jako $I_t = 1$. W szczególności fakt wystąpienia przekroczenia VaR w okresie $t - 1$ nie powinien zwiększać prawdopodobieństwa wystąpienia przekroczenia w okresie t . Jeżeli wcześniejsze przekroczenia wartości zagrożonej stano-

⁴ W całej pracy przyjęto uproszczoną notację, w której nie rozróżnia się, w miejscach niepozostawiających wątpliwości ze względu na kontekst, zmiennych losowych oraz ich wartości.

wią prognozę przyszłego wyjątku, to jest to wskazanie nieprawidłowo skonstruowanego modelu. Koncentracja wyjątków *VaR* w pewnych przedziałach czasowych, oznaczającą naruszenie zasady niezależności, pokazuje, że model *VaR* jest niereprezentatywny w tym sensie, że nie została w nim zawarta cała dostępna informacja rynkowa [Christoffersen 1998].

Interpretacja pierwszego kryterium odnosi się do sytuacji, kiedy ryzyko jest systematycznie niedoszacowane lub przeszacowane, co przekłada się na nieadekwatną do modelu liczbę wyjątków w długim okresie. Z punktu widzenia stabilności finansowej przedsiębiorstwa za szczególnie niebezpieczne należy uznać niespełnienie drugiego kryterium, oznaczającego możliwość występowania okresów grupowania się wyjątków. Przyczyną występowania dużej liczby przekroczeń w pewnych okresach, naprzemiennie z seriami strat mniejszych od prognozowanego zagrożenia jest zjawisko gromadzenia (persystencji) zmienności stopy zwrotu (*volatility clustering*). Gromadzenie zmienności charakteryzuje zwłaszcza rynki finansowe, ale może dotyczyć również innych rynków np. surowcowych i polega na naprzemiennym występowaniu okresów dużych i małych wahań stóp zwrotu⁵. Ekonomiczne wyjaśnienie tego zjawiska odnosi się do faktu, że ponadprzeciętnie wysokie lub niskie stopy zwrotu powodują gwałtowne, impulsywne reakcje inwestorów i napędzają kolejne duże zmiany cen. Z kolei spokojne okresy na rynku wykazują tendencję do przedłużania się aż do momentu wystąpienia zewnętrznego impulsu dużych wahań cen. W efekcie występuje autokorelacja zmiennych tworzących proces przekroczeń. Taka sytuacja powoduje istnienie przedłużających się okresów nieprzewidzianego ryzyka o dużej skali. Konsekwencją wystawienia przedsiębiorstwa w krótkim okresie na serię strat, przekraczających przewidywany poziom zagrożenia, może być utrata płynności finansowej. Skutki wystąpienia koncentracji wyjątków mogą być zatem trudniejsze do przezwyciężenia niż większa od przewidywanej liczba ponadprzeciętnych strat, rozłożonych równomiernie w długim przedziale czasowym.

Własności bezwarunkowego pokrycia oraz niezależności wyjątków stanowią dwa oddzielne kryteria oceny modelu ryzyka i powinny być traktowane w sposób rozłączny. Możliwa jest zarówno sytuacja gdy model spełnia kryterium związane z rozkładem bezwarunkowym, natomiast nie posiada własności niezależności wyjątków oraz sytuacja, gdy spełniony jest drugi postulat a pierwsza zasada zostaje naruszona. Za dobry model uznaje się taki, który ma zdolność generowania prognoz spełniających obydwa kryteria [por. np. Campbell 2005].

⁵ Formalną definicję zjawiska persystencji zmienności wprowadza się za pomocą procesu *GARCH* [Engle i Patton 2001].

Wykorzystując własności procesu Bernoulliego, spełnienie obydwóch kryteriów dla procesu $(I_t)_{t=1}^T$ można łącznie zawrzeć w zapisie

$$I_t \stackrel{iid}{\sim} Z - J(p), \quad (1.17)$$

$t = 1, \dots, T$, gdzie $Z - J(p)$ oznacza rozkład zero-jedynkowy z parametrem p . Zapis ten oznacza, że wyjątki VaR powinny mieć jednakowe, niezależne rozkłady zero-jedynkowe i stanowi kompletną charakteryzację prawidłowego modelu VaR na poziomie tolerancji p . Istnieje zatem równoważność między prognozami VaR stanowiącymi dobre oszacowanie kwantyla rozkładu straty a spełnieniem własności bezwarunkowego pokrycia i niezależności wyjątków.

W aktualnej literaturze powstało wiele propozycji testów sprawdzających jedną lub obydwie omawiane własności [por. np. Christoffersen 2012, Wang 2010]. Szczególnie dynamicznie rozwijającą się grupą testów, wykorzystywanych przy ocenie prognoz VaR , są testy autokorelacji przekroczeń. Ich celem jest wykrycie występowania skupień wyjątków VaR . Sprawdzają one zatem czy warunkowe prawdopodobieństwo przekroczenia VaR jest niezmienne w czasie.

1.3. Koherentne miary ryzyka rynkowego

1.3.1. Aksjomatyczna definicja ryzyka

Miarę ryzyka rynkowego określa się jako pewną funkcję rzeczywistą zmiennej losowej reprezentującej zysk z portfela. Wyodrębnia się dwie własności decydujące o zaliczeniu funkcji do miar ryzyka.

Definicja 1.3 [Follmer i Schied 2002]. Niech \mathcal{V} będzie niepustym zbiorem zmiennych losowych określonych na przestrzeni probabilistycznej (Ω, \mathcal{F}, P) . Mówimy, że funkcja $\rho : \mathcal{V} \rightarrow \mathbb{R} \cup \{+\infty\}$ jest miarą ryzyka jeżeli spełnia aksjomaty:

1. (Monotoniczność). Dla każdych $X, Y \in \mathcal{V}$, takich że $X \leq Y$

$$\rho(X) \geq \rho(Y).$$

2. (Niezmienność ze względu na translację). Dla każdej stałej $c \in \mathbb{R}$ i każdego $X \in \mathcal{V}$

$$\rho(X + c) = \rho(X) - c, \quad c \in \mathbb{R}.$$

Własność monotoniczności mówi, że jeżeli jeden z portfeli charakteryzuje się stale większym zyskiem, to ryzyko z nim związane jest mniejsze niż z drugim portfelem. Wymóg niezmienniczości ze względu na translację nazywany jest inaczej własnością portfela wolnego od ryzyka i mówi o tym, że dodanie do portfela pewnej kwoty zmniejsza ryzyko portfela o tę kwotę. Ryzyko zmiennej losowej może być zatem pomniejszone o początkową dopłatę a w szczególności może być wyzerowane jeżeli dopłata jest równa wartości miary ryzyka. Prostym przykładem miary ryzyka jest wartość oczekiwana zmiennej przeciwnej do wartości portfela $\rho(X) = E(-X)$ [por. np. Utkin 2010, s. 70–73].

Współczesne kierunki badań nad pomiarem ryzyka rynkowego wywodzą się z koncepcji wartości zagrożonej i stanowią uzupełnienie podstawowego modelu *VaR*, powstając często w odpowiedzi na krytykę związaną z niedoskonałościami tego modelu. Z punktu widzenia zastosowań do oceny ryzyka zjawisk nietypowych, badania koncentrują się na analizie zdarzeń ekstremalnych, charakteryzowanych przez ogony rozkładów prawdopodobieństwa, wykraczające poza poziom *VaR*. Prowadzi to do modeli opartych na mierze *expected shortfall* (*ES*) opisującej możliwe straty w przypadku wystąpienia wyjątku *VaR*. Z drugiej strony rozwój modeli *ES* odpowiada teoretycznym poszukiwaniom miary ryzyka spełniającej zbiór pewnych postulatów, wykraczających poza wymienione wyżej własności monotoniczności i niezmienniczości ze względu na translację, jak subaddytywność czy wypukłość. Dodatkowych postulatów nie spełnia model *VaR*, co może prowadzić do sytuacji w której, w skrajnych przypadkach, dywersyfikacja portfela składników narażonych na ryzyko zwiększa ryzyko całego portfela lub do sytuacji, kiedy niemożliwe jest jednoznaczne wyznaczenie wartości *VaR*, ze względu na istnienie wielu ekstremów lokalnych. Zbiór postulatów stawianych miarom ryzyka sformułowany został w formie definicji miary koherentnej.

Definicja 1.4 [Artzner et al. 1997, 1999]. Niech \mathcal{V} będzie niepustym zbiorem zmiennych losowych określonych na przestrzeni probabilistycznej (Ω, \mathcal{F}, P) . Wówczas odwzorowanie $\rho : \mathcal{V} \rightarrow \mathbb{R} \cup \{+\infty\}$ nazywamy koherentną miarą ryzyka, jeżeli spełnia aksjomaty definicji 1.3 oraz następujące warunki:

1. (Subaddytywność). Dla każdych $X, Y \in \mathcal{V}$

$$\rho(X + Y) \leq \rho(X) + \rho(Y).$$

2. (Dodatnia jednorodność). Dla każdego $a \geq 0$ oraz każdego $X \in \mathcal{V}$

$$\rho(aX) = a\rho(X).$$

Własność subaddytywności zapewnia, że dywersyfikacja portfela ma pozytywny wpływ, polegający na tym, że połączenie dwóch portfeli nie powoduje dodatkowego ryzyka. Dodatnia jednorodność stanowi graniczny przypadek subaddytywności, który informuje jak zmienia się ryzyko portfela przy braku efektu dywersyfikacji. Z własności tej wynika ponadto, że poziom ryzyka jest niezależny od zmian skali w jednostce pomiaru wartości portfela, czyli np. zmiana waluty nie ma znaczenia dla wyceny ryzyka.

W definicji (1.4) własność monotoniczności można zastąpić warunkiem, by miara ryzyka była niedodatnia dla zmiennych nieujemnych

$$X \geq 0 \Rightarrow \rho(X) \leq 0.$$

Z własności tej oraz z własności subaddytywności wynika, że miara ryzyka jest monotoniczna. Poza miarą $\rho(X) = E(-X)$, przykładem prostej miary koherentnej jest $\rho(X) = -X_{min}$, gdzie $X_{min} = \min\{X(\Omega)\}$.

Miara ryzyka spełniająca aksjomaty 1 i 2 jest miarą wypukłą. Stąd pierwsze dwa aksjomaty koherencji zastępowane są niekiedy własnością wypukłości

$$\rho(aX + (1-a)Y) \leq a\rho(X) + (1-a)\rho(Y), \quad a \in [0, 1]. \quad (1.18)$$

Ponieważ wypukłość nie pociąga za sobą dodatniej jednorodności, omawiane zbiory aksjomatów nie są tożsame. Miary ryzyka wypukłe, monotoniczne i niezmiennicze ze względu na przesunięcie nazywa się słabo koherentnymi [Embrechts, McNeil i Straumann 1999].

Powyższa definicja miary koherentnej nie zależy od rozkładów zmiennych losowych określonych na rozważanej przestrzeni mierzalnej (Ω, \mathcal{F}) , które z kolei zostały uwzględnione w określeniu tzw. scenariuszowej miary ryzyka.

Definicja 1.5 [Artzner et al. 1999]. Niech \mathcal{P} będzie niepustym zbiorem miar probabilistycznych (nazywanych również uogólnionymi scenariuszami) określonych na przestrzeni mierzalnej (Ω, \mathcal{F}) . Miarą scenariuszową na Ω dla zmiennej losowej X nazywamy funkcję postaci

$$\rho_{\mathcal{P}}(X) = \sup_{P \in \mathcal{P}} E_P(-X). \quad (1.19)$$

Istnieje ścisły związek pomiędzy pojęciem miary koherentnej a scenariuszowej. Można pokazać, że miara scenariuszowa określona definicją 1.5 jest miarą koherentną. Dokładnie aksjomaty koherencji są spełnione wtedy i tylko wtedy, kiedy przekrój zbiorów, na których określone są miary ze zbioru \mathcal{P} wynosi Ω . Ponadto miara ρ jest koherentna wtedy i tylko wtedy gdy istnieje taka

rodzina miar probabilistycznych \mathcal{P} , określona na (Ω, \mathcal{F}) , że zachodzi warunek (1.19).

Zbliżony zespół aksjomatów określa tzw. uogólnioną miarę odchylenia. W celu uogólnienia odchylenia standardowego, stanowiącego klasyczną miarę ryzyka, proponuje się pewien funkcjonał mający, przy dodatkowym warunku, własność koherencji.

Definicja 1.6 [Rockafellar, Uryasev i Zabaranin 2006]. Uogólniona miara ryzyka to funkcjonał $D : \mathcal{V} \rightarrow \mathbb{R}^+$ o następujących własnościach dla każdego $X, Y \in \mathcal{V}$

1. $D(X + a) = D(X)$, $a \in \mathbb{R}$.
2. $D(0) = 0$, $D(X) \geq 0$.
3. $D(bX) = bD(X)$, $b \geq 0$.
4. $D(X + Y) \leq D(X) + D(Y)$.

Uogólniona miara jest koherentna jeżeli dla każdego $X \in \mathcal{V}$ zachodzi $D(X) \leq EX - \inf X$. Jeżeli D jest funkcją zmiennej losowej $X - EX$, to odchylenie znika jedynie w przypadku $X = EX$, co odzwierciedla ideę pomiaru niepewności związanej z wartościami zmiennej X . Funkcjonał D porównuje się z normą nałożoną na zbiór wszystkich możliwych zmiennych losowych \mathcal{V} , w której pomija się wymóg symetrii. Trzy pierwsze aksjomaty uogólnionej miary odchylenia tworzą słabszy zespół warunków określający tzw. zgodną miarę ryzyka [por. np. Domański 2011, s. 155–157].

Model VaR , omówiony we wcześniejszej części pracy spełnia większość postulatów sformułowanych wyżej jako wymogi miary ryzyka. Pokazano, że VaR jest miarą dodatnio jednorodną, monotoniczną oraz niezmienniczą ze względu na translację [Artzner et al. 1999, Kusuoka 2001]. Można również wykazać, że VaR posiada dwie dodatkowe użyteczne własności. Jeżeli $\rho = VaR_p(X)$ to spełnione są aksjomaty:

1. (Współmonotoniczna addytywność). Dla każdych zmiennych $X, Y \in \mathcal{V}$ które spełniają własność współmonotoniczności, tj. dla każdych $\omega, \omega' \in \Omega$

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geq 0,$$

zachodzi

$$\rho(X + Y) = \rho(X) + \rho(Y).$$

2. (Niezmienniczość ze względu na rozkład⁶). Dla każdych $X, Y \in \mathcal{V}$ jeżeli $P(X \leq t) = P(Y \leq t)$ dla każdego $t \in \mathbb{R}$ to

$$\rho(X) = \rho(Y).$$

Wprost z definicji wynika, że miara VaR posiada własność niezmienniczości ze względu na rozkład w mocnym sensie. W przypadku VaR dwie zmienne losowe X i Y nie muszą mieć identycznych rozkładów, żeby zapewnić spełnienie warunku $\rho(X) = \rho(Y)$ lecz wystarcza jedynie lokalna zgodność rozkładów w punkcie odpowiadającym wybranemu poziomowi tolerancji. W szczególności zmienne o lekkich ogonach mogą charakteryzować się dokładnie takim samym poziomem VaR , co zmienne o grubych ogonach. Z kolei brak subaddytywności VaR oznacza, że ryzyko subportfeli rozpatrywanych osobno może być mniejsze niż portfela łączącego je w całość. Możliwe jest podanie przykładów instrumentów, które z bardzo niewielkim prawdopodobieństwem powodują duże straty, charakteryzując się niskim poziomem VaR i połączone w jeden portfel dają wyższy poziom VaR niż suma VaR dla subportfeli. Fakt ten może wpływać niekorzystnie na zarządzanie ryzykiem, stanowiąc bodziec do redukcji ryzyka poprzez dzielenie portfela [Szego 2002]. Ze względu na fakt, że VaR jest miarą niezmienniczą ze względu na rozkład ale nie spełnia postulatów koherencji naturalnym kierunkiem badań było poszukiwanie najmniejszej miary koherentnej oraz inwariantnej ze względu na rozkład, co doprowadziło do definicji ES , która jest najmniejszą koherentną i niezmienniczą ze względu na rozkład majorantą VaR .

1.3.2. Wprowadzenie pojęcia ES i innych miar koherentnych

W nurcie badań nad aksjomatyczną teorią pomiaru ryzyka wprowadzono dwie uśrednione miary TCE (średni niedobór nie mniejszy niż wartość zagrożona, ang. *tail conditional expectation*, nazywana również ETL od ang. *expected tail loss*) i WCE (najgorszy średni niedobór warunkowy, ang. *worst case expectation*). Rozważmy zmienną losową X określoną na przestrzeni probabilistycznej (Ω, \mathcal{F}, P) , spełniającą $E(\max\{0, -X\}) < \infty$. Ustalmy $p \in (0, 1)$.

⁶ Własność ta została wprowadzona pod oryginalnym terminem „law invariance” [Kusuoka 2001] i jako jej przybliżoną interpretację podaje się zasadę, że miara ryzyka może być jedynie szacowana na podstawie obserwacji statystycznych [Tasche 2002].

Definicja 1.7 [Artzner et al. 1999]. Średnim niedoborem zmiennej losowej X nie mniejszym niż wartość zagrożona na poziomie tolerancji p nazywamy liczbę

$$TCE_p(X) = E(-X | -X \geq VaR_p(X)). \quad (1.20)$$

Definicja (1.20) w ogólności nie dostarcza koherentnej miary ryzyka, nie spełniając postulatów subaddytywności, zatem zaproponowano jednocześnie miarę WCE .

Definicja 1.8 [Artzner et al. 1999]. Najgorszym średnim niedoborem warunkowym zmiennej losowej X na poziomie tolerancji p nazywamy liczbę

$$WCE_p(X) = -\inf\{E(X|A) : P(A) > p \wedge A \subset \Omega\}. \quad (1.21)$$

Miara WCE jest koherentną miarą ryzyka. Dla zmiennych losowych ciągłych, dla każdego poziomu tolerancji $p \in (0, 1)$ zachodzi równość $TCE_p(X) = WCE_p(X)$, co pociąga za sobą spełnienie warunków koherencji również przez miarę TCE . Dla zmiennych dyskretnych równość ta zachodzi na rynku dwustanowym. Bardziej ogólnie pokazano, że jeżeli Ω jest zbiorem skończonym, P jest prawdopodobieństwem jednostajnym na zbiorze Ω a X jest różnowartościową zmienną losową, to dla każdego $p \in (0, 1)$ zachodzi $TCE_p(X) = WCE_p(X)$ [Artzner et al. 1999]. Praktyczne zastosowanie tego twierdzenia jest jednak ograniczone ze względu na silne założenie o różnowartościowości zmiennej losowej. Jeżeli X jest zmienną losową określoną na skończonym zbiorze Ω , zachodzą następujące relacje

$$VaR_p(X) \leq TCE_p(X) \leq WCE_p(X) \quad (1.22)$$

[por. Utkin 2010, s. 91-97].

Badania teoretyczne związane z uzupełnieniem modelu VaR na gruncie teorii miar koherentnych prowadziły do poszukiwania najmniejszej koherentnej miary ryzyka dominującej VaR . Wykazano, że taka najmniejsza majoranta nie istnieje, dowodząc jednocześnie, że istnieje najmniejsza koherentna i niezmiennicza ze względu na rozkład miara ryzyka dominująca VaR [Delbaen 1998]. Kontynuując ten nurt badań zaproponowano miary $CVaR$ (warunkowa wartość zagrożona, ang. *conditional value-at-risk*, wprowadzana również pod nazwą $AVaR$ – przeciętna wartość zagrożona, ang. *average value-at-risk* lub skumulowana wartość zagrożona, ang. *accumulated value-at-risk*) i ES (średni niedobór, ang. *expected shortfall*), które w przypadku ciągłym sprowadzają się do formuły $E(-X | -X \geq VaR_p(X))$.

Definicja 1.9 [Pflug 2000]. Warunkową wartością zagrożoną zmiennej X na poziomie tolerancji p nazywamy liczbę⁷

$$CVaR_p(X) = \inf \left\{ \frac{E((X - s)^-)}{p} - s : s \in \mathbb{R} \right\}. \quad (1.23)$$

Definicja 1.10 [Acerbi, Nordio i Sirtori 2001]. Średnim niedoborem zmiennej losowej X na poziomie tolerancji p nazywamy liczbę

$$ES_p(X) = -\frac{1}{p} \left(E(X \mathbf{1}_{\{X \leq q^p(X)\}}) - q^p(X)(P(X \leq q^p(X)) - p) \right). \quad (1.24)$$

W przypadku ciągłym mamy $q_p(X) = q^p(X) = F_X^{-1}(p)$ oraz $P(X \leq F_X^{-1}(p)) = p$, zatem

$$ES_p(X) = -\frac{1}{p} E(X \mathbf{1}_{\{X \leq q^p(X)\}}) = E(-X | -X \geq VaR_p(X)). \quad (1.25)$$

Druga część wzoru (1.24), czyli $\frac{q^p(X)(P(X \leq q^p(X)) - p)}{p}$ potrzebna jest w przypadku zmiennych losowych dyskretnych, by suma prawdopodobieństw wartości mniejszych bądź równych $q^p(X)$ wynosiła jeden. Interpretacja miary ES , związana z formułą (1.24), mówi, że nadwyżka średniego niedoboru nad wartością zagrożoną jest równa stosunkowi średniej nadwyżki górnego kwantyla nad wartością zmiennej do poziomu tolerancji.

Innym sposobem definiowania ES jest wykorzystanie warunku

$$ES_p(X) = -\frac{1}{p} \int_0^p q_u(X) du \quad (1.26)$$

[Acerbi i Tasche 2002], który wynika bezpośrednio z definicji wartości oczekiwanej

$$ES_p(X) = E(-X | -X \geq VaR_p(X)) = -\frac{1}{p} \int_{-\infty}^{VaR_p(X)} X f(X) dX,$$

⁷ Formuła określająca miarę $CVaR$ została zaproponowana przez różnych autorów [Pflug 2000 i niezależnie Rockafellar i Uryasev 2000, 2002].

gdzie $f(X)$ jest gęstością zmiennej X . Po podstawieniu

$$\begin{aligned} X &= F_X^{-1}(p), \\ dX &= \frac{1}{F'(F_X^{-1}(p))} dp, \\ dX &= \frac{1}{f(X)}, \\ f(X)dX &= dp \end{aligned}$$

mamy

$$ES_p(X) = -\frac{1}{p} \int_0^p F_X^{-1}(u) du = -\frac{1}{p} \int_0^p q_u(X) du. \quad (1.27)$$

Pokazano, że zaproponowane niezależnie formuły $CVaR$ i ES są sobie równe

$$CVaR_p(X) = ES_p(X)$$

[Acerbi i Tasche 2002]. Dla zmiennej losowej X określonej na przestrzeni probabilistycznej (Ω, \mathcal{F}, P) , spełniającej $E(\max\{0, -X\}) < \infty$ zachodzą relacje

$$TCE_p(X) \leq WCE_p(X) \leq ES_p(X),$$

natomiast jeżeli X jest dodatkowo zmienną losową ciągłą to

$$TCE_p(X) = WCE_p(X) = ES_p(X).$$

Ostatni warunek implikuje spełnienie aksjomatów koherencji powyższych miar ryzyka dla zmiennych ciągłych. W przypadku zmiennych losowych dyskretnych przedstawione wyżej formuły definiują różne miary ryzyka i w tym przypadku kluczowe z punktu widzenia teorii wyceny ryzyka jest wskazanie miary koherentnej [Utkin 2010, s. 97-101].

Twierdzenie 1.1 [Acerbi i Tasche 2002]. Niech X będzie zmienną losową spełniającą $E(\max\{0, -X\}) < \infty$. Wtedy $ES_p(X)$ jest koherentną miarą ryzyka.

ES stanowi najbardziej restrykcyjną miarę ryzyka spośród powyższych propozycji, w tym sensie, że spełnia najobszerniejszy zbiór aksjomatów. Uznawana jest zatem za najlepiej teoretycznie uzasadniony sposób oceny ryzyka. Stanowi również podstawę szerszej klasy miar, wprowadzonej pod pojęciem spektralnej miary ryzyka [Acerbi 2002]. Określenie miary spektralnej jest

uogólnieniem poprzednich definicji, przedstawionych w niniejszym rozdziale, w którym przyjmuje się, że gęstość jest mnożona przez dopuszczalną funkcję awersji do ryzyka, umożliwiając wprowadzenie do modelu subiektywnych wag związanych z różnymi poziomami straty.

Niech $\mathcal{L}^1([0, 1])$ będzie przestrzenią, w której każdy element reprezentuje klasę funkcji, różniących się od siebie co najwyżej na podzbiornym zbiorze $[0, 1]$ miary zero. Określmy normę tej przestrzeni jako

$$\|\phi\| = \int_0^1 |\phi(p)| dp.$$

Definicja 1.11 [Acerbi 2002]. Spektralną miarą ryzyka zmiennej losowej X , utworzoną przez dopuszczalne spektrum $\phi \in \mathcal{L}^1([0, 1])$, nazywamy klasę miar postaci

$$M_\phi(X) = \int_0^1 VaR_p(X) \phi(p) dp, \quad (1.28)$$

gdzie funkcja $\phi : [0, 1] \rightarrow \mathbb{R}$, nazywana funkcją awersji do ryzyka miary M_ϕ , spełnia następujące warunki:

1. (Nieujemność.) $\phi(p) \geq 0$ dla każdego $p \in [0, 1]$.
2. (Unormowanie.) $\|\phi\| = \int_0^1 |\phi(p)| dp = \int_0^1 \phi(p) dp = 1$.
3. (Monotoniczność.) $\phi(p_1) \geq \phi(p_2)$ dla każdych $0 \leq p_1 \leq p_2 \leq 1$.

Dwa początkowe warunki oznaczają, że wagi związane z poszczególnymi wartościami zmiennej X są nieujemne i sumują się do jedności. Interpretacja ekonomiczna ostatniego z powyższych postulatów odzwierciedla awersję do ryzyka. Zgodnie z nim wyższe wagi przypisuje się do bardziej oddalonych od środka rozkładu wartości zmiennej X , czyli do większych strat. Im większa awersja do ryzyka, tym funkcja ϕ rośnie szybciej. Funkcja ϕ , nazywaną funkcją awersji, pełni rolę wag, które uśredniają wszystkie możliwe wartości zmiennej losowej X . Dla różnych funkcji ϕ_1, ϕ_2 , reprezentujących ten sam element ϕ należący do przestrzeni $\mathcal{L}^1([0, 1])$, otrzymuje się tę samą miarę M_ϕ .

Formuła określona definicją 1.11 nie wyklucza z klasy miar spektralnych miar charakteryzujących się neutralnym stosunkiem do ryzyka. Miarą neutralną, która nie zawiera wag związanych z poziomami straty jest ES , stanowiąca szczególny przypadek w klasie miar spektralnych. Zastąpienie nierówności słabej nierównością ostrą w warunku (3) zawęziłoby zakres dopuszczalnych miar wyłącznie do miar z wagami przypisującymi większy udział większym stratom [Wang 2010, s. 8–9].

Odpowiednie określenie funkcji ϕ , odzwierciedlające własność awersji do ryzyka, skutkuje uzyskaniem miary koherentnej.

Twierdzenie 1.2 [Acerbi 2002]. M_ϕ jest miarą koherentną wtedy i tylko wtedy, gdy spektrum ryzyka ϕ jest dopuszczalne w sensie definicji 1.11.

Posiłkując się wykładniczą funkcją użyteczności $U(x) = -e^{rx}$ oraz dobierając formułę, w taki sposób, by spełnione były postulaty (1-3) otrzymuje się ważoną wykładniczo funkcję awersji postaci

$$\phi(p) = \frac{Re^{1-p}}{1 - e^{-R}}, \quad (1.29)$$

gdzie $R > 0$. Funkcja ta zależy tylko od jednego parametru, który posiada interpretację ekonomiczną i nazywany jest współczynnikiem awersji do ryzyka. Dzięki takiemu określeniu funkcji ϕ wyższe wartości parametru R , oznaczające wyższą awersję do ryzyka, powodują, że wagi związane z wyższymi stratami rosną szybciej. Dla $R = 1$ funkcja ta jest w przybliżeniu stała. Spektralna miara ryzyka z wykładniczą funkcją awersji przyjmuje postać

$$M_\phi = \int_0^1 \frac{Re^{1-p}}{1 - e^{-R}} q_p(X) dp. \quad (1.30)$$

1.3.3. Kryteria oceny modelu ES

Ze względu na brak obserwowalności realizacji ES , ocena modeli ES , podobnie jak VaR nie może opierać się na typowych metodach, wykorzystujących porównywanie szeregów prognoz oraz zrealizowanych wartości zmiennej. Konieczne jest wykorzystanie procesów zawierających informacje o nieobserwowalnej zmiennej, np. konstruowanych w oparciu o porównanie prognoz ryzyka i zrealizowanego ciągu wartości stopy zwrotu. Ścisły związek między pojęciami VaR i ES ⁸ powoduje, że weryfikacja modeli VaR i ES częściowo opiera się na jednakowych procesach oraz kryteriach z nimi związanych. W przypadku obydwóch modeli rozważa się proces przekroczeń (1.16) postaci

$$I_t = \mathbf{1}_{\{-R_t > VaR_{t,p}(R_t)\}},$$

$t = 1, \dots, T$, w którym wartość 1 reprezentuje wyjątek VaR .

⁸ ES określana jest w literaturze jak rozwiązanie zadania optymalizacyjnego, w którym VaR jest parametrem [Wang 2010, s. 75–76].

Testowanie modeli VaR i ES na podstawie procesu (1.16) przebiega dwutorowo i polega na weryfikacji kryteriów bezwarunkowego i warunkowego pokrycia. Pierwsze z nich mówi, że prawdopodobieństwo wystąpienia wyjątku wynosi p i sprawdzane jest na podstawie rozkładu bezwarunkowego zmiennych I_t . W praktyce oznacza ono, że całkowity udział przekroczeń w badanym okresie powinien być równy p . Kryterium warunkowego pokrycia sprawdza prawdopodobieństwo przekroczenia w rozkładzie warunkowym, czyli dotyczy zależności między zmiennymi I_1, I_2, \dots, I_t , $t \in \{2, \dots, T\}$. Zakłada się, że zaobserwowane wartości przekroczeń nie powinny zawierać żadnej informacji o przyszłych stanach procesu, w szczególności wystąpienie wyjątku nie powinno zwiększać prawdopodobieństwa wyjątku w okresie następnym. Spełnienie kryterium warunkowego pokrycia przekłada się na brak skupień wyjątków VaR w krótkich okresach (rozdz. 1.2.2.).

Kryteria bazujące na procesie $(I_t)_{t=1}^T$ odnoszą się do częstości przekroczeń oraz ich rozłożenia w czasie, nie uwzględniają natomiast wielkości przekroczeń, zatem w celu weryfikacji modeli ES dodatkowo wprowadza się proces postaci

$$R_t^* = (R_t - ES_p(R_t))I_t, \quad (1.31)$$

który nazywany jest procesem reszt, ze względu na fakt, że jest formowany jako proces różnic między realizacją zmiennej losowej a jej wartością oczekiwaną (w przypadku zmiennych losowych ciągłych ES jest wartością oczekiwaną przekroczenia VaR). Dla dowolnego procesu $(R_t)_{t=1}^T$ proces $(R_t^*)_{t=1}^T$ stanowi fragment ciągu różnic martyngałowych, spełniający $E(R_t^* | \mathcal{F}_t) = 0$. Przy dodatkowym, typowym założeniu, że $(R_t)_{t=1}^T$ jest procesem $GARCH$, w którym $R_t = \sigma_t Z_t$, $Z_t \sim iid(0, 1)$, proces (1.31) można zapisać w postaci

$$R_t^* = \sigma_t (Z_t - ES_p(Z_t)) \mathbf{1}_{\{-Z_t > VaR_p(Z_t)\}}, \quad (1.32)$$

$t = 1, \dots, T$. Powyższa postać R_t^* sugeruje konstrukcję testu statystycznego weryfikującego model ES w oparciu o realizacje wystandaryzowanego procesu

$$R_t^{**} = \frac{R_t^*}{\sigma_t}, \quad (1.33)$$

$t = 1, \dots, T$. Przy przyjętych założeniach proces ten jest procesem niezależnych, jednakowych zmiennych losowych o zerowej wartości oczekiwanej i jednostkowej wariancji, czyli

$$R_t^{**} \sim iid(0, 1), \quad (1.34)$$

$t = 1, \dots, T$. Prawdopodobieństwo przyjęcia wartości mniejszej niż zero w procesie R_t^{**} wynosi $1 - p$.

Zakłada się, że odpowiednio skonstruowany model powinien odzwierciedlać powyższe własności. Stanowią one zatem podstawę konstruowania testów statystycznych do oceny modeli VaR i ES , wykorzystujących realizacje zmiennych losowych I_t , R_t , R_t^* i R_t^{**} , $t = 1, \dots, T$. Testowanie prawdopodobieństwa wystąpienia wartości R_t^{**} mniejszej niż zero jest równoważne z testowaniem bezwarunkowego prawdopodobieństwa przekroczenia w procesie $(I_t)_{t=1}^T$. Głównym uzupełnieniem procedury weryfikacyjnej w stosunku do testów VaR jest sprawdzenie hipotezy o wartości oczekiwanej R_t^* lub R_t^{**} [McNeil, Frey i Embrechts 2005].

Rozdział 2

Testy wartości zagrożonej (VaR) i oczekiwanego niedoboru (ES)

Wśród miar ryzyka wyróżnia się miary zmienności, wrażliwości i zagrożenia. Zgodnie z klasycznym podejściem w statystyce, do pomiaru ryzyka służą przede wszystkim miary zmienności jak wariancja i odchylenie standardowe. Istnieje bogata literatura poświęcona estymacji i testom wariancji oraz kowariancji, uwzględniająca zagadnienia o charakterze wielowymiarowym¹. Konkurencyjna koncepcja pomiaru ryzyka za pomocą miar wrażliwości (głównie tzw. współczynnika β), historycznie stanowiła uproszczenie modeli ryzyka opartych na relacji średnia - wariancja, omijające problem szacowania macierzy kowariancji². Weryfikacja hipotez statystycznych w ramach tego nurtu badań sprowadzała się, w dużej mierze, do oceny istotności parametrów regresji, co stanowi temat dobrze opisany w literaturze statystycznej i ekonometrycznej³.

¹ Przegląd metod estymacji i testów wariancji oraz macierzy kowariancji można znaleźć w literaturze statystycznej [np. Jędrzejczak 2012, Krzyśko 2009].

² Problem ten wynikał z potrzeby przetwarzania dużej ilości danych potrzebnych do precyzyjnego szacowania macierzy kowariancji, co nie ma znaczenia przy aktualnym stanie zaawansowania technologii komputerowej.

³ Przegląd metod estymacji i testowania współczynników regresji można znaleźć w literaturze ekonometrycznej [np. Greene 2011].

Podejścia oparte na badaniu zmienności lub wrażliwości nie dają obrazu potencjalnej wielkości straty, co powoduje, że w odpowiedzi na potrzeby uczestników rynku, dzisiejsza teoria oceny ryzyka wiąże się z badaniem poziomu zagrożenia i opiera się na kwantylowych miarach ryzyka. Według aksjomatycznej definicji ryzyka, miarą ryzyka jest funkcja, której wartości określają wielkość zagrożenia oraz która spełnia zespół aksjomatów koherencji. Zgodnie z aktualnymi kierunkami badań⁴, praca koncentruje się na miarach wartości zagrożonej, czyli na mierze *VaR* oraz, stanowiącej jej uzupełnienie, koherentnej mierze *ES*.

W niniejszym rozdziale dokonano przeglądu metod wnioskowania statystycznego, zaproponowanych w dotychczasowej literaturze do oceny modeli opartych na miarach *VaR* i *ES*. Przedstawione zostały możliwości modyfikacji klasycznych testów udziału przekroczeń *VaR*, skonstruowane z wykorzystaniem reguł Walda i mnożników Lagrange'a. Dzięki zastosowaniu zasady mnożników Lagrange'a uzyskano test, którego praktyczną przewagą jest określoność – zatem możliwość bezpośredniego stosowania – w przypadku szeregu, w którym nie zaobserwowano żadnego przekroczenia *VaR*. W szczególności ma to znaczenie przy testowaniu *VaR* na niskich poziomach tolerancji, takich jak zalecane obecnie w świetle Umowy Bazylejskiej III. Ponadto zaprezentowano własną propozycję testowania modelu *VaR* dla wielu poziomów tolerancji za pomocą wielowymiarowego testu sprawdzającego niezależność wektorów losowych. Dzięki wielowymiarowemu podejściu wykorzystano większą ilość informacji z próby oraz zapewniono elastyczność w zarządzaniu ryzykiem, polegającą na możliwości dostosowania testowanego obszaru rozkładu do potrzeb analizy. W kontekście modeli opartych na mierze *ES* zaproponowano przeprowadzenie standaryzacji zmiennej losowej objaśnianej w teście regresyjnym, co umożliwiło uzyskanie przybliżonej stacjonarności składnika losowego.

⁴ Przegląd badań dotyczących teorii pomiaru ryzyka można znaleźć w literaturze zagranicznej [np. Szego 2004, Hull 2015, Skoglund i Chen 2015] lub polskiej [np. Trzpiot 2010, Utkin 2010].

2.1. Testy *VaR* oparte na procesach Bernoulliego i Markowa

2.1.1. *Klasyczne testy bezwarunkowego rozkładu wyjątków VaR*

Wnioskowanie statystyczne przy ocenie modeli *VaR* oparte jest na procesie przekroczeń $(I_t)_{t=1}^T$ zdefiniowanym wzorem (1.16). Weryfikacji podlegają dwa postulaty stawiane modelom wartości zagrożonej, odnoszące się do warunkowego oraz bezwarunkowego rozkładu przekroczeń. W odniesieniu do rozkładu bezwarunkowego zmiennej losowej I_t , zgodnie z tzw. własnością bezwarunkowego pokrycia, przyjmuje się założenie, że prawdopodobieństwo przekroczenia prognozy wartości narażonej na ryzyko powinno być równe założonemu poziomowi tolerancji p , czyli $P(I_t = 1) = p$, $t = 1, \dots, T$. Testowanie prawdopodobieństwa przekroczenia *VaR* oparte jest na relacji liczby wyjątków do całkowitego czasu trwania obserwacji. W sytuacji gdy straty przekraczające *VaR* zdarzają się częściej niż wskazuje odsetek czasu wyznaczony przez prawdopodobieństwo p , można podejrzewać, że zastosowany model systematycznie niedoszacowuje ryzyko portfela. Sytuacja przeciwna, kiedy straty występują w mniejszej proporcji niż założony poziom tolerancji, sygnalizuje zbyt konserwatywny charakter modelu.

Hipoteza zerowa odpowiadająca kryteriom oceny modelu *VaR* ma postać

$$H_0 : I_t \stackrel{iid}{\sim} Z - J(p), \quad t = 1, \dots, T, \quad (2.1)$$

gdzie $Z - J(p)$ oznacza rozkład zero-jedynkowy z parametrem p , który równy jest założonemu prawdopodobieństwu przekroczenia prognozy *VaR* (poziomowi tolerancji). W testach opartych na rozkładzie bezwarunkowym przyjmuje się założenie o niezależności zmiennych zero-jedynkowych, czyli $I_t \stackrel{iid}{\sim} Z - J(\pi_1)$, $t = 1, \dots, T$, natomiast weryfikacji podlega parametr π_1 , t.j. $H_0 : \pi_1 = p$. Testy te mają zatem moc jedynie wobec hipotez alternatywnych postaci $H_1 : \pi_1 \neq p$, $H_1 : \pi_1 > p$ lub $H_1 : \pi_1 < p$.

Statystyka $T_1 = \sum_{t=1}^T I_t$, jako suma zmiennych o rozkładzie zero-jedynkowym, ma rozkład dwumianowy $B(T, \pi_1)$. Przy założeniu prawdziwości hipotezy zerowej rozkład ten ma postać $B(T, p)$:

$$T_1 = \sum_{t=1}^T I_t \sim B(T, p). \quad (2.2)$$

Przyjmując statystykę testową (2.2), stanowiącą sumę wyjątków VaR , można przeprowadzić test hipotezy H_0 korzystając bezpośrednio z rozkładu dwumianowego. Ponieważ rozkład dwumianowy jest rozkładem dyskretnym, uzyskanie testu o zadanym rozmiarze, wymaga zastosowania techniki randomizacji (rozdz. 3.1.2.).

Jeżeli suma przekroczeń VaR przewyższa wartość oczekiwaną rozkładu dwumianowego $EB(T, p) = Tp$ można podejrzewać, że model niedoszacowuje ryzyka. Wówczas, przeprowadzając test jednostronny, korzysta się z prawostronnego obszaru krytycznego odpowiadającego rozkładowi dwumianowemu $B(T, p)$. W przypadku sumy przekroczeń mniejszej od Tp podejrzewa się, że model przeszacowuje ryzyko oraz stosuje się lewostronny obszar krytyczny wyznaczony z rozkładu $B(T, p)$. Test dwumianowy T_1 , bazując na sumie przekroczeń VaR , nie uwzględnia możliwych zależności między kolejnymi wyjątkami. Test ten odnosi się zatem wyłącznie do bezwarunkowego rozkładu zmiennych I_t , $t = 1, \dots, T$.

W przypadku gdy liczba obserwacji T jest duża, korzystając z centralnego twierdzenia granicznego uzasadnione jest przybliżanie rozkładu Bernoulliego za pomocą rozkładu normalnego. Uzyskuje się statystykę Z_{uc} postaci

$$Z_{uc} = \frac{T_1 - Tp}{\sqrt{Tp(1-p)}}, \quad (2.3)$$

która ma asymptotyczny rozkład normalny $N(0, 1)$.

W innym teście zaproponowano przyjęcie założenia, że ciąg realizacji procesu (1.16) stanowi ciąg niezależnych prób Bernoulliego o tym samym nieznanym prawdopodobieństwie sukcesu [Kupiec 1995]. Kryterium oceny prognoz wartości zagrożonej sformułowane zostało w postaci testu ilorazu wiarygodności⁵ restrykcji parametru π_1 w ramach modelu $I_t \stackrel{iid}{\sim} Z - J(\pi_1)$, $t = 1, \dots, T$.

Wnioskowanie oparto na porównaniu empirycznego udziału przekroczeń $\hat{\pi}_1$, stanowiącego oszacowanie bezwarunkowego prawdopodobieństwa sukcesu π_1 , z parametrem p , równym poziomowi tolerancji prognoz VaR [por. Pipień 2006, s. 143-147]. Hipoteza zerowa postaci $H_0 : \pi_1 = p$, względem hipotezy alternatywnej $H_1 : \pi_1 \neq p$, sprawdzana jest za pomocą statystyki

$$LR_{uc} = -2 \log \frac{p^{T_1} (1-p)^{T_0}}{\hat{\pi}_1^{T_1} (1-\hat{\pi}_1)^{T_0}}, \quad (2.4)$$

⁵ Zasada konstrukcji testów ilorazu wiarygodności wprowadzona została w oparciu o metodę największej wiarygodności [Neyman i Pearson 1928].

gdzie T_0 jest liczbą okresów, dla których nie wystąpiło przekroczenie prognozy *VaR*, czyli $T_0 = T - T_1$, $\hat{\pi}_1 = \frac{T_1}{T}$. Wykorzystując asymptotyczne własności ilorazu wiarygodności, otrzymuje się, że, przy założeniu prawdziwości hipotezy zerowej statystyka LR_{uc} ma asymptotyczny rozkład $\chi^2_{(1)}$.

Innym testem bazującym na założeniu o procesie Bernoulliego jest test *TUFF* (ang. *time until first failure test*), stanowiący punkt wyjściowy do konstruowania procedur opartych na procesie odległości między wyjątkami *VaR* [Kupiec 1995]. Test ten bada prawdopodobieństwo przekroczenia π_1 w procesie wyjątków poprzez sprawdzenie czasu, który upłynął do wystąpienia pierwszego przekroczenia *VaR*. Odwrotność liczby jednostek czasu wykorzystywana jest jako oszacowanie z próby prawdopodobieństwa sukcesu w założonym modelu Bernoulliego. Oszacowanie to, za pomocą ilorazu wiarygodności, porównywane jest z przyjętym poziomem tolerancji. Hipotezę zerową, identyczną jak w teście Kupca, wobec takiej samej alternatywy, weryfikuje się za pomocą statystyki testowej postaci:

$$LR_{TUFF} = -2 \log \frac{p(1-p)^{V_1-1}}{\tilde{\pi}_1(1-\tilde{\pi}_1)^{V_1-1}}, \quad (2.5)$$

gdzie $\tilde{\pi}_1 = \frac{1}{V_1}$, V_1 – czas w dniach do pierwszego przekroczenia *VaR*. Statystyka LR_{TUFF} , przy założeniu prawdziwości hipotezy zerowej, ma asymptotyczny rozkład $\chi^2_{(1)}$ [Kupiec 1995].

2.1.2. *Modyfikacje podejścia do testowania bezwarunkowego rozkładu wyjątków VaR*

Przyjmując założenie o procesie Bernoulliego oraz stosując zasady Walda i mnożników Lagrange'a⁶ można skonstruować statystyki testowe, oparte na funkcji wiarygodności dla rozkładu Bernoulliego. Logarytmiczną funkcję wiarygodności dla szeregu $(I_t)_{t=1}^T$ można zapisać jako

$$\log L(\pi_1, I_1, I_2, \dots, I_T) = T_1 \log \pi_1 + (T - T_1) \log(1 - \pi_1),$$

skąd mamy, że estymatorem MNW parametru π_1 jest $\hat{\pi}_1 = \frac{T_1}{T}$. Estymatorem wariancji zmiennej T_1 , stanowiącej oszacowanie z próby liczby wyjątków $T\pi_1$,

⁶ Zasady te oparte są na porównaniu wartości szacowanego parametru do estymatora największej wiarygodności [Wald 1943] lub porównaniu maksymalnej wartości funkcji wiarygodności przy założeniu hipotezy zerowej i alternatywnej [Rao 1948, Silvey 1959]. Omówienie własności testów ilorazu wiarygodności, Walda i mnożników Lagrange'a można znaleźć w wielu pozycjach literatury [np. Engle 1984, Maddala 2008].

jest wyrażenie $T\hat{\pi}_1(1 - \hat{\pi}_1)$. Statystyka Walda dla hipotezy $H_0 : \pi_1 = p$, względem hipotezy alternatywnej $H_1 : \pi_1 \neq p$, ma zatem postać:

$$W_{uc} = \frac{(Tp - T_1)^2}{T\frac{T_1}{T}(1 - \frac{T_1}{T})} = \frac{T(Tp - T_1)^2}{T_1(T - T_1)}. \quad (2.6)$$

Przy założeniu prawdziwości hipotezy zerowej statystyka W_{uc} ma asymptotyczny rozkład $\chi^2_{(1)}$.

W celu uzyskania statystyki testu mnożnika Lagrange'a zauważmy, że funkcję oceny dla szeregu $(I_t)_{t=1}^T$ można zapisać jako

$$s(\pi_1) = \frac{\partial \log L(\pi_1, I_1, I_2, \dots, I_T)}{\partial \pi_1} = \frac{T_1 - T\pi_1}{\pi_1(1 - \pi_1)}.$$

Informacją Fishera jest:

$$\begin{aligned} I(\pi_1) &= -E \left(\frac{\partial^2 \log L(\pi_1, I_1, I_2, \dots, I_T)}{\partial \pi_1^2} \right) = \\ &= -E \left(\frac{-T\pi_1(1 - \pi_1) - (T_1 - T\pi_1)(1 - 2\pi_1)}{\pi_1^2(1 - \pi_1)^2} \right) \end{aligned}$$

Ponieważ $E(T_1 - T\pi_1) = 0$, to

$$I(\pi_1) = \frac{T}{\pi_1(1 - \pi_1)}.$$

Statystyka testu mnożnika Lagrange'a, dla hipotezy $H_0 : \pi_1 = p$, ma zatem postać:

$$LM_{uc} = \frac{s^2(p)}{I(p)} = \frac{(Tp - T_1)^2}{p^2(1 - p)^2} \cdot \frac{p(1 - p)}{T} = \frac{(Tp - T_1)^2}{Tp(1 - p)}. \quad (2.7)$$

Przy założeniu prawdziwości hipotezy zerowej statystyka LM_{uc} ma asymptotyczny rozkład $\chi^2_{(1)}$ i stanowi odpowiednik statystyki testowej (3.2), mającej asymptotyczny rozkład normalny. Przewagą testu LM_{uc} w stosunku do popularnie stosowanego testu Kupca LR_{uc} oraz testu opartego na statystyce Walda W_{uc} jest określoność w sytuacji całkowitego braku wyjątków *VaR*. Statystyki Kupca i Walda korzystają z oszacowania wariancji z próby opartego na empirycznym prawdopodobieństwie przekroczenia *VaR*. Przy braku przekroczeń oszacowanie to wynosi 0, co powoduje, że statystyka testu jest

nieokreślona. W teście mnożnika Lagrange'a estymatorem wariancji jest macierz informacji w punkcie wyznaczonym przez hipotezę zerową, co pozwala ominąć problem nieokreśloności statystyki testowej. Dzięki temu unika się konieczności stosowania alternatywnej wersji statystyki testowej dla przypadku $T_1 = 0$. Sytuacja braku przekroczeń jest często spotykana w praktyce, zwłaszcza gdy przyjmowany jest niski poziom tolerancji rzędu 1% lub 0,1%.

Opisane wyżej testy, bazujące na założeniu o procesie Bernoulliego, poddają weryfikacji parametr π_1 , mówiący o prawdopodobieństwie sukcesu, sprawdzając jego zgodność z przyjętym poziomem tolerancji. Koncentrują się zatem jedynie na bezwarunkowym rozkładzie przekroczeń *VaR*. Ich wadą jest pominięcie kwestii zależności wartości I_t , dla ustalonego $t \in \{2, \dots, T\}$, od historii procesu czyli $(I_1, \dots, I_{t-2}, I_{t-1})$. Procedury te nie mają zdolności wykrywania modeli, które prawidłowo odzwierciedlają prawdopodobieństwo przekroczenia *VaR*, ale nie odwzorowują zjawiska gromadzenia zmienności i których stosowanie może skutkować występowaniem serii następujących po sobie wyjątków. Z tego względu niezbędnym uzupełnieniem testów rozkładu bezwarunkowego są testy oparte na rozkładzie warunkowym, które koncentrują się na rozmieszczeniu wyjątków *VaR* w czasie.

2.1.3. Testy warunkowego rozkładu wyjątków *VaR*

Własność bezwarunkowego pokrycia odnosi się do częstotliwości występowania wyjątków *VaR*, natomiast drugi postulat, zwany własnością warunkowego pokrycia lub niezależności, ocenia sposób rozłożenia wyjątków w próbie. Występowanie nierównomiernego rozkładu może sygnalizować istnienie zależności między kolejnymi przekroczeniami prognozowanego zagrożenia.

W kontekście testowania autokorelacji, lub szerzej – zależności, naturalnymi kandydatami są: test Ljunga-Boxa, Pearsona lub testy nieparametryczne oparte na rozkładzie serii. Niezadowalająca moc tych testów (oraz innych popularnych w ekonometrii testów autokorelacji jak np. test *LM*) w przypadku stosowania do zero-jedynkowego procesu przekroczeń $(I_t)_{t=1}^T$ była czynnikiem stymulującym rozwój badań, dotyczących wnioskowania statystycznego przy wyborze modeli ryzyka. Głównym problemem przy konstrukcji testów autokorelacji jest sformułowanie hipotezy alternatywnej. Jeden z rozwijanych nurtów badań wywodzi się z propozycji oceny prognoz przedziałowych z wykorzystaniem procesu Bernoulliego, względem hipotezy alternatywnej skonstruowanej w oparciu o proces Markowa [Christoffersen 1998].

Test Christoffersena bazuje na ogólnym kryterium oceny prognoz przedziałowych w szeregu czasowym. Uwzględniając fakt, że podstawową przyczyną nieskuteczności prognoz poza próbę jest błędny model, kryterium oceny prognoz zostało uniezależnione od założeń związanych ze sprawdzanym modelem bazowym.

Założmy, że $(X_t)_{t=1}^T$ jest obserwowanym szeregiem czasowym a $([L_{t|t-1}(p), U_{t|t-1}(p)])_{t=H+1}^T$, $1 \leq H < T$, jest ciągiem prognoz przedziałowych na poziomie istotności p z okresu $t-1$ na okres t , gdzie $L_{t|t-1}(p)$ oznacza dolną a $U_{t|t-1}(p)$ górną granicę przedziału ufności. Zdefiniujmy I_t jako zmienną indykatorkową, wskazującą czy zaobserwowana wartość X_t znalazła się wewnątrz prognozy przedziałowej

$$I_t = \begin{cases} 0, & \text{gdy } X_t \in [L_{t|t-1}(p), U_{t|t-1}(p)], \\ 1, & \text{gdy } X_t \notin [L_{t|t-1}(p), U_{t|t-1}(p)], \end{cases} \quad (2.8)$$

$t = H + 1, \dots, T$, $1 \leq H < T$. Przy powyższych oznaczeniach prognozę przedziałową nazywa się efektywną względem zbioru informacji Ω_{t-1} jeżeli dla każdego $t = H + 1, \dots, T$ zachodzi

$$E(I_t | \Omega_{t-1}) = p. \quad (2.9)$$

Definicja warunkowej efektywności prognozy opiera się na zbiorze Ω_{t-1} wszystkich informacji dostępnych w okresie $t-1$, co pozwala skonstruować test efektywności prognozy bez odwoływania się do jakichkolwiek założeń o użytym modelu. Z punktu widzenia zastosowań ekonomicznych jest to użyteczna cecha, ponieważ omija założenia o rozkładach, często kwestionowane w naukach ekonomicznych. W szczególności powyższa własność jest przydatna w kontekście oceny prognoz VaR, gdzie bazowy szereg czasowy jest z reguły niestacjonarny z powodu zmieniającej się w czasie wariancji procesu, co utrudnia stosowanie typowych metod ekonometrycznych.

Definicja prognozy efektywnej (2.9) przekłada się na operacyjne kryterium testowe za pomocą procesu Bernoulliego. Ustalmy $t \in \{H + 1, \dots, T\}$, $1 \leq H < T$. Niech $\Omega_{t-1} = (I_1, I_2, \dots, I_{t-1})$. Wtedy równanie

$$E(I_t | \Omega_{t-1}) = E(I_t | I_1, I_2, \dots, I_{t-1}) = p$$

jest równoważne warunkowi (1.17) postaci

$$I_t \stackrel{iid}{\sim} Z - J(p),$$

$t = H + 1, \dots, T$.

Wykorzystując fakt, że $VaR_p(R_t)$ definiuje się jako kwantyl rzędu p rozkładu stopy zwrotu R_t , prognozy *VaR* można traktować jako prognozy przedziałowe postaci $(-\infty, U_{t|t-1}(p)]$ (w przypadku krótkiej pozycji *VaR* definiuje się jako kwantyl rzędu $1 - p$, co daje przedział postaci $[L_{t|t-1}(p), +\infty)$, natomiast jeżeli strata jest rozumiana zarówno jako niższa jak i wyższa od oczekiwań stopa zwrotu, otrzymuje się przedział $[L_{t|t-1}(p), U_{t|t-1}(p)]$).

Kryterium testowe (1.17), zwane własnością warunkowego pokrycia, implikuje dwa warunki: własność bezwarunkowego pokrycia $E(I_t) = p$, $t = H + 1, \dots, T$, oraz niezależność przekroczeń. Pierwszy z nich można weryfikować za pomocą testu Kupca (lub innych opisanych wyżej testów bezwarunkowego pokrycia), natomiast w odniesieniu do drugiego warunku proponuje się test sformułowany w języku warunkowych prawdopodobieństw przejścia w jednorodnym łańcuchu Markowa. Przyjmuje się założenie, że proces $(I_t)_{t=1}^T$ stanowi fragment jednorodnego łańcucha Markowa o dwóch stanach, gdzie 1 oznacza wyjątek a 0 jego brak. Rozważa się proces o rozkładzie początkowym (π_0, π_1) oraz macierzy przejścia:

$$\Pi = \begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix},$$

gdzie π_{ij} oznacza prawdopodobieństwo wystąpienia stanu j w okresie t pod warunkiem, że w okresie $t - 1$ wystąpił stan i , $t = 2, \dots, T$.

Funkcja wiarygodności dla powyższego procesu ma postać

$$L(\Pi, I_1, I_2, \dots, I_T) = \pi_{01}^{T_{01}} (1 - \pi_{01})^{T_{00}} \pi_{11}^{T_{11}} (1 - \pi_{11})^{T_{10}},$$

gdzie T_{ij} – liczba obserwacji dla których po stanie i nastąpił stan j . Maksymalizując funkcję wiarygodności otrzymuje się estymator MNW macierzy przejścia

$$\hat{\Pi} = \begin{bmatrix} \hat{\pi}_{00} & \hat{\pi}_{01} \\ \hat{\pi}_{10} & \hat{\pi}_{11} \end{bmatrix} = \begin{bmatrix} \frac{T_{00}}{T_{0\bullet}} & \frac{T_{01}}{T_{0\bullet}} \\ \frac{T_{10}}{T_{1\bullet}} & \frac{T_{11}}{T_{1\bullet}} \end{bmatrix},$$

$$T_{0\bullet} = T_{00} + T_{01}, T_{1\bullet} = T_{10} + T_{11}.$$

Niezależność wyjątków *VaR* implikuje, że prawdopodobieństwo znalezienia się procesu w okresie t w stanie 1, oznaczającym przekroczenie, nie powinno zależeć od faktu, czy w okresie $t - 1$ proces był w stanie 1 czy 0. Analogicznie, prawdopodobieństwo osiągnięcia przez proces stanu 0 w okresie t nie powinno zależeć od stanu procesu w poprzednim okresie. Warunki te można zapisać następującymi formułami:

$$\pi_{11} = \pi_{01} \wedge \pi_{10} = \pi_{00}. \quad (2.10)$$

Ponieważ π_{10} jest funkcją π_{11} oraz π_{00} jest funkcją π_{01} , to zachodzi implikacja $\pi_{11} = \pi_{01} \Rightarrow \pi_{10} = \pi_{00}$. Wystarczy zatem testować hipotezę sformułowaną w języku warunkowych prawdopodobieństw przejścia jako $H_0 : \pi_{01} = \pi_{11}$, która odpowiada macierzy przejść postaci

$$\Pi = \begin{bmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{bmatrix}.$$

Hipoteza alternatywna jest w tym przypadku sformułowana jako $H_1 : \pi_{01} \neq \pi_{11}$. Przy założeniu hipotezy zerowej, funkcja wiarygodności przyjmuje postać:

$$L(\Pi, I_1, I_2, \dots, I_T) = (1 - \pi_1)^{T_{0\bullet}} \pi_1^{T_{1\bullet}}.$$

Hipoteza zerowa sprawdzana jest za pomocą statystyki skonstruowanej w oparciu o metodę ilorazu wiarygodności

$$LR_{ind} = -2 \log \frac{\hat{\pi}_1^{T_{1\bullet}} (1 - \hat{\pi}_1)^{T_{0\bullet}}}{\hat{\pi}_{01}^{T_{01}} (1 - \hat{\pi}_{01})^{T_{00}} \hat{\pi}_{11}^{T_{11}} (1 - \hat{\pi}_{11})^{T_{10}}}. \quad (2.11)$$

Statystyka LR_{ind} , przy założeniu prawdziwości hipotezy zerowej, ma asymptotyczny rozkład $\chi^2_{(1)}$ [Christoffersen 1998].

W celu jednoczesnej weryfikacji kryteriów bezwarunkowego pokrycia i niezależności, możliwe jest połączenie powyższego testu niezależności LR_{ind} Christoffersena z testem LR_{uc} Kupca. Funkcja wiarygodności dla hipotezy zerowej $H_0 : \pi_{01} = \pi_{11} = p$ ma postać:

$$L(p, I_1, I_2, \dots, I_T) = (1 - p)^{T_{0\bullet}} p^{T_{1\bullet}},$$

zatem iloraz wiarygodności w łącznym teście warunkowego pokrycia formułowany jest jako

$$LR_{cc} = -2 \log \frac{(1 - p)^{T_{0\bullet}} p^{T_{1\bullet}}}{\hat{\pi}_{01}^{T_{01}} (1 - \hat{\pi}_{01})^{T_{00}} \hat{\pi}_{11}^{T_{11}} (1 - \hat{\pi}_{11})^{T_{10}}}. \quad (2.12)$$

Uwzględniając że $LR_{cc} = LR_{uc} + LR_{ind}$ mamy, przy założeniu hipotezy zerowej, że statystyka LR_{cc} ma asymptotyczny rozkład $\chi^2_{(2)}$ [Christoffersen 1998].

Warunki (2.10), sformułowane identycznie jak w teście Markowa, związane z prawdopodobieństwem przekroczenia *VaR* w zależności od stanu procesu w okresie poprzednim, można również weryfikować w oparciu o statystykę χ^2 w standardowym teście niezależności Pearsona. W tym ujęciu,

zamiast funkcji wiarygodności dla procesu Markowa, wyróżnia się zmienne losowe I_t, I_{t-1} oraz ich możliwe wartości 0 i 1, odpowiadające stanom w tym procesie. Hipotezę zerową formułuje się jako niezależność tych zmiennych, natomiast hipoteza alternatywna stwierdza ich zależność. Dla każdej z tych zmiennych proces $(I_t)_{t=1}^T$ generuje $T - 1$ obserwacji. W oparciu o dwudzielczą tablicę kontyngencyjną (tab. 2.1), wykorzystując poprzednie oznaczenia, wprowadzone dla procesu $(I_t)_{t=1}^T$, wyznacza się statystykę testową postaci

$$Q_{ind}^{X^2} = \sum_{i=0}^1 \sum_{j=0}^1 \frac{(T_{ij} - \hat{T}_{ij})^2}{\hat{T}_{ij}}, \quad (2.13)$$

gdzie $\hat{T}_{ij} = \frac{T_{i\bullet} T_{\bullet j}}{T-1}$, $i, j = 0, 1$. Przy założeniu prawdziwości hipotezy zerowej, mówiącej o niezależności zmiennych, statystyka ta ma asymptotyczny rozkład $\chi_{(1)}^2$.

Tab. 2.1: Tablica kontyngencyjna do testu niezależności *VaR* opartego na statystyce $Q_{ind}^{X^2}$ Pearsona

I_{t-1}	I_t		$T_{i\bullet}$
	0	1	
0	T_{00}	T_{01}	$T_{0\bullet} = T_{00} + T_{01}$
1	T_{10}	T_{11}	$T_{1\bullet} = T_{10} + T_{11}$
$T_{\bullet j}$	$T_{\bullet 1} = T_{00} + T_{10}$	$T_{\bullet j} = T_{01} + T_{11}$	$T - 1$

Źródło: Opracowane własne.

Sposób sformułowania hipotez alternatywnych w testach LR_{ind} i $Q_{ind}^{X^2}$ Pearsona powoduje, że umożliwiają one sprawdzenie autokorelacji tylko pierwszego rzędu oraz uwzględniają jedynie wpływ poprzedniego stanu procesu $(I_t)_{t=1}^T$ na wystąpienie wyjątku *VaR*, nie biorąc pod uwagę możliwego wpływu innych zmiennych egzogenicznych. Problemy te rozwiązuje dynamiczny test kwantylowy (test *DQ*) oparty na modelu regresji liniowej [Engle i Manganelli 2004]. Test *DQ* stanowi modyfikację popularnego w ekonometrii uogólnionego testu autokorelacji mnożnika Lagrange'a, bazującego na regresji reszt z równania regresji względem opóźnień tych reszt. W dynamicznym teście kwantylowym zaproponowano wykorzystanie procesu zmiennych postaci $Hit_t = I_t - p$, czyli

$$Hit_t = \begin{cases} 1 - p, & \text{gdy } -R_t > VaR_p(R_t), \\ -p, & \text{gdy } -R_t \leq VaR_p(R_t), \end{cases} \quad (2.14)$$

$t = 1, \dots, T$. Wówczas jeżeli zachodzi warunek $E(I_t) = p$, to $E(Hit_t) = 0$ oraz jeżeli przekroczenia VaR nie są skorelowane to dla każdego $\omega_t \in \Omega_t$ zachodzi $E(Hit_{t+1}\omega_t) = 0$. Łączny test dwóch powyższych własności oparty został na regresji

$$Hit_t = \beta_0 + \sum_{g=1}^G \beta_{1,g} I_{t-g} + \sum_{h=1}^H \beta_{2,h} f(I_{t-h}, I_{t-h-1}, \dots, I_1, R_{t-h}, R_{t-h-1}, \dots, R_1) + \epsilon_t, \quad (2.15)$$

gdzie f jest funkcją zmiennych należących do zbioru Ω_{t-1} , więc takich, o których informacja jest dostępna w okresie t , $t = K, \dots, T$, $K = \max\{G, H\} + 1$. Wśród takich zmiennych uwzględnia się w szczególności opóźnienia zmiennej I_t , ale również proponuje się włączanie poprzednich prognoz VaR , wcześniejszych wartości stopy zwrotu, kwadratów stopy zwrotu, traktowanych jako estymatory zrealizowanej zmienności lub innych oszacowań zmienności np. wartości zmienności implikowanej.

Oznaczmy $\beta_1 = (\beta_{1,1}, \beta_{1,2}, \dots, \beta_{1,G})$, $\beta_2 = (\beta_{2,1}, \beta_{2,2}, \dots, \beta_{2,H})$, $\beta = (\beta_0, \beta_1, \beta_2)'$ a \mathbf{X} niech będzie macierzą zmiennych objaśniających w regresji (2.15). Procedura testowa polega na badaniu istotności powyższej regresji za pomocą statystyki mnożnika Lagrange'a lub Walda. W zaproponowanym przez autorów procedury teście Walda hipoteza zerowa postaci $H_0 : \beta = 0$ sprawdzana jest, względem hipotezy $H_1 : \beta \neq 0$, za pomocą statystyki

$$DQ = \frac{\hat{\beta}' \mathbf{X}' \mathbf{X} \hat{\beta}}{p(1-p)}, \quad (2.16)$$

gdzie $\hat{\beta}$ stanowi oszacowanie MNK wektora parametrów w regresji względem zmiennej Hit_t . Statystyka DQ ma, przy założeniu prawdziwości hipotezy zerowej, asymptotyczny rozkład $\chi^2_{(g+h+1)}$ [Engle i Manganelli 2004].

Ze względu na fakt, że I_t jest zmienną zero-jedynkową, jako alternatywę do powyższej postaci regresji liniowej, proponuje się wykorzystanie regresji logistycznej i testowanie, za pomocą ilorazu wiarygodności, parametrów równania

$$P(I_t = 1) = \frac{e^{\mathbf{X}'\beta}}{1 + e^{\mathbf{X}'\beta}}, \quad (2.17)$$

$t = K, \dots, T$. Test ilorazu wiarygodności hipotezy $H_0 : \frac{e^\alpha}{1+e^\alpha} = p \wedge \beta_1 = 0 \wedge \beta_2 = 0$ względem $H_1 : \frac{e^\alpha}{1+e^\alpha} \neq p \vee \beta_1 \neq 0 \vee \beta_2 \neq 0$ ma postać:

$$DQ_{logist} = -2 \log \frac{L(p, 0, 0, I_1, I_2, \dots, I_T)}{L(\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2, I_1, I_2, \dots, I_T)}, \quad (2.18)$$

gdzie logarytmiczna funkcja wiarygodności wyraża się wzorem

$$\begin{aligned} \log L(\beta_0, \beta_1, \beta_2, I_1, I_2, \dots, I_T) &= \\ &= \sum_{t=1}^T I_t \log \frac{e^{\mathbf{X}'\beta}}{1 + e^{\mathbf{X}'\beta}} + \sum_{t=1}^T (1 - I_t) \log \left(1 - \frac{e^{\mathbf{X}'\beta}}{1 + e^{\mathbf{X}'\beta}}\right), \end{aligned} \quad (2.19)$$

natomiast $\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_2$ są estymatorami MNW parametrów równania (2.17) [Berkowitz, Christoffersen i Pelletier 2011].

Zastosowanie teorii martyngałów pozwala na inne podejście do testowania kryteriów wyznaczających dobre prognozy *VaR*. Warunki bezwarunkowego pokrycia oraz niezależności zmiennych w procesie $(I_t)_{t=H+1}^T$ są spełnione jeżeli proces $(Hit_t)_{t=H+1}^T$ jest ciągiem przyrostów martyngałowych, czyli kiedy

$$E(Hit_t | \Omega_{t-1}) = 0, \quad (2.20)$$

gdzie zbiór Ω_{t-1} jest zbiorem informacji dostępnych w okresie $t - 1$, $t = H + 1, \dots, T$, $1 \leq H < T$ [Berkowitz, Christoffersen i Pelletier 2011].

Warunek (2.20) w łatwy sposób przekłada się na kryterium testowe. Niech t będzie ustalonym okresem. Własność ciągu przyrostów martyngałowych implikuje, że proces $(Hit_t)_{t=H+1}^T$ jest nieskorelowany dla wszystkich rzędów opóźnień. Dla każdej zmiennej $X_{t-h} \in \Omega_{t-1}$ zachodzi $E(Hit_t \otimes X_{t-h}) = 0$, dla każdego $h = 1, 2, \dots, H$. W szczególności, przyjmując $X_{t-h} = Hit_{t-1}$, mamy

$$E(Hit_t Hit_{t-1}) = 0, \quad (2.21)$$

co sprowadza się, tak jak w przypadku statystyki LR_{ind} , do testowania autokorelacji pierwszego rzędu. Przyjmując bardziej ogólnie $X_{t-h} = Hit_{t-h}$ otrzymujemy warunek braku autokorelacji dowolnego rzędu

$$E(Hit_t Hit_{t-h}) = 0, \quad h = 1, 2, \dots, H. \quad (2.22)$$

Zapis ten koresponduje z omawianą wcześniej własnością niezależności procesu przekroczeń *VaR*, natomiast zasada bezwarunkowego pokrycia wynika z własności procesu przyrostów martyngałowych poprzez zastosowanie prawa iterowanej wartości oczekiwanej. Warunkowa wartość oczekiwana $E(Hit_t | \Omega_{t-1}) = 0$ przekłada się na wartość bezwarunkową $E(I_t) = p$ [Hurlin i Tokpavi 2007].

Naturalnym sposobem testowania hipotezy $H_0 : E(Hit_t Hit_{t-1}) = 0$, opartej na warunku (2.21), względem hipotezy alternatywnej postaci $H_1 : E(Hit_t Hit_{t-1}) \neq 0$, jest wykorzystanie współczynników autokorelacji z próby. W tym celu proponuje się zastosowanie testu współczynników autokorelacji Ljunga-Boxa⁷, którego sprawdzian ma postać

$$LB = T(T + 2) \sum_{h=1}^H \frac{\hat{\rho}_h^2}{T - h}, \quad (2.23)$$

gdzie ρ_h są współczynnikami autokorelacji rzędu h , $h = 1, \dots, H$ w procesie Hit_t , $1 \leq H < T$. Statystyka testu LB ma, przy założeniu hipotezy zerowej, rozkład χ_H^2 [Ljung i Box 1978].

Wykorzystując fakt, że obserwacje procesu $(I_t)_{t=1}^T$, opisującego występowanie lub nie przekroczeń *VaR*, tworzą serie zer lub jedynek, do sprawdzania niezależności zmiennych tworzących ten proces można wykorzystać metody nieparametryczne, oparte na rozkładzie serii. Zaletą metod nieparametrycznych jest to, że umożliwiają one testowanie niezależności wobec szerokiego spektrum hipotez alternatywnych, uwzględniających różne formy oraz rzędy korelacji.

Do weryfikacji hipotezy o niezależności w ramach metod nieparametrycznych służy test serii Walda-Wolfowitza [1940]. W ogólności hipoteza zerowa tego testu sformułowana jest w terminach zgodności rozkładów, przy czym hipoteza ta przekłada się na operacyjne kryterium sprawdzania niezależności występowania zer i jedynek w obserwowanym procesie. Idea testu serii opiera się na założeniu, że jeżeli obserwacje oznaczone zerami i jedynekami pochodzą z tego samego rozkładu, to ich występowanie powinno być niezależne a wówczas serie zer i jedynek powinny podlegać rozkładowi liczby serii. Wykorzystując ten rozkład, możliwe jest zatem skonstruowanie testu dokładnego dla próby skończonej, którego statystyką jest zaobserwowana liczba serii, t.j. $WW = k$. Zgodnie z teorią kombinatoryki, przy założeniu niezależności zmiennych, prawdopodobieństwo wystąpienia k serii we fragmencie procesu o długości T , składającym się z T_0 zer i T_1 jedynek wynosi

⁷ W literaturze ekonomicznej można znaleźć propozycję zastosowania testu Ljunga-Boxa do oceny modeli *VaR* oraz jego porównanie z innymi testami *VaR* na podstawie rzeczywistych danych bankowych [Berkowitz, Christoffersen i Pelletier 2011].

$$P(k|T_0, T_1) = \begin{cases} \frac{2C_{T_0-1}^{k/2-1} C_{T_1-1}^{k/2-1}}{C_T^{T_0}}, & k \text{ parzyste} \\ \frac{C_{T_0-1}^{(k-1)/2} C_{T_1-1}^{(k-1)/2-1} + C_{T_0-1}^{(k-1)/2-1} C_{T_1-1}^{(k-1)/2}}{C_T^{T_0}}, & k \text{ nieparzyste} \end{cases}, \quad (2.24)$$

gdzie C_n^m jest liczbą kombinacji m elementów spośród n elementów [Wald i Wolfowitz 1940]. Jak wynika ze wzoru (2.24), rozkład liczby serii jest zależny zarówno od całkowitej liczby obserwacji T jak i liczby jedynek T_1 , oznaczających w przypadku procesu $(I_t)_{t=1}^T$ liczbę przekroczeń *VaR*.

2.2. Testy oparte na procesie odległości wyjątków *VaR*

2.2.1. Testy braku pamięci

Odrębną klasę testów, służących statystycznej ocenie modeli *VaR*, stanowią testy oparte na procesie odległości między wyjątkami. Wykorzystują one transformację bazowego procesu $(I_t)_{t=1}^T$ w proces zmiennych $(V_t)_{t=1}^T$ zdefiniowanych następująco

$$V_i = t_i - t_{i-1}, \quad (2.25)$$

gdzie t_i oznacza okres i -tego przekroczenia *VaR*, $i = 1, \dots, N$ a N jest liczbą przekroczeń. Przyjmując założenie, że proces $(I_t)_{t=1}^T$ generowany jest zgodnie ze schematem Bernoulliego, zaproponowano test LR_{TUFF} , który opiera się na czasie, upływającym do momentu wystąpienia pierwszego przekroczenia *VaR*, czyli V_1 [Kupiec 1995]. Odwrotność tego czasu traktowana jest jako oszacowanie prawdopodobieństwa sukcesu w założonym modelu Bernoulliego (test ten omówiony został w rozdz. 2.1.1.). Uogólnieniem powyższej propozycji jest test autokorelacji pierwszego rzędu oparty na wszystkich odległościach między przekroczeniami w badanym szeregu [Haas 2001]. Przy założeniu modelu $I_t \stackrel{iid}{\sim} Z - J(\pi_i)$ dla każdego podokresu $t = t_{i-1}, \dots, t_i$, gdzie $t_0 = 1$, umożliwia on łączne testowanie hipotezy $H_0 : I_t \stackrel{iid}{\sim} Z - J(p)$, względem hipotezy alternatywnej $H_1 : \pi_i \neq p$ dla pewnego $i = 1, \dots, N$. Hipoteza alternatywna w powyższej postaci obejmuje zarówno zależność między zmiennymi, skutkującą zmiennością parametru procesu, jak i bezwarunkowe prawdopodobieństwo przekroczenia *VaR* różne od p . Statystyka testu Haasa stanowi naturalne uogólnienie testu LR_{TUFF} i ma postać

$$LR_{Haas} = -2 \log \frac{p(1-p)^{V_1-1}}{\tilde{\pi}_1(1-\tilde{\pi}_1)^{V_1-1}} + \sum_{i=2}^N -2 \log \frac{p(1-p)^{V_i-1}}{\tilde{\pi}_i(1-\tilde{\pi}_i)^{V_i-1}}, \quad (2.26)$$

gdzie $\tilde{\pi}_i = \frac{1}{V_i}$, $i = 1, \dots, N$. Statystyka LR_{Haas} , przy założeniu prawdziwości hipotezy zerowej, ma asymptotyczny rozkład $\chi^2_{(N)}$.

Transformacja procesu wyjątków VaR w proces odległości umożliwia konstruowanie testów wykorzystujących własności rozkładów geometrycznego i wykładniczego, stosowanych w modelach statystycznych do opisu czasu trwania zjawisk lub czasu oczekiwania na zdarzenie [Christoffersen i Peltier 2004]. Idea stojąca za wykorzystaniem tych rozkładów opiera się na spostrzeżeniu, że zjawisko grupowania wyjątków VaR skutkuje dużą liczbą bardzo krótkich lub bardzo długich odległości między przekroczeniami. Przy założeniu poprawnego modelu VaR , czas upływający między przekroczeniami powinien charakteryzować się brakiem pamięci a jego przeciętna wartość powinna wynosić $\frac{1}{p}$. Wówczas otrzymuje się dyskretny rozkład czasu między przekroczeniami

$$f_{geo}(V_i, p) = (1 - p)^{V_i - 1} p, \quad (2.27)$$

odpowiadający funkcji prawdopodobieństwa rozkładu geometrycznego, gdzie p jest poziomem tolerancji VaR . Funkcja hazardu, interpretowana jako intensywność występowania wyjątków, korespondująca z powyższym rozkładem, jest funkcją stałą

$$\lambda_{geo}(V_i) = \frac{f_{geo}(V_i, p)}{S_{geo}(V_i - 1, p)} = \frac{(1 - p)^{V_i - 1} p}{1 - \sum_{j=0}^{V_i - 2} (1 - p)^j p}, \quad (2.28)$$

gdzie S_{geo} jest funkcją przeżycia zmiennej V_i , odpowiadającą rozkładowi geometrycznemu.

Przyjmując założenie, że fakt wystąpienia wyjątku nie powinien nieść za sobą informacji dotyczącej prawdopodobieństwa wystąpienia wyjątku w kolejnych okresach, weryfikacja hipotezy o niezależności zmiennych I_t może zostać zastąpiona testowaniem braku pamięci zmiennych V_i . Korzystając z faktu, że rozkład wykładniczy jest ciągłym odpowiednikiem rozkładu geometrycznego oraz jedynym ciągłym rozkładem charakteryzującym się brakiem pamięci, skonstruowano test hipotezy zerowej, zakładającej rozkład wykładniczy, względem hipotezy alternatywnej, która obejmuje szerszą klasę rozkładów. Możliwe są różne warianty testu w zależności od rodziny rozkładów stanowiącej alternatywę.

Rodziny rozkładów, które zagnieżdżają w sobie rozkład wykładniczy są rozkłady Weibulla i gamma. W przypadku rozkładu Weibulla funkcja gęstości ma postać

$$f_W(V_i, a, b) = a^b b V_i^{b-1} e^{-(aV_i)^b}, \quad (2.29)$$

gdzie $a, b > 0$, i zawiera przypadek rozkładu wykładniczego dla $b = 1$. Zatem hipoteza zerowa tego testu przyjmuje postać $H_0 : b = 1$.

Funkcję hazardu Weibulla można przedstawić wzorem

$$\lambda_W(V) = \frac{f_W(V_i, a, b)}{S_W(V_i - 1, a, b)} = a^b b V^{b-1}. \quad (2.30)$$

Malejąca funkcja hazardu odpowiada wartościom $b < 1$, które oznaczają dużą liczbę krótkich odległości między przekroczeniami w okresach o zwiększonej zmienności procesu stopy zwrotu oraz dużą liczbę długich odległości w okresach uspokojenia na rynku. Test Weibulla wymaga oszacowania funkcji wiarygodności w modelu bez restrykcji oraz w modelu z restrykcją $b = 1$. Dla modelu bez restrykcji wymagana jest jedynie maksymalizacja funkcji wiarygodności względem jednego parametru, ponieważ dla ustalonej wartości b , oszacowanie parametru a wyznacza się ze wzoru

$$\hat{a} = \left(\frac{N - C_1 - C_N}{\sum_{i=1}^N V_i^b} \right)^{\frac{1}{b}}.$$

Podobnie, jak w przypadku gęstości Weibulla, dla $b = 1$ funkcja gęstości rozkładu wykładniczego zagnieżdżona jest w funkcji gęstości rozkładu gamma:

$$f_\Gamma(V_i, a, b) = \frac{a^b V_i^{b-1} e^{-(aV_i)}}{\Gamma(b)}, \quad (2.31)$$

gdzie $a, b > 0$. Dla rozkładu gamma nie istnieje jawna postać funkcji hazardu, natomiast zjawisko gromadzenia zmienności można opisać korzystając ze znajomości dwóch pierwszych momentów i współczynnika zmienności, który wynosi $\frac{SD(V_i)}{E(V_i)} = \frac{1}{b}$. W celu oszacowania funkcji wiarygodności w modelu bez ograniczeń konieczne jest poszukiwanie maksimum funkcji wiarygodności ze względu na dwa parametry a i b .

Test restrykcji parametru b w powyższych modelach Weibulla i gamma, skonstruowany w oparciu o metodę ilorazu wiarygodności ma postać odpowiednio

$$LR_{Weib} = -2 \log \frac{L(\hat{a}, 1, V_1, V_2, \dots, V_N)}{L(\hat{a}_W, \hat{b}_W, V_1, V_2, \dots, V_N)} \quad (2.32)$$

lub

$$LR_{gamma} = -2 \log \frac{L(\hat{a}, 1, V_1, V_2, \dots, V_N)}{L(\hat{a}_\Gamma, \hat{b}_\Gamma, V_1, V_2, \dots, V_N)}, \quad (2.33)$$

gdzie \hat{a}_W , \hat{b}_W , \hat{a}_Γ , \hat{b}_Γ są estymatorami MNW parametrów a i b w rozkładach odpowiednio Weibilla i gamma natomiast \hat{a} jest oszacowaniem MNW przy restrykcji $b = 1$. Uwzględniając fakt cenzurowania pierwszej oraz ostatniej odległości w procesie odstępów pomiędzy przekroczeniami, logarytmiczna funkcja wiarygodności dla próby V_1, V_2, \dots, V_N określona jest jako

$$\log L(a, b, V_1, V_2, \dots, V_N) = C_1 \log S(V_1) + (1 - C_1) \log f(V_1) + \sum_{i=2}^{N-1} \log f(V_i) C_N \log S(V_N) + (1 - C_N) \log f(V_N), \quad (2.34)$$

gdzie C_i przyjmuje wartość 1, jeżeli odstęp V_i jest cenzurowany oraz 0 w przeciwnym wypadku, $i = 1, 2, \dots, N$, f jest gęstością a S – funkcją przeżycia zmiennej V_i . Statystyki testowe LR_{Weib} i LR_{gamma} mają, przy założeniu prawdziwości hipotezy zerowej, asymptotyczny rozkład $\chi^2_{(1)}$ [Christoffersen i Pelletier 2004].

2.2.2. Testy parametrów regresji z rozkładem wykładniczym

Przedstawione wyżej testy oparte na procesie odległości między przekroczeniami VaR nie uwzględniają kolejności przekroczeń VaR , co z kolei brane jest pod uwagę w teście $EACD$ (ang. *exponential autoregressive conditional duration* – wykładniczy, autoregresyjny, warunkowy odstęp), opartym na podejściu regresyjnym [Engle i Russel 1998]. Test $EACD$ sprawdza niezależność przekroczeń VaR , wykorzystując regresję przekroczeń względem ich wcześniejszych wartości, w której składnik losowy pochodzi z rozkładu wykładniczego

$$E_{i-1}(V_i) = a + bV_{i-1}, \quad (2.35)$$

przy czym $a, b > 0$, $i = 2, \dots, N$. Ze względu na założenie o rozkładzie wykładniczym, rozważa się funkcję gęstości odstępów między przekroczeniami V_i postaci:

$$f_{EACD}(V_i, a, b) = \frac{1}{a + bV_{i-1}} e^{-\frac{V_i}{a + bV_{i-1}}}, \quad (2.36)$$

gdzie dla $b = 0$, otrzymuje się rozkład wykładniczy jako szczególny przypadek. Używając współczynnik zmienności do opisanego gromadzenia zmienności, korzysta się ze wzoru $\frac{SD(V_i)}{E(V_i)} = \frac{1}{\sqrt{1-2\alpha^2}}$. Zbiór regresorów w teście $EACD$ może być rozszerzany poprzez dołączenie informacji o wcześniejszych prognozach VaR lub zmienności z poprzedzających okresów. Test $EACD$ wymaga wyznaczenia funkcji wiarygodności w przypadku modelu z restrykcjami oraz

modelu bez restrykcji. Podobnie jak w testach opartych na modelach Weibulla i gamma proponuje się wykorzystanie ilorazu wiarygodności postaci (2.34). Statystyka *EACD*, oparta na ilorazie wiarygodności z restrykcją dotyczącą jednego parametru ma, przy założeniu prawdziwości hipotezy zerowej, asymptotyczny rozkład $\chi^2_{(1)}$ [Engle i Russel 1998].

2.3. Testy bazujące na gęstości rozkładów

2.3.1. Testy zgodności

Testy *VaR* konstruowane zgodnie z zasadą sprawdzania warunków bezwarunkowego pokrycia oraz niezależności bazują na subiektywnym wyborze poziomu tolerancji, którego test dotyczy. Według standardowego podejścia model generujący prognozy *VaR* podlega weryfikacji ex-post dla ustalonego poziomu tolerancji $p \in (0, 1)$, co wiąże się z definicją *VaR* jako kwantyla ustalonego rzędu p rozkładu stopy zwrotu. Rozwinięciem powyższego podejścia jest przyjęcie ogólnego kryterium oceny modelu *VaR*, mówiącego, że model należy uznać za dobry jeżeli posiada zdolność produkowania trafnych prognoz *VaR* jednocześnie dla różnych poziomów tolerancji. W skrajnej postaci powyższe podejście prowadzi do testowania trafności prognoz dla wszystkich poziomów tolerancji z przedziału $(0,1)$, czyli porównywania funkcji gęstości.

Sprawdzanie zgodności funkcji gęstości w testach *VaR* bazuje na transformacji Rosenblatta⁸. Jeżeli R_t jest zmienną losową reprezentującą stopę zwrotu z rozważanego portfela w okresie t a F_{R_t} dystrybuantą tej zmiennej, to *VaR* można utożsamiać z kwantylem rzędu p rozkładu stopy zwrotu czyli $VaR_p(R_t) = F_{R_t}^{-1}(p)$, $t = 1, \dots, T$. Oznaczmy \hat{F}_{R_t} estymator dystrybuanty uzyskany za pomocą ocenianego modelu *VaR*. Ustalmy $t \in \{1, \dots, T\}$. Niech f_{R_t} i \hat{f}_{R_t} będą odpowiednio funkcją gęstości i estymatorem gęstości zmiennej R_t . Zgodnie z transformacją Rosenblatta,

$$U_t = \hat{F}_{R_{t-1}}(R_t), \quad (2.37)$$

zmienna U_t ma rozkład charakteryzowany funkcją gęstości postaci $\frac{f_{R_t}(\hat{F}_{R_t}(R_t))}{\hat{f}_{R_t}(\hat{F}_{R_t}(R_t))}$. Przy założeniu prawdziwości hipotezy zerowej, mówiącej o zgodności

⁸ Transformacja Rosenblatta umożliwia, przy spełnieniu założenia o zgodności funkcji gęstości, otrzymanie zmiennej losowej charakteryzującej się rozkładem równomiernym [Rosenblatt 1952].

rozkładów, estymator \hat{f}_{R_t} odpowiada gęstości f_{R_t} i otrzymuje się gęstość rozkładu jednostajnego na przedziale (0,1)

$$U_t \sim U_{(0,1)}. \quad (2.38)$$

Uogólniając rozważania dotyczące jednego okresu t na zbiór $t = 1, \dots, T$ zakłada się, że $(\hat{F}_{R_t})_{t=1}^T$ oraz $(\hat{f}_{R_t})_{t=1}^T$ są odpowiednio ciągami estymatorów dystrybuant i gęstości $(F_{R_t})_{t=1}^T$ i $(f_{R_t})_{t=1}^T$. Zakładając zgodność ciągu estymatorów $(\hat{f}_{R_t})_{t=1}^T$ z ciągiem gęstości $(f_{R_t})_{t=1}^T$ otrzymuje się, że ciąg transformacji Rosenblatta $(U_t = \hat{F}_{R_{t-1}}(R_t))_{t=1}^T$ jest ciągiem niezależnych zmiennych losowych o rozkładzie jednostajnym na przedziale (0,1), tzn.

$$U_t \stackrel{iid}{\sim} U_{(0,1)}, \quad (2.39)$$

$t = 1, \dots, T$, [Crnkovic i Drachman 1995, 1996].

Testowanie zgodności z rozkładem równomiernym, wynikające z powyższego warunku, może być przeprowadzone za pomocą wielu testów zgodności⁹. Propozycje występujące w literaturze w kontekście testowania VaR obejmują zastosowanie nieparametrycznego testu Kuipera [Crnkovic i Drachman 1995, 1996]¹⁰, testu Kołmogorowa-Smirnova, Andersona-Darlinga lub Cramera-von Misesa [Barbachan, Ornelas i Farias 2006]. Statystyki powyższych testów bazują na mierze odległości między dystrybuantą empiryczną a teoretyczną. Statystyka testu Kuipera wykorzystuje sumę odległości postaci

$$D_{Kui} = \max_{t=1, \dots, T} \{\hat{F}_{U_t}(U_t) - F_{U(0,1)}(U_t)\} + \max_{t=1, \dots, T} \{F_{U(0,1)}(U_t) - \hat{F}_{U_t}(U_t)\},$$

statystyka Kołmogorowa-Smirnova oparta na jest maksimum rozpiętości między dystrybuantami

$$D_{KS} = \max_{t=1, \dots, T} \{\hat{F}_{U_t}(U_t) - F_{U(0,1)}(U_t)\},$$

podobnie jak statystyka Andersona-Darlinga

$$D_{AD} = \sqrt{T} \max_{t=1, \dots, T} \frac{\{\hat{F}_{U_t}(U_t) - F_{U(0,1)}(U_t)\}}{\sqrt{w(t)}},$$

⁹ Literatura dotycząca testów zgodności jest bardzo obszerna [np. Domański 1990, s. 58–78, Domański i Pruska 2000, s. 166–195].

¹⁰ Autorzy tej propozycji argumentują, że test Kuipera jest jednakowo wrażliwy na wszystkie wartości zmiennej losowej, w przeciwieństwie do testu Kołmogorowa-Smirnova, który jest bardziej czuły w okolicach mediany. Poza weryfikacją zgodności z rozkładem jednostajnym, Ci sami autorzy proponują testowanie niezależności zmiennych U_t za pomocą testu BDS [Crnkovic i Drachman 1995, 1996].

gdzie w jest funkcją wagową, za którą najczęściej przyjmuje się $w(t) = \frac{1}{\sqrt{F_{U(0,1)}(U_t)(1-\hat{F}_{U_t}(U_t))}}$, natomiast test Cramera von Misesa opiera się na sumowaniu kwadratów odchyień

$$D_{CM} = \sum_{t=1}^T \left(F_{U(0,1)}(U_t) - \frac{2t-1}{2T} \right)^2 + \frac{1}{12T}.$$

Możliwe jest również posługiwanie się wieloma modyfikacjami powyższych statystyk.

Jako uzupełnienie formalnych procedur testowania zgodności rozkładów proponuje się statystyczną ocenę kryterium (2.39) za pomocą metod graficznych, które pozwalają uzyskać wskazówki odnośnie przyczyn odrzucenia modelu. Odrzucenie modelu *Var* może następować ze względu na niespełnienie założenia o jednostajności rozkładu lub założenia o niezależności zmiennych. W przypadku każdej z przyczyn, z punktu widzenia zarządzania ryzykiem, istotne może być określenie natury niezgodności jak np. heteroskedastyczność lub struktura autokorelacji stopy zwrotu. W celu dokonania graficznej oceny warunku (2.39) proponuje się histogram, wykres prognozy przedziałowej (ocena zgodności z rozkładem jednostajnym) oraz korelogramy zmiennych $U_t - \bar{U}$, $(U_t - \bar{U})^2$, $(U_t - \bar{U})^3$, $(U_t - \bar{U})^4$, pokazujące strukturę zależności poprzez momenty do 4 rzędu (ocena niezależności) [Diebold, Gunther i Tay 1998].

2.3.2. Testy ilorazu wiarygodności

Nieparametryczne testy zgodności wykorzystywane do sprawdzania własności (2.38) wymagają dużej liczby obserwacji aby zapewnić satysfakcjonującą moc testu, zatem w kolejnych badaniach zaproponowano transformację zmiennych U_t umożliwiającą zastosowanie testu parametrycznego [Berkowitz 2001]. Rozważmy zmienną Z_t określoną wzorem

$$Z_t = \Phi^{-1}(U_t) = \Phi^{-1}(\hat{F}_{R_t}(R_t)), \quad (2.40)$$

gdzie Φ oznacza dystrybuantę rozkładu normalnego. Przy założeniu prawdziwości hipotezy zerowej, mówiącej o zgodności \hat{f}_{R_t} z f_{R_t} , zachodzi

$$Z_t \stackrel{iid}{\sim} N(0, 1), \quad t = 1, \dots, T. \quad (2.41)$$

Przekształcenie (2.40), popularnie stosowane w technikach generowania zmiennych losowych, pozwala otrzymać zmienne o rozkładzie normalnym

$N(0, 1)$, dzięki czemu możliwe jest stosowanie obszernego zakresu metod statystycznych związanych z założeniem normalności. Ponadto przekształcenie to, przy pewnych założeniach, pozwala interpretować otrzymane zmienne jako wyjściowy szereg czasowy, co zawiera się w twierdzeniu mówiącym, że jeżeli h jest gęstością zmiennej Z_t a ϕ gęstością rozkładu normalnego standardowego, to $\log \frac{f_{R_t}(R_t)}{\hat{f}_{R_t}(R_t)} = \log \frac{h(Z_t)}{\phi(Z_t)}$. z twierdzenia tego wynika, że rozważana transformacja zachowuje charakter rozbieżności w stosunku do rozkładu teoretycznego, tzn. jeżeli $f_{R_t} > \hat{f}_{R_t}$ ($f_{R_t} < \hat{f}_{R_t}$) na pewnym przedziale, to na tym przedziale $h(Z_t) > \phi(Z_t)$ ($h(Z_t) < \phi(Z_t)$) [Berkowitz 2001].

Statystyczna weryfikacja warunku (2.41) może przybierać wiele form, od testowania momentów, poprzez testy normalności, po testy niezależności. W szczególności, przyjmując model postaci

$$Z_t - \mu = \rho_1(Z_{t-1} - \mu) + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2) \quad (2.42)$$

można testować hipotezę zerową $H_0 : \mu = 0, \rho_1 = 0, \sigma = 1$ względem hipotezy alternatywnej obejmującej modele autoregresji pierwszego rzędu z zastosowaniem metody ilorazu wiarygodności. Logarytmiczna funkcja wiarygodności dla modelu (2.42) ma postać

$$\begin{aligned} \log L(\mu, \sigma, Z_1, Z_2, \dots, Z_T) = & \\ = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log \frac{\sigma^2}{1 - \rho_1^2} + \frac{(Z_1 - \mu/(1 - \rho_1))^2}{2\sigma^2/(1 - \rho_1^2)} + & \\ - \frac{T-1}{2} \log(2\pi) - \frac{T-1}{2} \log \sigma^2 - \sum_{t=2}^T \frac{(Z_t - \mu - \rho_1 Z_{t-1})^2}{2\sigma^2}. & \end{aligned} \quad (2.43)$$

Na gruncie zasady ilorazu wiarygodności, testowanie hipotezy zerowej obejmującej łącznie parametry μ, σ i ρ_1 wykonuje się za pomocą statystyki

$$LR_B = -2(\log L(0, 1, 0, Z_1, Z_2, \dots, Z_T) + \log L(\hat{\mu}, \hat{\sigma}, \hat{\rho}_1, Z_1, Z_2, \dots, Z_T)), \quad (2.44)$$

która ma asymptotyczny rozkład $\chi_{(3)}^2$.

Test autokorelacji, sformułowany z wykorzystaniem powyższego modelu, polega na weryfikacji hipotezy zerowej $H_0 : \rho_1 = 0$ wobec hipotezy alternatywnej $H_1 : \rho_1 \neq 0$ za pomocą statystyki

$$LR_B^* = -2(\log L(\hat{\mu}, \hat{\sigma}, 0, Z_1, Z_2, \dots, Z_T) + \log L(\hat{\mu}, \hat{\sigma}, \hat{\rho}_1, Z_1, Z_2, \dots, Z_T)) \quad (2.45)$$

o rozkładzie $\chi_{(1)}^2$.

Uogólnienie modelu (2.42) pozwala uwzględnić hipotezy alternatywne mówiące o autokorelacji wyższych rzędów. Ogólny model, obejmujący szeroką klasę zależności, przyjmuje postać

$$Z_t = \mu + \rho_1 Z_{t-1} + \rho_2 Z_{t-2} + \dots + \rho_H Z_{t-H} + \gamma_1 Z_{t-1}^2 + \gamma_2 Z_{t-2}^2 \dots + \gamma_G Z_{t-G}^2 + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma). \quad (2.46)$$

W najszerszej wersji testu weryfikuje się hipotezę zerową łączną $H_0 : \mu = 0 \wedge \rho_1 = \rho_2 \dots = \rho_H = \gamma_1 = \gamma_2 \dots = \gamma_G = 0 \wedge \sigma = 1, \forall t \in 1, 2, \dots, T$, co oznacza nałożenie $G + H + 2$ restrykcji i korzystanie z rozkładu $\chi_{(G+H+2)}^2$ [Berkowitz 2001].

W zależności od dostępnej liczby danych możliwe są dalsze uogólnienia modelu, obejmujące dynamikę wyższych rzędów. Objęcie hipotezą alternatywną szerokiej klasy modeli zbliża test do testu nieparametrycznego, sprawdzającego wszelkie odstępstwa od modelu niezależnych zmiennych losowych o rozkładzie normalnym. Ponadto możliwe jest wprowadzenie do modelu zmiennych egzogenicznych.

2.3.3. Testy gęstości spektralnej

Jednym ze sposobów badania okresowości i korelacji w ekonomicznych procesach stochastycznych jest wykorzystanie metod teorii sygnałów rozwijanej w fizyce. W teorii sygnałów, zapis zmienności ustalonego parametru fizycznego traktuje się jako funkcję czasu i dąży się do opracowania metod matematycznych, które pozwolą dokonać dekompozycji i poznać pewne własności tej funkcji. W tym celu wprowadza się funkcję próbkującą, która poprzez zadaną operację matematyczną z funkcją badaną daje wynik zależny od obecności szczególnych własności funkcji badanej. Wykorzystuje się operację mnożenia, przy której szczególna postać jednego z czynników wymusza zerowanie

całego iloczynu. Operacja ta wykonywana jest w całej dziedzinie. Najbardziej ogólną postacią wzoru opisującego powyższą procedurę jest

$$S = \int_0^T X(t)Y(t)dt, \quad (2.47)$$

gdzie X jest funkcją badaną, Y – funkcją próbkującą a interwał całkowania T powinien być co najmniej równy spodziewanemu okresowi badanego przebiegu. Jeżeli funkcje X i Y są wobec siebie funkcjami ortogonalnymi, to w przypadku braku równości tożsamościowej, wynik takiego całkowania jest równy zeru. Własność tę wykorzystuje się w analizie widmowej, której celem jest odnalezienie sygnałów o określonej częstotliwości w sygnale łącznym¹¹.

Podstawą analizy widmowej jest twierdzenie Fouriera, mówiące, że każdy przebieg okresowy da się wyrazić jako kombinację liniową elementarnych przebiegów okresowych, jakimi są funkcje $\sin(x)$ i $\cos(x)$. Jeżeli zatem X jest funkcją okresową, to da się ją przybliżyć za pomocą wzoru

$$X(t) = \sum_{k=-\infty}^{\infty} a_k \cos k\omega_0 t + b_k \sin k\omega_0 t, \quad (2.48)$$

gdzie a_k i b_k , $k \in \mathbb{Z}$ noszą nazwę współczynników Fouriera. Wartości współczynników Fouriera otrzymuje się za pomocą całek

$$a_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t) \cos k\omega_0 t dt, \quad k \in \mathbb{Z}, \quad (2.49)$$

$$b_k = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} X(t) \sin k\omega_0 t dt, \quad k \in \mathbb{Z}. \quad (2.50)$$

¹¹ W naukach ekonomicznych szczególnym przykładem operacji próbkowania jest wyznaczanie wartości funkcji autokorelacji ρ_X procesu stochastycznego $(X(t))_{t=-\infty}^{\infty}$. Funkcję autokorelacji można przedstawić jako całkę z iloczynu funkcji badanej i próbkującej. W przypadku autokorelacji rzędu h funkcją próbkującą nie jest funkcja trygonometryczna lecz ta sama funkcja przesunięta w czasie o h , $h \in \mathbb{R}$,

$$\rho_X(h) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T (X(t) - E(X(t)))(X(t-h) - E(X(t-h)))dt.$$

Dla $h = 0$ funkcja autokorelacji przyjmuje wartość maksymalną. Jeżeli analizowana funkcja jest funkcją okresową, to funkcja autokorelacji przyjmuje wartość maksymalną dla h równemu jej okresowi lub wielokrotności tego okresu.

Proces wyznaczania współczynników a_k i b_k nazywany jest prostym przekształceniem (transformacją) Fouriera, natomiast proces sumowania wszystkich składowych funkcji X nazywany jest odwrotnym przekształceniem Fouriera. W procesie rozkładu funkcji okresowej otrzymuje się składowe, których częstotliwości ω są wielokrotnościami częstotliwości podstawowej ω_0 . Prosta transformata Fouriera przenosi sygnał z dziedziny czasu w dziedzinę częstotliwości, natomiast transformata odwrotna z dziedziny częstotliwości w dziedzinę czasu.

Reprezentacja sygnału w dziedzinie częstotliwości za pomocą współczynników a_k i b_k nazywana jest widmem. Jeżeli w widmie nie jest obecny składnik odpowiadający danej częstotliwości to wśród wyrazów a_k i b_k nie ma żadnego wyrażenia podniesionego do kwadratu. Pod wszystkimi całkami są iloczyny $\cos k\omega_0 t$ ani $\sin k\omega_0 t$ o różnych częstotliwościach, co przy odpowiednio długich interwałach całkowania daje wynik zero. Jeżeli częstotliwość funkcji próbującej jest równa częstotliwości którejś ze składowych obecnych w funkcji analizowanej, wówczas całka jest różna od zera a jej wartość proporcjonalna do współczynnika a_k lub b_k .

Korzystając z teorii spektralnej w analizie ryzyka, proponuje się konstrukcję testu autokorelacji $VarR$ opartą na pojęciu gęstości spektralnej (widmowej), która jest transformatą Fouriera funkcji autokowariancji procesu przekroczeń $(I_t)_{t=1}^T$ [Durlauf 1991]. Wykorzystuje się fakt, że transformacja Fouriera jest przekształceniem bezstratnym, w tym sensie że gęstość spektralna zawiera tę samą informację o strukturze czasowej procesu co wyjściowa funkcja autokorelacji. W terminach zawartości informacji funkcje gęstości spektralnej i autokorelacji są tożsame, natomiast każda z nich w inny sposób eksponuje własności szeregu czasowego¹². Dzięki temu możliwe jest konstruowanie testów opartych na gęstości spektralnej, charakteryzujących się wyższą mocą wobec pewnych klas hipotez alternatywnych, dotyczących struktury korelacyjnej procesu.

Rozważa się hipotezę zerową mówiącą, że proces $(I_t)_{t=1}^T$ jest fragmentem ciągu różnic martyngałowych $H_0 : E(I_t | \Omega_{t-1}) = p$, gdzie Ω_t jest zbiorem informacji dostępnym w momencie t a $p \in (0, 1)$ jest ustalonym poziomem tolerancji. Hipoteza ta równoważna jest stwierdzeniu, że funkcje autokorelacji $\rho_I(h)$ oraz autokowariancji $\sigma_I(h)$ są równe zero dla wszystkich rzędów opóźnień h .

¹² Omówienie spektralnych metod analizy korelacji z zastosowaniem do oceny ryzyka można znaleźć w literaturze polskiej [Małecka 2015] i zagranicznej [Berkowitz, Christoffersen i Pelletier 2011] oraz w zamieszczonych tam odniesieniach.

Koncepcja testu polega porównywaniu funkcji gęstości spektralnej f_I , stanowiącej transformację funkcji autokowariancji $\sigma_I(h)$,

$$f_I(\omega) = \frac{1}{2\Pi} \sum_{h=-\infty}^{\infty} \sigma_I(h)e^{-ih\omega} \quad (2.51)$$

do teoretycznej funkcji gęstości spektralnej procesu różnic martyngałowych. Do wyznaczenia gęstości spektralnej zgodnej z hipotezą zerową wykorzystuje się asymptotyczne własności funkcji autokorelacji z próby. Przy założeniu hipotezy zerowej¹³ zachodzi:

1. $\hat{\rho}_I(h) \xrightarrow{d} 0$,
2. $\sqrt{T}\hat{\rho} \xrightarrow{d} N(0, \mathbb{I})$, gdzie \mathbb{I} jest macierzą jednostkową stopnia H , $\hat{\rho}$ jest H -elementowym wektorem autokorelacji a symbol \xrightarrow{d} oznaczają słabą zbieżność [Durlauf 1991].

Z powyższych własności wynika, że przy założeniu hipotezy zerowej gęstość spektralna jest funkcją stałą

$$f_I(\omega) = \frac{\sigma_I(0)}{2\Pi} = \frac{p(1-p)}{2\Pi}. \quad (2.52)$$

Dystrybuanta odpowiadająca funkcji gęstości spektralnej jest funkcją liniową postaci

$$F_I(\lambda) = \int_{-\Pi}^{\lambda} f_I(\omega)d\omega = \frac{\sigma_I(0)}{2\Pi}(\lambda + \Pi) = \frac{p(1-p)}{2\Pi}(\lambda + \Pi). \quad (2.53)$$

Testowanie hipotezy H_0 opiera się na badaniu zbieżności funkcji losowej będącej estymatorem gęstości spektralnej. Estymatorem gęstości spektralnej (2.51) jest funkcja

$$P_T(\omega) = \frac{1}{2\Pi} \sum_{h=-(T-1)}^{T-1} \hat{\sigma}_I(h)e^{-ih\omega}. \quad (2.54)$$

¹³ Wyprowadzenie asymptotycznych własności funkcji autokorelacji z próby przeprowadza się przy pewnych dość ogólnych warunkach związanych z momentami zmiennych losowych tworzących rozważany proces stochastyczny [warunki te można znaleźć w pracy Durlauf 1991].

Miarą odległości między gęstością spektralną a teoretyczną funkcją stałą jest funkcja losowa dana wzorem

$$\gamma_T(\omega) = \frac{1}{2\Pi} \left(\sum_{h=-(T-1)}^{T-1} \hat{\sigma}_I(h) e^{-ih\omega} - \sigma_I(0) \right). \quad (2.55)$$

Ponieważ powyższa formuła opisuje funkcję losową, nie można jej wykorzystać bezpośrednio jako statystykę testową. Funkcja ta nie jest punktowo zbieżna ze względu na brak zgodności estymatorów dla poszczególnych częstotliwości ω . Przy założeniu prawdziwości hipotezy zerowej skumulowane wartości tej funkcji dążą do zera, zatem jako miarę rozbieżności obserwacji w stosunku do hipotezy zerowej można wykorzystać funkcję

$$\Gamma_T(\lambda) = \int_0^\lambda \left(P_T(\omega) - \frac{\hat{\sigma}_I(0)}{2\Pi} \right) d\omega, \quad \lambda \in [0, \Pi]. \quad (2.56)$$

Jako podstawę konstrukcji testu statystycznego proponuje się zmodyfikowaną postać powyższego wyrażenia

$$\begin{aligned} U_T(t) &= \sqrt{2T} \int_0^{\Pi t} \left(\frac{P_T(\omega)}{\hat{\sigma}_I(0)} - \frac{1}{2\Pi} \right) d\omega = \\ &= \frac{\sqrt{2}}{\Pi} \sum_{h=1}^{T-1} \sqrt{T} \hat{\rho}_I(h) \frac{\sin h\Pi t}{h}, \quad t \in [0, 1], \end{aligned} \quad (2.57)$$

w której przeprowadzono odwzorowanie λ w Πt , $t \in [0, 1]$, przemnożono funkcję przez $\sqrt{2}$ oraz dokonano normalizacji, dzieląc przez $\frac{\sigma_I(0)}{\sqrt{T}}$. Dzięki temu otrzymano formułę wyrażoną jako funkcja znormalizowanej względem \sqrt{T} autokorelacji z próby¹⁴. W stosunku do wyrażenia (2.56), zmodyfikowana statystyka jest odporna na zjawisko heteroskedastyczności [Durlauf 1991].

Przy założeniu prawdziwości hipotezy H_0 zachodzi słaba zbieżność $U_T(t) \xrightarrow{d} B(t)$, $t \in [0, 1]$ gdzie $B(t)$ jest mostem Browna¹⁵ na $[0, 1]$. Opierając się na powyższej zbieżności możliwe jest przeprowadzenie testu z wykorzystaniem kilku statystyk, które odwzorowują funkcję losową $U_T(t)$ w zmienną losową. Propozycje występujące w literaturze obejmują: statystykę Andersona-Darlinga

$$SD_{AD} = \int_0^1 \frac{U(t)^2}{t(1-t)} dt, \quad (2.58)$$

¹⁴ Operacje te ułatwiają wyprowadzenie asymptotycznego rozkładu statystyki testowej [dowód – patrz Durlauf 1991].

¹⁵ Most Browna jest procesem Winera warunkowym względem $W(1) = 0$, czyli $B(t) = W(t) \{W(1) = 0\}$, $t \in [0, 1]$.

statystykę Cramera von Misesa

$$SD_{CVM} = \int_0^1 U(t)^2 dt, \quad (2.59)$$

statystykę Kołmogorowa-Smirnova

$$SD_{KS} = \sup_{t \in [0,1]} |U(t)| \quad (2.60)$$

i statystykę Kuipera

$$SD_{Kui} = \sup_{0 \leq s, t \leq 1} |U(t) - U(s)| \quad (2.61)$$

[por. np. Durlauf 1991, Deo 2000].

2.4. Wielowymiarowe testy *VaR*

2.4.1. Test *Ljunga-Boxa* dla wielu poziomów *VaR*

Jako podejście pośrednie między testami *VaR* dla jednego poziomu tolerancji a testami sprawdzającymi zgodność całej funkcji gęstości zaproponowano wielowymiarowy test *VaR*, służący testowaniu modelu ryzyka jednocześnie na wielu poziomach tolerancji [Hurlin i Tokpavi 2007]. Jednowymiarowe testy, bazując na procesie $(I_t)_{t=1}^T$ zdefiniowanym jako proces wyjątków *VaR* na ustalonym poziomie tolerancji p (rozdz. 2.1.), pozwalają jedynie na weryfikację modelu ograniczoną do sprawdzania trafności prognoz dla jednego kwantyla rozkładu stopy zwrotu. Z drugiej strony zastosowanie testów wykorzystujących całą funkcję gęstości lub dystrybuantę (rozdz. 2.3.), implikuje przyłożenie jednakowej wagi do prognozowania kwantyli z ogonów jak i bliskich środka rozkładu. Z punktu widzenia zarządzania ryzykiem takie podejście nie jest wystarczająco elastyczne, ponieważ najczęściej przedmiotem zainteresowania jest szacowanie kwantyli pochodzących z ogonów rozkładów. W szczególności podejście wykorzystujące cały rozkład zmiennej nie może być stosowane do weryfikacji modeli wywodzących się z teorii wartości ekstremalnych. Rozwiązaniem jest skorzystanie z testów wielowymiarowych, które umożliwiają ograniczenie analizy do wybranego obszaru funkcji gęstości zmiennej losowej, a zatem dostosowanie jej do potrzeb zarządzania ryzykiem. Ponadto zastosowanie testu wielowymiarowego umożliwia bezpośrednią kontrolę rozmiaru testu, co stanowi o jego przewadze w stosunku do wielokrotnego testowania *VaR* na wybranych poziomach tolerancji.

Wielowymiarowy test *VaR* powstał jako uogólnienie testów autokorelacji Boxa-Pierce i Ljunga-Boxa. Rozważmy proces $(X_t)_{t=1}^T$ $X_t \sim N(0, \sigma^2)$. Załóżmy, że r_h jest współczynnikiem autokorelacji rzędu h , zmiennych X_t czyli

$$\hat{r}_h = \frac{\sum_{t=h+1}^T X_t X_{t-h}}{\sum_{t=1}^T X_t^2},$$

$h = 1, 2, \dots, H$, $1 \leq H < T$. Korzystając z wielowymiarowej normalności wektora losowego (r_1, r_2, \dots, r_H) [Anderson 1942] o wartości oczekiwanej zero, wariancji $\text{var}(r_h) = \frac{T-h}{T(T+2)}$ i kowariancji $\text{cov}(r_g, r_h) = 0$, $g, h = 1, 2, \dots, H$, $g \neq h$, można pokazać, że statystyka

$$LB = T(T+2) \sum_{h=1}^H \frac{\hat{r}_h^2}{T-h}, \quad (2.62)$$

dla dużych T ma rozkład χ_H^2 .

Przyjmując przybliżenie $\text{var}(r_h) = \frac{1}{T}$, zaproponowano test dla hipotezy zerowej $H_0 : r_h = 0$, $h = 1, 2, \dots, H$ postaci [Box i Pierce 1970]

$$BP = T \sum_{h=1}^H \hat{r}_h^2. \quad (2.63)$$

Dalsze badania pokazały, że duże rozbieżności między empirycznym rozkładem statystyki testowej (2.63) a teoretycznym rozkładem χ_H^2 wynikają przede wszystkim z przybliżenia (2.62) za pomocą formuły (2.63). Zaproponowano zatem testowanie hipotezy $H_0 : r_h = 0$, $h = 1, 2, \dots, H$ z wykorzystaniem statystyki postaci (2.62) [Ljung i Box 1978].

Konstrukcja powyższych testów opiera się na założeniu normalności zmiennych w szeregu czasowym, co jest niezgodne z faktycznym rozkładem I_t . W przypadku innych niż normalny rozkładów zmiennych losowych korzysta się z wyników badań dotyczących rozkładu współczynników autokorelacji z próby, pokazujących że asymptotyczna normalność wektora (r_1, r_2, \dots, r_H) nie wymaga normalności zmiennych, potrzebne jest jedynie założenie o skończonej wariancji [Anderson i Walker 1964].

Rozważając proces przekroczeń $(I_t)_{t=1}^T$, badanie autokorelacji z wykorzystaniem powyższego testu wymaga transformacji $Hit_t = I_t - p$, która pozwala otrzymać zerową wartość oczekiwaną $E(Hit_t) = 0$. Wówczas statystyka testu Ljunga-Boxa przyjmuje postać

$$LB = T(T+2) \sum_{h=1}^H \frac{\hat{\rho}_h^2}{T-h}, \quad (2.64)$$

gdzie ρ_h są współczynnikami autokorelacji rzędu h , $h = 1, 2, \dots, H$ w procesie $(Hit_t)_{t=1}^T$.

Wielowymiarowy test VaR stanowi uogólnienie powyższego testu Ljunga-Boxa. Przyjmijmy oznaczenie $I_t(p)$ dla zmiennej indykatorowej mówiącej, czy w okresie t , $t = 1, \dots, T$, wystąpiło przekroczenie VaR na poziomie tolerancji p . Zasady bezwarunkowego pokrycia oraz niezależności zmiennych w procesie $(I_t(p))$ zachodzą jeżeli $Hit_t(p) = I_t(p) - p$ jest ciągiem przyrostów martyngałowych, czyli kiedy

$$E(Hit_t(p)|\Omega_{t-1}) = 0,$$

$t = H + 1, \dots, T$, $1 \leq H < T$, gdzie zbiór Ω_{t-1} informacji dostępnych w okresie $t - 1$ może zawierać nie tylko przeszłe wartości procesu przekroczeń czyli $(Hit_1(p), Hit_2(p), \dots, Hit_{t-1}(p))$, ale również wartości każdej zmiennej dostępne w okresie $t - 1$, takie jak przeszłe prognozy VaR lub przekroczenia VaR związane z innym niż p poziomem tolerancji np. $(Hit_1(q), Hit_2(q), \dots, Hit_{t-1}(q))$, $q \in (0, 1)$, $q \neq p$.

Niech $t \in H + 1, \dots, T$ będzie ustalonym okresem. Własności ciągu przyrostów martyngałowych implikują, że dla każdej zmiennej $X_{t-h} \in \Omega_{t-1}$ zachodzi $E(Hit_t(p)) \otimes X_{t-h} = 0$, w szczególności z prawa iterowanej wartości oczekiwanej wynika, że $E(Hit_t(p)) = 0$ oraz dla każdego $h = 1, \dots, t - 1$

$$E(Hit_t(p)Hit_{t-h}(p)) = 0. \quad (2.65)$$

Jeżeli $Hit_{t-h}(q) \in \Omega_{t-1}$, to dla każdego $p, q \in (0, 1)$, $p \neq q$ i dla każdego $h = 1, 2, \dots, t - 1$

$$E(Hit_t(p)Hit_{t-h}(q)) = 0. \quad (2.66)$$

Równanie (2.66) stanowi podstawę wielowymiarowego uogólniania testu Ljunga-Boxa dla różnych poziomów tolerancji.

Niech $\mathbf{p} = (p_1, p_2, \dots, p_m)$ będzie ciągiem m różnych poziomów tolerancji z przedziału $(0, 1)$, interesujących z punktu widzenia zarządzania ryzykiem. Oznaczmy $Hit_t = (Hit_t(p_1), Hit_t(p_2), \dots, Hit_t(p_m))$ wektor zawierający informację o wyjątkach VaR w okresie t dla poziomów tolerancji p_1, p_2, \dots, p_m . Zgodnie z własnością niezależności dla procesu wektorowego $(Hit_t)_{t=1}^T$ zachodzi

$$cov(Hit_t Hit_{t-h}) = E(Hit_t Hit_{t-h}') = \delta_h \mathbf{V}, \quad (2.67)$$

gdzie \mathbf{V} jest $m \times m$ -wymiarową, symetryczną, niezerową macierzą a δ_h jest skalar, takim że

$$\delta_h = \begin{cases} 1 & \text{dla } h = 0, \\ 0 & \text{dla } h \neq 0. \end{cases} \quad (2.68)$$

Korzystając z powyższej własności, hipotezę zerową formułuje się w terminach współczynników autokorelacji jako

$$H_0 : cov(Hit_t Hit_{t-h}) = 0 \quad \forall h = 1, 2, \dots, H \quad (2.69)$$

lub w terminach niezależności zmiennych

$$H_0 : E(Hit_t(p_i) Hit_{t-h}(p_j)) = 0 \\ \forall h = 1, 2, \dots, H, \quad \forall i, j = 1, 2, \dots, m. \quad (2.70)$$

Oznaczmy \hat{C}_h empiryczną macierz kowariancji wektora Hit_t dla ustalonego rzędu opóźnienia $h \in \{1, 2, \dots, H\}$.

$$\hat{C}_h = [\hat{c}_{ijh}] = \sum_{t=h+1}^T Hit_t Hit'_{t-h}.$$

Niech $\hat{R}_h = D\hat{C}_hD$, gdzie D jest macierzą diagonalną, posiadającą na przekątnej odchylenia standardowe zmiennych $Hit_t(p_i)$, czyli $\sqrt{\hat{c}_{ii0}}$ dla $i = 1, 2, \dots, m$. Wielowymiarowa statystyka testu autokorelacji ma postać

$$LB_H = T \sum_{h=1}^H (\text{vec} \hat{R}_h)' (\hat{R}_0^{-1} \times \hat{R}_0^{-1}) (\text{vec} \hat{R}_h), \quad (2.71)$$

gdzie vec dla $m \times m$ -wymiarowej macierzy oznacza wektor, w którym element (i, j) macierzy znajduje się na pozycji $(j-1)m + 1$. Korzystając z badań dotyczących wielowymiarowych testów autokorelacji otrzymuje się, że przy założeniu prawdziwości hipotezy zerowej statystyka (2.71) ma asymptotyczny rozkład $\chi^2_{(Hm^2)}$ [Li i McLeod 1981].

Przeprowadzenie testu LB_H wymaga wyboru parametrów m i H . Ze względu na przeznaczenie testu rozważa się wartości $m \geq 2$, co oznacza, że weryfikacja modelu ryzyka uwzględnia co najmniej 2 poziomy tolerancji *VaR*. Zwiększanie m wiąże się ze wzrostem mocy testu. Z drugiej strony, badanie ryzyka koncentruje się zwykle na lewym ogonie rozkładu stopy zwrotu, co powoduje, że wybrane poziomy tolerancji p_i , $i = 1, 2, \dots, m$, są bliskie zeru i pochodzą z wąskiego zakresu. Wybór m może stwarzać problemy obliczeniowe, gdyż wyznaczenie wartości statystyki LB_H możliwe jest jedynie pod warunkiem, że macierz \hat{R}_0 jest nieosobliwa. Wykorzystanie wielu zbliżonych poziomów tolerancji zwiększa prawdopodobieństwo, że macierz \hat{R}_0 będzie osobliwa. Zatem zwiększanie m przy jednoczesnym ograniczeniu badania do

określonego zakresu kwantyli rozkładu może prowadzić do niemożliwości obliczenia wartości statystyki. Autorzy testu sugerują wykorzystanie dwóch lub trzech poziomów tolerancji, czyli $m = 2, 3$ [Hurlin i Tokpavi 2007].

Wartość H wpływa na szybkość zbieżności rozkładu statystyki LB_H do teoretycznego rozkładu $\chi^2_{(Hm^2)}$. Jak pokazały badania, wybór H powinien korespondować zarówno z wartością T jak i m . Eksperymenty symulacyjne, badające rozbieżność między rzeczywistym a teoretycznym rozkładem pokazały wyraźną tendencję statystyki Kołmogorova-Smirnova do odrzucania hipotezy o zgodności rozkładów dla dużych H . Dla prób o rozmiarach między $T = 250$ a $T = 1000$ obserwacji autorzy sugerują wartości $H = 1, 2, \dots, 5$ [Hurlin i Tokpavi 2007].

2.4.2. *Propozycja zastosowania wielowymiarowego testu macierzy korelacji*

Innym sposobem weryfikacji hipotezy o autokorelacji przekroczeń dla wielu różnych poziomów tolerancji VaR może być zastosowanie wielowymiarowych metod opartych na macierzy korelacji. Rozważmy test niezależności podwektorów losowych [por. np. Rencher 2002]. Podstawą tego testu jest założenie, że wektor wszystkich obserwowanych zmiennych losowych można podzielić na podwektory, w których każdy zawiera powiązane ze sobą zmienne. Weryfikacja hipotezy o niezależności polega na sprawdzeniu, czy zmienne znajdujące się w określonym podwektorze są niezależne od wszystkich zmiennych należących do innych podwektorów. Konstrukcja testu opiera się na zasadzie ilorazu wiarygodności i wykorzystuje estymatory macierzy korelacji podzielonej na bloki odpowiadające poszczególnym podwektorom.

Zastosowanie testu macierzy korelacji do oceny modelu opartego na VaR wymaga przyjęcia, że wektor wszystkich obserwowanych zmiennych zawiera ciągi przekroczeń VaR na różnych poziomach tolerancji, we wszystkich obserwowanych momentach. Następnie należy utworzyć podwektory losowe grupujące przekroczenia VaR na różnych poziomach tolerancji, które wystąpiły w tym samym czasie.

Niech \mathbf{X} będzie wektorem losowym o rozkładzie $\mathbf{X} \sim \mathbf{N}_n(\mu, \Sigma)$. Załóżmy, że wektor n zmiennych losowych \mathbf{X} został podzielony na $H + 1$ podwektorów składających się z pewnych powiązanych zmiennych

$$\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_H). \quad (2.72)$$

Wtedy macierz kowariancji Σ można zapisać w postaci blokowej

$$\Sigma = \begin{bmatrix} \Sigma_{00} & \Sigma_{01} & \dots & \Sigma_{0H} \\ \Sigma_{10} & \Sigma_{11} & \dots & \Sigma_{1H} \\ \vdots & \vdots & & \vdots \\ \Sigma_{H0} & \Sigma_{H1} & \dots & \Sigma_{HH} \end{bmatrix}, \quad (2.73)$$

gdzie m_h jest liczbą zmiennych należących do podwektora X_h , $h = 0, 1, \dots, H$, $m_1 + m_2 + \dots + m_H = n$. Przy powyższych oznaczeniach $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_H$ nie reprezentuje próby prostej lecz podział wektora losowego. Rozważmy hipotezę zerową mówiącą o niezależności podwektorów $\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_H$, tj. $H_0 : \Sigma_{gh} = \mathbf{0}$, $g, h \in \{0, \dots, H\}$, $g \neq h$, którą równoważnie można zapisać

$$H_0 : \Sigma = \begin{bmatrix} \Sigma_{00} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \Sigma_{11} & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \Sigma_{HH} \end{bmatrix}. \quad (2.74)$$

W celu zastosowania omawianego testu do weryfikacji modelu opartego na mierze *VaR*, przyjmijmy jak poprzednio oznaczenie $I_t(p)$, $t = 1, \dots, T$, $p \in (0, 1)$ dla zmiennych tworzących proces przekroczeń *VaR* na poziomie tolerancji p oraz załóżmy, że $\mathbf{p} = (p_1, p_2, \dots, p_m)$ jest ciągiem m poziomów tolerancji rozważanych w badaniu. Niech $Hit_t(p) = I_t(p) - p$. W ocenie modelu *VaR* wektor obserwowanych zmiennych ma postać $Hit = (Hit_t(p_1), Hit_t(p_2), \dots, Hit_t(p_{m_0}), \dots, Hit_{t-H}(p_1), Hit_{t-H}(p_2), \dots, Hit_{t-H}(p_{m_H}))$. Zmienne te można pogrupować w podwektory $Hit_{t-h} = (Hit_{t-h}(p_1), Hit_{t-h}(p_2), \dots, Hit_{t-h}(p_{m_h}))$, $h = 0, 1, \dots, H$, $1 \leq H < T$, obejmujące jeden okres oraz wszystkie rozważane w tym okresie poziomy tolerancji.

Wynikająca z kryteriów oceny modeli *VaR* własność ciągu przyrostów martyngałowych implikuje

$$H_0 : cov(Hit_t Hit_{t-h}) = 0 \quad \forall h = 1, 2, \dots, H, \quad (2.75)$$

co za pomocą macierzy kowariancji wektora Hit można przedstawić jako

$$H_0 : cov(HitHit) = \begin{bmatrix} cov(Hit_t Hit_t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & cov(Hit_{t-1} Hit_{t-1}) & \dots & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & cov(Hit_{t-H} Hit_{t-H}) \end{bmatrix}. \quad (2.76)$$

Powyższy zapis odpowiada hipotezie zerowej (2.74), gdzie przyjmuje się, że $\Sigma_{hh} = cov(Hit_{t-h}Hit_{t-h})$, $h = 0, 1, \dots, H$.

Statystyka testu ilorazu wiarygodności dla powyższej hipotezy ma postać

$$LR_H^1 = \frac{\det \tilde{\mathbf{R}}}{\det \hat{\mathbf{R}}_{00} \det \hat{\mathbf{R}}_{11} \cdots \det \hat{\mathbf{R}}_{HH}}, \quad (2.77)$$

gdzie $\hat{\mathbf{R}}_{hh}$, $h = 0, 1, \dots, H$ jest macierzą korelacji z próby dla podwektora Hit_{t-h} a $\tilde{\mathbf{R}}$ jest estymatorem macierzy korelacji dla wektora Hit przy założeniu, że zachodzi H_0 , czyli $\mathbf{R}_{gh} = \mathbf{0}$, $g, h = 0, 1, \dots, H$, $g \neq h$. Można pokazać, że dla $H = 1$ powyższa statystyka ma rozkład Wilksa [Rencher 2002, s. 259-261]. W przypadku braku normalności obserwowanego wektora losowego, analogicznie jak w teście Ljunga-Boxa, korzysta się z twierdzeń asymptotycznych. Przy $H > 1$ korzysta się z przybliżania rozkładem χ^2 rozkładu zmodyfikowanej statystyki

$$LR_H = (N - 1)c \ln LR_H^1, \quad (2.78)$$

gdzie

$$c = 1 - \frac{1}{12b(N - 1)}(2a_3 + 3a_2),$$

$$b = \frac{a_2}{2}, \quad a_2 = n^2 - \sum_{h=0}^H m_h^2, \quad a_3 = n^3 - \sum_{h=1}^H m_h^3,$$

natomiast N jest liczebnością próby.

Liczba parametrów w macierzy σ wynosi $\frac{1}{2}N(N + 1)$ natomiast przy założeniu hipotezy H_0 liczba parametrów w każdej z macierzy Σ_{hh} jest równa $\frac{1}{2}m_h(m_h + 1)$. Daje to całkowitą liczbę parametrów w modelu z restrykcjami $\frac{1}{2}\sum_{h=1}^H m_h(m_h + 1)$. Stąd liczba stopni swobody, równa różnicy między liczbą parametrów w modelu bez restrykcji a liczbą parametrów w modelu z restrykcjami, wynosi

$$\frac{1}{2}N(N + 1) - \frac{1}{2}\sum_{h=0}^H m_h(m_h + 1) = \frac{a_2}{p} = b.$$

Rozkład statystyki LR_H można zatem przybliżyć rozkładem χ_b^2 [Rencher 2002, s. 261-262].

W typowych sytuacjach związanych z oceną ryzyka można się spodziewać, że wybór analizowanych poziomów tolerancji nie zależy od rzędu opóźnienia zmiennej reprezentującej przekroczenia VaR . W przypadku gdy w każdym

wektorze Hit_{t-h} rozważane są te same poziomy tolerancji wszystkie badane podwektory mają tę samą liczbę elementów, $m_0 = m_1 = \dots = m_H = m$. Wtedy we wzorze (2.78) odpowiednie wyrażenia upraszczają się do postaci

$$b = \frac{a_2}{2}, \quad a_2 = n^2 - (H + 1)m^2, \quad a_3 = n^3 - (H + 1)m^3.$$

2.5. Testy *ES*

2.5.1. Testy parametryczne

Testy *ES* proponowane w literaturze różnią się między sobą pod względem sposobu sformułowania hipotez statystycznych, co powoduje, że weryfikują modele ryzyka pod różnymi aspektami. Najbardziej restrykcyjnym podejściem jest test wielkości przekroczenia (ang. exception magnitude), sprawdzający model *ES* poprzez zgodność empirycznego i teoretycznego rozkładu stopy zwrotu [Berkowitz 2001]. Stanowi on modyfikację testu LR_B (rozdz. 2.3.2.), wykorzystującego transformację stopy zwrotu R_t w zmienną $Z_t = \Phi^{-1}(\hat{F}_{R_t}(R_t))$. Podejście zaproponowane przez Berkowitza umożliwia, za pomocą powyższej transformacji, weryfikację zgodności empirycznego rozkładu zmiennej R_t z rozkładem teoretycznym poprzez testowanie istotności parametrów w modelu regresji ze składnikiem losowym o rozkładzie normalnym (przy założeniu prawdziwości hipotezy zerowej). Zastosowanie powyższej procedury testowej, w jej klasycznej wersji, prowadzi do odrzucenia modelu ryzyka jeżeli wystąpi niezgodność rozkładów zarówno w ogonach jak i w obszarze środka rozkładu. W celu opracowania procedury koncentrującej się na ogonie rozkładów zaproponowano konstrukcję testu opartego na cenzurowanej funkcji wiarygodności. Dzięki temu otrzymano test mający zastosowanie do sprawdzania zgodności największych odchyłeń stopy zwrotu z modelem *ES*. Obszar krytyczny wyznaczony dla funkcji gęstości ogona rozkładu wykorzystywany jest jako przybliżenie obszaru krytycznego dla średniej wielkości przekroczenia ponad ustalony poziom tolerancji tj. $E(R_t | R_t < VaR_{t,p})$.

W teście wielkości przekroczeń VaR , podobnie jak w teście Berkowitza dla całego rozkładu stopy zwrotu, rozważa się model

$$Z_t - \mu = \rho_1(Z_{t-1} - \mu) + \epsilon_t, \quad \epsilon \sim N(0, \sigma^2), \quad (2.79)$$

gdzie $Z_t = \Phi^{-1}(\hat{F}_{R_t}(R_t))$, $t = 1, \dots, T$, jest zmienną losową otrzymaną w wyniku podwójnej transformacji. W pierwszym kroku transformacji, przy założeniu poprawnej specyfikacji funkcji gęstości stopy zwrotu w modelu *ES*,

otrzymuje się zmienną o rozkładzie równomiernym $U_{(0,1)}$, którą w drugim kroku przekształca się w zmienną o rozkładzie normalnym. W odpowiednio skonstruowanym modelu powinien zatem zachodzić warunek (2.41) postaci $Z_t \stackrel{iid}{\sim} N(0, 1)$, $t = 1, \dots, T$. W podstawowym podejściu proponuje się założenie $\rho_1 = 0$ i testowanie hipotezy zerowej $H_0 : \mu = 0, \sigma = 1$ względem hipotezy alternatywnej $H_0 : \mu \neq 0, \sigma \neq 1$.

W celu zawężenia procedury jedynie do badania ogona rozkładu stóp zwrotu, wykorzystuje się wartości zmiennej Z_t wykraczające poza ustalony próg. Definiuje się proces przekroczeń zmiennej Z_t ponad próg $\Phi^{-1}(p)$,

$$Z_t^* = \begin{cases} 0, & \text{gdy } Z_t \geq \Phi^{-1}(p), \\ Z_t, & \text{gdy } Z_t < \Phi^{-1}(p), \end{cases} \quad (2.80)$$

$t = 1, \dots, T$, gdzie p jest poziomem tolerancji przyjętym przy szacowaniu parametrów modelu ES.

W rozważanym modelu cenzurowana logarytmiczna funkcja wiarygodności parametrów μ i σ dla próby $Z_1^*, Z_2^*, \dots, Z_T^*$ ma postać

$$\begin{aligned} \ln L(\mu, \sigma, Z_1^*, Z_2^*, \dots, Z_T^*) &= \\ &= \sum_{t=1}^T \ln \phi\left(\frac{Z_t^* - \mu}{\sigma}\right) \mathbf{1}_{(Z_t^* = Z_t)} + \sum_{t=1}^T \ln\left(1 - \Phi\left(\frac{\Phi^{-1}(p) - \mu}{\sigma}\right)\right) \mathbf{1}_{(Z_t^* = 0)} = \\ &= \sum_{t=1}^T \left(-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(Z_t^* - \mu)^2\right) \mathbf{1}_{(Z_t^* = Z_t)} \\ &\quad + \sum_{t=1}^T \ln\left(1 - \Phi\left(\frac{\Phi^{-1}(p) - \mu}{\sigma}\right)\right) \mathbf{1}_{(Z_t^* = 0)}. \end{aligned} \quad (2.81)$$

Statystyka testowa oparta na funkcji wiarygodności powyższej postaci weryfikuje zgodność rozkładów na podstawie wybranego ogona. Pierwszy składnik wyrażenia jest funkcją wiarygodności obserwacji najbardziej oddalonych od środka rozkładu, pochodzących z rozkładu normalnego, uznanych za kluczowe z punktu widzenia ryzyka, natomiast drugi składnik może być interpretowany jako czynnik normalizujący, wynikający z cenzurowania próby.

Według zasady ilorazu wiarygodności, testowanie hipotezy zerowej obejmującej parametry μ i σ wykonuje się za pomocą statystyki

$$LR_B^{ES} = -2(\ln L(0, 1, Z_1^*, Z_2^*, \dots, Z_T^*) - \ln L(\hat{\mu}, \hat{\sigma}, Z_1^*, Z_2^*, \dots, Z_T^*)), \quad (2.82)$$

która ma asymptotyczny rozkład $\chi_{(2)}^2$ [Berkowitz 2001].

Analogicznie jak w teście Berkowitza dla całego rozkładu, rozszerzeniem powyższego podejścia jest badanie autokorelacji największych przekroczeń poprzez test istotności parametru ρ_1 w rozważanym modelu. Ogólna wersja testu może zawierać współczynniki autokorelacji wyższych rzędów. Podejście to może być również modyfikowane poprzez rozszerzenie modelu (2.79) o inne zmienne objaśniające, jak przeszłe poziomy *VaR* czy oszacowania wcześniejszej zmienności. Statystyka testowa (2.82) opisuje wszelkiego rodzaju rozbieżności między wynikami próby a zakładanymi wartościami parametrów μ i σ .

Test wielkości przekroczeń LR_B^{ES} stanowi pośrednie podejście pomiędzy testami weryfikującymi modele *VaR* i *ES* w sposób punktowy, w oparciu o wybraną statystykę jak średnia wielkość przekroczenia czy kwantyl rozkładu a testami badającymi zgodność całego rozkładu. Omówiony test jest testem zgodności dla ogona rozkładu, oparty jest więc na większej ilości informacji w stosunku do testów porównujących prognozowaną oraz zrealizowaną średnią wielkość największych przekroczeń, co pozwala uzyskać większą moc. Z drugiej strony, z konstrukcji testu wynika, że ma on jedynie moc wykrywania niezgodności rozkładów poprzez dwa pierwsze momenty μ i σ oraz ewentualne współczynniki autokorelacji. Zastosowanie tej procedury jest niezależne od rozkładu zmiennej R_t . Innym rozwiązaniem, omijającym problem założeń o rozkładzie stopy zwrotu jest wykorzystanie nieparametrycznych metod bootstrapowych opisanych niżej.

Ze względu na definicję *ES* jako wartości oczekiwanej ogona rozkładu (w przypadku zmiennych losowych ciągłych) naturalnym sposobem weryfikacji modelu *ES* jest porównywanie oszacowań tej miary ze średnią wartością skrajnych obserwacji. Problem konstrukcji takiego testu polega na nieznanym rozkładzie odpowiedniej statystyki. Jako możliwe rozwiązanie zaproponowano wykorzystanie techniki punktu siodłowego¹⁶, umożliwiającej przybliżanie rozkładów sum zmiennych losowych z wykorzystaniem funkcji generujących momenty i kumulanty odpowiednich rozkładów oraz ich pochodnych [Wong 2008]. Metoda Wonga oryginalnie opracowana została dla rozkładu normalnego stopy zwrotu, możliwe jest jednak jej stosowanie dla dowolnych rozkładów. W odróżnieniu od testu wielkości wyjątków LR_B^{ES} , sprawdzającego zgodność rozkładów, test oparty na technice punktu siodłowego dedykowany jest weryfikacji hipotezy o wartości miary *ES*.

¹⁶ Technika punktu siodłowego umożliwia przybliżanie rozkładu sumy zmiennych losowych bez wykorzystania twierdzeń granicznych [Lugannani i Rice 1980, jej omówienie i rozszerzenie można znaleźć w pracy Butler i Wood 2004].

Rozważmy stopę zwrotu R_t pochodzącą z rozkładu normalnego o dystrybucie Φ i gęstości ϕ , założmy ponadto, że rozkład ten jest niezmienny w czasie i oznaczmy $R = R_t$, $t = 1, \dots, T$. Wówczas można zapisać $ES_p(R_t) = ES_p(R)$, $t = 1, \dots, T$ oraz sformułować hipotezę zerową jako $H_0 : ES_p(R) = ES_p^{(0)}(R)$ wobec hipotezy alternatywnej $H_1 : ES_p(R) > ES_p^{(0)}(R)$. W przypadku testowania hipotezy zakładającej rozkład normalny stopy zwrotu przyjmuje się $ES_p^{(0)}(R) = -\mu_X$, gdzie μ_X jest wartością oczekiwaną odpowiednich skrajnych obserwacji rozkładu normalnego.

Określmy zmienną losową X za pomocą wzoru

$$P(X \leq x) = P(R \leq x | R < q(p)) = \frac{P(R \leq x)}{P(R < q(p))}, \quad x < q(p), \quad (2.83)$$

gdzie $q(p) = \Phi^{-1}(p)$ jest kwantylem rozkładu normalnego rzędu p . Zmienna X reprezentuje zatem stopę zwrotu pod warunkiem, że $R < q(p)$ a oczekiwany niedobór ES na poziomie tolerancji p wyraża się jako $-E(X)$. Funkcją gęstości X jest

$$f_X(x) = \frac{1}{p}\phi(x), \quad (2.84)$$

natomiast dystrybuantą

$$F_X(x) = \frac{1}{p}\Phi(x), \quad (2.85)$$

dla $x < q(p)$. Za pomocą funkcji generującej momenty łatwo uzyskuje się wartość oczekiwaną i wariancję zmiennej X . Funkcja generująca momenty zmiennej X ma postać [Wong 2008]

$$M_X(t) = \frac{1}{p}e^{\frac{t^2}{2}}\phi(q(p) - t), \quad (2.86)$$

a jej pochodne dane są wzorami

$$M'_X(t) = tM_X(t) - \frac{1}{p}e^{q(p)t}\phi(q(p)), \quad (2.87)$$

$$M''_X(t) = tM'_X(t) + M_X(t) - \frac{1}{p}e^{q(p)t}q(p)\phi(q(p)), \quad (2.88)$$

$$M_X^{(k)}(t) = tM_X^{(k-1)}(t) + (k-1)M_X^{(k-2)}(t) - \frac{1}{p}e^{q(p)t}q(p)^{(k-1)}\phi(q(p)), \quad k \geq 3. \quad (2.89)$$

Wówczas wartość oczekiwana i wariancja zmiennej X dane są odpowiednio wzorami

$$\mu_X = E(X) = -\frac{\phi q(p)}{p} = -\frac{e^{-\frac{q(p)}{2}}}{p\sqrt{2\Pi}}, \quad (2.90)$$

$$\sigma_X^2 = \text{var}(X) = 1 - \frac{q(p)\phi(q(p))}{p} - \mu_X^2 = 1 - \frac{q(p)e^{-\frac{q(p)}{2}}}{p\sqrt{2\Pi}} - \mu_X^2. \quad (2.91)$$

Ponieważ $ES_p(R) = -E(X)$, to średnia z próby $-\bar{X}$ jest naturalnym estymatorem tej miary i jest ona wykorzystywana jako statystyka testowa S , weryfikująca hipotezę o wartości *ES*, czyli $S = -\bar{X}$. Wnioskowanie statystyczne na podstawie S wymaga znajomości rozkładu tej statystyki lub przynajmniej odpowiedniej p -wartości. Wykorzystanie faktu, że $\frac{\sqrt{N}(\bar{X}-\mu_X)}{\sigma_X}$ ma asymptotyczny rozkład normalny wydaje się nieodpowiednie w świetle niewielkiej liczby wyjątków *VaR* występujących w praktyce, proponuje się zatem szacowanie prawdopodobieństwa $P(\bar{X} \leq \bar{x})$ z wykorzystaniem przybliżonego rozkładu sumy zmiennych losowych, otrzymanego techniką punktu siodłowego. Za pomocą funkcji generującej kumulanty można zapisać funkcję gęstości średniej \bar{X} jako

$$f_{\bar{X}}(\bar{x}) = \frac{N}{2\Pi} \int_{-\infty}^{\infty} e^{N(K_X(it) - it\bar{x})} dt, \quad (2.92)$$

gdzie $K_X(t) = \ln M_X(t)$ jest funkcją generującą kumulanty zmiennej X . Wówczas

$$P(\bar{X} > \bar{x}) = \int_{\bar{x}}^{q(p)} f_{\bar{X}}(u) du = \frac{1}{2\Pi i} \int_{s-i\infty}^{s+i\infty} e^{N(K_X(t) - t\bar{x})} t^{-1} dt, \quad (2.93)$$

gdzie s jest punktem siodłowym spełniającym

$$K'_X(s) = \bar{x} \quad (2.94)$$

[Wong 2008]. Można pokazać, że jeżeli s spełnia (2.94), to, z wykorzystaniem tego punktu, możliwe jest przybliżenie powyższego prawdopodobieństwa $P(\bar{X} > \bar{x})$, dzięki czemu uzyskuje się formułę stanowiącą przybliżenie p -wartości w teście dla *ES*:

$$P(\bar{X} \leq \bar{x}) = \begin{cases} \Phi(\xi) - \phi(\xi)\left(\frac{1}{\eta} - \frac{1}{\xi} + \mathbf{O}(N^{-\frac{3}{2}})\right) & \text{dla } \bar{x} < q(p), \\ 1 & \text{dla } \bar{x} \geq q(p), \end{cases} \quad (2.95)$$

gdzie $\eta = s\sqrt{NK_X''(s)}$, $\xi = \text{sgn}(s)\sqrt{2N(s\bar{x} - K_X(s))}$, $\bar{x} \neq \mu_x$ [Lugannani i Rice 1980].

Chociaż przedstawiona wyżej procedura zakłada normalność¹⁷ rozkładu stopy zwrotu R , możliwe jest przeprowadzenie testu opartego na technice punktu siodłowego dla innych rozkładów, wymaga to jednak dodatkowych obliczeń związanych z wyznaczeniem funkcji generujących momenty, kumulanty, ich pochodnych oraz zastosowaniem wzoru (2.93). Alternatywnym rozwiązaniem w przypadku przyjęcia założenia o innym rozkładzie jest wykorzystanie transformacji przekształcającej zmienne o ustalonym rozkładzie F_{R_t} , $t = 1, 2, \dots, T$, w zmienne o rozkładzie normalnym, analogicznie jak w testach LR_B (rozdz. 2.3.2.) oraz LR_B^{ES} .

Przewagą testu S opartego na technice punktu siodłowego w stosunku do testu LR_B^{ES} jest wykorzystanie przybliżenia dla rozkładu sumy zmiennych losowych z pominięciem przejścia asymptotycznego, co umożliwia stosowanie procedury przy niewielkich rozmiarach próby. Jest to szczególnie istotne w przypadku testów ES , które opierają się wyłącznie na obserwacjach przekraczających VaR , stanowiących niewielką część całkowitej liczby obserwacji. Test S możliwy jest do przeprowadzenia w sytuacji zaobserwowania tylko jednego wyjątku VaR .

W przypadku zmiennych losowych ciągłych, definicja miary ES sprowadza się do wartości oczekiwanej, zatem można rozważać testowanie modelu ES za pomocą regresji liniowej. Wykorzystuje się tu analogię pomiędzy modelem regresji liniowej a pojęciem warunkowej wartości oczekiwanej. Wartość oczekiwana różnicy między stopą zwrotu a wartością miary ES , pod warunkiem, że stopa zwrotu przekracza VaR , wynosi zero, co wynika bezpośrednio z definicji miary ES . Przy poprawnym modelu ES wszelkie regresje objaśniające tę różnicę za pomocą dostępnych zmiennych są statystycznie nieistotne. Proponuje się zatem konstrukcję testu w oparciu o badanie istotności regresji różnicy między stopą zwrotu a oszacowaniem ES względem wektora \mathbf{X}_{t-1} , zawierającego zmienne dostępne w okresie $t - 1$, w szczególności opóźnienia stopy zwrotu od 1 do H :

$$-R_t - \hat{ES}_p(R_t) = a + b\mathbf{X}_{t-1} + \epsilon_t, \text{ gdzie } R_t < -V\hat{a}R_p(R_t), \quad (2.96)$$

$t = H + 1, \dots, T$, $1 \leq H < T$. Powyższą regresję definiuje się jedynie dla stóp zwrotu przekraczających VaR , pozostałe obserwacje usuwane są z próby. Zakłada się, że ϵ_t są niezależnymi zmiennymi losowymi o jednakowym rozkładzie oraz są niezależne od \mathbf{X}_{t-1} . Weryfikacja poprawności modelu ES może polegać na sprawdzaniu istotności parametru a ($H_0 : a = 0$) parametru b

¹⁷ Sposób wyznaczania punktu siodłowego s w przypadku rozkładu normalnego można znaleźć w literaturze [Wong 2008].

($H_0 : b = 0$) lub obu parametrów łącznie ($H_0 : a = 0 \wedge b = 0$) za pomocą standardowej statystyki Fischera, oznaczanej w tym przypadku F_{Ch} [Christoffersen 2012, s. 308–309].

Wnioskowanie statystyczne na podstawie regresji (2.96) może być obciążone błędami ze względu na niespełnienie założeń o składniku losowym ϵ_t . W ogólnym przypadku, zarówno zmienna R_t jak i $ES_p(R_t)$ posiadają rozkłady charakteryzujące się zmiennością w czasie, co przekłada się na brak stacjonarności składnika losowego ϵ_t i może skutkować brakiem zbieżności rozkładu statystyki F_{Ch} do rozkładu Fischera. Podstawowym problemem może być heteroskedastyczność wynikająca ze zmienności wariancji zmiennej R_t . Ze względu na możliwość popełniania dużych błędów pierwszego rodzaju, wynikających z różnicy między rzeczywistym a teoretycznym rozkładem statystyki testowej, w niniejszej pracy proponuje się przeprowadzenie standaryzacji zmiennej objaśniającej względem odchylenia standardowego σ_t . W tym celu należy wykorzystać estymator $\hat{\sigma}_t$. Procedura standaryzacji zmiennej objaśnianej prowadzi do regresji postaci

$$\frac{-R_t - \hat{E}S_p(R_t)}{\hat{\sigma}_t} = a^* + b^* \mathbf{X}_{t-1} + \epsilon_t^*, \text{ gdzie } R_t < -V\hat{a}R_p(R_t), \quad (2.97)$$

$t = H + 1, \dots, T$, $1 \leq H < T$. Zakładając stacjonarność zmiennych objaśniających \mathbf{X}_{t-1} otrzymuje się, że ϵ_t^* jest zmienną losową posiadającą rozkład w przybliżeniu stacjonarny, $t = H + 1, \dots, T$. Tak samo jak w przypadku testu opartego na regresji (2.96), zmodyfikowany test F_{Ch}^* , oparty na równaniu (2.97), wykorzystuje statystykę Fischera.

Parametryczny charakter powyższych testów pozwala spodziewać się wyższej mocy niż w przypadku procedur nieparametrycznych. Z drugiej strony konieczność przyjęcia założenia o rozkładzie stopy zwrotu nakłada ograniczenia na ich stosowanie w praktyce. Niezbędne jest rozważenie błędów wynikających z przyjęcia rozkładu niezgodnego z rzeczywistym, co może być skomplikowane i przemawia za poszukiwaniem metod nieparametrycznych.

2.5.2. Testy nieparametryczne

Przedstawione wyżej testy parametryczne miary *ES* bazują na transformacjach wyjściowego procesu stopy zwrotu lub procesu przekroczeń, prowadzących do znanej postaci parametrycznej rozkładu. Bezpośrednie zastosowanie statystyki testowej, będącej naturalną miarą odległości między wartością *ES* zgodną z hipotezą zerową a jej empiryczną realizacją nie jest możliwe, ponieważ rozkład tej statystyki jest nieznan. W celu przeprowadzenia testu

opartego na odległości między zakładaną a empiryczną wartością ES , problem nieznanego rozkładu statystyki testowej można rozwiązać poprzez wykorzystanie techniki testu bootstrapowego [Efron i Tibshirani 1993].

Założmy, że $(R_t)_{t=1}^T$ jest fragmentem ściśle stacjonarnego procesu stochastycznego reprezentującego obserwacje logarytmicznej stopy zwrotu. Naturalną miarą rozbieżności między modelem ES a obserwacjami stopy zwrotu jest statystyka

$$V = \frac{\sum_{t=1}^T (R_t - (-\hat{E}S_p(R_t))) \mathbf{1}_{\{R_t < VaR_p(R_t)\}}}{\sum_{t=1}^T \mathbf{1}_{\{R_t < VaR_p(R_t)\}}}. \quad (2.98)$$

Ponieważ w ogólnym przypadku, gdy dopuszcza się zmienność wariancji procesu w czasie, szereg obserwacji (R_1, R_2, \dots, R_T) nie stanowi próby pochodzącej z jednakowego rozkładu, rozważa się podejście oparte na warunkowych resztach procesu przekroczeń

$$U_t = \frac{R_t - (-\hat{E}S_p(R_t))}{\hat{\sigma}_t}, \quad (2.99)$$

dla $-R_t > VaR_p(R_t)$, gdzie zakładany jest ogólny model stopy zwrotu postaci

$$R_t = \mu_t + \sigma_t Z_t, \quad (2.100)$$

Z_t są niezależnymi, jednakowo rozłożonymi zmiennymi o wartości oczekiwanej zero i jednostkowej wariancji, $t = 1, \dots, T$ [McNeil i Frey 2000].

Przy założeniach $\hat{\mu}_t = \mu_t$, $\hat{\sigma}_t = \sigma_t$ i $\hat{F}_{Z_t}^{-1}(p) = F_{Z_t}^{-1}(p)$ mamy

$$\begin{aligned} U_t &= \frac{R_t - (-\hat{E}S_{t,p})}{\sigma_t} = \\ &= \frac{\mu_t + \sigma_t Z_t - (\hat{\mu}_t + \hat{\sigma}_t E(Z_t | Z_t < \hat{F}_{Z_t}^{-1}(p)))}{\sigma_t} = \\ &= \frac{\mu_t + \sigma_t Z_t - (\mu_t + \sigma_t E(Z_t | Z_t < F_{Z_t}^{-1}(p)))}{\sigma_t} = \\ &= Z_t - E(Z_t | Z_t < F_{Z_t}^{-1}(p)), \end{aligned} \quad (2.101)$$

gdzie $F_{Z_t}^{-1}(p)$ jest kwantylem rzędu p zmiennej losowej Z_t .

Wprowadzenie dodatkowej standaryzacji postaci

$$\tilde{U}_t = \frac{U_t}{\hat{v}ar(Z_t | Z_t < F_{Z_t}^{-1}(p))} \quad (2.102)$$

umożliwia, przy założeniu $\hat{var}(Z_t|Z_t < F_{Z_t}^{-1}(p)) = var(Z_t|Z_t < F_{Z_t}^{-1}(p))$, otrzymanie niezależnych reszt warunkowych o jednakowym rozkładzie, zerowej wartości oczekiwanej i jednostkowej wariancji¹⁸ [McNeil i Frey 2000].

Niech t_i oznacza okres i -tego wyjątku *VaR*, gdzie $i = 1, \dots, N$, N jest liczbą wyjątków. Oznaczając wartość oczekiwaną zmiennych $\tilde{U}_{(i)} = \tilde{U}_{t_i}$ przez μ , t.j. $\mu = E(\tilde{U}_{(i)})$, weryfikację modelu *ES* przeprowadza się w formie testu hipotezy $H_0 : \mu = 0$. Rozważa się statystykę

$$U = \frac{\tilde{U} - \mu}{S/\sqrt{N}}, \quad (2.103)$$

gdzie

$$\mu = 0,$$

$$\tilde{U} = \frac{1}{N} \sum_{i=1}^N \tilde{U}_{(i)},$$

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\tilde{U}_{(i)} - \tilde{U})^2}.$$

W celu weryfikacji hipotezy H_0 należy wygenerować M prób bootstrapowych $U_{1,j}^B, U_{2,j}^B, \dots, U_{N,j}^B$, $j = 1, \dots, M$, według rozkładu bootstrapowego, w którym każda obserwacja z próby $\tilde{U}_{(1)}, \tilde{U}_{(2)}, \dots, \tilde{U}_{(N)}$ ma jednakowe prawdopodobieństwo inkluzji do próby bootstrapowej, wynoszące $\frac{1}{N}$. Na podstawie j -tej wartości próby bootstrapowej oblicza się

$$\tilde{U}_{j,B} = \frac{1}{N} \sum_{i=1}^N U_{i,j}^B,$$

$$S_{j,B} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (U_{i,j}^B - \tilde{U}_{j,B})^2}$$

oraz

$$U_{j,B} = \frac{\tilde{U}_{j,B} - \tilde{U}_{j,B}}{S_{j,B}/\sqrt{N}}. \quad (2.104)$$

¹⁸ Analityczne postaci $var(Z_t|Z_t < F_{Z_t}^{-1}(p))$ w przypadku, gdy $Z_t \sim N(0, 1)$, $Z_t \sim S_\nu$ oraz w przypadku, gdy $Z_t - a|Z_t > a \sim GPD(\xi, \beta)$ można znaleźć w literaturze [np. McNeil i Frey 2000, Righi i Ceretta 2013].

P -wartość w teście bootstrapowym wyznacza się jako

$$\hat{p}_N(U) = \frac{1 + \sum_{j=1}^M \mathbf{1}_{\{|U_{j,B}| > |U|\}}}{1 + M},$$

gdzie wartość 1 jest dodawana w celu uniknięcia zerowych p -wartości¹⁹.

Omówiony powyżej test bootstrapowy może zostać wykonany z pominięciem standaryzacji (2.102). Wówczas procedurę przeprowadza się na podstawie reszt postaci (2.99).

Słabością rozważanych testów jest fakt, że zależą one od wyznaczonych na podstawie pewnego modelu oszacowań VaR . Skuteczność modelu VaR przekłada się na wynik testu, ponieważ w procedurze testowej uwzględniane są jedynie te wyniki z próby (R_1, R_2, \dots, R_T) , które wykraczają poza oszacowaną wartość VaR . Przy złym modelu VaR , udział takich obserwacji może znacznie różnić się od zakładanego odsetka populacji, dla którego wyznacza się średnią wartość przekroczenia. Wprowadza się zatem modyfikację powyższych testów, wykorzystującą kwantyl z próby zamiast oszacowania VaR . W ten sposób w zmodyfikowanym teście uwzględniane są wartości zmiennej, które w zaobserwowanej próbie wystąpiły w jednym ze $\frac{1}{p}$ skrajnych przypadków. Uwzględniając tę modyfikację otrzymuje się statystykę postaci

$$V^* = \frac{\sum_{t=1}^T D_t \mathbf{1}_{\{D_t < D_p\}}}{\sum_{t=1}^T \mathbf{1}_{\{D_t < D_p\}}}, \quad (2.105)$$

gdzie $D_t = R_t - (-\hat{E}S_p(R_t))$, $t = 1, \dots, T$, a D_p oznacza kwantyl empiryczny rzędu p szeregu (D_1, D_2, \dots, D_T) . W literaturze proponuje się również połączenie statystyk V i V^* oraz weryfikację modelu ES na podstawie $\frac{|V| + |V^*|}{2}$ [Embrechts, Kaufmann i Patie 2005].

Statystyka V^* , stanowiąc odpowiednik miary odległości (2.98), podobnie jak ona, nie jest poddawana standaryzacji ze względu na odchylenie standardowe zmiennych R_t , $t = 1, \dots, T$, które może ulegać zmianom w czasie. Uwzględniając problem heteroskedastyczność w bazowym procesie stopy zwrotu, można rozważać przeprowadzenie testu wprowadzając zmienną

$$U_t^* = \frac{R_t - (-\hat{E}S_p(R_t))}{\hat{\sigma}_t}, \quad (2.106)$$

¹⁹ Omówienie metody bootstrap oraz testów bootstrapowych można znaleźć w literaturze statystycznej polskiej [np. Kończak 2011, Pekasiewicz 2012, Domański et al. 2014] lub zagranicznej [np. Chernick 2008].

dla wartości spełniających warunek $D_t < D_p$, czyli $U_{(j)}^* = U_{t_j}$, $j = 1, \dots, N^*$, gdzie t_j jest numerem j -tej obserwacji, która mniejsza jest od kwantyla empirycznego rzędu p , natomiast N^* jest liczbą takich obserwacji. W oparciu o otrzymaną próbę $(U_{(1)}^*, U_{(2)}^*, \dots, U_{(N^*)}^*)$ można przeprowadzić test bootstrapowy analogicznie jak dla zmiennych \tilde{U}_t oraz U_t . Oznaczmy $\mu^* = E(U_j^*)$. Wówczas statystyka testowa przyjmuje postać

$$U^* = \frac{\bar{U}^* - \mu^*}{S^*/\sqrt{N^*}}, \quad (2.107)$$

gdzie

$$\mu^* = 0,$$

$$\bar{U}^* = \frac{1}{N^*} \sum_{j=1}^{N^*} U_{(j)}^*$$

oraz

$$S^* = \sqrt{\frac{1}{N^* - 1} \sum_{j=1}^{N^*} (U_{(j)}^* - \bar{U}^*)^2}.$$

Rozdział 3

Ocena własności statystycznych testów wartości zagrożonej (VaR) i oczekiwanego niedoboru (ES)

Rozwój złożonych, nieliniowych modeli statystycznych, związanych z oceną ryzyka rynkowego, doprowadził do szerokiego wykorzystania metody największej wiarygodności, w oparciu o którą skonstruowana została większość testów rozważanych w niniejszym badaniu. W przypadku testów opartych na tej metodzie przeważnie znane są jedynie rozkłady asymptotyczne. Często konkurencyjne procedury testowe są asymptotycznie równoważne, różnią się jednak obliczeniowo oraz posiadają inne własności dla prób skończonych. Dla niewielkich prób korzystanie z rozkładów asymptotycznych może prowadzić do dużych błędów, co zwłaszcza ma miejsce w sytuacji, gdy rozmiar próby rozumiany jest jako liczba wyjątków VaR , która zwykle jest mała.

W celu oceny prawdopodobieństwa błędu pierwszego oraz drugiego rodzaju, w niniejszym rozdziale przeprowadzono analizę porównawczą rozmiaru i mocy testów stosowanych w ocenie modeli ryzyka rynkowego. Analiza własności testów przeprowadzona została z wykorzystaniem metody Monte Carlo. Symulacyjnie wyznaczono kwantyle rozkładów statystyk testowych

dla prób skończonych oraz oszacowano rozmiar testów¹. Jako wstęp do badania symulacyjnego przedstawiono problem odpowiedniej konstrukcji eksperymentów Monte Carlo, ze szczególnym uwzględnieniem eksperymentów służących ocenie mocy testów autokorelacji.

W rozważanych testach autokorelacji *VaR* badanie mocy można przeprowadzać na podstawie różnych rodzajów eksperymentów, wykorzystujących generowanie zmiennych zależnych. W analizach związanych z zarządzaniem ryzykiem zwykle przyjmowane są eksperymenty reprezentujące procesy finansowe, bazujące na modelu *GARCH*. Ze względu na uniwersalny charakter miar *VaR* i *ES* oraz szeroki zakres zastosowań, wykraczający poza rynek finansowy, w niniejszej pracy zaproponowano rozszerzenie analizy mocy testów z wykorzystaniem innych schematów autokorelacji. Rozważane były eksperymenty pozwalające na generowanie danych skorelowanych o brzegowym rozkładzie gamma, które mogą być wykorzystane do odzwierciedlenia zjawiska zwanego gromadzeniem wariancji. Analizowano struktury autokorelacji pierwszego rzędu oraz wyższych rzędów z wygasającymi wykładniczo współczynnikami autokorelacji. Rozważano również generator bazujący na własności Markowa. Dla poszczególnych schematów uwzględniono różną siłę korelacji. Przedstawiono związek między określoną strukturą autokorelacji szeregów czasowych a wyborem testów statystycznych stosowanych w analizie ryzyka.

W wyniku analizy rozmiaru i mocy wyłoniono grupy testów gwarantujące najwyższy stopień kontroli nad błędem pierwszego rodzaju oraz najmniejszy błąd drugiego rodzaju. Podsumowując analizę porównawczą w ramach poszczególnych grup testów, dokonano wyboru testów optymalnych w świetle przeprowadzonego badania symulacyjnego, gdzie test optymalny rozumiany był jako test, który cechuje się przewagą nad innymi w sensie dokładności – czyli zachowania rozmiaru zgodnego z założonym rozmiarem nominalnym – oraz w sensie skuteczności – czyli wyższej mocy wobec rozważanych hipotez alternatywnych niż inne testy o podobnym rozmiarze. Przy wyborze procedur optymalnych w pewnych przypadkach kierowano się dodatkowo praktycznymi aspektami, związanymi z możliwością aplikacji do danych rynkowych.

¹ Empiryczne kwantyle badanych rozkładów nie zostały zamieszczone w pracy ze względu na jej objętość i przejrzystość wywodu, są jednak dostępne na życzenie u autorki.

3.1. Projektowanie eksperymentów do oceny rozmiaru

3.1.1. *Eksperymenty oparte na procesie Bernoulliego*

Zgodnie z kryteriami weryfikacji modelu ryzyka (rozdz. 1.2.2.), podstawą statystycznej oceny systemów zarządzania ryzykiem są testy autokorelacji. Hipoteza zerowa takich testów zakłada brak autokorelacji, zatem uzyskanie modelu z nią zgodnego następuje poprzez generowanie ciągu realizacji niezależnych zmiennych losowych. Przy ocenie modeli ryzyka, rozumianego jako wartość zagrożona, zmienne losowe podlegające testowaniu mają w większości charakter zero-jedynkowy² i oznaczają przekroczenie dopuszczalnego poziomu ryzyka lub jego brak. W tym przypadku model zgodny z hipotezą zerową otrzymuje się poprzez wygenerowanie T wyników niezależnych prób Bernoulliego z prawdopodobieństwem sukcesu p , równym poziomowi tolerancji założonemu przy szacowaniu VaR . W ten sposób otrzymuje się ciąg realizacji procesu wyjątków zgodny zarówno z hipotezą o braku autokorelacji przekroczeń, jak również z hipotezą o bezwarunkowym prawdopodobieństwie przekroczenia na poziomie p .

Duża liczba testów stosowanych w analizie szeregów czasowych skonstruowana jest w oparciu o metodę ilorazu wiarygodności i znane są jedynie asymptotyczne rozkłady statystyk testowych. Przy założeniu nieskończonej wielkości próby rozkład ilorazu wiarygodności dąży do rozkładu χ^2 . Dla niewielkich prób korzystanie z tego rozkładu do wyznaczenia obszaru krytycznego może prowadzić do dużych błędów. Zwłaszcza ma to miejsce w sytuacji, gdy rozmiar próby rozumiany jest jako liczba wyjątków. Przyjęcie poziomu tolerancji 5% powoduje, że dla rocznego ciągu dziennych obserwacji wartość oczekiwana liczby wyjątków wynosi 12,5, natomiast przy poziomie tolerancji 1% tylko 2,5. W przypadku tak małego rozmiaru próby oraz braku analitycznej postaci rozkładu statystyki testowej dla prób skończonych możliwe jest wykorzystanie techniki testu Monte Carlo [Dwass 1957]. Technika ta oparta jest na założeniu, że rozkład statystyki testowej, przy założeniu prawdziwości hipotezy zerowej, dla skończonej próby, może zostać wyznaczony symulacyjnie. Umożliwia to otrzymanie tzw. testu dokładnego czyli testu, którego rzeczywisty rozmiar jest dokładnie równy założonemu nominalnemu poziomowi istotności.

² Zmienne inne niż zero-jedynkowe występują w testach ES oraz pewnych testach VaR , w których wykonuje się np. regresję zmiennej losowej oznaczającej przekroczenie VaR względem wielkości przekroczenia lub innych zmiennych z okresów poprzednich.

3.1.2. Zastosowanie metody Monte Carlo do uzyskania testu dokładnego

Założmy, że rozważana jest hipoteza zerowa

$$H_0 : \theta_0 \in \Theta_0$$

a obszar krytyczny jest postaci $\{S \geq c\}$. W ogólności rozkład statystyki testowej S , przy założeniu prawdziwości hipotezy zerowej, zależy od nieznanego parametru θ_0 , jednak przyjęcie założenia, że tak nie jest obejmuje szeroką klasę testów i pozwala na wyznaczenie testu dokładnego. W przypadku, gdy w całej podprzestrzeni Θ_0 rozkład S wyznaczony jest jednoznacznie, można zapisać

$$P_\theta(S \leq x) = F_S(x) \quad \forall \theta \in \Theta_0,$$

gdzie F_S jest dystrybuantą statystyki testowej S przy założeniu prawdziwości hipotezy zerowej. Oznaczmy $P \equiv P_\theta, \theta \in \Theta_0$. Stała c dobierana jest w ten sposób by

$$P(S \geq c) = 1 - F(c) + P(S = c) \leq \alpha,$$

gdzie α jest ustalonym poziomem istotności testu, $0 < \alpha < 1$.

Rozważmy sytuację, że rozkład S przy założeniu prawdziwości hipotezy zerowej jest niemożliwy do wyprowadzenia analitycznie, natomiast może zostać wyznaczony symulacyjnie. Niech S_1, S_2, \dots, S_N będzie próbą niezależnych zmiennych losowych³ o jednakowym rozkładzie, zgodnym z rozkładem statystyki S .

Technika testu Monte Carlo polega na zastąpieniu dystrybuanty teoretycznej F_S dystrybuantą empiryczną bazującą na próbie S_1, S_2, \dots, S_N :

$$\hat{F}_{S,N}(x) \equiv \hat{F}_N(x, \mathbf{S}) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(S_i \leq x)}, \quad (3.1)$$

gdzie $\mathbf{S} = (S_1, S_2, \dots, S_N)$. Funkcja przeżycia odpowiadająca dystrybuancie empirycznej dana jest wzorem:

$$\hat{G}_{S,N}(x) \equiv \hat{G}_N(x, \mathbf{S}) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(S_i \geq x)}. \quad (3.2)$$

³ W ogólności założenie o niezależności zmiennych losowych może być osłabione i zastąpione założeniem, że możliwa jest zamiana kolejności zmiennych tzn. wektor losowy $(S_1, S_2, \dots, S_N) \sim (S_{k_1}, S_{k_2}, \dots, S_{k_N})$ dla każdej permutacji (k_1, k_2, \dots, k_N) ciągu $(1, 2, \dots, N)$ [Dufour 2006].

Dla ciągłego rozkładu F powyższa empiryczna funkcja przeżycia umożliwia wyznaczenie p -wartości dla statystyki testowej S_0 ze wzoru

$$\hat{p}_N(S_0) = \frac{N\hat{G}_N(S_0) + 1}{N + 1}$$

[Dufour 2006].

W przypadku rozkładów ciągłych spełnione jest założenie $P(S_i = S_j) = 0$, $i \neq j$, $i, j = 0, 1, \dots, N$, które jest podstawą dowodzenia, że powyższa metoda testu Monte Carlo umożliwia otrzymanie testu na ustalonym poziomie istotności. Dla rozkładów dyskretnych, dla których założenie to nie jest spełnione, konieczne jest wykorzystanie techniki testu zrandomizowanego, która gwarantuje, że nawet w przypadku, gdy istnieje niezerowe prawdopodobieństwo wystąpienia dwóch jednakowych wartości w próbie $\mathbf{S} = S_1, S_2, \dots, S_N$, prawdopodobieństwo że w ciągu zrandomizowanych statystyk rangowych z tej próby wszystkie wartości będą różne wynosi 1. Dzięki temu test zrandomizowany dla rozkładu dyskretnego statystyki testowej zapewnia zachowanie założonego poziomu istotności [Dufour 2006].

Załóżmy, że z każdą statystyką S_j , $j = 1, 2, \dots, N$ związana jest zmienna losowa U_j , $j = 1, 2, \dots, N$, taka że U_1, U_2, \dots, U_N są niezależnymi zmiennymi losowymi pochodzącymi z rozkładu równomiernego $U(0, 1)$ na przedziale $(0, 1)$, ciąg $\mathbf{U} = (U_1, U_2, \dots, U_N)$ jest niezależny od \mathbf{S} . W ciągu par postaci

$$W_j = (S_j, U_j), \quad j = 1, 2, \dots, N,$$

można wprowadzić następującą relację porządku leksykograficznego

$$(S_i, U_i) \leq (S_j, U_j) \Leftrightarrow (S_i < S_j \vee (S_i = S_j \text{ i } U_i \leq U_j)).$$

Korzystając z powyższej relacji porządku definiuje się zrandomizowaną pseudo-dystrybuantę empiryczną

$$\tilde{F}_{S,N}(x) \equiv \tilde{F}_N(x, U_0, \mathbf{S}, \mathbf{U}) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{((S_i, U_i) \leq (x, U_0))}, \quad (3.3)$$

gdzie U_0 jest zmienną losową pochodzącą z rozkładu $U(0, 1)$, niezależną od \mathbf{S} i \mathbf{U} . Funkcja \tilde{F}_N zachowuje własności dystrybuanty rozkładu prawdopodobieństwa za wyjątkiem faktu, że w punktach skokowych może nie być prawostronnie ciągła.

Za pomocą funkcji przeżycia odpowiadającej dystrybuancie $\tilde{F}_{S,N}$, t.j.

$$\tilde{G}_{S,N}(x) \equiv \tilde{G}_N(x, U_0, \mathbf{S}, \mathbf{U}) = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{((S_i, U_i) \geq (x, U_0))}, \quad (3.4)$$

definiuje się p -wartość dla testu Monte Carlo [Dufour 2006]

$$\hat{p}_N(S_0) = \frac{N\tilde{G}_N(S_0) + 1}{N + 1}.$$

Powyższa procedura gwarantuje otrzymanie testu dokładnego dla prób skończonych w przypadku nieciągłego rozkładu statystyki testowej. W analizie ryzyka sytuacja taka zachodzi, gdy test oparty jest na zero-jedynkowym procesie wyjątków.

3.2. Projektowanie eksperymentów do oceny mocy

3.2.1. Eksperymenty wykorzystujące proces *GARCH*

Badanie mocy testów autokorelacji wymaga projektowania eksperymentów wykorzystujących generowanie zmiennych skorelowanych, przy czym możliwe jest zastosowanie różnych schematów autokorelacji. Przy weryfikacji modeli ryzyka ocenie podlega zero-jedynkowy proces przekroczeń *VaR* lub, w przypadku modeli *ES*, proces wielkości wyjątków. W większości testów *VaR* i *ES* podprzestrzeń wyznaczona przez hipotezę alternatywną zawiera wszelkie procesy stochastyczne charakteryzujące się korelacją przekroczeń w czasie. Projekt odpowiedniego eksperymentu symulacyjnego może być skonstruowany na dwa sposoby: może polegać na bezpośrednim generowaniu przebiegu procesu przekroczeń *VaR* lub odwoływać się do bazowego procesu stopy zwrotu. W pierwszym przypadku generator zero-jedynkowego procesu przekroczeń powinien uwzględnić odpowiednią strukturę autokorelacji. W drugim przypadku wykorzystuje się fakt, że bezpośrednią przyczyną skorelowania przekroczeń jest zjawisko gromadzenia wariancji. Zatem możliwe jest wykorzystanie algorytmów generowania zmiennych zależnych, które odzwierciedlają naprzemienne występowanie okresów niskiej i wysokiej wariancji procesu. Traktując kwadraty stopy zwrotu jako estymatory wariancji, można stosować generatory, które zakładają ustaloną strukturę korelacji kwadratu zmiennej losowej. Wówczas proces wyjątków otrzymuje się jako

dodatnią różnicę między wartościami stopy zwrotu a przyjętym poziomem VaR .

W literaturze przedmiotu stosuje się różne metody projektowania eksperymentów służących ocenie mocy testów autokorelacji przekroczeń VaR . Proponowane są eksperymenty, w których ustala się, że dane dotyczące stopy zwrotu pochodzą z rozkładu normalnego o stałej wariancji bądź z rozkładu t-Studenta o stałej liczbie stopni swobody. Następnie proces wyjątków zgodny z hipotezą alternatywną otrzymywany jest poprzez przyjęcie do estymacji wartości zagrożonej modeli heteroskedastycznych [Lopez 1999]. Niedopasowanie modelu stopy zwrotu i modelu wartości zagrożonej gwarantuje, że wypadkowy proces wyjątków cechuje się autokorelacją zmiennych.

Ze względu na stosowanie modeli ryzyka do danych finansowych, przyjmuje się często, że proces generujący realizacje stopy zwrotu powinien spełniać pewne ogólne postulaty, wynikające z badań dotyczących przebiegu procesów finansowych. Bez względu na częstotliwość obserwowania wartości zmiennej, zakłada się, że model powinien uwzględniać dużą zmienność wariancji procesu z okresu na okres. Ponadto zakłada się, że wahania zmienności cechują się tzw. „persystencją”, przejawiającą się w trwających wiele okresów skupieniach wysokiego lub niskiego poziomu zmienności procesu. Powyższe postulaty znajdują odzwierciedlenie w procesach klasy $GARCH$, przez co modele te są często stosowane do generowania wartości stopy zwrotu [np. Fiszeder 2009].

Podstawowy model $GARCH(q,r)$ [Bollerslev 1986], stanowiący uogólnienie modelu $ARCH(q,r)$ [Engle 1982], dany jest wzorami:

$$R_t = \sigma_t Z_t, \quad Z_t \sim N(0, 1),$$

$$\sigma_t^2 = \omega_1 + \sum_{i=1}^r \alpha_i R_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2,$$

gdzie $\omega_1, \alpha_i, \beta_j$ są parametrami procesu, $\omega_1 > 0, \alpha_i \geq 0, \beta_j \geq 0, i = 1, 2, \dots, q, j = 1, 2, \dots, r$. Dla $q = 1, r = 1$ otrzymuje się proces $GARCH(1,1)$, w którym wariancja warunkowa wyraża się wzorem

$$\sigma_t^2 = \omega_1 + \alpha_1 R_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (3.5)$$

W modelu $GARCH(1,1)$, za pomocą sterowania wartościami parametrów, można osiągnąć określoną wartość współczynnika autokorelacji ρ_h rzędu h kwadratów stóp zwrotu, korzystając ze wzorów:

$$\rho_1 = \alpha_1 + \frac{\alpha_1^2 \beta_1}{1 - 2\alpha_1 \beta_1 - \beta_1^2}, \quad (3.6)$$

$$\rho_h = \rho_1(\alpha_1 + \beta_1)^{h-1}, \quad h = 1, 2, \dots \quad (3.7)$$

[por. np. Franses, van Dijk 2008]. W celu otrzymania procesu przekroczeń o ustalonej strukturze autokorelacji, przyjmuje się następnie poziom wartości zagrożonej odpowiadający modelowi homoskedastycznemu lub modelowi uwzględniającemu zależność w czasie wariancji, w którym ustala się błędny schemat autokorelacji. Można to osiągnąć poprzez wygenerowanie oszacowań *VaR* np. z modelu średniej ruchomej lub modelu *GARCH* o nieprawidłowych parametrach [Lopez 1999, Campbell 2005].

Inną możliwością jest generowanie procesu stopy zwrotu z modelu *GARCH* o ustalonej wariancji a następnie przyjęcie poziomu *VaR* odpowiadającego systematycznie zaniżonej zmienności, stanowiącej pewien ustalony procent zmienności wyjściowej [Małecka 2013]. Eksperyment taki skutkuje autokorelacją przekroczeń, której jednak nie można określić liczbowo, za pomocą wartości współczynnika autokorelacji.

Ze względu na dwa kryteria oceny modelu wartości zagrożonej – udział przekroczeń oraz ich zależność w czasie (rozd. 1.2.2.) – sugeruje się by projekt eksperymentu symulacyjnego traktował obydwa kryteria rozłącznie. W ten sposób ocena mocy testu autokorelacji jest uniezależniona od spełnienia kryterium związanego z liczbą przekroczeń. W literaturze proponuje się zatem eksperymenty pozwalające, za pomocą odpowiedniego doboru parametrów, kontrolować obydwa kryteria. Przyjmując model *GARCH* o ustalonych parametrach możliwe jest kontrolowanie autokorelacji kwadratów zmiennej losowej. Dzięki ustaleniu oszacowań *VaR* na poziomie kwantyla z próby kontroli podlega również liczba wyjątków [Piontek 2013]. Alternatywnym podejściem jest przyjęcie wartości zagrożonej na poziomie wyznaczonego analitycznie kwantyla bezwarunkowego rozkładu stopy zwrotu, co ma tę przewagę, że pozwala uzyskać spójny teoretyczny model procesu przekroczeń.

Zamiast ustalania *VaR* na poziomie kwantyla z próby lub rozkładu teoretycznego, możliwe jest przyjęcie eksperymentu, którego projekt zakłada, że procesem generującym wartości stopy zwrotu jest proces *GARCH*, natomiast oszacowania *VaR* uzyskiwane są metodą symulacji historycznej [Lopez 1999, Christoffersen i Pelletier 2004, Campbell 2005, Berkowitz, Christoffersen i Pelletier 2011]. Przyjęcie takiego modelu w eksperymencie symulacyjnym uniemożliwia pomiar i kontrolę stopnia autokorelacji. Ponadto założenie, że w wyniku zastosowania metody symulacji historycznej do szacowania wartości zagrożonej, otrzymany zostanie model zgodny z hipotezą alternatywną, implikuje negatywną ocenę tej metody.

3.2.2. Propozycja eksperymentu opartego na procesie BGAR

W literaturze poświęconej zarządzaniu ryzykiem przyjmuje się, że główna trudność związana z kontrolą ryzyka rynkowego wynika ze zjawiska gromadzenia zmienności. Zjawisko to może być przyczyną występowania sekwencji przekroczeń VaR , niebezpiecznych z punktu widzenia płynności finansowej przedsiębiorstw. W związku z tym do oceny mocy testów VaR proponuje się eksperymenty konstruowane w ten sposób, by odzwierciedlały zjawisko gromadzenia zmienności. W tym celu konieczne jest zastosowanie algorytmów pozwalających generować dane o zadanej strukturze autokorelacji, pochodzących z określonego rozkładu. Wykorzystanie generatorów zmiennych o rozkładzie brzegowym χ^2 umożliwi uzyskanie szeregu kwadratów stopy zwrotu. Odzwierciedlenie gromadzenia zmienności można otrzymać poprzez projekt eksperymentu, gwarantujący pozytywną korelację kwadratów stopy zwrotu w czasie.

Reprezentacja zjawiska gromadzenia zmienności możliwa jest poprzez wykorzystanie procesu autoregresyjnego $BGAR(1)$ [Lewis, McKenzie i Hugs 1989]. Proces ten charakteryzuje się brzegowym rozkładem gamma i określoną strukturą autokorelacji, odpowiadającą procesowi $AR(1)$. Generowanie wartości zmiennych z procesu $BGAR(1)$ oparte jest na transformacji beta-gamma i wykorzystuje własności rozkładów gamma i beta.

Niech $B(m, n)$ oznacza zmienną losową pochodzącą z rozkładu beta $\beta(m, n)$ z parametrami $m > 0$ i $n > 0$. Funkcja gęstości tej zmiennej ma postać

$$f_{B(m,n)}(x) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} x^{m-1} (1-x)^{n-1}, \quad 0 \leq x \leq 1, \quad m > 0, \quad n > 0,$$

gdzie Γ oznacza funkcję gamma. Niech $G(k, \theta)$ oznacza zmienną losową pochodzącą z rozkładu gamma $\Gamma(k, \theta)$ z parametrem kształtu $k > 0$ i parametrem skali $\theta > 0$. Funkcja gęstości tej zmiennej ma postać

$$f_{G(k,\theta)}(x) = \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x}, \quad x \geq 0, \quad k > 0, \quad \theta > 0.$$

Transformacja beta gamma wykorzystuje fakt, że suma dwóch niezależnych zmiennych losowych pochodzących z rozkładów gamma o takich samych parametrach skali $G'(k', \theta)$ i $G''(k'', \theta)$ daje zmienną losową o rozkładzie gamma:

$$G(k' + k'', \theta) = G'(k', \theta) + G''(k'', \theta).$$

Zmienną losową o rozkładzie beta można zapisać za pomocą niezależnych zmiennych losowych pochodzących z rozkładu gamma w następujący sposób:

$$B(m, n) = \frac{G'(m, \theta)}{G'(m, \theta) + G'(n, \theta)}.$$

Wiadomo ponadto, że zmienna $B(m, n)$ jest niezależna od sumy występującej w mianowniku $G(m + n, \theta) = G'(m, \theta) + G'(n, \theta)$. z powyższego przedstawienia zmiennej o rozkładzie beta wynika od razu, że zmienną pochodzącą z rozkładu gamma można uzyskać jako iloczyn

$$G(m, \theta) = B(m, n)G(m + n, \theta).$$

Powyższy zapis pokazuje, że możliwe jest zarówno zwiększanie jak i redukcja parametru kształtu w rozkładzie gamma za pomocą odpowiednio: dodawania niezależnej zmiennej losowej o rozkładzie gamma lub mnożenia przez niezależną zmienną losową o rozkładzie beta.

Konstrukcja procesu $BGAR(1)$ o rozkładzie brzegowym gamma i strukturze korelacji opisywanej parametrem r opiera się na generowaniu szeregu wartości niezależnych zmiennych losowych z rozkładów gamma i beta oraz wykorzystuje przekształcenie

$$\begin{aligned} Y_n(k, \theta) &= B_n(kr, k\bar{r})Y_{n-1}(k, \theta) + B'_n(k\bar{r}, kr)G_n(k, \theta) = \\ &= B_n(kr, k\bar{r})Y_{n-1}(k, \theta) + G'_n(k\bar{r}, \theta), \end{aligned} \quad (3.8)$$

gdzie $0 \leq r \leq 1$, $\bar{r} = 1 - r$, $n = 1, 2, \dots$, $G'_n(k\bar{r}, \theta)$ są niezależnymi zmiennymi z rozkładu $\Gamma(k\bar{r}, \theta)$, $B_n(kr, k\bar{r})$ są niezależnymi zmiennymi losowymi z rozkładu $\beta(kr, k\bar{r})$ oraz ciągi te są niezależne między sobą. Jeżeli zmienna $Y_{n-1}(k, \theta, r)$ będzie pochodziła z rozkładu $\Gamma(k, \theta)$, to przemnożenie jej przez $B_n(kr, k\bar{r})$ daje zmienną o rozkładzie $\Gamma(kr, \theta)$, a po dodaniu $G'_n(k\bar{r}, \theta)$ otrzymuje się zmienną o rozkładzie $\Gamma(k, \theta)$. z wykorzystaniem momentów rozkładu beta otrzymuje się, że współczynniki autokorelacji ρ_h rzędu h , $h = 1, 2, \dots$ W powyższym procesie wynoszą:

$$\rho_h = \text{Corr}(Y_t, Y_{t-h}) = r^{|h|}, \quad h = 0, \pm 1, \pm 2, \dots$$

Powyższa procedura zatem pozwala otrzymać stacjonarny proces Markowa $BGAR(1)$, składający się ze zmiennych o rozkładzie brzegowym gamma, charakteryzujących się wygasającymi wykładniczo współczynnikami autokorelacji.

3.2.3. Propozycja eksperymentu opartego na procesie BGMA

Reprezentacja zjawiska gromadzenia zmienności możliwa jest do uzyskania, podobnie jak w przypadku opisanego powyżej procesu $BGAR(1)$, za pomocą procesu $BGMA(1)$ o brzegowym rozkładzie gamma i strukturze korelacji odpowiadającej schematowi $MA(1)$. Jeżeli $G_n(k, \theta)$ jest ciągiem wartości niezależnych, jednakowo rozłożonych zmiennych losowych z rozkładu gamma $\Gamma(k, \theta)$ a $B_n(k\bar{s}, k\bar{s})$, $\bar{s} = 1 - s$, jest ciągiem niezależnych, jednakowo rozłożonych zmiennych losowych z rozkładu beta $\beta(k\bar{s}, k\bar{s})$, proces $BGMA(1)$ określa się następująco

$$Y_n = G_n(k, \theta) + B_n(k\bar{s}, k\bar{s})G_{n-1}(k, \theta), \quad 0 \leq s \leq 1. \quad (3.9)$$

Otrzymany proces jest procesem stacjonarnym. Ponieważ prawa strona równania (3.9) jest sumą zmiennych losowych o rozkładach gamma o parametrach skali θ i parametrach kształtu k oraz $k\bar{s}$, to brzegowy rozkład procesu jest rozkładem $\Gamma(k(1+s), \theta)$. z punktu widzenia projektowania eksperymentów symulacyjnych wygodnie jest określić parametry procesu bazowego w terminach pożądaných parametrów procesu docelowego, można zatem przyjąć następującą parametryzację procesów wyjściowych: $G_n(\frac{k}{1+s}, \theta)$ i $B_n(\frac{k\bar{s}}{1+s}, \frac{k\bar{s}}{1+s})$. Powyższa reparametryzacja pozwala osiągnąć proces $BGMA(1)$ o rozkładzie brzegowym $\Gamma(k, \theta)$. Przyjmując notację $B_n = B_n(\frac{k\bar{s}}{1+s}, \frac{k\bar{s}}{1+s})$, $G_n = G_n(\frac{k}{1+s}, \theta)$, $B = B(\frac{k\bar{s}}{1+s}, \frac{k\bar{s}}{1+s})$, $G = G(\frac{k}{1+s}, \theta)$, mamy $cov(Y_n, Y_{n-1}) = cov(B_n G_{n-1}, G_{n-1}) = E(B)V(G)$. Zatem otrzymana struktura autokorelacji ma postać

$$\rho_h = \begin{cases} \frac{s}{1+s}, & |h| = 1, \\ 0, & |h| > 1. \end{cases}$$

Konstrukcja procesu $BGMA(1)$ pozwala otrzymać pełen zakres współczynników korelacji możliwy dla procesu $MA(1)$, czyli $0 \leq \rho_1 \leq 0,5$.

3.2.4. Propozycja eksperymentu opartego na procesie Markowa

W odróżnieniu od przedstawionych wyżej sposobów otrzymania ciągu przekroczeń poprzez generowanie wartości stopy zwrotu, których kwadraty są pozytywnie skorelowane w czasie, proces przekroczeń można uzyskać w sposób bezpośredni, odwołując się do własności Markowa. Wykorzystanie tej własności pozwala skonstruować model eksperymentu symulacyjnego, który

bezpośrednio generuje wartości procesu przekroczeń zamiast dwustopniowej procedury, w której wyjątki uzyskuje się jako dodatnie różnice między wartościami stopy zwrotu a przyjętym poziomem VaR .

W badaniu ryzyka rozumianego jako wartość zagrożona rozważa się jednorodny łańcuch Markowa o dwóch stanach, gdzie 1 oznacza wyjątek a 0 jego brak. Przyjmując założenie, że wartości zero-jedynkowej zmiennej losowej opisującej przekroczenia pochodzą z jednorodnego łańcucha Markowa o rozkładzie początkowym (π_0, π_1) oraz macierzy przejścia:

$$\begin{bmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{bmatrix},$$

zgodność z hipotezą alternatywną w teście autokorelacji wyjątków VaR otrzymuje się poprzez zagwarantowanie, że $\pi_{01} \neq \pi_{11}$. Dla ustalonych π_0 i π_{11} , wartości parametrów π_1 i π_{10} wynikają z własności prawdopodobieństwa, natomiast odpowiedni dobór π_{01} , a co za tym idzie również π_{00} , pozwala zagwarantować stacjonarność procesu. Dobór prawdopodobieństw w rozkładzie bezwarunkowym wyznaczonym przez rozkład początkowy (π_0, π_1) umożliwia ustalenie liczby przekroczeń na założonym poziomie.

3.3. Wyniki symulacyjnej oceny rozmiaru i mocy testów VaR

3.3.1. Ocena rozmiaru testów VaR

Badanie rozmiaru testów VaR opiera się na eksperymentach wykorzystujących proces Bernoulliego. Za ich pomocą generuje się trajektorie procesu przekroczeń zgodne z hipotezą zerową. Eksperymenty te pozwalają na otrzymanie danych zgodnych z kryteriami oceny modelu VaR dotyczącymi zarówno rozkładu bezwarunkowego jak i warunkowego. Dzięki niezależnemu przeprowadzaniu prób Bernoulliego otrzymuje się realizacje zmiennych niezależnych natomiast sterowanie parametrem procesu gwarantuje odpowiedni udział przekroczeń VaR . W niniejszym badaniu ocenę rozmiaru testów przeprowadzono w dwóch krokach. W pierwszym z nich symulacyjnie wyznaczono rozmiar, korzystając z teoretycznych rozkładów statystyk testowych, natomiast w drugim kroku dokonano graficznego porównania kwantyli empirycznych i teoretycznych.

W celu symulacyjnego wyznaczenia rozmiaru rozważanych testów VaR wygenerowano T wyników niezależnych prób Bernoulliego z prawdopodobień-

stwem sukcesu p , równym poziomowi tolerancji założonemu przy szacowaniu *VaR*. W ten sposób otrzymano ciąg realizacji procesu $\{I_t\}_{t=1}^T$, cechującego się wartością oczekiwaną p oraz brakiem autokorelacji. Otrzymany ciąg był zatem zgodny z hipotezą zerową testów dotyczących rozkładu bezwarunkowego, mówiącą o prawdopodobieństwie przekroczenia *VaR* na poziomie p oraz z hipotezami zerowymi pozostałych testów, mówiącymi o braku autokorelacji przekroczeń. Przyjęto $p = 0,05$. W regresji liniowej i regresji logistycznej w testach DQ i DQ_{logist} jako zmienne objaśniające przyjęto wartości procesu przekroczeń z pięciu okresów poprzedzających okres t , $t = 6, \dots, T$. W teście LB uwzględniono autokorelację zmiennej I_t , $t = 6, \dots, T$, do rzędu piątego włącznie.

W przypadku testów opartych na porównywaniu prognozowanej i empirycznej funkcji gęstości lub spektralnej gęstości stopy zwrotu, wartości R_t , $t = 1, \dots, T$ wygenerowano z procesu $GARCH(1,1)$ postaci

$$\begin{aligned} R_t &= \sigma_t Z_t, & Z_t &\sim N(0, 1), \\ \sigma_t^2 &= \omega_1 + \alpha_1 R_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned} \quad (3.10)$$

z parametrami⁴ $\omega_1 = 0,000001$, $\alpha = 0,14$, $\beta_1 = 0,85$. W celu uzyskania wartości zmiennej U_t , $t = 1, \dots, T$ przeprowadzono transformację Rosenblatta (rozdz. 2.3.2.) za pomocą dystrybuanty rozkładu normalnego z o wartości oczekiwanej zero i odchyleniu standardowym wynikającym z procesu (3.3.1.). Transformacji w zmienne Z_t , $t = 1, \dots, T$ dokonano z wykorzystaniem wartości parametrów przyjętych w procesie generującym dane.

W wielowymiarowych testach LB_H i LR_H badano przekroczenia *VaR* jednocześnie dla dwóch poziomów tolerancji $p_1 = 0,01$ i $p_2 = 0,05$. Procesy przekroczeń dla obydwóch poziomów tolerancji otrzymano w dwustopniowej procedurze. W pierwszym kroku wygenerowano wartości stopy zwrotu R_t , $t = 1, \dots, T$ z procesu $GARCH(1,1)$ danego wzorem (3.3.1.), natomiast w drugim kroku, z wykorzystaniem odpowiednich parametrów procesu generującego dane, wyznaczono oszacowania *VaR* jako kwantyle rzędów 0,01 i 0,05 zmiennych R_t , $t = 1, \dots, T$.

Badanie symulacyjne przeprowadzono dla poziomu istotności $\alpha = 5\%$. Wszystkie eksperymenty powtórzono dla długości szeregów czasowych $T = 250, 500, 750, 1000$, wykonując każdorazowo 10000 replikacji⁵. Wyjątkiem

⁴ Parametry procesu $GARCH(1,1)$ przyjęto na podstawie badania wstępnego dokonanego dla sześciu indeksów giełdowych [Małecka 2011].

⁵ Taka liczba replikacji zagwarantowała otrzymanie dokładnych oszacowań, zwłaszcza dla dłuższych szeregów, gdzie różnice przy powtarzaniu eksperymentów widoczne były dopiero na trzecim miejscu po przecinku.

w odniesieniu do długości szeregów był test Haasa, gdzie rozmiar próby rozumiany jest nie jako liczba obserwacji wartości zmiennych $R_t, t = 1, 2, \dots, T$ lub równoważnie $I_t, t = 1, 2, \dots, T$, lecz jako liczba wyjątków N . Ponieważ w teście Haasa asymptotyczny rozkład statystyki testowej zależy od tak rozumianej wielkości próby, odpowiednie eksperymenty skonstruowano w ten sposób by uzyskać odpowiednią liczbę N równą wartości oczekiwanej liczby przekroczeń $E(\sum_{t=1}^T I_t) = Tp$.

Graficzna ocena rozbieżności między empirycznymi a teoretycznymi rozkładami statystyk testowych polegała na porównaniu kształtów empirycznych i asymptotycznych dystrybuant oraz funkcji gęstości oraz opierała się na symulacyjnym wyznaczeniu wartości tych statystyk i dystrybuant rozkładów dla ustalonych, skończonych rozmiarów próby. Korzystając z modelu zgodnego z hipotezą zerową wygenerowano $M = 10000$ realizacji statystyk testowych S_1, S_2, \dots, S_M . Następnie wyznaczono dystrybuantę empiryczną ze wzoru

$$\hat{F}_M(x) = \frac{1}{M} \sum_{i=1}^M \mathbf{1}_{S_i \leq x}.$$

W pierwszym etapie oceny rozmiaru testów *VaR* symulacyjnie wyznaczono częstość odrzuceń, przy założeniu hipotezy zerowej, dla grupy testów Z_{uc} , LR_{uc} , W_{uc} , LM_{uc} i LR_{TUFF} . Grupa ta objęła procedury sprawdzające udział przekroczeń w bezwarunkowym rozkładzie wyjątków *VaR*⁶. Badanie przeprowadzono dla testów, w przypadku których korzysta się z asymptotycznych rozkładów statystyk testowych. Otrzymane oszacowania rozmiaru testów *VaR* różniły się nieznacznie od nominalnego poziomu 5%, ponadto widoczna była zbieżność do oczekiwanego poziomu wraz z wydłużaniem szeregu (tab. 3.1). W żadnym przypadku wyznaczony symulacyjnie rozmiar nie przekraczał 8%.

Spośród wszystkich asymptotycznych testów udziału przekroczeń, oszacowania rozmiaru najbliższe nominalnemu poziomowi 5% otrzymano dla testu Z_{uc} , w którym wykorzystywana jest asymptotyczna zbieżność rozkładu dwumianowego do rozkładu normalnego. Spośród testów opartych na statystyce χ^2 największą dokładnością, w odniesieniu do rozmiaru, charakteryzował się, zaproponowany w niniejszej pracy, test LM_{uc} , zbudowany w oparciu o regułę mnożników Lagrange'a. Zgodność wyznaczonych empirycznie oraz asymptotycznych rozkładów statystyk testowych, należących do rozważanej grupy

⁶ Analiza rozmiaru testów rozkładu bezwarunkowego nie dotyczy testu dwumianowego o statystyce T_1 , w przypadku której korzysta się z rozkładu dokładnego.

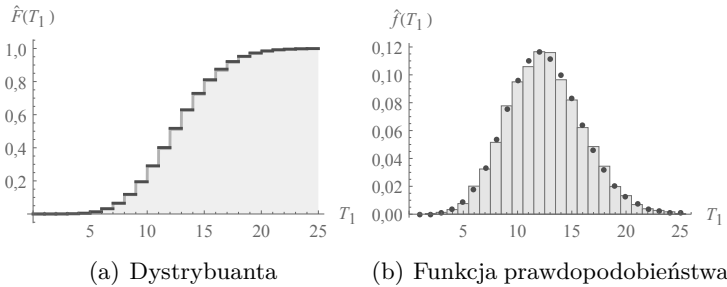
testów, potwierdzono następnie na podstawie analizy graficznej, opartej na porównaniu empirycznych i asymptotycznych kształtów funkcji gęstości i dystrybuant⁷ (rys. 3.2–3.6).

Tab. 3.1: Empiryczny rozmiar testów opartych na procesie Bernoulliego

Test	Długość szeregu			
	250	500	750	1000
Z_{uc}	0,043	0,048	0,054	0,051
LR_{uc}	0,063	0,057	0,056	0,053
W_{uc}	0,074	0,068	0,063	0,057
LM_{uc}	0,039	0,049	0,044	0,047
LR_{TUFF}	0,066	0,059	0,061	0,060

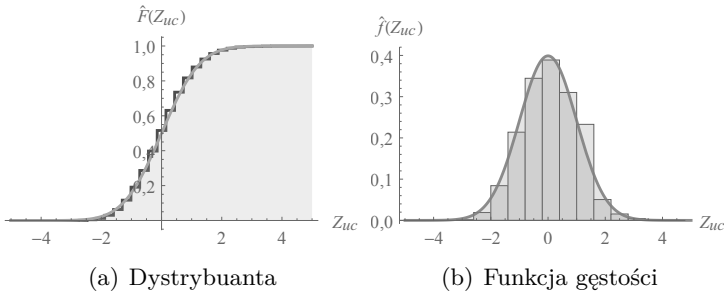
Uwaga 1: Wyniki przedstawione w tablicy nie uwzględniają testu dwumianowego T_1 , w przypadku którego korzysta się z dokładnego rozkładu statystyki testowej dla próby skończonej. Rozkład empiryczny statystyki T_1 na tle rozkładu dwumianowego przedstawia rys. 3.1.

Uwaga 2: Wszystkie wyniki badań prezentowane w tablicach oraz na wykresach zamieszczonych w pracy, o ile nie zaznaczono inaczej, stanowią opracowanie własne autorki.

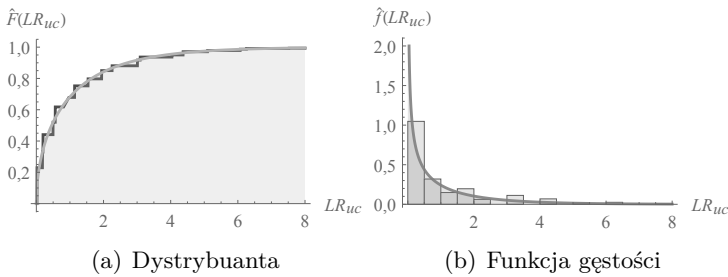


Rys. 3.1: Empiryczny i asymptotyczny rozkład statystyki testu T_1

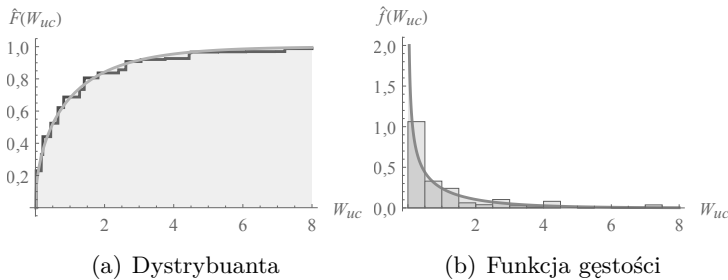
⁷ Na wszystkich rysunkach zamieszczonych w niniejszym rozdziale, zacienione pole na wykresie odpowiada rozkładowi asymptotycznemu. Prezentowane rysunki dotyczą długości próby $T = 250$. Wyjątkiem jest rysunek przedstawiający empiryczny rozkład statystyki WW na tle rozkładu liczby serii, który zaprezentowany został dla 1000 obserwacji, gdyż dla 250 obserwacji oczekiwana liczba przekroczeń to 12,5, co implikuje niewielką liczbę wartości przyjmowanych w rozkładzie serii.



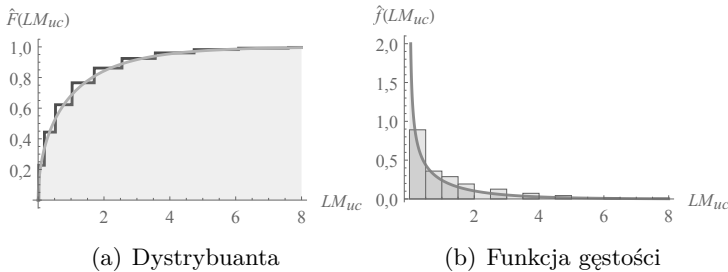
Rys. 3.2: Empiryczny i asymptotyczny rozkład statystyki testu Z_{uc}



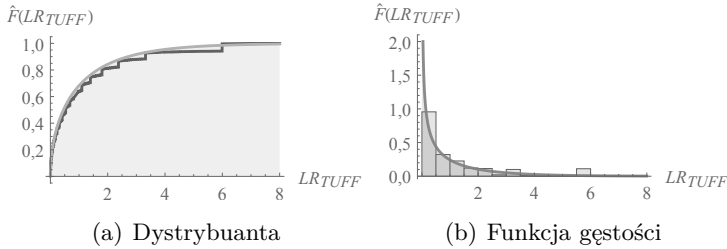
Rys. 3.3: Empiryczny i asymptotyczny rozkład statystyki testu LR_{uc}



Rys. 3.4: Empiryczny i asymptotyczny rozkład statystyki testu W_{uc}



Rys. 3.5: Empiryczny i asymptotyczny rozkład statystyki testu LM_{uc}

Rys. 3.6: Empiryczny i asymptotyczny rozkład statystyki testu LR_{TUFF}

W kolejnych etapach badania rozmiaru rozważano testy autokorelacji przekroczeń oparte na warunkowym rozkładzie wyjątków *VaR*. Analizę porównawczą przeprowadzono w podziale na grupy testów: bazujących na procesie Markowa, wykorzystujących proces odległości wyjątków *VaR*, zgodności, gęstości spektralnej oraz wielowymiarowych testów *VaR*.

Tab. 3.2: Empiryczny rozmiar testów opartych na procesie Markowa

Test	Długość szeregu			
	250	500	750	1000
LR_{ind}	0,071	0,083	0,122	0,137
DQ	0,062	0,061	0,055	0,051
DQ_{logist}	0,027	0,046	0,063	0,075
LB	0,062	0,054	0,047	0,050
$Q_{ind}^{x^2}$	0,044	0,039	0,035	0,039

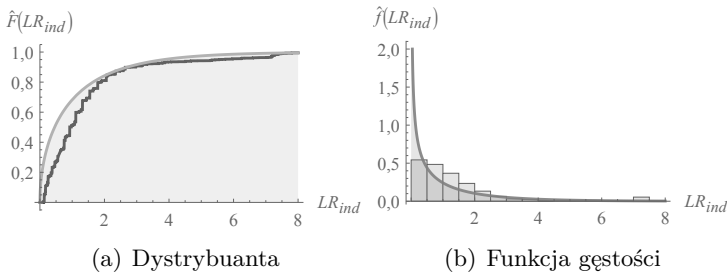
Uwaga: Wyniki przedstawione w tabelicy nie uwzględniają testu *WW* Walda Wolfowiza, w przypadku którego korzysta się z dokładnego rozkładu statystyki testowej dla próby skończonej. Rozkład empiryczny statystyki *WW* na tle rozkładu serii przedstawia rys. 3.12.

Wśród asymptotycznych testów LR_{ind} , DQ , DQ_{logist} , LB i $Q_{ind}^{x^2}$, których konstrukcja oparta została na procesie Markowa⁸, największą dokładność zaobserwowano dla klasycznego testu autokorelacji LB Ljunga-Boxa, opartego na współczynnikach autokorelacji (rys. 3.10) oraz dla testu DQ zbudowanego na zasadzie weryfikacji istotności regresji względem odpowiedniego kwantyla rozkładu stopy zwrotu (rys. 3.8, tab. 3.2). Empiryczny rozmiar tych testów

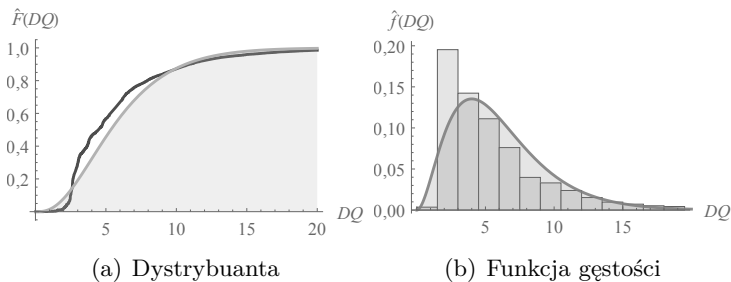
⁸ Analiza rozmiaru testów opartych na procesie Markowa nie dotyczy testu serii o statystyce *WW*, w przypadku której korzysta się z rozkładu dokładnego.

nieznacznie różnił się od poziomu nominalnego i zbiegał do niego z wydłużaniem szeregu czasowego.

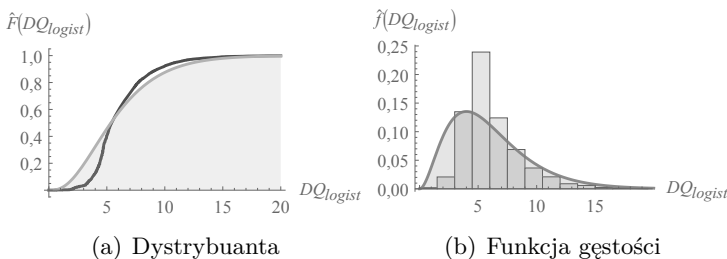
Drugi z testów regresyjnych – DQ_{logist} , wykorzystujący regresję logistyczną, charakteryzował się mniejszą dokładnością (rys. 3.9) oraz mniejszą stabilnością oszacowań rozmiaru w zależności od rozmiaru próby. Rozmiar testu $Q_{ind}^{\chi^2}$ Pearsona był systematycznie niedoszacowany (rys. 3.11). Z kolei rozmiar testu LR_{ind} , popularnie stosowanego do oceny modeli VaR , był systematycznie przeszacowany i najbardziej odległy od założonego poziomu nominalnego (rys. 3.7). Uzyskane symulacyjnie oszacowania rozmiaru tego testu przekraczały 10%.



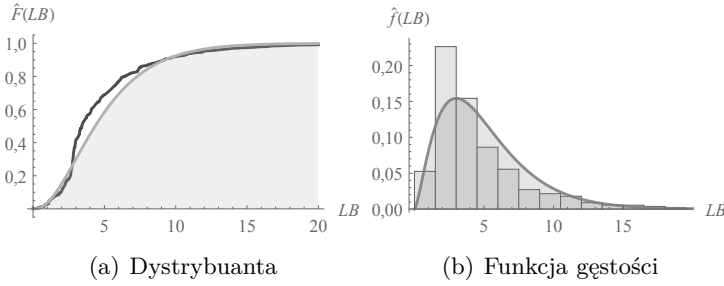
Rys. 3.7: Empiryczny i asymptotyczny rozkład statystyki testu LR_{ind}



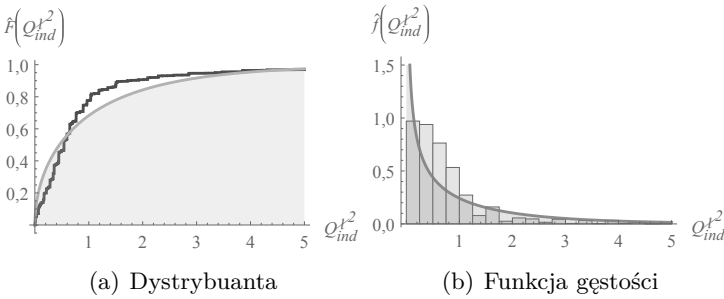
Rys. 3.8: Empiryczny i asymptotyczny rozkład statystyki testu DQ



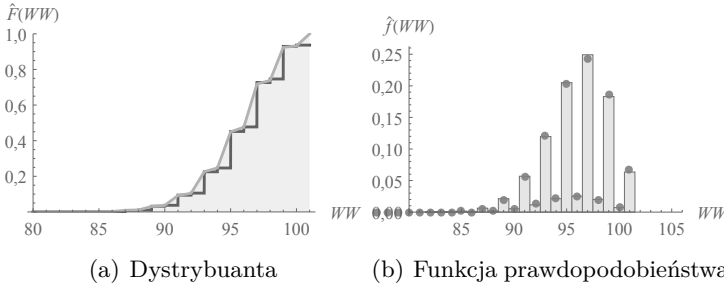
Rys. 3.9: Empiryczny i asymptotyczny rozkład statystyki testu DQ_{logist}



Rys. 3.10: Empiryczny i asymptotyczny rozkład statystyki testu *LB*



Rys. 3.11: Empiryczny i asymptotyczny rozkład statystyki testu Q_{ind}^2



Rys. 3.12: Empiryczny i asymptotyczny rozkład statystyki testu *WW*

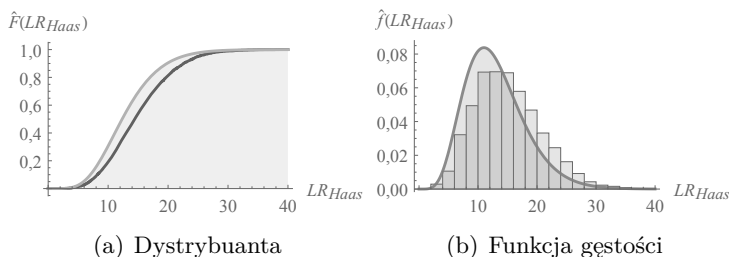
Dalsze wyniki badania symulacyjnego pokazały, że testy LR_{Haas} , LR_{Weib} , LR_{gamma} i $EACD$, należące do kolejnej grupy, zbudowane w oparciu o proces odległości wyjątków *VaR*, charakteryzowały się mniejszą dokładnością niż testy poprzedniej grupy (tab. 3.3). W ich przypadku częstość odrzucania hipotezy zerowej była z reguły zawyżona, świadcząc o większym od założonego błędzie I rodzaju. Oszacowania rozmiaru najbardziej zbliżone do nominalnego poziomu 5% otrzymano dla testów LR_{Weib} i LR_{gamma} , skonstruowanych

w oparciu o ideę sprawdzania własności braku pamięci z wykorzystaniem rozkładu wykładniczego (rys. 3.14, 3.15). W szczególności największą dokładność miał test LR_{Weib} , gdzie wyznaczony symulacyjnie rozmiar nie przekraczał 10% dla żadnej długości szeregu. Symulacyjne oszacowania rozmiaru dla różnych długości szeregu nie potwierdziły jednak zbieżności omawianych statystyk do rozkładu asymptotycznego wraz ze zwiększaniem próby.

Tab. 3.3: Empiryczny rozmiar testów opartych na procesie odległości wyjątków VaR

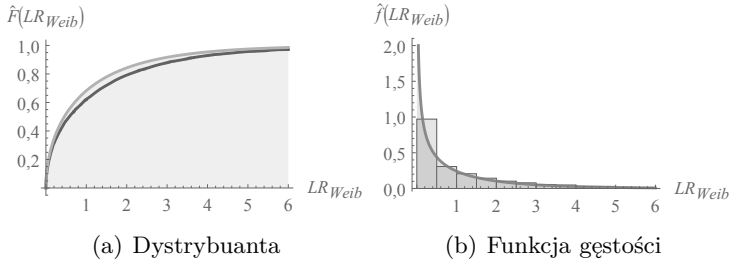
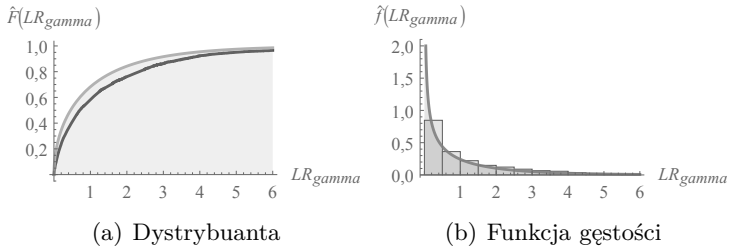
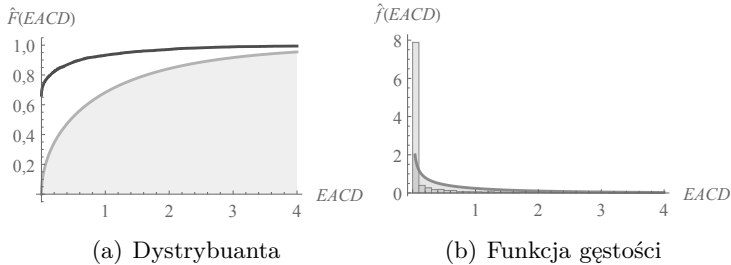
Test	Długość szeregu			
	250	500	750	1000
LR_{Haas}	0,127	0,133	0,165	0,189
LR_{Weib}	0,077	0,074	0,089	0,100
LR_{gamma}	0,086	0,088	0,111	0,116
$EACD$	0,007	0,009	0,011	0,012

Wyraźną rozbieżność kwantyli empirycznych z rozkładem asymptotycznym otrzymano dla testów $EACD$ i LR_{Haas} . Oszacowania rozmiaru dla regresyjnego testu $EACD$, wykorzystującego rozkład wykładniczy, wskazały na zdecydowanie konserwatywny charakter, skutkujący niewielką częstością odrzucania hipotezy zerowej (rys. 3.16). W przypadku tego testu ma miejsce nietypowa sytuacja, gdzie hipoteza zerowa (parametr regresji równy zero) leży dokładnie na granicy przestrzeni wszystkich dopuszczalnych parametrów⁹ ($[0, +\infty)$). Taka konstrukcja hipotez może skutkować brakiem zbieżności statystyki testowej do granicznego rozkładu χ_1^2 .



Rys. 3.13: Empiryczny i asymptotyczny rozkład statystyki testu LR_{Haas}

⁹ W omawianej sytuacji zamiast rozkładu χ_1^2 można rozważyć wykorzystanie mieszaniny rozkładów χ_1^2 oraz rozkładu jednopunktowego skoncentrowanego w zerze.

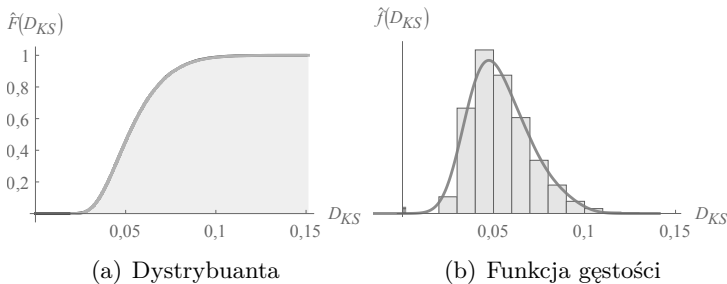
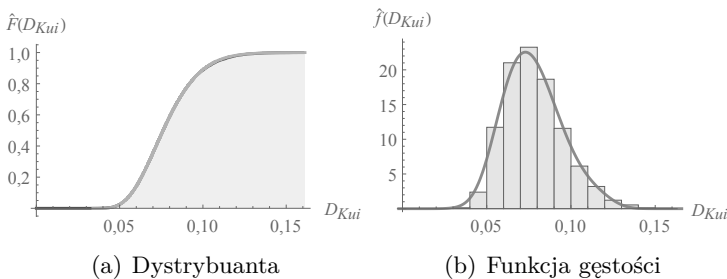
Rys. 3.14: Empiryczny i asymptotyczny rozkład statystyki testu LR_{Weib} Rys. 3.15: Empiryczny i asymptotyczny rozkład statystyki testu LR_{gamma} Rys. 3.16: Empiryczny i asymptotyczny rozkład statystyki testu $EACD$

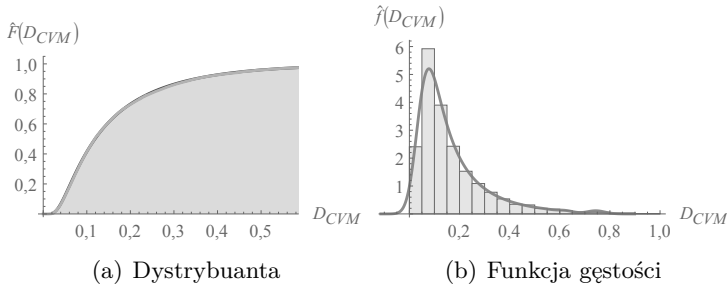
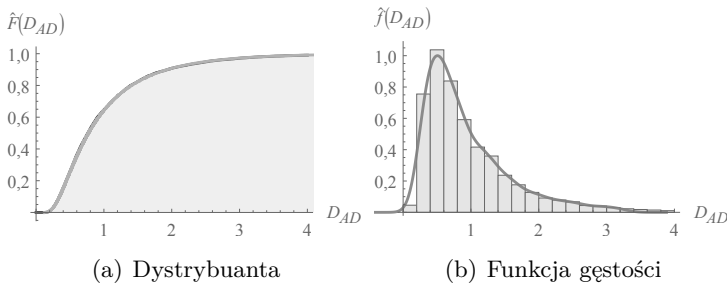
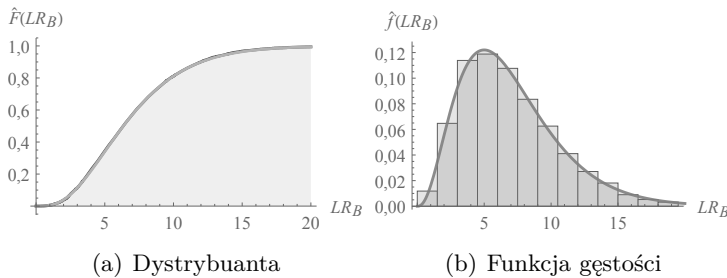
W przeciwieństwie do statystyki $EACD$, w przypadku testu LR_{Haas} zaobserwowano jego liberalny charakter, skutujący zbyt częstym odrzucaniem hipotezy zerowej przy korzystaniu z rozkładu granicznego (rys. 3.13). Konstrukcja tego testu polega na sumowaniu statystyk, z których każda ma rozkład χ_1^2 , w wyniku czego otrzymuje się rozkład χ_N^2 . Warunkiem zbieżności całej sumy jest zbieżność poszczególnych składników, co może nie zachodzić, gdyż szacowanie parametrów w modelu bez restrykcji dla każdego składnika sumy wykorzystuje jedynie czas oczekiwania na kolejne przekroczenie VaR . W praktyce często jest to tylko kilka-kilkanaście obserwacji, co może skutkować brakiem normalności oszacowań MNW. Efektem może być problem ze zbieżnością całej sumy do rozkładu granicznego.

Testy D_{KS} , D_{Kui} , D_{CVM} , D_{AD} i LR_B , zaliczane do następnej rozważanej grupy – testów zgodności – charakteryzowały się oszacowaniami rozmiaru bliskimi nominalnemu poziomowi 5% (tab. 3.4). W wszystkich przypadkach częstość odrzucania modelu zgodnego z hipotezą zerową mieściła się między 4,6% a 5,7%. Analiza graficzna rozkładów statystyk testowych pokazała pokrywanie się, w przybliżeniu, empirycznych i asymptotycznych funkcji gęstości i dystrybuant (rys. 3.17–3.21).

Tab. 3.4: Empiryczny rozmiar testów zgodności

Test	Długość szeregu			
	250	500	750	1000
D_{KS}	0,055	0,054	0,049	0,049
D_{Kui}	0,048	0,046	0,046	0,049
D_{CVM}	0,046	0,052	0,048	0,052
D_{AD}	0,052	0,051	0,047	0,047
LR_B	0,047	0,057	0,046	0,050

Rys. 3.17: Empiryczny i asymptotyczny rozkład statystyki testu D_{KS} Rys. 3.18: Empiryczny i asymptotyczny rozkład statystyki testu D_{Kui}

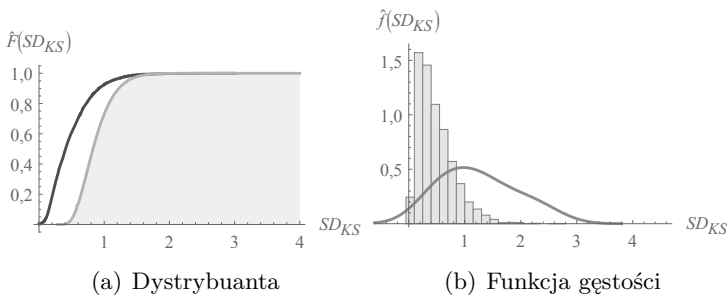
Rys. 3.19: Empiryczny i asymptotyczny rozkład statystyki testu D_{CVM} Rys. 3.20: Empiryczny i asymptotyczny rozkład statystyki testu D_{AD} Rys. 3.21: Empiryczny i asymptotyczny rozkład statystyki testu LR_B

W dalszym etapie badania symulacyjnego badano testy autokorelacji przekroczeń *VaR* – SD_{KS} , $SD_{K_{ui}}$, SD_{CVM} i SD_{AD} – wykorzystujące spektralną funkcję gęstości. W przypadku wszystkich testów tej grupy zaobserwowano przesunięcie w rozkładu asymptotycznego w stosunku do empirycznego w kierunku dodatnim, świadczące o konserwatywnym charakterze testów (rys. 3.22–3.25). Przesunięcie rozkładów spowodowało, że otrzymane oszacowania rozmiaru testów były systematycznie zaniżone – najczęściej nie przekraczały lub nieznacznie przekraczały 4% (tab. 3.5). Zaobserwowana niedokładność i konserwatywny charakter testów sugerują, że zastosowanie rozkładów

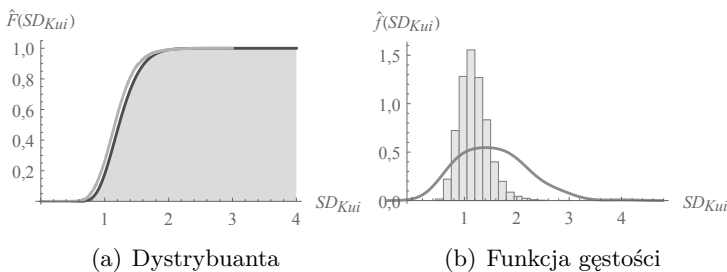
asymptotycznych w tej grupie testów prowadzi do zbyt rzadkiego odrzucania hipotezy zerowej. Może to skutkować błędną akceptacją modelu w przypadku faktycznego wystąpienia autokorelacji przekroczeń VaR . Spośród całej grupy testów gęstości spektralnej największą zgodnością rozkładów charakteryzował się test SD_{AD} , wykorzystujący statystykę Andersona-Darlinga.

Tab. 3.5: Empiryczny rozmiar testów gęstości spektralnej

Test	Długość szeregu			
	250	500	750	1000
SD_{KS}	0,028	0,029	0,033	0,032
SD_{Kui}	0,028	0,031	0,033	0,035
SD_{CVM}	0,033	0,037	0,039	0,041
SD_{AD}	0,035	0,040	0,043	0,045

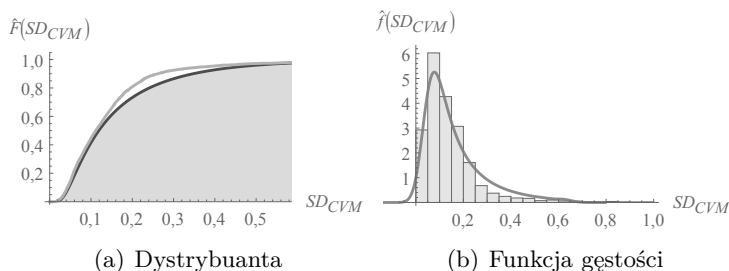
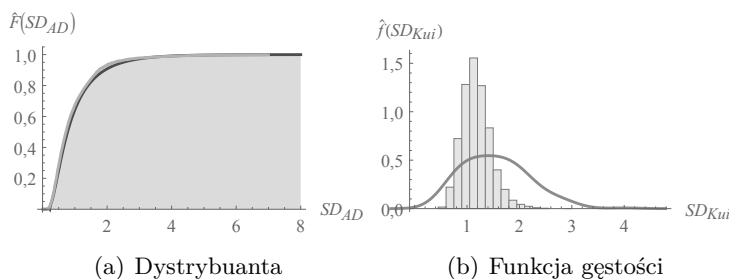


Rys. 3.22: Empiryczny i asymptotyczny rozkład statystyki testu SD_{KS}



Rys. 3.23: Empiryczny i asymptotyczny rozkład statystyki testu SD_{Kui}

Ostatnią grupą testów VaR , analizowaną pod kątem rozmiaru, były testy wielowymiarowe LB_H i LR_H , pozwalające na jednoczesne sprawdzanie autokorelacji przekroczeń w modelach VaR dla wielu poziomów tolerancji. W przypadku obydwóch testów wielowymiarowych oszacowania rozmiaru przekraczały 10% dla szeregów liczących 250 obserwacji (tab. 3.6). Różnice

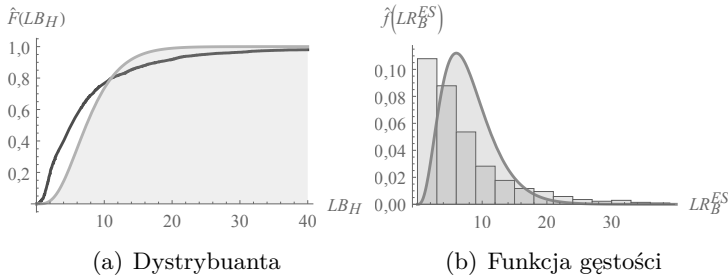
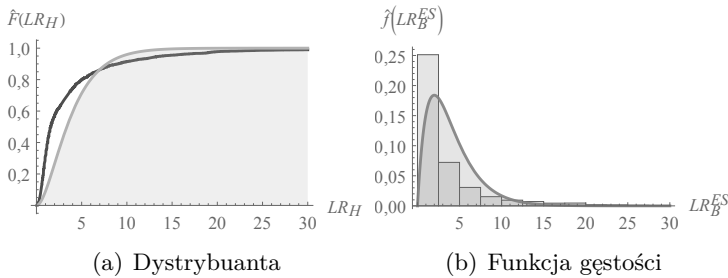
Rys. 3.24: Empiryczny i asymptotyczny rozkład statystyki testu SD_{CVM} Rys. 3.25: Empiryczny i asymptotyczny rozkład statystyki testu SD_{AD}

między rozkładami empirycznymi i asymptotycznymi dla szeregów o długości 250 potwierdziła analiza graficzna (rys. 3.26–3.27). Jednak w rozważanej grupie testów zaobserwowano wyraźną zbieżność rozkładów empirycznych do rozkładów asymptotycznych przy wydłużaniu szeregów czasowych.

Tab. 3.6: Empiryczny rozmiar wielowymiarowych testów *VaR*

Test	Długość szeregu			
	250	500	750	1000
LB_H	0,199	0,126	0,105	0,092
LR_H	0,102	0,092	0,095	0,087

Porównanie testów wielowymiarowych pokazało, że wykorzystanie, zaproponowanego w niniejszej pracy do oceny modelu *VaR*, testu LR_H , porównującego macierze korelacji, pozwoliło zbliżyć rozmiar empiryczny do przyjętego poziomu 5%. Dla najkrótszych rozważanych szeregów – o długości 250 obserwacji – wielowymiarowy test LB_H charakteryzował się rozmiarem bliskim 20%, natomiast zastosowanie testu LR_H umożliwiło zmniejszenie błędu, polegającego na nieprawidłowym odrzucaniu hipotezy zerowej, do ok. 10%

Rys. 3.26: Empiryczny i asymptotyczny rozkład statystyki testu LB_H Rys. 3.27: Empiryczny i asymptotyczny rozkład statystyki testu LR_H

W wyniku porównania wszystkich rozważanych w niniejszym badaniu testów autokorelacji VaR wyciągnięto wniosek, że najlepszym dopasowaniem rozmiaru empirycznego do założonego poziomu istotności charakteryzują się: klasyczna statystyka Ljunga-Boxa LB oraz dynamiczny test kwantylowy DQ , oparty na podejściu regresyjnym. Poza typowymi testami autokorelacji, największą dokładność w odniesieniu do błędu I rodzaju stwierdzono dla testów zgodności, w których weryfikacja modelu VaR dokonywana jest w oparciu o dopasowanie całych rozkładów prawdopodobieństwa. Zarówno dla testu LR_B opartego o momenty jak i dla nieparametrycznych testów D_{KS} , $D_{K_{ui}}$, D_{CVM} , D_{AD} otrzymane oszacowania rozmiaru były bliskie 5%.

3.3.2. Ocena mocy testów VaR

Ocena mocy testów odnoszących się do bezwarunkowego rozkładu przekroczeń – T_1 , Z_{uc} , LR_{uc} , W_{uc} , LM_{uc} , i LR_{TUFF} – przeprowadzona została z wykorzystaniem procesu Bernoulliego z parametrem $\pi_1 \neq p$, zatem prawdopodobieństwo przekroczenia VaR ustalono na poziomie niezgodnym z hipotezą H_0 . Przyjmowano π_1 kolejno na poziomach 0,01; 0,03; 0,07; 0,09.

Symulacyjna ocena mocy pozostałych testów – odnoszących się do warunkowego rozkładu przekroczeń – wymagała generowania danych charakte-

ryzujących się określoną strukturą zależności w czasie. W niniejszej pracy rozważano cztery schematy autokorelacji odzwierciedlające zjawisko gromadzenia zmienności oraz, wynikające z niego, skorelowanie zmiennych tworzących proces wyjątków *VaR* (rozd. 3.2.). Dla każdego schematu analizowano trzy warianty cechujące się różną siłą zależności, zatem łącznie rozważano 12 struktur korelacji.

W trzech początkowych schematach zastosowano dwustopniową procedurę, której pierwszy krok polegał na generowaniu wartości stopy zwrotu z procesu cechującego się gromadzeniem zmienności. Jako miarę gromadzenia zmienności przyjęto korelację kwadratów stopy zwrotu. Wykorzystywano zatem generatory zmiennych losowych umożliwiające uzyskanie zmiennych, których kwadraty mają zadaną strukturę korelacji. Kolejne warianty w każdym ze schematów były tak dobierane, by współczynnik autokorelacji kwadratów ρ_1 wynosił 0,1; 0,3 i 0,5. W drugim kroku procedury przyjęto oszacowanie *VaR* na poziomie kwantyla rzędu $p = 0,05$ bezwarunkowego rozkładu stopy zwrotu. Niedopasowanie modelu *VaR* oraz modeli generujących wartości stopy zwrotu pozwoliło uzyskać proces przekroczeń cechujący się autokorelacją, o udziale przekroczeń zgodnym z przyjętym poziomem tolerancji. W przypadku testów dla dwóch poziomów tolerancji – LB_H i LR_H – oszacowania *VaR* przyjmowano na poziomach kwantyli rzędów $p_1 = 0,01$ i $p_2 = 0,05$.

W pierwszym schemacie do odzwierciedlenia zjawiska gromadzenia zmienności wykorzystano proces *GARCH*(1,1) postaci (3.3.1.) z parametrami $\omega_1 = 0,000001$, $\beta_1 = 0,85$. Wartość parametru α_1 ustalana była w taki sposób, by zapewnić, że współczynnik autokorelacji ρ_1 kwadratów stopy zwrotu $R_t, t = 1, 2, \dots$ przyjmie ustaloną wartość, wynoszącą kolejnych wariantach 0,1; 0,3; 0,5.

Drugi schemat rozważany w pracy wykorzystywał proces *BGAR*(1)

$$Y_n(k, \theta, r) = B_n(kr, k\bar{r})Y_{n-1}(k, \theta) + G'_n(k\bar{r}, \theta),$$

o rozkładzie brzegowym gamma $\Gamma(\frac{1}{2}, 2)$. Wartości zmiennych $G'_n(k\bar{r}, \theta)$ i $B_n(kr, k\bar{r})$, $\bar{r} = 1 - r$, generowano niezależnie z rozkładów odpowiednio gamma $\Gamma(k\bar{r}, \theta)$ i beta $\beta(kr, k\bar{r})$, natomiast wartości $Y_{n-1}(k, \theta, r)$ pochodziły z rozkładu $\Gamma(k, \theta)$. Wartość parametru k przyjęto na poziomie $k = \frac{1}{2}$, parametru θ na poziomie $\theta = 2$, natomiast parametr r , odpowiadający za siłę autokorelacji kwadratów zmiennej, przyjmował w kolejnych wariantach wartości 0,1; 0,3; 0,5.

W procesie $BGMA(1)$, wykorzystanym w kolejnym eksperymencie,

$$Y_n = G_n\left(\frac{k}{1+s}, \theta\right) + B_n\left(\frac{ks}{1+s}, \frac{k\bar{s}}{1+s}\right)G_n\left(\frac{k}{1+s}, \theta\right)$$

taki sam jak wyżej rozkład brzegowy gamma uzyskano poprzez generowanie w sposób niezależny wartości $G_n\left(\frac{k}{1+s}, \theta\right)$ i $B_n\left(\frac{ks}{1+s}, \frac{k\bar{s}}{1+s}\right)$, $\bar{s} = 1-s$, z rozkładów odpowiednio $\Gamma\left(\frac{k}{1+s}, \theta\right)$ i $\beta\left(\frac{ks}{1+s}, \frac{k\bar{s}}{1+s}\right)$ z parametrami $k = \frac{1}{2}, \theta = 2$. W celu otrzymania odpowiednich wartości współczynnika autokorelacji kwadratów przyjmowano $s = \frac{\rho_1}{1-\rho_1}$, gdzie wartość ρ_1 wynosiła kolejno 0,1; 0,3; 0,5.

W ostatnim schemacie autokorelacji zastąpiono dwustopniową procedurę otrzymania procesu wyjątków VaR bezpośrednim generatorem procesu wyjątków, opartym na własności Markowa. Ponieważ w tak skonstruowanym eksperymencie nie generuje się wartości stopy zwrotu, konieczne było przyjęcie miary zależności nie odwołującej się do procesu $(R_t)_{t=1}^T$. Jako miarę zależności między kolejnymi przekroczeniami przyjęto π_{11} w macierzy przejść łańcucha Markowa. Ustalono rozkład początkowy $(\pi_0; \pi_1) = (0, 95; 0, 5)$. Wszystkie analizowane warianty spełniały $\pi_{11} \neq \pi_{01}$, co gwarantowało istnienie zależności między przekroczeniem VaR a wartością procesu w kolejnym okresie. Rozważano parametr π_{11} o wartościach 0,1; 0,2; 0,3, natomiast wartość π_{01} przyjmowana była na poziomie zapewniającym stacjonarność procesu. W rozkładzie stacjonarnym prawdopodobieństwo wystąpienia wyjątku było zgodne z założonym poziomem tolerancji¹⁰.

W przypadku wielu rozważanych testów korzysta się z asymptotycznych rozkładów statystyk testowych. W większości testów rozmiar próby odpowiada liczbie przekroczeń VaR , która jest z reguły niewielka, zatem korzystanie z rozkładów granicznych mogłoby powodować dużą niedokładność analiz. Z tego względu dla testów, w przypadku których nie jest znany dokładny rozkład statystyki dla próby skończonej, symulacyjną ocenę mocy przeprowadzono przy zastosowaniu techniki testu Monte Carlo (rozdz. 3.1.2.). Zgodnie z algorytmem testu Monte Carlo rozkłady statystyk testowych wyznaczano symulacyjne. Dla rozkładów dyskretnych zastosowano randomizację testów, umożliwiającą dokładne zachowanie rozmiaru.

Badanie symulacyjne przeprowadzono dla poziomu istotności $\alpha = 5\%$. Eksperymenty powtarzano dla szeregów o długości $T = 250, 500, 750, 1000$, wykonując każdorazowo 10000 replikacji¹¹. W teście LR_{Haas} , gdzie rozmiar

¹⁰ Reprezentacja procesów $GARCH(q,r)$, $BGAR(1)$, $BGMA(1)$ i Markowa oraz ich podstawowe własności podane zostały w rozdz. 3.2.

¹¹ Taka liczba replikacji zagwarantowała otrzymanie dokładnych oszacowań, zwłaszcza

próby rozumiany jest jako liczba wyjątków N , podobnie jak przy badaniu rozmiaru, eksperymenty skonstruowano w ten sposób by uzyskać odpowiednią liczbę N , równą wartości oczekiwanej liczby przekroczeń $E(\sum_{t=1}^T I_t) = Tp$.

W pierwszym etapie badania związanego z oceną mocy testów *VaR* przeprowadzono porównanie testów T_1 , Z_{uc} , LR_{uc} , W_{uc} , LM_{uc} i LR_{TUFF} , odnoszących się do bezwarunkowego rozkładu wyjątków, służących sprawdzaniu udziału przekroczeń *VaR* (tab. 3.7-3.10). Symulacyjna ocena mocy testów tej grupy pokazała, że najwyższą mocą cechowały się testy T_1 i Z_{uc} , oparte bezpośrednio na rozkładzie dwumianowym lub na zbieżności rozkładu dwumianowego do rozkładu normalnego. Szczególnie dużą przewagę tych testów (przejawiającą się w niektórych eksperymentach niemalże dwukrotnie wyższą mocą empiryczną – tab. 3.9) w stosunku do testów opartych na rozkładzie χ^2 zaobserwowano dla krótkich szeregów czasowych. Testy T_1 i Z_{uc} były w stanie wykrywać błędne modele ryzyka z ok. 70% skutecznością na podstawie szeregów 750 obserwacji (w przypadku niedoszacowania ryzyka, t.j. $p = 0,7$ lub $p = 0,9$) lub na podstawie 500 obserwacji (w przypadku przeszacowania ryzyka, t.j. $p = 0,1$ lub $p = 0,3$).

Tab. 3.7: Empiryczna moc na podstawie eksperymentu opartego na schemacie Bernoulliego, $p = 0,01$

Test	Długość szeregu			
	250	500	750	1000
T_1	0,99	1,00	1,00	1,00
Z_{uc}	0,99	1,00	1,00	1,00
LR_{uc}	0,98	1,00	1,00	1,00
W_{uc}	0,99	1,00	1,00	1,00
LM_{uc}	0,97	1,00	1,00	1,00
LR_{TUFF}	0,28	0,34	0,33	0,33

Wśród testów bazujących na statystyce o rozkładzie χ^2 testy LM_{uc} i W_{uc} charakteryzowały się wyższą częstością odrzucania błędnej hipotezy od standardowej statystyki LR_{uc} . Szczegółowe wnioski dotyczące mocy tych testów różniły się znacząco w zależności od hipotezy alternatywnej. W przypadku niedoszacowania ryzyka ($p = 0,7$ lub $p = 0,9$), co uznawane jest za szczególnie niebezpieczne z punktu widzenia zarządzania ryzykiem w praktyce, największe oceny mocy uzyskano dla testu LM_{uc} , którego konstrukcję zaproponowano w niniejszej pracy, w oparciu o regułę regułę mnożników Lagrange'a

dla dłuższych szeregów, gdzie różnice przy powtarzaniu eksperymentów widoczne były dopiero na trzecim miejscu po przecinku.

Tab. 3.8: Empiryczna moc na podstawie eksperymentu opartego na schemacie Bernoulliego, $p = 0,03$

Test	Długość szeregu			
	250	500	750	1000
T_1	0,45	0,72	0,86	0,94
Z_{uc}	0,47	0,72	0,87	0,94
LR_{uc}	0,37	0,62	0,80	0,90
W_{uc}	0,42	0,67	0,83	0,92
LM_{uc}	0,28	0,56	0,76	0,89
LR_{TUFF}	0,07	0,06	0,07	0,07

Tab. 3.9: Empiryczna moc na podstawie eksperymentu opartego na schemacie Bernoulliego, $p = 0,07$

Test	Długość szeregu			
	250	500	750	1000
T_1	0,37	0,62	0,74	0,86
Z_{uc}	0,39	0,60	0,74	0,86
LR_{uc}	0,21	0,46	0,62	0,76
W_{uc}	0,16	0,39	0,55	0,72
LM_{uc}	0,31	0,53	0,67	0,80
LR_{TUFF}	0,06	0,07	0,07	0,07

Tab. 3.10: Empiryczna moc na podstawie eksperymentu opartego na schemacie Bernoulliego, $p = 0,09$

Test	Długość szeregu			
	250	500	750	1000
T_1	0,81	0,97	1,00	1,00
Z_{uc}	0,82	0,97	1,00	1,00
LR_{uc}	0,67	0,93	0,99	1,00
W_{uc}	0,57	0,91	0,98	1,00
LM_{uc}	0,77	0,96	0,99	1,00
LR_{TUFF}	0,08	0,08	0,08	0,09

(tab. 3.9-3.10). Test ten osiągał moc rzędu 70% dla szeregów 750 obserwacji, natomiast testy LR_{uc} i W_{uc} dla 1000 obserwacji. Z kolei w przypadku przeszacowania ryzyka ($p = 0,1$ lub $p = 0,3$) najwyższą mocą charakteryzował się test W_{uc} , skonstruowany w myśl zasady Walda (tab. 3.7-3.8). Osiągał on

empiryczną częstość odrzucania błędnej hipotezy na poziomie 70% przy 500 obserwacjach, natomiast testy LR_{uc} i LM_{uc} przy 750 obserwacjach.

Bardzo niskie oszacowania mocy, odbiegające od pozostałych testów, zaobserwowano dla testu LR_{TUFF} , który wykorzystuje niewielką część informacji z próby, gdyż jego kryterium testowe zbudowane jest jedynie na podstawie czasu upływającego do pierwszego przekroczenia VaR . Otrzymane oszacowania mocy dla tego testu były we wszystkich eksperymentach niższe od 40%, niezależnie od długości szeregu.

W kolejnym etapie badania oceniano moc testów odnoszących się do warunkowego rozkładu przekroczeń VaR , służących wykrywaniu autokorelacji. Pierwszą grupą analizowanych testów autokorelacji były testy LR_{ind} , DQ , DQ_{logist} , LB , $Q_{ind}^{x^2}$ i WW , skonstruowane w oparciu o zero-jedynkowy proces wyjątków VaR , zgodny z łańcuchem Markowa (tab. 3.11-3.14).

Tab. 3.11: Empiryczna moc na podstawie eksperymentu opartego na procesie $GARCH$

Test	ρ_1	Długość szeregu			
		250	500	750	1000
LR_{ind}	0,1	0,10	0,16	0,20	0,22
	0,3	0,23	0,42	0,58	0,69
	0,5	0,26	0,46	0,60	0,71
DQ	0,1	0,28	0,41	0,52	0,64
	0,3	0,50	0,82	0,95	0,99
	0,5	0,49	0,94	0,99	1,00
DQ_{logist}	0,1	0,25	0,34	0,41	0,49
	0,3	0,59	0,83	0,94	0,97
	0,5	0,75	0,96	0,99	1,00
LB	0,1	0,28	0,43	0,56	0,65
	0,3	0,56	0,82	0,92	0,97
	0,5	0,60	0,82	0,91	0,96
$Q_{ind}^{x^2}$	0,1	0,12	0,16	0,19	0,20
	0,3	0,21	0,36	0,47	0,54
	0,5	0,22	0,42	0,55	0,64
WW	0,1	0,07	0,12	0,14	0,18
	0,3	0,13	0,27	0,40	0,48
	0,5	0,14	0,30	0,44	0,56

Wnioski dotyczące wyboru testów o największej mocy różniły się znacząco w zależności od procesu generującego dane. Otrzymane wyniki pokazały, że rozważane testy charakteryzowały się największą skutecznością w przypadku danych pochodzących z procesu *GARCH*. Ograniczając rozważania do tego eksperymentu, największe oszacowania mocy otrzymano dla klasycznego testu autokorelacji *LB* oraz testów kwantylowych *DQ* i *DQ_{logist}*. Skuteczność wykrywania błędnych modeli rzędu 80% była w przypadku tych testów osiągnięta dla szeregów 500 obserwacji (i współczynnika korelacji $\rho_1 = 0.3$). Zdecydowanie niższą empiryczną moc otrzymano dla popularnie stosowanego w analizach ryzyka testu *LR_{ind}*. Skuteczność rzędu 70% w tym przypadku osiągnięta została dopiero dla szeregu 1000 obserwacji (i współczynnika korelacji $\rho_1 = 0.3$). Najniższymi oszacowaniami mocy charakteryzowały się testy niezależności Pearsona $Q_{ind}^{\chi^2}$ i serii Walda-Wolfowitza *WW*.

Tab. 3.12: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGAR*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
<i>LR_{ind}</i>	0,1	0,07	0,11	0,12	0,14
	0,3	0,30	0,49	0,67	0,76
	0,5	0,63	0,87	0,96	0,99
<i>DQ</i>	0,1	0,10	0,11	0,14	0,17
	0,3	0,32	0,50	0,68	0,76
	0,5	0,63	0,88	0,97	0,99
<i>DQ_{logist}</i>	0,1	0,09	0,09	0,10	0,11
	0,3	0,30	0,42	0,56	0,66
	0,5	0,63	0,85	0,95	0,99
<i>LB</i>	0,1	0,09	0,13	0,16	0,17
	0,3	0,34	0,55	0,72	0,81
	0,5	0,70	0,91	0,98	1,00
$Q_{ind}^{\chi^2}$	0,1	0,27	0,40	0,52	0,59
	0,3	0,75	0,94	0,98	0,99
	0,5	0,93	0,99	1,00	1,00
<i>WW</i>	0,1	0,19	0,33	0,43	0,54
	0,3	0,67	0,90	0,97	0,99
	0,5	0,87	0,99	1,00	1,00

W przypadku danych pochodzących z procesów *BGAR*, *BGMA* lub łańcucha Markowa najmocniejszymi testami okazały się testy niezależności Pearsona $Q_{ind}^{\chi^2}$ i serii Walda-Wolfowitza *WW*, które osiągały skuteczność rzędu 70%-80% (przy współczynniku korelacji $\rho_1 = 0.3$) już dla najkrótszych szeregów, składających się z 250 obserwacji (tab. 3.12-3.14). Pozostałe testy z reguły były w stanie wykryć błędny model ryzyka dopiero przy szeregu 750 obserwacji (moc empiryczna ok. 70% przy współczynniku korelacji $\rho_1 = 0.3$).

Tab. 3.13: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGMA*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
<i>LR_{ind}</i>	0,1	0,07	0,10	0,12	0,13
	0,3	0,28	0,50	0,68	0,79
	0,5	0,67	0,93	0,99	1,00
<i>DQ</i>	0,1	0,08	0,11	0,14	0,16
	0,3	0,29	0,50	0,68	0,79
	0,5	0,64	0,92	0,99	1,00
<i>DQ_{logist}</i>	0,1	0,09	0,08	0,10	0,12
	0,3	0,29	0,45	0,59	0,71
	0,5	0,66	0,92	0,98	1,00
<i>LB</i>	0,1	0,08	0,11	0,15	0,17
	0,3	0,31	0,54	0,73	0,82
	0,5	0,69	0,94	0,99	1,00
$Q_{ind}^{\chi^2}$	0,1	0,27	0,40	0,52	0,60
	0,3	0,83	0,97	0,99	1,00
	0,5	0,99	1,00	1,00	1,00
<i>WW</i>	0,1	0,18	0,32	0,43	0,55
	0,3	0,74	0,95	0,99	1,00
	0,5	0,97	1,00	1,00	1,00

Drugą rozważaną grupą testów autokorelacji były testy *LR_{Haas}*, *LR_{Weib}*, *LR_{gamma}* i *EACD*, bazujące na procesie odległości między wyjątkami *VaR* (tab. 3.15-3.18). Wyniki badania pokazały, że, tak jak większość testów poprzedniej grupy, testy te charakteryzowały się stosunkowo wysoką mocą w eksperymentach z procesem *GARCH*. W tym przypadku rozważane testy osiągały z reguły moc rzędu 70% przy szeregach 500 obserwacji (i współczynniku korelacji $\rho_1 = 0.3$). Najwyższą skutecznością wyróżniał się regresyjny

test z rozkładem wykładniczym *EACD*, który był w stanie wykrywać błędne modele już dla najkrótszych rozważanych szeregów, natomiast przy 500 obserwacjach (i współczynniku korelacji $\rho_1 = 0.3$) miał empiryczną moc powyżej 80%.

Tab. 3.14: Empiryczna moc na podstawie eksperymentu opartego na łańcuchu Markowa

Test	π_{11}	Długość szeregu			
		250	500	750	1000
<i>LR_{ind}</i>	0,1	0,10	0,13	0,17	0,19
	0,2	0,36	0,59	0,76	0,86
	0,3	0,64	0,89	0,97	0,99
<i>DQ</i>	0,1	0,10	0,15	0,19	0,22
	0,2	0,36	0,59	0,74	0,85
	0,3	0,61	0,87	0,97	0,99
<i>DQ_{logist}</i>	0,1	0,10	0,11	0,13	0,16
	0,2	0,35	0,53	0,66	0,77
	0,3	0,61	0,87	0,96	0,99
<i>LB</i>	0,1	0,10	0,16	0,21	0,24
	0,2	0,39	0,64	0,79	0,88
	0,3	0,67	0,90	0,98	0,99
<i>Q²_{ind}</i>	0,1	0,20	0,28	0,35	0,43
	0,2	0,58	0,80	0,91	0,96
	0,3	0,82	0,96	0,99	1,00
<i>WW</i>	0,1	0,14	0,23	0,29	0,37
	0,2	0,48	0,74	0,87	0,94
	0,3	0,75	0,94	0,99	1,00

Test *EACD* charakteryzował się wyższą od innych mocą również w pozostałych eksperymentach. W przypadku eksperymentów opartych na procesach *BGAR* i *BGMA* wyraźnie spadały oszacowania mocy dla testów *LR_{Weib}* i *LR_{gamma}*, sprawdzających własność braku pamięci, natomiast testy *LR_{H_{aa}s}* i *EACD* ciągle pozostawały stosunkowo skuteczne. Z kolei w eksperymentach opartych na łańcuchu Markowa spadała moc wszystkich testów. Najskuteczniejsze również w tym przypadku procedury *LR_{H_{aa}s}* i *EACD* były w stanie wykrywać błędne modele dopiero z korelacją minimum $\rho_1 = 0.5$ i przy szeregach liczących co najmniej 500 obserwacji.

Tab. 3.15: Empiryczna moc na podstawie eksperymentu opartego na procesie *GARCH*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
<i>LR_{Haas}</i>	0,1	0,35	0,49	0,61	0,69
	0,3	0,45	0,60	0,74	0,83
	0,5	0,47	0,67	0,80	0,88
<i>LR_{Weib}</i>	0,1	0,05	0,13	0,21	0,29
	0,3	0,32	0,68	0,86	0,93
	0,5	0,44	0,75	0,87	0,93
<i>LR_{gamma}</i>	0,1	0,03	0,06	0,09	0,13
	0,3	0,17	0,55	0,76	0,88
	0,5	0,28	0,70	0,86	0,92
<i>EACD</i>	0,1	0,31	0,31	0,42	0,51
	0,3	0,59	0,82	0,93	0,97
	0,5	0,70	0,88	0,96	0,98

Tab. 3.16: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGAR*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
<i>LR_{Haas}</i>	0,1	0,24	0,34	0,40	0,46
	0,3	0,55	0,76	0,87	0,94
	0,5	0,73	0,93	0,98	1,00
<i>LR_{Weib}</i>	0,1	0,03	0,02	0,02	0,02
	0,3	0,13	0,31	0,46	0,57
	0,5	0,51	0,85	0,94	0,98
<i>LR_{gamma}</i>	0,1	0,02	0,01	0,01	0,01
	0,3	0,10	0,25	0,37	0,50
	0,5	0,47	0,83	0,93	0,98
<i>EACD</i>	0,1	0,24	0,15	0,14	0,15
	0,3	0,52	0,65	0,78	0,87
	0,5	0,85	0,96	0,99	1,00

Tab. 3.17: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGMA*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
<i>LR_{H_{aa}s}</i>	0,1	0,24	0,32	0,39	0,46
	0,3	0,58	0,82	0,91	0,96
	0,5	0,85	0,98	1,00	1,00
<i>LR_{W_{ei}b}</i>	0,1	0,03	0,02	0,02	0,02
	0,3	0,09	0,28	0,42	0,56
	0,5	0,49	0,93	0,99	1,00
<i>LR_{gamma}</i>	0,1	0,03	0,01	0,01	0,01
	0,3	0,06	0,20	0,33	0,49
	0,5	0,41	0,92	0,99	1,00
<i>EACD</i>	0,1	0,24	0,14	0,14	0,14
	0,3	0,50	0,66	0,82	0,90
	0,5	0,95	0,99	1,00	1,00

Tab. 3.18: Empiryczna moc na podstawie eksperymentu opartego na łańcuchu Markowa

Test	π_{11}	Długość szeregu			
		250	500	750	1000
<i>LR_{H_{aa}s}</i>	0,1	0,18	0,25	0,30	0,35
	0,2	0,42	0,61	0,72	0,80
	0,3	0,60	0,82	0,91	0,96
<i>LR_{W_{ei}b}</i>	0,1	0,03	0,03	0,02	0,01
	0,3	0,05	0,10	0,15	0,19
	0,3	0,17	0,43	0,58	0,72
<i>LR_{gamma}</i>	0,1	0,03	0,02	0,01	0,01
	0,2	0,04	0,06	0,09	0,13
	0,3	0,14	0,36	0,50	0,67
<i>EACD</i>	0,1	0,27	0,15	0,14	0,13
	0,2	0,32	0,36	0,46	0,54
	0,3	0,45	0,67	0,76	0,81

Ze względu na konieczność wykorzystania wartości stopy zwrotu do przeprowadzenia testów D_{KS} , D_{Kui} , D_{CVM} , D_{AD} , należących do kolejnej grupy – testów zgodności – nie możliwe było przeprowadzenie oceny mocy na podstawie eksperymentu opartego na procesie Markowa, w którym bezpośrednio generowane są przekroczenia *VaR*. Ocena mocy przeprowadzono na podstawie eksperymentów wykorzystujących procesy *GARCH*, *BGAR* i *BGMA* (tab. 3.19-3.21). Tak jak w przypadku poprzednich grup testów wyższą moc zaobserwowano przy eksperymencie opartym na procesie *GARCH* niż dla danych wygenerowanych przez procesy *BGAR* i *BGMA*.

Porównanie poszczególnych testów zgodności pokazało, że największą mocą charakteryzował się test LR_B , który wykorzystuje informację o oszacowaniach rozkładu stopy zwrotu oraz sprawdza dopasowanie rozkładów poprzez ich momenty (tab. 3.19-3.21). Empiryczna moc tego testu była powtarzalna we wszystkich rodzajach eksperymentów i wyższa niż dla wcześniejszych grup testów autokorelacji *VaR*. Skuteczność rzędu 70%, dochodząca nawet do 90%, osiągana była w większości eksperymentów już dla najkrótszych szeregów, liczących 250 obserwacji (przy współczynniku autokorelacji $\rho_1 = 0.3$).

Tab. 3.19: Empiryczna moc na podstawie eksperymentu opartego na procesie *GARCH*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
D_{KS}	0,1	0,08	0,08	0,09	0,10
	0,3	0,32	0,43	0,51	0,58
	0,5	0,63	0,79	0,87	0,91
D_{Kui}	0,1	0,14	0,16	0,18	0,20
	0,3	0,54	0,64	0,71	0,78
	0,5	0,79	0,90	0,95	0,98
D_{CVM}	0,1	0,08	0,08	0,09	0,10
	0,3	0,34	0,44	0,52	0,60
	0,5	0,64	0,78	0,86	0,91
D_{AD}	0,1	0,12	0,12	0,13	0,14
	0,3	0,48	0,58	0,67	0,75
	0,5	0,75	0,89	0,95	0,98
LR_B	0,1	0,21	0,21	0,24	0,24
	0,3	0,63	0,68	0,71	0,73
	0,5	0,81	0,87	0,89	0,89

Tab. 3.20: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGAR*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
D_{KS}	0,1	0,07	0,08	0,08	0,08
	0,3	0,15	0,15	0,16	0,15
	0,5	0,28	0,29	0,32	0,29
D_{Kui}	0,1	0,06	0,06	0,07	0,07
	0,3	0,14	0,14	0,15	0,15
	0,5	0,30	0,30	0,33	0,32
D_{CVM}	0,1	0,07	0,08	0,07	0,08
	0,3	0,14	0,14	0,16	0,14
	0,5	0,26	0,26	0,29	0,28
D_{AD}	0,1	0,08	0,07	0,07	0,07
	0,3	0,16	0,15	0,17	0,14
	0,5	0,29	0,29	0,32	0,29
LR_B	0,1	0,14	0,22	0,36	0,46
	0,3	0,88	0,99	1,00	1,00
	0,5	1,00	1,00	1,00	1,00

Tab. 3.21: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGMA*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
D_{KS}	0,1	0,08	0,07	0,07	0,07
	0,3	0,13	0,13	0,13	0,12
	0,5	0,19	0,20	0,20	0,19
D_{Kui}	0,1	0,07	0,07	0,07	0,07
	0,3	0,12	0,12	0,12	0,13
	0,5	0,20	0,21	0,22	0,22
D_{CVM}	0,1	0,07	0,07	0,08	0,08
	0,3	0,13	0,12	0,12	0,12
	0,5	0,16	0,18	0,18	0,18
D_{AD}	0,1	0,08	0,07	0,07	0,07
	0,3	0,14	0,12	0,13	0,11
	0,5	0,19	0,19	0,19	0,19
LR_B	0,1	0,15	0,22	0,37	0,45
	0,3	0,92	1,00	1,00	1,00
	0,5	1,00	1,00	1,00	1,00

W przypadku pozostałych testów tej grupy – D_{KS} , D_{Kui} , D_{CVM} i D_{AD} – zaobserwowano dużą wrażliwość, większą niż dla innych grup testów, na rodzaj procesu generującego dane. Częstość odrzucania błędnej hipotezy była porównywalna z poprzednimi grupami testów w eksperymencie opartym na procesie *GARCH* (ok. 70% osiągane przy szeregu 750 obserwacji i $\rho_1 = 0.3$ lub przy szeregu 250 obserwacji i $\rho_1 = 0.5$). Najskuteczniejsze w tej grupie były testy D_{Kui} i D_{AD} . Jednak empiryczna moc tych testów była zdecydowanie niższa w przypadku eksperymentów opartych na innych procesach, gdzie dla najdłuższych szeregów osiągała dopiero ok. 30%.

Badanie symulacyjne przeprowadzone w grupie testów gęstości spektralnej – SD_{KS} , SD_{Kui} , SD_{CVM} , SD_{AD} – tak jak w przypadku poprzedniej grupy testów pokazało dużą wrażliwość na rodzaj procesu generującego dane, jednak charakter zależności był odmienny (tab. 3.22-3.25). Testy gęstości spektralnej charakteryzowały się wyjątkową, w porównaniu z innymi testami, skutecznością w przypadku danych pochodzących z procesów *BGAR* i *BGMA*, gdzie empiryczna moc tych testów z reguły przewyższała 70% już dla szeregów 250 obserwacji (przy współczynniku autokorelacji $\rho_1 = 0.3$).

Tab. 3.22: Empiryczna moc na podstawie eksperymentu opartego na procesie *GARCH*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
SD_{KS}	0,1	0,12	0,20	0,25	0,28
	0,3	0,23	0,46	0,59	0,76
	0,5	0,23	0,54	0,71	0,84
SD_{Kui}	0,1	0,11	0,17	0,21	0,23
	0,3	0,22	0,43	0,57	0,69
	0,5	0,23	0,51	0,72	0,84
SD_{CVM}	0,1	0,13	0,18	0,21	0,25
	0,3	0,25	0,43	0,55	0,71
	0,5	0,26	0,53	0,68	0,82
SD_{AD}	0,1	0,13	0,25	0,30	0,34
	0,3	0,28	0,58	0,72	0,84
	0,5	0,30	0,67	0,81	0,91

Niewiele niższe oszacowania mocy otrzymano przy danych pochodzących z łańcucha Markowa, gdzie 70% skuteczność była osiągalna dla 500 obserwacji (przy współczynniku autokorelacji $\rho_1 = 0.3$). Z kolei testy gęstości spektralnej charakteryzowały się niższą skutecznością wykrywania błędnych

modeli ryzyka niż testy innych grup gdy obserwacje pochodziły z procesu *GARCH*, typowego dla danych finansowych. W tym przypadku z reguły 70% oszacowania mocy były obserwowane dopiero przy 1000 obserwacji. Analiza porównawcza poszczególnych testów należących do grupy testów gęstości

Tab. 3.23: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGAR*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
SD_{KS}	0,1	0,25	0,40	0,47	0,55
	0,3	0,74	0,91	0,97	1,00
	0,5	0,93	0,99	1,00	1,00
SD_{Kui}	0,1	0,18	0,25	0,32	0,39
	0,3	0,65	0,85	0,94	0,98
	0,5	0,88	0,99	1,00	1,00
SD_{CVM}	0,1	0,26	0,40	0,46	0,55
	0,3	0,72	0,92	0,98	1,00
	0,5	0,93	1,00	1,00	1,00
SD_{AD}	0,1	0,22	0,41	0,46	0,53
	0,3	0,71	0,92	0,98	0,99
	0,5	0,93	1,00	1,00	1,00

Tab. 3.24: Empiryczna moc na podstawie eksperymentu opartego na procesie *BGMA*

Test	ρ_1	Długość szeregu			
		250	500	750	1000
SD_{KS}	0,1	0,28	0,37	0,49	0,59
	0,3	0,82	0,97	1,00	1,00
	0,5	0,98	1,00	1,00	1,00
SD_{Kui}	0,1	0,20	0,23	0,31	0,40
	0,3	0,69	0,91	0,99	1,00
	0,5	0,97	1,00	1,00	1,00
SD_{CVM}	0,1	0,28	0,37	0,48	0,58
	0,3	0,80	0,97	1,00	1,00
	0,5	0,99	1,00	1,00	1,00
SD_{AD}	0,1	0,24	0,39	0,48	0,56
	0,3	0,80	0,97	1,00	1,00
	0,5	0,99	1,00	1,00	1,00

spektralnej pokazała, że z reguły najwyższą częstość odrzuceń błędnych modeli otrzymywano dla testu SD_{AD} opartego na statystyce Andersona-Darlinga.

Tab. 3.25: Empiryczna moc na podstawie eksperymentu opartego na łańcuchu Markowa

Test	π_{11}	Długość szeregu			
		250	500	750	1000
SD_{KS}	0,1	0,18	0,27	0,35	0,43
	0,2	0,56	0,78	0,88	0,97
	0,3	0,81	0,95	0,99	1,00
SD_{Kui}	0,1	0,11	0,14	0,20	0,26
	0,2	0,44	0,66	0,80	0,92
	0,3	0,72	0,89	0,98	0,99
SD_{CVM}	0,1	0,17	0,25	0,32	0,37
	0,2	0,55	0,79	0,89	0,97
	0,3	0,82	0,95	0,99	1,00
SD_{AD}	0,1	0,16	0,27	0,33	0,40
	0,2	0,54	0,80	0,89	0,95
	0,3	0,83	0,95	0,99	1,00

Przeprowadzenie oceny i porównania mocy testów wielowymiarowych LB_H i LR_H możliwe było na podstawie trzech rodzajów eksperymentów Monte Carlo – $GARCH$, $BGAR$ i $BGMA$ – w których z odpowiednich procesów generowano wartości stopy zwrotu. Po przyjęciu założeń dotyczących modelu ryzyka, pozwoliło to na otrzymanie oszacowań VaR na różnych poziomach tolerancji oraz utworzenie szeregów przekroczeń odpowiadających tym poziomom.

Otrzymane oszacowania dla testów wielowymiarowych pokazały, że ich moc jest porównywalna dla różnych rodzajów procesów generujących dane (choć nieco wyższa dla procesów $BGAR$ i $BGMA$ niż dla procesu $GARCH$) oraz porównywalna z najlepszymi z wcześniej omawianych testów (tab. 3.26–3.28). Empiryczna moc rzędu 70% w przypadku testów wielowymiarowych z reguły osiągnięta była dla długości szeregu wynoszącej 500 obserwacji (przy współczynniku autokorelacji $\rho_1 = 0.3$).

Porównanie wyników badania pomiędzy testem LB_H i zaproponowanym w niniejszej pracy do oceny modeli VaR , wielowymiarowym testem dla macierzy korelacji LR_H pokazało, że wykorzystanie tego drugiego podejścia pozwoliło uzyskać poprawę mocy. Częstość odrzuceń uzyskana na podstawie statystyki LR_H była w większości przypadków wyższa niż przy wykorzysta-

niu testu LB_H . W szczególności dużą (wielokrotną) poprawę mocy uzyskano szeregów – o długości 250 obserwacji. W przypadku eksperymentów $BGAR$ i $BGMA$ pozwoliło to zaobserwować dla takich szeregów 70% skuteczność (przy współczynniku autokorelacji $\rho_1 = 0.3$).

Tab. 3.26: Empiryczna moc na podstawie eksperymentu opartego na procesie $GARCH$

Test	ρ_1	Długość szeregu			
		250	500	750	1000
LB_H	0,1	0,05	0,32	0,94	0,53
	0,3	0,06	0,73	0,98	0,92
	0,5	0,07	0,85	0,99	0,97
LR_H	0,1	0,18	0,27	0,34	0,39
	0,3	0,35	0,58	0,70	0,81
	0,5	0,43	0,66	0,78	0,85

Tab. 3.27: Empiryczna moc na podstawie eksperymentu opartego na procesie $BGAR$

Test	ρ_1	Długość szeregu			
		250	500	750	1000
LB_H	0,1	0,06	0,24	0,92	0,42
	0,3	0,11	0,81	1,00	0,98
	0,5	0,20	0,99	1,00	1,00
LR_H	0,1	0,22	0,34	0,43	0,51
	0,3	0,67	0,88	0,96	0,99
	0,5	0,91	0,99	1,00	1,00

Tab. 3.28: Empiryczna moc na podstawie eksperymentu opartego na procesie $BGMA$

Test	ρ_1	Długość szeregu			
		250	500	750	1000
LB_H	0,1	0,07	0,23	0,93	0,42
	0,3	0,11	0,86	1,00	1,00
	0,5	0,19	1,00	1,00	1,00
LR_H	0,1	0,23	0,34	0,45	0,53
	0,3	0,75	0,93	0,99	1,00
	0,5	1,00	1,00	1,00	1,00

3.3.3. Wybór optymalnych testów *VaR*

Analiza porównawcza oszacowań rozmiaru i mocy przeprowadzona w ramach poszczególnych grup testów pozwoliła wyłonić procedury gwarantujące najwyższy stopień kontroli nad błędem pierwszego rodzaju oraz jednocześnie charakteryzujące się najwyższą skutecznością wykrywania błędnych modeli ryzyka. W szczególności zaobserwowano przypadki, w których, dzięki zaproponowanym w pracy modyfikacjom można uzyskać dokładniejsze zachowanie rozmiaru lub poprawę mocy.

W wyniku analizy testów udziału przekroczeń *VaR* można wyciągnąć wniosek, że najdokładniejsze w kontekście rozmiaru jak i najskuteczniejsze w sensie mocy wykrywania błędnych modeli były testy oparte bezpośrednio na rozkładzie dwumianowym (T_1) lub jego zbieżności do rozkładu normalnego (Z_{uc}). Badanie symulacyjne pokazało, że zbieżność statystyki dwumianowej do rozkładu normalnego jest szybsza niż statystyki ilorazu wiarygodności, w popularnym teście Kupca (LR_{uc}), do rozkładu χ^2 .

W odniesieniu do testów opartych na statystyce χ^2 , korzystne zarówno z punktu widzenia praktycznego jak i z punktu widzenia własności statystycznych, okazało się wykorzystanie zaproponowanych w niniejszej pracy statystyk bazujących na zasadach Walda oraz mnożników Lagrange'a. Dzięki zastosowaniu zasady mnożników Lagrange'a uzyskano test LM_{uc} , którego praktyczną przewagą jest określoność – zatem możliwość bezpośredniego stosowania – w przypadku szeregu, w którym nie zaobserwowano żadnego przekroczenia *VaR*. W szczególności ma to znaczenie przy testowaniu *VaR* na niskich poziomach tolerancji, takich jak zalecane obecnie w świetle Umowy Bazylejskiej III. Ponadto badanie symulacyjne pokazało przewagę testu mnożników Lagrange'a w odniesieniu do wykrywania modeli niedoszacowujących ryzyka.

Spośród testów autokorelacji *VaR*, testami, które zapewniały zarówno dokładność, rozumianą jako odpowiedni poziom błędu I rodzaju, jak i wysoką skuteczność wykrywania nieprawidłowych modeli, były: klasyczny test Ljunga-Boxa LB oparty na współczynnikach autokorelacji oraz dynamiczny test kwantylowy DQ oparty na podejściu regresyjnym. Testy te miały najwyższą skuteczność w przypadku danych generowanych przez proces $GARCH$, co w szczególności uzasadnia ich stosowanie w analizach finansowych, jednak ich moc pozostawała na zadowalającym poziomie również wobec innych klas hipotez alternatywnych. Ze względu na duży błąd pierwszego rodzaju, odrzucono, popularnie stosowany w ocenie autokorelacji przekroczeń *VaR*, test Christoffersena LR_{ind} .

Wysoką mocą, stabilną wobec różnych klas hipotez alternatywnych, wśród wszystkich rozważanych testów wyróżnił się test *EACD* oparty na regresji z rozkładem wykładniczym. Wyniki badania pokazały jednak rozmiękanie się rozmiaru tego testu z założonym poziomem istotności, co może wynikać z faktu, że jego hipoteza zerowa leży dokładnie na granicy przestrzeni wszystkich dopuszczalnych parametrów. Zatem przy stosowaniu tego testu należałoby uwzględnić korektę rozkładu asymptotycznego lub korzystać z kwantyli empirycznych.

Inne testy wykorzystujące w swojej konstrukcji rozkład wykładniczy i własność braku pamięci – testy Weibulla LR_{Weib} i gamma LR_{gamma} – charakteryzowały się asymptotyczną zbieżnością do założonego rozmiaru oraz zadowalającą skutecznością w przypadku danych generowanych przez proces *GARCH*, przy czym nieco lepsze własności zaobserwowano dla statystyki LR_{Weib} . Jednak ich moc spadała w odniesieniu do hipotez alternatywnych obejmujących inne procesy.

Uwzględnienie w eksperymencie symulacyjnym szerokiej klasy procesów stochastycznych pozwoliło na wyciągnięcie wniosku, że ważnym uzupełnieniem analiz ryzyka może być zastosowanie testów opartych na gęstości spektralnej, wywodzących się z nauk fizycznych. Pokazano ich wysoką czułość na błędne modele w przypadku danych generowanych przez inne procesy niż proces *GARCH*, gdzie moc popularnych testów *VaR* z reguły spada. Najlepszym spośród testów gęstości spektralnej, w kontekście oszacowań rozmiaru i mocy, był test SD_{AD} oparty na statystyce Andersona-Darlinga. Jednak ze względu na przesunięcie rozkładu asymptotycznego w stosunku do empirycznego, należałoby wykluczyć stosowanie rozkładów asymptotycznych przy przeprowadzaniu testów tej grupy.

Analizę testów *VaR*, konstruowanych w oparciu o proces przekroczeń dla jednego ustalonego poziomu tolerancji, uzupełniono propozycją testu wielowymiarowego LR_H , sprawdzającego niezależność wektorów losowych. Podejście wielowymiarowe pozwoliło na jednoczesne testowanie modelu dla wielu poziomów tolerancji, dzięki czemu możliwe było wykorzystanie większej ilości informacji z próby. Ponadto konstrukcja zaproponowanego testu umożliwiła dużą elastyczność w zarządzaniu ryzykiem dzięki możliwości dostosowania testowanego obszaru rozkładu do potrzeb analizy. W porównaniu z wcześniejszym wielowymiarowym testem *VaR*, opartym na uogólnieniu statystyki Ljunga-Boxa LB_H , wykorzystanie testu zaproponowanego w niniejszej pracy pozwoliło dwukrotnie zmniejszyć błąd pierwszego rodzaju oraz wielokrotnie poprawić moc przy krótkich szeregach czasowych.

W odniesieniu do propozycji restrykcyjnej oceny modeli *VaR* poprzez weryfikację zgodności całego rozkładu, przeprowadzone badanie wykazało dokładność oraz największą skuteczność wykrywania błędnych modeli dla procedury LR_B , opartej na transformacji Rosenblatta i sprawdzaniu zgodności momentów. Jednak ze względu na zależność tej procedury od założeń dotyczących postaci parametrycznej rozkładu stopy zwrotu, jej stosowanie w praktyce może być utrudnione. Wyciąganie zasadnych wniosków wymagałoby sprawdzenia odporności na ryzyko estymacyjne, związane z niespełnieniem założeń o rozkładzie prawdopodobieństwa.

Pozostałe testy zgodności, oparte na porównywaniu wartości transformowanych zmiennych z rozkładem równomiernym, cechowały się dużą dokładnością w sensie oszacowań rozmiaru. Jednak ich moc była na poziomie porównywalnym z innymi grupami testów jedynie w przypadku danych pochodzących z procesów *GARCH*. Testy te nie były skuteczne wobec innych hipotez alternatywnych, co powoduje, że ich stosowanie wymaga znajomości procesu generującego dane lub uzupełnienia innymi procedurami testowymi. Spośród testów zgodności z rozkładem równomiernym, największą częstością wykrywania błędnych modeli cechował się test D_{AD} oparty na statystyce Andersona-Darlinga.

3.4. Wyniki symulacyjnej oceny rozmiaru i mocy testów *ES*

3.4.1. Ocena rozmiaru testów *ES*

Badanie rozmiaru testów *ES* oparto na eksperymentach w których wartości stopy zwrotu R_t , $t = 1, \dots, T$ generowano z procesu *GARCH*(1,1) postaci

$$\begin{aligned} R_t &= \sigma_t Z_t, & Z_t &\sim N(0, 1), \\ \sigma_t^2 &= \omega_1 + \alpha_1 R_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \end{aligned} \quad (3.11)$$

z parametrami¹² $\omega_1 = 0,000001$, $\alpha = 0,14$, $\beta_1 = 0,85$. Zgodność z hipotezą zerową osiągnięto poprzez wyznaczanie oszacowań *ES* jako wartości oczekiwanej uciętych rozkładów $R_t|_{\{-\infty, VaR_p(R_t)\}}$, $t = 1, \dots, T$ z wykorzystaniem parametrów procesu generującego dane.

¹² Parametry procesu *GARCH*(1,1) przyjęto na podstawie badania wstępnego dokonanego dla sześciu indeksów giełdowych [Małecka 2011].

W celu uzyskania wartości zmiennej U_t , $t = 1, \dots, T$ w teście wielkości wyjątków LR_B^{ES} przeprowadzono transformację Rosenblatta (rozdz. 2.3.2.) za pomocą dystrybuanty rozkładu normalnego o wartości oczekiwanej zero i odchyleniu standardowym wynikającym z procesu (3.4.1.). Transformacji w zmienne Z_t , $t = 1, \dots, T$ dokonano z wykorzystaniem wartości parametrów przyjętych w procesie generującym dane. Parametry te zastosowano również do uzyskania wystandaryzowanych reszt U_t , \hat{U}_t i U_t^* w testach bootstrapowych. Ocenę rozmiaru testów przeprowadzono w dwóch krokach. W pierwszym z nich symulacyjnie wyznaczono rozmiar, korzystając z teoretycznych rozkładów statystyk testowych, natomiast w drugim kroku dokonano graficznego porównania kwantyli empirycznych i teoretycznych.

Badanie symulacyjne przeprowadzono dla poziomu istotności $\alpha = 5\%$. Wszystkie eksperymenty powtórzono dla długości prób $T = 250, 500, 750, 1000$, wykonując każdorazowo 10000 replikacji¹³.

Przeprowadzone badanie symulacyjne pokazało, że rozmiar większości testów dedykowanych modelom ES zbliżony jest do nominalnego poziomu 5% (tab. 3.29). Za wyjątkiem testu bootstrapowego U^* i testu F_{Ch} otrzymane oszacowania rozmiaru dla wszystkich rozważanych długości szeregów mieściły się w przedziale od 4,7% do 5,6%. W przypadku testu LR_B^{ES} zgodność rozkładów empirycznych i asymptotycznych potwierdzona została na podstawie analizy graficznej, przedstawiającej porównanie otrzymanych empirycznie i asymptotycznych kształtów funkcji gęstości i dystrybuant (rys. 3.28).

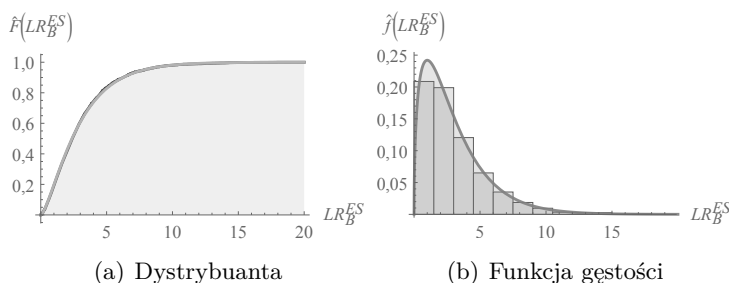
Tab. 3.29: Empiryczny rozmiar testów *ES*

Test	Długość szeregu			
	250	500	750	1000
LR_B^{ES}	0,051	0,052	0,051	0,053
S	0,056	0,047	0,049	0,050
F_{Ch}	0,156	0,200	0,251	0,299
F_{Ch}^*	0,050	0,050	0,054	0,052
U	0,035	0,043	0,046	0,047
U^*	0,103	0,112	0,125	0,142

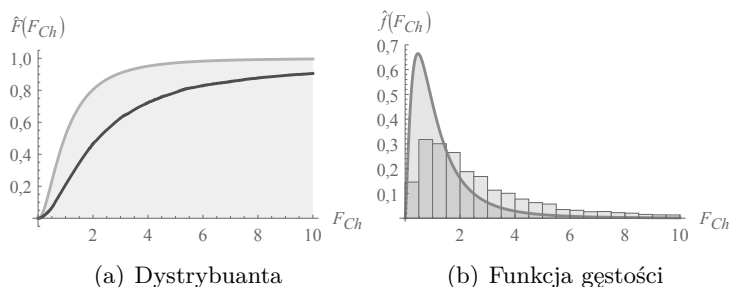
¹³ Taka liczba replikacji zagwarantowała otrzymanie dokładnych oszacowań, zwłaszcza dla dłuższych szeregów, gdzie różnice przy powtarzaniu eksperymentów widoczne były dopiero na trzecim miejscu po przecinku.

Porównanie takie nie było możliwe dla testów nieparametrycznych oraz testu *S*, w którym zamiast rozkładu asymptotycznego proponuje się wyznaczanie przybliżonych *p*-wartości w oparciu o technikę punktu siodłowego.

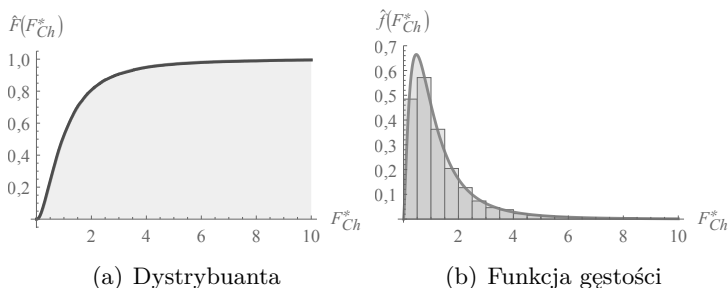
Oceny rozmiaru wyznaczone dla testów regresyjnych F_{Ch} i zmodyfikowanej wersji F_{Ch}^* pokazały, że zaproponowana standaryzacja pozwoliła na uzyskanie zgodności z rozkładem teoretycznym oraz, tym samym, redukcję błędu pierwszego rodzaju. Częstość odrzuceń dla F_{Ch} była ponad dwa razy większa od nominalnego poziomu istotności. Różnicę między empirycznymi kwantylami a odpowiadającymi im wartościami rozkładu teoretycznego F_{Ch} potwierdziła graficzna ocena kształtów dystrybuant i gęstości (rys. 3.29). Ponadto w przypadku testu F_{Ch} , opartego na regresji, gdzie może występować problem braku stacjonarności składnika losowego, wyniki nie potwierdziły zbieżności rozkładu statystyki testowej do teoretycznego rozkładu Fischera. W przypadku zmodyfikowanej statystyki F_{Ch}^* częstość odrzuceń była w przybliżeniu równa nominalnemu poziomowi 5% oraz kształt rozkładu w przybliżeniu pokrył się z kształtem odpowiadającym teoretycznemu rozkładowi Fischera (rys. 3.30).



Rys. 3.28: Empiryczny i asymptotyczny rozkład statystyki testu LR_B^{ES}



Rys. 3.29: Empiryczny i asymptotyczny rozkład statystyki testu F_{Ch}

Rys. 3.30: Empiryczny i asymptotyczny rozkład statystyki testu F_{Ch}^*

Poza testem regresyjnym F_{Ch} , największą rozbieżnością między empirycznym rozmiarem testu a założonym poziomem istotności zaobserwowano dla testu bootstrapowego U^* , który weryfikuje hipotezę o wartości ES za pomocą średniej z próby, wyznaczonej na podstawie ustalonego udziału skrajnych obserwacji. W przypadku testu U , dla którego otrzymano mniejszą częstość popełniania błędu I rodzaju, średnia obliczana jest dla obserwacji przekraczających poziom VaR .

3.4.2. Ocena mocy testów ES

W celu symulacyjnego wyznaczenia mocy testów ES wykorzystano eksperymenty, w których jako proces generujący dane przyjęto $GARCH(1,1)$ dany wzorem (3.4.1.), natomiast oszacowania ES wyznaczono z użyciem błędnych parametrów. W przypadku każdego z rozważanych testów przeprowadzono cztery warianty eksperymentu symulacyjnego. W pierwszym wariantcie do wyznaczenia ES przyjęto model homoskedastyczny, w którym odchylenie standardowe zmiennych R_t , $t = 1, \dots, T$ ustalono na poziomie odchylenia standardowego rozkładu bezwarunkowego. W wariantcie tym uzyskano oszacowania ES zgodne z wartością oczekiwaną uciętego rozkładu bezwarunkowego zmiennych $R_t|_{(-\infty, VaR_p(R_t))}$, $t = 1, \dots, T$, nie uwzględniające zmienności procesu w czasie. Pozostałe trzy warianty uwzględniały zmienność procesu w czasie, jednak w każdym momencie t , $t = 1, \dots, T$ odchylenie standardowe zmiennej R_t przyjmowane było na zaniżonym poziomie σ_t^* , stanowiącym ustalony procent rzeczywistego odchylenia standardowego σ_t . Uzyskane w ten sposób poziomy ES odzwierciedlały zjawisko heteroskedastyczności ale były systematycznie niedoszacowane. W kolejnych eksperymentach przyjmowano wartości odchylenia standardowego wynoszące 90%, 70% i 50% rzeczywistej wartości tego parametru.

Podobnie jak w przypadku testów *VaR*, ze względu na nieznaną dokładność rozkładów statystyk dla prób skończonych, symulacyjną ocenę mocy przeprowadzono przy zastosowaniu techniki testu Monte Carlo (rozdz. 3.1.2.). Dzięki temu zagwarantowano dokładny rozmiar testów i uzyskano porównywalność otrzymanych oszacowań mocy. Zgodnie z algorytmem testu Monte Carlo rozkłady statystyk testowych wyznaczano symulacyjne. Badanie symulacyjne przeprowadzono dla poziomu istotności $\alpha = 5\%$. Wszystkie eksperymenty powtórzono dla długości prób $T = 250, 500, 750, 1000$, wykonując każdorazowo 10000 replikacji¹⁴.

W pierwszym etapie oceny mocy testów *ES* porównanie mocy przeprowadzono na podstawie eksperymentu, w którym wartości stopy zwrotu generowano z procesu *GARCH* a następnie oszacowania *ES* uzyskiwano na podstawie błędnego modelu homoskedastycznego (tab. 3.30). W wyniku takiego eksperymentu otrzymuje się naprzemiennie niedoszacowane i przeszacowane wartości ryzyka. Z tego względu w tym etapie pominięto test *S* oparty na technice punktu siodłowego, który, ze uwagi na swoją konstrukcję, ma charakter jednostronny i służy jedynie wykrywaniu modeli niedoszacowujących ryzyka.

Tab. 3.30: Empiryczna moc na podstawie eksperymentu opartego procesie homoskedastycznym

Test	Długość szeregu			
	250	500	750	1000
LR_B^{ES}	0,73	0,85	0,94	0,96
F_{Ch}^*	0,09	0,15	0,17	0,20
U	0,38	0,41	0,45	0,52
U^*	0,18	0,23	0,26	0,26

Najwyższe oceny mocy, rzędu 70% lub więcej dla wszystkich długości szeregu, uzyskano w przypadku testu LR_B^{ES} , porównującego empiryczną i teoretyczną funkcję gęstości uciętą do zakresu wartości przekraczających *VaR*. Niższe oceny mocy, rzędu ok. 50% zaobserwowano dla testu bootstrapowego U . Najniższymi oszacowaniami charakteryzowały się testy F_{Ch}^* – wykorzystujący podejście regresyjne¹⁵ – oraz test bootstrapowy U^* , których empiryczna moc nie przekraczała 30%.

¹⁴ Taka liczba replikacji zagwarantowała otrzymanie dokładnych oszacowań, zwłaszcza dla dłuższych szeregów, gdzie różnice przy powtarzaniu eksperymentów widoczne były dopiero na trzecim miejscu po przecinku.

¹⁵ W badaniu mocy, ze względu na dużą niedokładność testu, stwierdzoną przy badaniu rozmiaru, nie uwzględniono statystyki F_{Ch} , zastępując ją zmodyfikowaną wersją F_{Ch}^* .

Wyniki testu F_{Ch}^* pokazały, że zaproponowana regresja liniowa nie była skuteczna w sensie objaśniania różnic między błędnymi oszacowaniami ES a wartościami stopy zwrotu w przypadku przekroczenia VaR . W rozważanym eksperymencie wyjątki VaR pojawiały się najczęściej w okresach o zwiększonej zmienności rynkowej – kiedy następował wzrost lub spadek stóp zwrotu natomiast dynamika oszacowań ES była niewystarczająca by zapewnić odpowiednie dostosowanie prognoz ryzyka. Sytuacje takie mogły być poprzedzone zarówno okresami spokojnymi jak i mogły występować w efekcie trwałych zawirusowań na rynku. Skutkiem tego mógł być brak istotności regresji opartej na przeszłych wartościach stopy zwrotu i oszacowaniach ES .

Zbyt niska częstość odrzuceń w teście bootstrapowym U^* może być powodowana sytuacjami, kiedy w zaobserwowanym szeregu wystąpiło niewiele przekroczeń VaR – poniżej wartości oczekiwanej. W takim przypadku test U^* , w przeciwieństwie do U przyjmuje zbyt dużą liczbę obserwacji należących do ogona rozkładu, wskutek czego ich średnia jest zaniżona (co do wartości bezwzględnej). Wynikająca z niej wartość statystyki testowej jest w konsekwencji niewystarczająca do odrzucenia błędnej hipotezy zerowej.

Drugi etap oceny mocy testów ES przeprowadzono na podstawie eksperymentu, w którym wartości stopy zwrotu generowane były z procesu $GARCH$ a następnie oszacowania ES uzyskiwano przyjmując model o zaniżonym odchyleniu standardowym (tab. 3.31). Otrzymane rezultaty pokazały, że największą częstość odrzuceń błędnego modelu pozwalają uzyskać metody bootstrapowe U i U^* , oparte na oszacowaniu ES z próby oraz rozkładzie statystyki testowej otrzymanym symulacyjnie przez wielokrotne replikowanie próby. Wyniki badania w szczególności wskazały na szybki wzrost skuteczności tych testów przy oddalaniu się od hipotezy zerowej. Efektem tego były oszacowania mocy rzędu 90% już dla najkrótszych rozważanych szeregów przy $\sigma_t^* = 0,7\sigma_t$.

Porównanie ocen mocy dla dwóch testów bootstrapowych U i U^* pokazało, że wyższą częstość odrzuceń otrzymywano dla testu U^* , opartego na ustalonym udziale obserwacji należących do ogona rozkładu. W przeciwieństwie do U^* , test U podlega ryzyku związanemu z błędną estymacją modelu VaR . Z tego powodu, ze względu na systematycznie zaniżoną zmienność, skutkującą zaniżonym oszacowaniem VaR , test U oparty był, w rozważanym eksperymencie, na zbyt dużej liczbie obserwacji leżących blisko środka rozkładu. Zbliżenie do centralnej części rozkładu działało w tym przypadku w kierunku spadku mocy testu. Efekt ten miał jednak znaczenie jedynie w pobliżu hipotezy zerowej ($\sigma_t^* = 0,09\sigma_t$).

Tab. 3.31: Empiryczna moc na podstawie eksperymentu opartego na procesie *GARCH*

Test	σ_t^*	Długość szeregu			
		250	500	750	1000
LB_B^{ES}	$0,9\sigma_t$	0,63	0,77	0,85	0,89
	$0,7\sigma_t$	0,65	0,81	0,88	0,92
	$0,5\sigma_t$	0,90	0,98	0,99	1,00
S	$0,9\sigma_t$	0,35	0,52	0,64	0,71
	$0,7\sigma_t$	0,54	0,74	0,85	0,91
	$0,5\sigma_t$	0,87	0,97	0,99	1,00
F_{Ch}^*	$0,9\sigma_t$	0,06	0,05	0,05	0,05
	$0,7\sigma_t$	0,05	0,05	0,05	0,06
	$0,5\sigma_t$	0,06	0,04	0,05	0,06
U	$0,9\sigma_t$	0,18	0,29	0,36	0,44
	$0,7\sigma_t$	0,92	1,00	1,00	1,00
	$0,5\sigma_t$	1,00	1,00	1,00	1,00
U^*	$0,9\sigma_t$	0,48	0,72	0,88	0,91
	$0,7\sigma_t$	1,00	1,00	1,00	1,00
	$0,5\sigma_t$	1,00	1,00	1,00	1,00

Podobnie dużą skutecznością jak testy bootstrapowe charakteryzowały się testy S i LR_B^{ES} , które osiągnęły częstotliwość odrzucania błędnej hipotezy rzędu 70% dla 500 obserwacji i $\sigma_t^* = 0,7\sigma_t$ oraz niemalże 90% dla 250 obserwacji i $\sigma_t^* = 0,5\sigma_t$. W szczególności, w przypadku testu LR_B^{ES} , empiryczna moc rzędu 70% przy 500 obserwacjach dla wszystkich wariantów eksperymentu potwierdziła dużą czułość tego testu na modele niedoszacowujące ryzyka.

Najniższą mocą charakteryzował się test regresyjny F_{Ch}^* , gdzie częstotliwość wykrywania błędnych modeli była bliska założonemu rozmiarowi testu. Rozważany na tym etapie badania eksperyment pokazał zatem możliwość obciążenia testu. W eksperymencie tym przekroczenia VaR występowały zarówno w okresach spokojnych jak i charakteryzujących się dużą zmiennością. W przypadku wystandaryzowanych danych, będących podstawą testu F_{Ch}^* , wielkość wyjątków w każdej sytuacji generowana była przez ten sam system, niezależnie od zjawiska gromadzenia zmienności występującego w danych wyjściowych. Z tej przyczyny dostępne informacje o stopie zwrotu lub prognozach ES z okresów poprzedzających mogły nie mieć mocy objaśniania, w sposób statystycznie istotny, wielkości przekroczeń VaR .

3.4.3. Wybór optymalnych testów ES

W porównaniu z weryfikacją modeli VaR , testowanie modeli ES jest zadaniem bardziej wymagającym z punktu widzenia teorii statystyki. Mimo powszechnej akceptacji tej miary ze względu na jej ugruntowanie w aksjomatycznej teorii ryzyka, testowanie ES wciąż pozostaje tematem dyskusyjnym [np. Carver 2013, Carver 2014, Hull i White 2014]. Głównym wyzwaniem w tym zakresie jest brak znajomości odpowiednich rozkładów statystyk testowych lub przynajmniej p -wartości, potrzebnych do podejmowania decyzji. ES definiowana jest jako wartość oczekiwana obserwacji należących do ogona rozkładu, zatem jej oszacowanie nie może być bezpośrednio poddane procedurze testowej, gdyż rozkład średniej skrajnych obserwacji z próby w ogólnym przypadku nie jest znany.

Przeprowadzone badanie pozwoliło wybrać dwie statystyki testowe, które w różny sposób omijają problem nieznaności rozkładów oraz charakteryzują się najlepszymi własnościami statystycznymi. Symulacyjna analiza rozmiaru i mocy pokazała, że test S , oparty na aproksymacji p -wartości za pomocą rozwinięcia funkcji generującej momenty w szereg Taylora, oraz test U , wykorzystujący metodę bootstrapową, gwarantują zarówno dokładność jak i skuteczność, rozumianą jako moc wykrywania niepoprawnych modeli.

Wybór testu U , jako testu optymalnego spośród dwóch procedur bootstrapowych U i U^* podyktowany został przede wszystkim symulacyjną oceną rozmiaru rozważanych metod. Zawyżone wyniki rozmiaru zaobserwowane dla testu U^* pokazały możliwość zbyt częstego odrzucania hipotezy zerowej w przypadku korzystania z bootstrapowego oszacowania rozkładu tej statystyki. Z drugiej strony test U , wykorzystujący w swojej konstrukcji oszacowanie poziomu VaR , może prowadzić do błędnych decyzji w przypadku gdy poziom ten zostanie nieprawidłowo wyznaczony. Zatem praktyczne stosowanie testu U wymaga gwarancji prawidłowej oceny VaR lub przeprowadzenia dodatkowej analizy ryzyka estymacyjnego, związanego z niewłaściwymi wynikami estymacji VaR .

Tak samo jak w przypadku testu bootstrapowego U^* , analiza rozmiaru pokazała konieczność przeprowadzenia modyfikacji lub odrzucenia procedury regresyjnej F_{Ch} . Zastąpienie statystyki F_{Ch} zmodyfikowaną wersją F_{Ch}^* , gdzie zapewniona została przybliżona stacjonarność składnika losowego równania regresji, przyniosło efekt zachowania odpowiedniego rozmiaru testu, jednak badanie mocy wykazało nieskuteczność omawianej procedury w wykrywaniu błędnych modeli ES .

Wybrane w wyniku porównania rozmiaru i mocy procedury S i U , weryfikujące model ES poprzez ocenę zaobserwowanych wartości tej miary, uzupełniono parametrycznym testem LR_B^{ES} , który ocenia zgodność rozkładów prawdopodobieństwa. Procedura ta, analogicznie do testu LR_B dla miary VaR , bada dopasowanie momentów rozkładu teoretycznego i empirycznego, jest zatem bardziej restrykcyjna w swojej konstrukcji niż testy S i U . W przypadku miary ES , w przeciwieństwie do VaR , ocena zgodności rozkładów ograniczona jest do skrajnych wartości zmiennej losowej. Wyniki badania symulacyjnego, podobnie jak w przypadku statystyki LR_B , pokazały zgodność rozmiaru z założonym rozmiarem nominalnym oraz wysoką moc testu LR_B^{ES} . Jednak ze względu na wykorzystywanie założeń o parametrycznej postaci rozkładu, statystyka ta może być narażona na ryzyko estymacyjne. Jej praktyczne stosowanie wymaga zatem sprawdzenia odporności na błędną specyfikację rozkładu stopy zwrotu.

Rozdział 4

Weryfikacja modeli ryzyka na przykładzie szeregów empirycznych

Podstawowym obszarem zastosowań modeli ryzyka, opartych na rozważanych w niniejszej pracy miarach, jest rynek finansowy. Jednak ze względu na uniwersalny charakter pojęć *VaR* i *ES*, mogą one być również stosowane do oceny ryzyka związanego z innego rodzaju działalnością gospodarczą. Prezentowany w niniejszym rozdziale pracy zakres zastosowań modeli opartych na *VaR* i *ES* uwzględnia rynek finansowy oraz towarowy. Pokazane zostały przykłady zastosowania procedur testowych omawianych we wcześniejszych częściach pracy do rzeczywistych szeregów czasowych, obejmujących indeksy cen akcji, surowców oraz zbóż.

Przykład wykorzystania testów statystycznych oparty został na szeregach prognoz stóp zwrotu, przekroczeń *VaR* i oszacowań *ES*, wynikających z zastosowania konkurencyjnych modeli szeregów czasowych. W przypadku wszystkich rozważanych indeksów giełdowych zastosowano jednolitą procedurę estymacji wartości *VaR* i wartości *ES*. W ramach parametrycznych metod estymacji wykorzystano zarówno homoskedastyczne modele *VaR*, ze składnikiem losowym pochodzącym z rozkładów normalnego i t-Studenta, jak i modele heteroskedastyczne, które uwzględniają zmienność wariancji rozkładu w czasie. Ze względu na możliwość występowania różnych zachowań inwestorów w sytuacji wzrostów oraz spadków na rynku, uwzględniono również modele heteroskedastyczne z asymetrią zmienności. Rozważano parametryczne mo-

dele *GARCH* oraz modele *POT* wywodzące się z teorii wartości ekstremalnych. Umożliwiło to przewidywanie poziomu zagrożenia z wykorzystaniem rozkładów klasy Pareto, opisujących grube ogony oraz prognozowanie straty w sytuacjach ekstremalnych wahań, kiedy przekraczany jest poziom *VaR*. W obszarze statystyki nieparametrycznej zastosowano metody polegające na bezpośrednim szacowaniu kwantyla, omijające problem przyjmowania założeń o klasie rozkładów prawdopodobieństwa zmiennej losowej. Wykorzystano podejście klasycznej symulacji historycznej oraz symulacji filtrowanej, uwzględniającej zmienność momentów rozkładu stopy zwrotu w czasie.

W prezentowanej analizie empirycznej wykorzystano wnioski płynące z poprzednich rozdziałów pracy, dotyczące wyboru testów statystycznych do weryfikacji modeli ryzyka. Przeprowadzono obszerną procedurę testową bazującą na metodach reprezentujących wszystkie omawiane w poprzednich rozdziałach grupy testów *VaR* i *ES*, wybranych na podstawie przeprowadzonego badania symulacyjnego. Pokazano odmienność polskiego oraz amerykańskiego rynku kapitałowego, odnoszącą się do asymetrii zachowań inwestorów oraz wykazano potrzebę stosowania metod analizy ryzyka dopasowanych do specyfiki badanego rynku.

4.1. Opis badania empirycznego

4.1.1. Opisowa analiza szeregów czasowych

Weryfikacja modeli *VaR* i *ES*, przeprowadzona na przykładzie danych empirycznych, objęła szeregi czasowe S&P500, WIG20, GSCI Wheat, Gold Bullion LBM¹. Wybór szeregów dokonany został w sposób celowy i został poddyktowany chęcią uwzględnienia przykładów zarówno z rynku finansowego jak i towarowego oraz uzyskania porównywalności wyników z wcześniejszymi badaniami empirycznymi dotyczącymi oceny modeli *VaR* i *ES* [np. Hamidi et al. 2010, Kjellson 2013, Małecka 2014].

¹ S&P500 jest indeksem w skład którego wchodzi 500 firm o największej kapitalizacji, notowanych na amerykańskich giełdach New York Stock Exchange i NASDAQ, natomiast WIG20 to indeks giełdowy 20 największych spółek akcyjnych notowanych na warszawskiej Giełdzie Papierów Wartościowych. Gold Bullion LBM jest indeksem cen spot złota, tworzonym na podstawie transakcji z rynku OTC, nadzorowanego przez London Bullion Market Association. Rynek ten stanowi największe, pod względem wartości rozliczanych transakcji, światowe centrum handlu złotem i srebrem. Indeks S&P GSCI Wheat reprezentuje ceny spot pszenicy na amerykańskiej giełdzie Chicago Mercantile Exchange i uznawany jest za punkt odniesienia do ocen zmian światowych cen rynku zboża.

Badanie przeprowadzono dla dziennych logarytmicznych stóp zwrotu z okresu od 1 stycznia 1995 do 31 grudnia 2013, co pozwoliło otrzymać ok. 5000 obserwacji w każdym szeregu czasowym. Przedział czasowy dobrano w ten sposób, by dostępne były w nim dane dzienne dla każdego z rozważanych indeksów. Długość okna, na podstawie którego estymowano parametry modeli ustalono na 1000 obserwacji. Każdorazowo po przesunięciu okna dokonywano prognozy *VaR* i *ES* na jeden dzień do przodu. Korekty ocen parametrów dokonywano co pięć obserwacji.

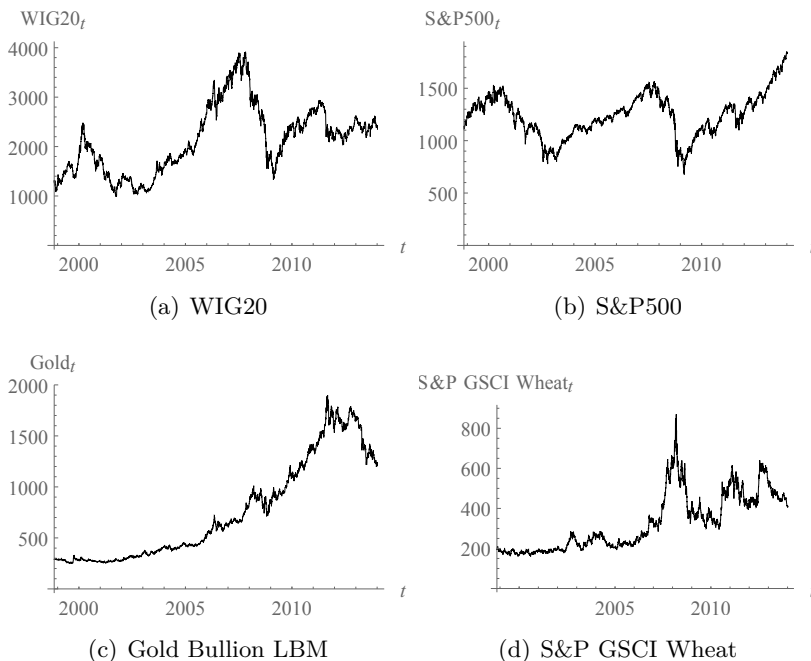
Tab. 4.1: Wartości statystyk opisowych dla stopy zwrotu, 1.01.1995–31.06.2013

Indeks	Średnia	Odch. std.	Min.	Max.	Skośność	Kurtoza
WIG20	0,00014	0,016	-0,097	0,082	-0,14	5,63
S&P500	0,00013	0,013	-0,095	0,110	-0,17	10,80
Gold Bullion LBM	0,00036	0,011	-0,102	0,074	-0,35	9,21
S&P GSCI Wheat	0,00019	0,019	-0,100	0,088	0,04	5,16

Przeprowadzona w pierwszym kroku badania empirycznego analiza wartości podstawowych statystyk opisowych rozważanych indeksów pokazała, że indeksy WIG20 i S&P500, obrazujące ceny akcji, charakteryzowały się najniższą średnią stopą zwrotu (tab. 4.1). Średnia dla indeksu WIG20, była wyższa niż dla indeksu S&P500. Z wyższą średnią dla indeksu WIG20 związane było wyższe odchylenie standardowe. Ekonomiczna interpretacja tego zjawiska nawiązuje do różnic między gospodarkami o krótszej historii rynku kapitałowego, gdzie możliwe jest osiągnięcie wysokich zysków kosztem podwyższonego ryzyka, w porównaniu z rynkami kapitałowymi o długiej tradycji.

Najwyższą średnią stopą zwrotu charakteryzował się indeks Gold Bullion LBM, obrazujący ceny złota. Odchylenie standardowe opisujące zmienność cen złota było mniejsze niż wartości odchylenia standardowego dla indeksów rynków kapitałowych, jednak wysoka stopa zwrotu pociągnęła za sobą koszt w postaci bardzo silnej asymetrii lewostronnej. Wskaźnik asymetrii dla rozważanego indeksu osiągnął wartość kilkakrotnie wyższą (co do wartości bezwzględnej) niż dla pozostałych indeksów. Jego poziom świadczy o występowaniu wysokich odchyleń stopy zwrotu w kierunku ujemnym, co potwierdza również najniższa wartość minimum otrzymana dla indeksu Gold Bullion LBM. Stopa zwrotu dla indeksu S&P GSCI Wheat, obrazującego ceny pszenicy znajdowała się przeciętnie pomiędzy stopą zwrotu indeksów rynku kapitałowego a indeksu rynku złota, związana była jednak z najwyższym odchyleniem standardowym, które świadczy o dużych przeciętnych odchyleniach od wartości średniej.

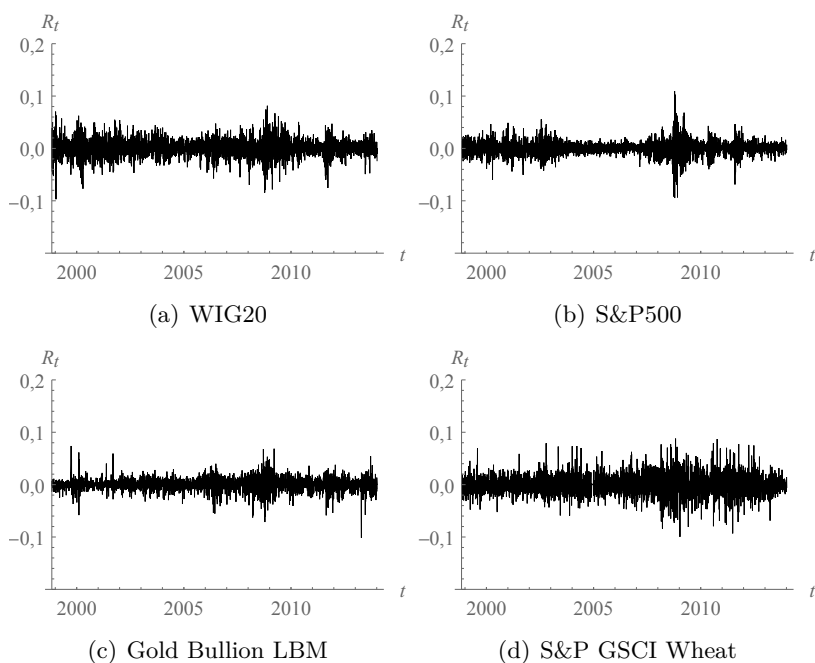
Podobnie jak w przypadku cen złota, indeksy rynków kapitałowych charakteryzowały się asymetrią lewostronną, co pokazuje typowe dla rynków akcji zjawisko występowania większych odchyleń stopy zwrotu w kierunku ujemnym niż dodatnim. Zgodnie z interpretacją ekonomiczną zachowań inwestycyjnych, wiąże się to gwałtownymi reakcjami inwestorów w przypadku spadków cen. Odchylenia stopy zwrotu w górę zdarzały się częściej lecz były niższe, co obrazuje umiarkowane reakcje inwestorów na wzrosty cen. W przypadku rynku pszenicy nie zaobserwowano wyraźnej asymetrii stopy zwrotu. Stopy zwrotu wszystkich rozważanych indeksów cechowały się leptokurtycznością, która świadczy o dużej koncentracji wartości stóp zwrotu wokół średniej. Największe wartości kurtozy zaobserwowano dla indeksów S&P500 oraz Gold Bullion LBM.



Rys. 4.1: Wartości indeksów, 1.01.1995–31.06.2013

Na rysunkach obrazujących historyczne wartości indeksów rynków kapitałowych WIG20 i S&P500 wyraźnie zaznaczyły się spadki związane z recesją w początku obecnego stulecia oraz kryzysem finansowym zapoczątkowanym w roku 2007 (rys. 4.1 a, b). Pomiedzy dwoma wymienionymi okresami, dla obu indeksów widoczna była tendencja wzrostowa, która dla indeksu S&P500 utrzymywała się również od roku 2009 do 2013. Wpływ kryzysu finansowego

2007/2008 był mniej widoczny w przypadku kształtowania się cen na rynku złota (rys. 4.1 c), na którym widoczna była długotrwała tendencja wzrostowa trwająca do roku 2012. Wykres indeksu S&P GSCI Wheat pozwolił zaobserwować wyraźne ożywienie w roku 2007 oraz gwałtowny spadek cen po tym czasie, począwszy od roku 2008 (rys. 4.1 d).

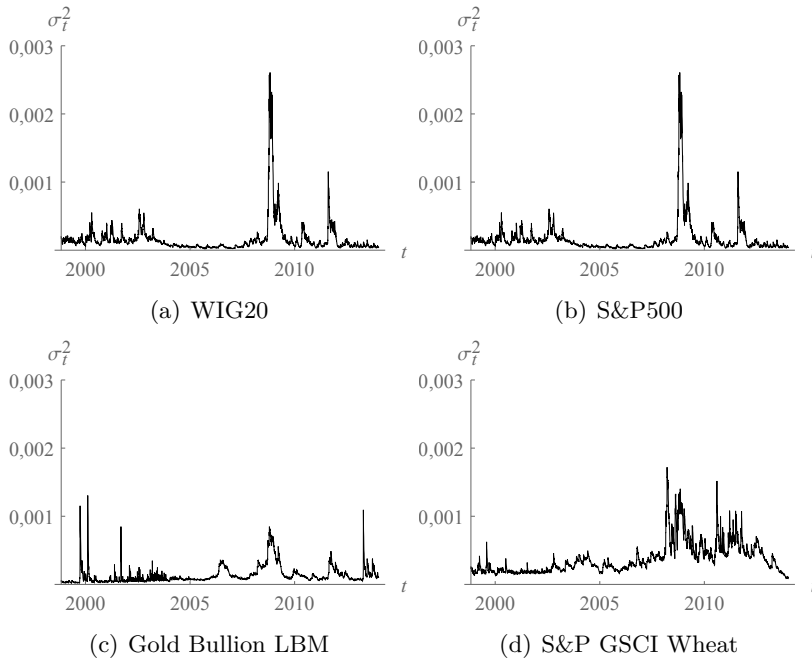


Rys. 4.2: Stopy zwrotu z indeksów, 1.01.1995–31.06.2013

Historyczny przebieg zmienności dla analizowanych indeksów pokazany został na rysunkach obrazujących wartości stopy zwrotu (rys. 4.2) oraz wariancję oszacowaną na podstawie modelu $GARCH(1,1)$ (rys. 4.3). W przypadku wszystkich rozważanych indeksów wykresy uwidoczniły zjawisko gromadzenia zmienności polegające na naprzemiennym występowaniu okresów spokojnych oraz okresów cechujących się gwałtownymi zmianami cen. Zjawisko to wyjaśniane jest impulsywnymi reakcjami inwestorów, będącymi następstwem niespodziewanych zmian na rynku. Nieprzewidziane zmiany wywołują okresy wzmożonej zmienności, po których następuje stopniowe uspokojenie sytuacji rynkowej aż do momentu wystąpienia kolejnego impulsu dużych zmian cen.

Graficzna analiza danych potwierdziła przeciętnie wyższą zmienność indeksów WIG20 oraz S&P GSCI Wheat od pozostałych dwóch rozważanych

indeksów. Szczególnie wysoki wzrost zmienności stóp zwrotu indeksów rynków kapitałowych związany z kryzysem finansowym zobrazowany został na rysunkach pokazujących zmienność indeksów WIG20 i S&P500.

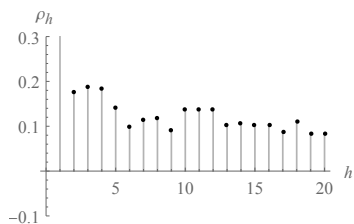


Rys. 4.3: Zmienność stóp zwrotu, 1.01.1995–31.06.2013

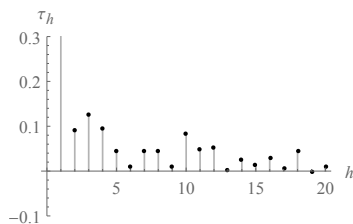
Gromadzenie zmienności uznaje się za jeden z głównych problemów w zarządzaniu ryzykiem rynkowym ze względu na fakt, że w przypadku zastosowania nieodpowiedniego modelu ryzyka, zjawisko to może prowadzić do szeregu nieprzewidywanych, następujących po sobie strat. Liczbowym odzwierciedleniem gromadzenia zmienności są wartości funkcji autokorelacji oraz autokorelacji częściowej dla kwadratów stóp zwrotu, które wyznaczono w kolejnym kroku analizy szeregów czasowych (rys. 4.4 a, b, 4.5 a, b, 4.6 a, b, 4.7 a, b). Większe wartości autokorelacji oraz autokorelacji częściowej zaobserwowano w przypadku indeksów rynków kapitałowych niż dla rynków towarowych.

Zastosowanie modeli $GARCH(1,1)$ do wyznaczenia ocen zmienności oraz przeprowadzenie standaryzacji stóp zwrotu z wykorzystaniem tych ocen pozwoliło częściowo wyeliminować zjawisko autokorelacji kwadratów stóp zwrotu (rys. 4.4 c, d, 4.5 c, d, 4.6 c, d, 4.7 c, d). Dzięki procedurze standaryzacji otrzymano szeregi reprezentujące obserwacje pochodzące z rozkładów w przybliżeniu niezależnych i niezmiennych w czasie. W przypadku indeksu S&P

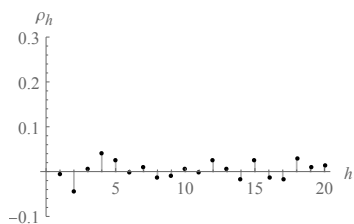
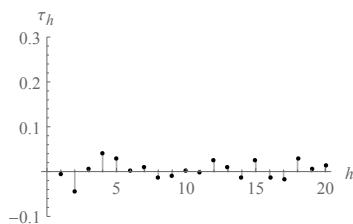
GSCI Wheat zjawisko autokorelacji stóp zwrotu zostało niemalże całkowicie wyeliminowane dla wszystkich rzędów autokorelacji.



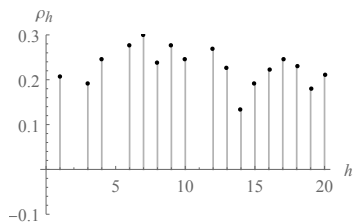
(a) ACF dla kwadratów stopy zwrotu



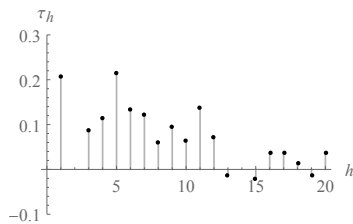
(b) PACF dla kwadratów stopy zwrotu

(c) ACF dla kwadratów reszt z modelu $GARCH(1,1)$ (d) PACF dla kwadratów reszt z modelu $GARCH(1,1)$

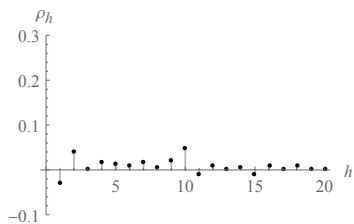
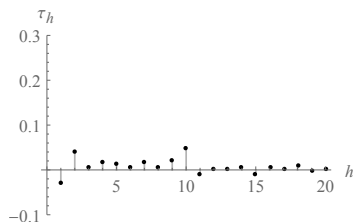
Rys. 4.4: Funkcje autokorelacji i autokorelacji częściowej dla indeksu WIG20



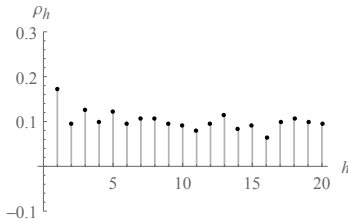
(a) ACF dla kwadratów stopy zwrotu



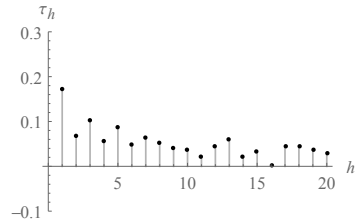
(b) PACF dla kwadratów stopy zwrotu

(c) ACF dla kwadratów reszt z modelu $GARCH(1,1)$ (d) PACF dla kwadratów reszt z modelu $GARCH(1,1)$

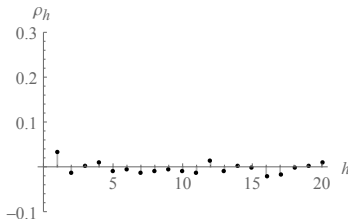
Rys. 4.5: Funkcje autokorelacji i autokorelacji częściowej dla indeksu S&P500



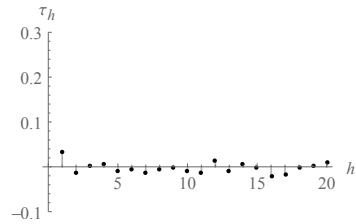
(a) ACF dla kwadratów stopy zwrotu



(b) PACF dla kwadratów stopy zwrotu

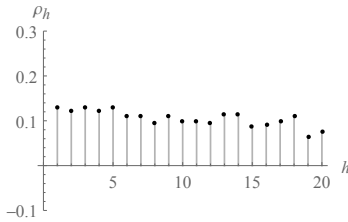


(c) ACF dla kwadratów reszt z modelu $GARCH(1,1)$

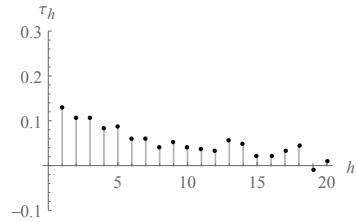


(d) PACF dla kwadratów reszt z modelu $GARCH(1,1)$

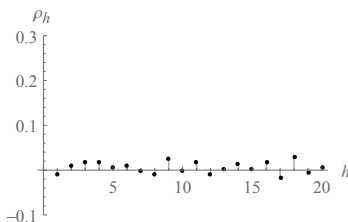
Rys. 4.6: Funkcje autokorelacji i autokorelacji częściowej dla indeksu Gold Bullion LBM



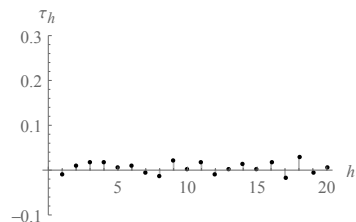
(a) ACF dla kwadratów stopy zwrotu



(b) PACF dla kwadratów stopy zwrotu



(c) ACF dla kwadratów reszt z modelu $GARCH(1,1)$



(d) PACF dla kwadratów reszt z modelu $GARCH(1,1)$

Rys. 4.7: Funkcje autokorelacji i autokorelacji częściowej dla indeksu S&P GSCI Wheat

4.1.2. Zastosowane metody wnioskowania statystycznego

Przykład wykorzystania testów statystycznych dla rozważanych szeregów czasowych oparty został na oszacowaniach VaR i ES otrzymanych za pomocą konkurencyjnych metod estymacji. Analiza porównawcza uwzględniła zarówno modele dotyczące bezwarunkowego jak i warunkowego rozkładu stóp zwrotu.

Metody parametryczne wymagały przyjęcia założeń dotyczących postaci funkcyjnej rozkładu stopy zwrotu. W badaniu rozważano rozkłady: normalny $N(\mu, \sigma)$ i t-studenta S_v . Przy założeniu $R_t \sim N(\mu, \sigma)$ estymatory VaR i ES miały postać odpowiednio

$$\begin{aligned} \hat{VaR}_p(R_t) &= \hat{\mu} + \hat{\sigma}\Phi^{-1}(p), \\ \hat{ES}_p(R_t) &= \hat{\mu} + \hat{\sigma}\frac{\phi(\Phi^{-1}(p))}{p}, \end{aligned}$$

gdzie ϕ i Φ oznaczają odpowiednio funkcję gęstości i dystrybuantę rozkładu normalnego. Przy założeniu rozkładu t-Studenta $R_t \sim S_v$ odpowiednie wzory przyjęły postać

$$\begin{aligned} \hat{VaR}_p(R_t) &= \hat{\mu} + \hat{\sigma}\sqrt{\frac{v-2}{v}}F_{S_v}^{-1}(p), \\ \hat{ES}_p(R_t) &= \hat{\mu} + \hat{\sigma}\frac{f_{S_v}(F_{S_v}^{-1}(p))}{p}\frac{v + (F_{S_v}^{-1}(p))^2}{v-1}, \end{aligned}$$

gdzie f_{S_v} i F_{S_v} oznaczają odpowiednio funkcję gęstości i dystrybuantę rozkładu t-Studenta S_v .

W celu uwzględniania ryzyka związanego z grubymi ogonami rozkładów, wykorzystano teorię wartości ekstremalnych (EVT, ang. *extreme value theory*). W pierwszym kroku przeprowadzono estymację VaR i ES za pomocą metody POT (ang. *pick-over-thresholds*) przy założeniu niezmienności rozkładu stopy zwrotu.

Metoda POT umożliwia uzyskanie parametrycznego estymatora ogona rozkładu zmiennej losowej bazując na zbieżności do uogólnionego rozkładu Pareto (GPD, generalized Pareto distribution). Jeżeli X jest zmienną o nieznanym rozkładzie F , to rozważa się zmienną $Y = X - u$ reprezentującą przekroczenia ponad ustalony poziom u . Zmienna Y ma rozkład dany dystrybuantą

$$F_u(y) = P(X - u < y | X > u) = \frac{F(y+u) - F(u)}{1 - F(u)}, \quad y > 0. \quad (4.1)$$

W metodzie *POT* wykorzystuje się zbieżność rozkładu Y dla $y > 0$ do rozkładu GPD o dystrybuancie postaci

$$G_{\xi, \beta(y)} = \begin{cases} 1 - (1 + \xi \frac{y}{\beta})^{-1/\xi} & \text{dla } \xi \neq 0, \quad \xi y > -\beta, \\ 1 - e^{-(y/\beta)} & \text{dla } \xi = 0, \quad y \geq 0. \end{cases}$$

Przekształcając wzór (4.1) otrzymuje się $F(x) = (1 - F(u))F_u(x - u) + F(u)$ dla $x > u$. Wykorzystując dystrybuantę empiryczną $\hat{F}(u) = \frac{n - n_u}{n}$, gdzie n_u jest liczbą przekroczeń ponad u można wyznaczyć estymator dystrybuanty F dla $x \geq u$, zwany estymatorem ogona rozkładu zmiennej X

$$\hat{F}(x) = \begin{cases} 1 - \frac{n_u}{n} (1 + \hat{\xi} \frac{x-u}{\hat{\beta}})^{-1/\hat{\xi}} & \text{dla } x > u \text{ i } \xi \neq 0, \\ 1 - \frac{n_u}{n} e^{-(x-u)/\hat{\beta}} & \text{dla } x > u \text{ i } \xi = 0. \end{cases} \quad (4.2)$$

gdzie $\hat{\xi}$ i $\hat{\beta}$ są estymatorami parametrów odpowiednio ξ i β . Dzięki oszacowaniu parametrów $\hat{\xi}$ i $\hat{\beta}$ uzyskuje się estymator kwantyla rzędu s , $s \in (0, 1)$, zmiennej X oraz, ze wzoru na wartość oczekiwaną uogólnionego rozkładu Pareto można wyznaczyć estymator wartości oczekiwanej ogona rozkładu zmiennej X :

$$\hat{F}^{-1}(s) = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\frac{n}{n_u} (1 - s)^{-\xi} - 1 \right), \quad (4.3)$$

$$\hat{E}(X|X > F^{-1}(s)) = \frac{\hat{F}^{-1}(s)}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \hat{\xi}}, \quad (4.4)$$

dla $s \neq 0$ [McNeil, Frey i Embrechts 2005].

W niniejszym badaniu oszacowania VaR i ES dla zmiennej R_t za pomocą metody *POT* uzyskano przyjmując, że rozkład R_t jest niezmienny w czasie oraz $X = -R_t$, $t = 1, \dots, T$, $n = T$, $s = 1 - p$. W przypadku $p = 0,05$ przyjęto u na poziomie kwantyla rzędu 0,94 natomiast w przypadku $p = 0,01$ na poziomie kwantyla rzędu 0,98. Wykorzystano zatem estymatory postaci

$$Va\hat{R}_p(R_t) = u + \frac{\hat{\beta}}{\hat{\xi}} \left(\frac{T}{T_u}(p)^{-\xi} - 1 \right), \quad (4.5)$$

$$E\hat{S}_p(R_t) = \frac{Va\hat{R}_p(R_t)}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \hat{\xi}}. \quad (4.6)$$

W przypadku modeli z warunkową heteroskedastycznością rozważano procesy $GARCH(1, 1)$ oraz $GJR - GARCH(1, 1)$ z rozkładami: normalnym i t-Studenta. Przyjęto model $R_t = \mu_t + \sigma_t Z_t$, $Z_t \sim iid.$ z równaniami wariancji

postaci

$$\sigma_t^2 = \omega_1 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad (4.7)$$

$$\sigma_t^2 = \omega_1 + (\alpha_1 + \gamma_i \mathbf{1}_{\{\epsilon_{t-1} > 0\}}) \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (4.8)$$

Przyjmując założenie o składniku losowym pochodzącym z rozkładu normalnego $Z_t \sim N(0, 1)$, $t = 1, \dots, T$, zastosowano estymatory VaR i ES odpowiednio postaci

$$\begin{aligned} \widehat{VaR}_p(R_t) &= \hat{\mu}_t + \hat{\sigma}_t \Phi^{-1}(p), \\ \widehat{ES}_p(R_t) &= \hat{\mu}_t + \hat{\sigma}_t \frac{\phi(\Phi^{-1}(p))}{p}, \end{aligned}$$

gdzie ϕ i Φ oznaczają odpowiednio funkcję gęstości i dystrybuantę rozkładu normalnego. Przy założeniu rozkładu t-Studenta $R_t \sim t_v$ odpowiednie wzory przyjęły postać

$$\begin{aligned} \widehat{VaR}_p(R_t) &= \hat{\mu}_t + \hat{\sigma}_t \sqrt{\frac{v-2}{v}} t_v^{-1}(p), \\ \widehat{ES}_p(R_t) &= \hat{\mu}_t + \hat{\sigma}_t \frac{f_{S_v}(F_{S_v}^{-1}(p)) v + (F_{S_v}^{-1}(p))^2}{p v - 1}, \end{aligned}$$

gdzie f_{S_v} i F_{S_v} oznaczają odpowiednio funkcję gęstości i dystrybuantę rozkładu t-Studenta S_v .

W kolejnym kroku badania wykorzystano modele opracowane na gruncie teorii wartości ekstremalnych w połączeniu z modelami heteroskedastycznymi $GARCH(1,1)$ i $GJR-GARCH(1,1)$. Założono jak wyżej model wartości oczekiwanej stopy zwrotu $R_t = \mu_t + \sigma_t Z_t$, $Z_t \sim iid$, gdzie wariancja σ_t pochodziła z procesów $GARCH(1,1)$ lub $GJR-GARCH(1,1)$. W celu otrzymania estymatorów VaR i ES , posługując się metodą POT , wykorzystano zbieżność rozkładu przekroczeń zmiennej Z_t ponad próg u do uogólnionego rozkładu Pareto. Przy powyższych założeniach estymatory VaR i ES przyjęły odpowiednio postać

$$\widehat{VaR}_p(R_t) = \hat{\mu}_t + \hat{\sigma}_t \left(u + \frac{\hat{\beta}}{\hat{\xi}} \left(\frac{T}{T_u}(p)^{-\xi} - 1 \right) \right), \quad (4.9)$$

$$\widehat{ES}_p(R_t) = \hat{\mu}_t + \hat{\sigma}_t \left(\frac{\widehat{VaR}_p(R_t)}{1 - \hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \hat{\xi}} \right). \quad (4.10)$$

gdzie $\hat{\xi}$ i $\hat{\beta}$ są estymatorami parametrów odpowiednio ξ i β uogólnionego rozkładu Pareto, T_u jest liczbą przekroczeń zmiennej Z_t ponad u natomiast

$\hat{\mu}_t$ i $\hat{\sigma}_t$ otrzymano na podstawie odpowiedniego modelu *GARCH* [McNeil i Frey 2000].

Parametry modeli bezwarunkowych oraz *GARCH*(1,1) i *GJR-GARCH*(1,1) otrzymano metodą quasi-największej wiarygodności [np. Madala 2008, s. 154-157]. W celu oszacowania parametrów uogólnionego rozkładu Pareto w metodzie *POT* wykorzystano ważoną prawdopodobieństwami metodę momentów [np. Małecka, Pekasiewicz 2013].

Pośród metod nieparametrycznych wykorzystano metodę symulacji historycznej oraz filtrowanej symulacji historycznej. Metody te nie wymagają żadnych założeń dotyczących rozkładu stopy zwrotu i polegają na wyznaczeniu *VaR* i *ES* jako odpowiednio kwantyla rozkładu empirycznego oraz średniej arytmetycznej rozkładu empirycznego uciętego do przedziału $(-\infty, \hat{VaR}_p(R_t))$. Estymatorem *VaR* w przypadku symulacji historycznej jest

$$\hat{VaR}_p(R_{T+1}) = -R_{(Tp)},$$

gdzie (R_1, R_2, \dots, R_T) jest szeregiem czasowym obserwacji stopy zwrotu a $X_{(k)}$ oznacza statystykę pozycyjną z próby rzędu k zmiennej losowej X . Nieparametrycznym estymatorem *ES* jest statystyka

$$\hat{ES}_p(R_t) = -\frac{\sum_{t=1}^T R_t \mathbf{1}_{\{-R_t \geq \hat{VaR}_p(R_t)\}}}{\sum_{t=1}^T \mathbf{1}_{\{-R_t \geq \hat{VaR}_p(R_t)\}}}$$

zatem

$$\hat{ES}_p(R_t) = -\frac{1}{Tp} \sum_{t=1}^T R_t \mathbf{1}_{\{-R_t \geq \hat{VaR}_p(R_t)\}} = -\frac{1}{Tp} \sum_{t=1}^{Tp} R_{(t)}.$$

W metodzie symulacji filtrowanej wyznaczenie statystyk z próby, stanowiących estymatory *VaR* i *ES*, dokonywane jest na podstawie szeregu oszacowanych reszt \hat{Z}_t z modelu stopy zwrotu postaci $R_t = \mu_t + \sigma_t Z_t$, $Z_t \sim iid$. Zakłada się ponadto, że wariancja procesu podlega procesowi *GARCH*. Filtracja polega na wykorzystaniu oszacowań $\hat{\mu}_t$ i $\hat{\sigma}_t$ otrzymanych na podstawie odpowiedniego modelu *GARCH* do przeprowadzenia transformacji zmiennej R_t w Z_t

$$\hat{Z}_t = \frac{R_t - (-\hat{\mu}_t)}{\hat{\sigma}_t}, \quad (4.11)$$

dla $-R_t > \hat{VaR}_p(R_t)$. Transformacja ta pozwala otrzymać szereg stanowiący w przybliżeniu zbiór zmiennych niezależnych o jednakowym rozkładzie [Barone-Adesi et al. 1999].

Przy założonym modelu kwantyl zmiennej R_t dany jest wzorem

$$F_{R_t}^{-1}(p)(R_t) = \mu_t + \sigma_t F_{R_t}^{-1}(p)(Z_t),$$

gdzie $F_{R_t}^{-1}(p)(Z_t)$ jest kwantylem rzędu p zmiennej Z_t , natomiast wartość oczekiwana

$$E(R_t | R_t < F_{R_t}^{-1}(p)(R_t)) = \mu_t + \sigma_t E(Z_t | Z_t < F_{R_t}^{-1}(p)(Z_t)).$$

Stąd

$$\begin{aligned} VaR_{t,p} &= \hat{\mu}_t + \hat{\sigma}_t \hat{F}_{R_t}^{-1}(p)(\hat{Z}_t), \\ \hat{E}S_{t,p} &= \hat{\mu}_t + \hat{\sigma}_t \hat{E}(\hat{Z}_t | \hat{Z}_t < \hat{F}_{R_t}^{-1}(p)(\hat{Z}_t)), \end{aligned}$$

gdzie $\hat{F}_{R_t}^{-1}(p)(\hat{Z}_t)$ i $\hat{E}(\hat{Z}_t | \hat{Z}_t < \hat{F}_{R_t}^{-1}(p)(\hat{Z}_t))$ otrzymano klasyczną metodą symulacji historycznej.

W niniejszym badaniu, dokonując filtracji, zakładano kolejno, że wariancja procesu R_t podlega procesowi $GARCH(1, 1)$ oraz $GJR - GARCH(1, 1)$, uwzględniającemu dodatkowo zjawisko asymetrii zmienności rozumiane jako różne reakcje zmienności na wartość stopy zwrotu z okresu poprzedniego w przypadku stóp ujemnych i dodatnich. W modelach $GARCH(1, 1)$ i $GJR - GARCH(1, 1)$ przyjęto specyfikację wariancji daną odpowiednio wzorami (4.7) i (4.8).

Ponieważ w celu otrzymania $\hat{\mu}_t$ i $\hat{\sigma}_t$ posłużono się parametrycznymi modelami klasy $GARCH$, cała procedura wyznaczania estymatorów filtrowanej symulacji historycznej miała charakter semi-parametryczny. W obydwóch metodach symulacji historycznej przyjęto $T = 250$.

Poziom tolerancji przy szacowaniu VaR ustalono na 5%, natomiast w przypadku weryfikacji hipotezy dotyczącej wielokrotnych poziomów VaR uwzględniono dodatkowo VaR na poziomie tolerancji 1%.

Wymienione wyżej metody estymacji modeli ryzyka zostały poddane ocenie z wykorzystaniem rozważanych w niniejszej pracy metod wnioskowania statystycznego. W prezentowanej analizie empirycznej wykorzystano wnioski płynące z poprzednich rozdziałów pracy, dotyczące wyboru optymalnych testów statystycznych do weryfikacji modeli opartych na VaR i ES . Z każdej grupy testów VaR wybrano reprezentujące tę grupę procedury, kierując się wynikami badania symulacyjnego, dotyczącego rozmiaru i mocy testów. Wyodrębniono testy o rozmiarze najbardziej zbliżonym do rozmiaru nominalnego a następnie spośród nich wybrano procedury, które w danej grupie, w świetle wyników badania symulacyjnego, charakteryzowały się najwyższą mocą.

W przypadku testów udziału przekroczeń VaR , jako testy o rozmiarze bliskim nominalnemu i wysokiej mocy, wykorzystano test dwumianowy T_1 oraz test oparty na zbieżności rozkładu dwumianowego do rozkładu normalnego Z_{uc} . Spośród testów opartych na statystyce χ^2 wybrano test zbudowany na zasadzie mnożników Lagrange'a LM_{uc} , zarówno ze względu na wskazania dotyczące rozmiaru jak i mocy, pokazujące jego przewagę nad testami W_{uc} i LR_{uc} . Test LR_{TUFF} został odrzucony ze względu na niską moc.

W grupie testów autokorelacji VaR , opartych na procesie Markowa, wybrano statystyki testowe LB – opartą na współczynnikach autokorelacji – oraz DQ – opartą na regresji kwantylowej. Odrzucono popularnie stosowany w analizach ryzyka test LR_{ind} , który, jak pokazały wyniki badania symulacyjnego, charakteryzował się największym błędem pierwszego rodzaju. Spośród dwóch testów opartych na podejściu regresyjnym, odrzucono test DQ_{logist} , zarówno ze względu na błąd pierwszego jak i drugiego rodzaju. Ze względu na niskie oszacowania mocy przy danych generowanych przez proces $GARCH$ odrzucono również testy $Q_{ind}^{\chi^2}$ Pearsona i WW Walda-Wolfowitza.

Spośród metod wnioskowania opartych na procesie odległości między przekroczeniami, na podstawie wyników dotyczących rozmiaru, odrzucone zostały testy LR_{Haas} i $EACD$. Pierwszy z nich cechował się wyższym błędem pierwszego rodzaju niż inne testy danej grupy. W przypadku statystyki $EACD$ zaobserwowano wyraźną rozbieżność między empirycznym a teoretycznym rozmiarem statystyki testowej, polegającą na przesunięciu w kierunku dodatnim. Rozważając dwa testy oparte na sprawdzaniu braku pamięci LR_{Weib} i LR_{gamma} , zarówno ze względu na oszacowania rozmiaru jak i mocy, wybrano test wykorzystujący rozkład Weibulla.

W grupie testów zgodności otrzymano zbliżone oszacowania rozmiaru, natomiast największymi oszacowaniami mocy charakteryzowały się testy LR_B i D_{AD} , co zdecydowało o wykorzystaniu ich w badaniu empirycznym. Spośród testów opartych na funkcji gęstości spektralnej najmniejszą rozbieżność w stosunku do rozmiaru nominalnego zaobserwowano dla statystyki Andersona-Darlinga, zatem w analizie empirycznej wykorzystano test SD_{AD} , oparty na tej statystyce.

W celu weryfikacji modeli opartych na mierze VaR z jednoczesnym uwzględnieniem wielu poziomów tolerancji wykorzystano zaproponowany w niniejszej pracy test LR_H . Test ten został wybrany ze względu na przewagę nad testem LB_H , stwierdzoną zarówno na podstawie badania rozmiaru jak i mocy.

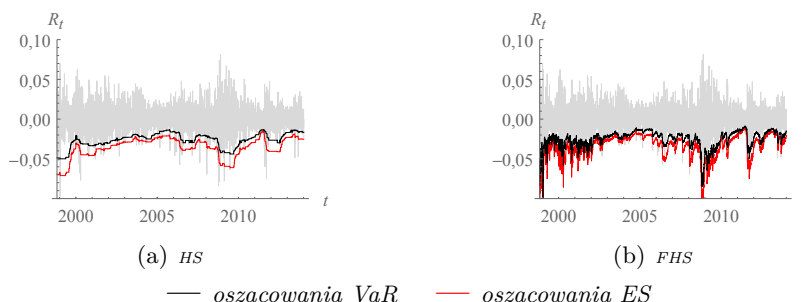
W przypadku miary ES , do oceny modeli wykorzystano testy parametryczne LR_B^{ES} i S , uwzględniając odmienny charakter hipotez weryfikowa-

nych za pomocą tych procedur (zgodność całych ogonów rozkładów lub tylko oszacowań ES). Spośród nieparametrycznych testów ES dokonano wyboru procedury opartej na statystyce U , uwzględniając wskazania badania symulacyjnego dotyczące zarówno rozmiaru jak i mocy. Odrzucono testy regresyjne ze względu na niedokładność (test F_{Ch}) lub niską moc (test F_{Ch}^*).

4.2. Wyniki badania empirycznego dla rynku finansowego

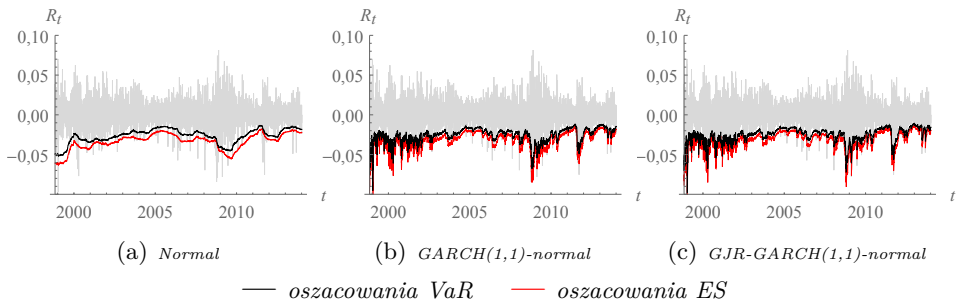
4.2.1. Ocena modeli VaR i ES dla indeksu WIG20

W pierwszym etapie związanym z oceną ryzyka rozważanych indeksów przeprowadzono estymację miar ryzyka VaR i ES z wykorzystaniem modeli symulacji historycznej i filtrowanej symulacji historycznej dla indeksu WIG20 (rys. 4.8 a, b). Analiza graficzna wyników pokazała niewielką dynamikę ocen miar ryzyka w modelu symulacji historycznej w porównaniu do rynkowej dynamiki stóp zwrotu. Wykres przedstawiający oszacowania VaR i ES w modelu symulacji historycznej obrazuje długie okresy powtarzających się ocen ryzyka, po których następują skokowe zmiany. Wykres ten odzwierciedla konstrukcję modelu symulacji historycznej, w którym zmiany ocen szacowanych miar następują jedynie w przypadku zdarzenia polegającego na włączeniu do próby, przy przesunięciu okna estymacji, nowej obserwacji należącej do ogona rozkładu. Zastosowanie filtracji polegającej na wyznaczeniu ocen VaR i ES dla reszt z modelu $GARCH(1,1)$ a następnie przeprowadzeniu transformacji tych ocen w oceny VaR i ES dla wyjściowego szeregu stopy zwrotu pozwoliło zwiększyć dynamikę oszacowań ryzyka.

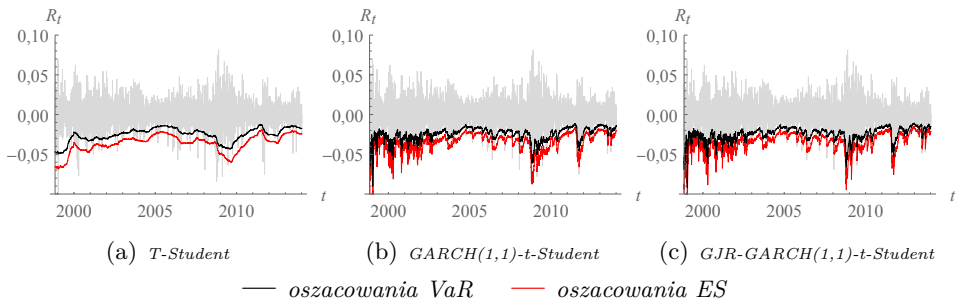


Rys. 4.8: Oszacowania VaR i ES w modelach symulacji historycznej dla indeksu WIG20, 1.01.1995–31.06.2013

Kolejne etapy badania wykorzystywały parametryczne estymatory miar VaR i ES oraz założenia kolejno o rozkładzie normalnym (rys. 4.9 a, b, c) i t-Studenta (rys. 4.10 a, b, c). Na każdym etapie wyznaczano miary ryzyka w modelach: homoskedastycznym, $GARCH(1,1)$ i $GJR-GARCH(1,1)$. Analiza graficzna pokazała różnice w dynamice ocen miar VaR i ES , wynikającą z zastosowania modeli homoskedastycznych lub heteroskedastycznych. W porównaniu do wcześniej stosowanych modeli symulacji historycznej wszystkie rodzaje modeli parametrycznych charakteryzowały się szybszym dostosowywaniem poziomu szacowanego ryzyka do zmienności rynkowej. Przyjęcie założenia o rozkładzie t-Studenta skutkowało otrzymaniem większych (co do wartości bezwzględnej) ocen miar ryzyka, zwłaszcza ES , niż w przypadku rozkładu normalnego.



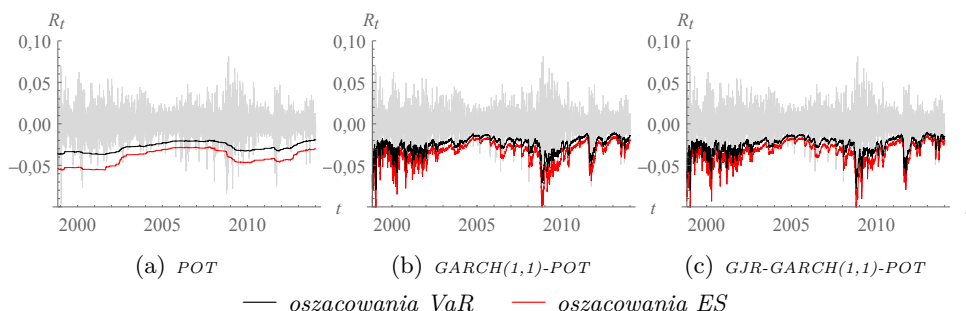
Rys. 4.9: Oszacowania VaR i ES w modelach z rozkładem normalnym dla indeksu WIG20, 1.01.1995–31.06.2013



Rys. 4.10: Oszacowania VaR i ES w modelach z rozkładem t-Studenta dla indeksu WIG20, 1.01.1995–31.06.2013

Oszacowane w kolejnym kroku wartości VaR i ES , z wykorzystaniem teorii wartości ekstremalnych i modelu POT , charakteryzowały się bardzo wolnym tempem dostosowania do zmienności rynkowych stóp zwrotu (rys. 4.11 a). Analiza graficzna potwierdziła zatem trudności w szybkim wykorzystaniu no-

wych informacji rynkowych przy zastosowaniu modelu, w którym oszacowanie parametrów uogólnionego rozkładu Pareto wymaga dużej próby wyjściowej, z której następnie wykorzystuje się jedynie obserwacje skrajne. Połączenie metody *POT* z modelami zmienności *GARCH(1,1)* i *GJR-GARCH(1,1)* pozwoliło uzyskać znaczną poprawę dynamiki ocen miary ryzyka (rys. 4.11 b, c).



Rys. 4.11: Oszacowania *VaR* i *ES* w modelach *POT* dla indeksu WIG20, 1.01.1995–31.06.2013

Finalną fazą badania było przeprowadzenie oceny modeli ryzyka dla indeksu WIG20 przy wykorzystaniu wybranych testów *VaR* i *ES*. Pierwsza grupa przeprowadzonych testów – T_1 , Z_{uc} , LM_{uc} – miała na celu sprawdzenie zgodności udziału przekroczeń *VaR* z 5% poziomem tolerancji. W przypadku niemalże wszystkich rozważanych modeli zastosowane statystyki testowe wskazały na brak podstaw do odrzucenia hipotezy zerowej, mówiącej że udział przekroczeń *VaR* wyniósł 5% (tab. 4.2). Wyniki badania pokazały zatem, że zarówno modele homoskedastyczne jak i heteroskedastyczne były skuteczne w kontekście szacowania poziomu wartości zagrożonej *VaR* w tym sensie, że liczba obserwacji przekraczających *VaR* zgodna była z założoną. Wniosek powyższy był niezależny od wyboru rozkładu przyjętego w modelach parametrycznych.

Drugą grupą zastosowanych testów były testy autokorelacji – DQ , LB , LR_{Weib} i SD_{AD} , które, badając warunkowy rozkład wyjątków *VaR*, sprawdzały czy zastosowanie weryfikowanych modeli nie powoduje występowania serii następujących po sobie przekroczeń. Analiza *p*-wartości pokazała, że najbardziej skuteczne w eliminowaniu autokorelacji przekroczeń *VaR* były modele z asymetrią zmienności: *GJR-GARCH(1,1)-normal*, *GJR-GARCH(1,1)-t-Student* i *GJR-GARCH(1,1)-POT* (tab. 4.3). W szczególności model z rozkładem *t*-Studenta nie został odrzucony przez żaden z rozważanych testów.

Tab. 4.2: P -wartości w testach rozkładu bezwarunkowego przekroczeń VaR dla indeksu WIG20, 1.01.1995–31.06.2013

Model	T_1	Z_{uc}	LM_{uc}
<i>HS</i>	0,43	0,41	0,82
<i>FHS</i>	0,29	0,31	0,61
<i>Normal</i>	0,95	0,95	0,09
<i>GARCH(1, 1)-normal</i>	0,94	0,95	0,11
<i>GJR-GARCH(1, 1)-normal</i>	0,80	0,81	0,38
<i>T-Student</i>	0,54	0,56	0,88
<i>GARCH(1, 1)-t-Student</i>	0,60	0,62	0,77
<i>GJR-GARCH(1, 1)-t-Student</i>	0,27	0,28	0,56
<i>POT</i>	0,99	0,99	0,02*
<i>GARCH(1, 1)-POT</i>	0,54	0,56	0,88
<i>GJR-GARCH(1, 1)-POT</i>	0,57	0,59	0,82

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Na podstawie statystyki wielowymiarowego testu LR_H , pozwalającej na bardziej szczegółową ocenę modeli VaR , z jednoczesnym uwzględnieniem poziomów tolerancji 1% i 5%, nie stwierdzono podstaw do odrzucenia modeli $GJR-GARCH(1,1)$ -normal i $GJR-GARCH(1,1)$ -t-Student.

Tab. 4.3: P -wartości w testach rozkładu warunkowego przekroczeń VaR dla indeksu WIG20, 1.01.1995–31.06.2013

Model	DQ	LB	LR_{Weib}	SD_{AD}	LR_H
<i>HS</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>FHS</i>	0,00*	0,00*	0,03*	0,00*	0,00*
<i>Normal</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>GARCH(1, 1)-normal</i>	0,02*	0,01*	0,08	0,00*	0,00*
<i>GJR-GARCH(1, 1)-normal</i>	0,18	0,10	0,33	0,01*	0,12
<i>T-Student</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>GARCH(1, 1)-t-Student</i>	0,02*	0,00*	0,04*	0,00*	0,00*
<i>GJR-GARCH(1, 1)-t-Student</i>	0,37	0,27	0,28	0,08	0,14
<i>POT</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>GARCH(1, 1)-POT</i>	0,01*	0,00*	0,01*	0,00*	0,04*
<i>GJR-GARCH(1, 1)-POT</i>	0,10	0,04*	0,20	0,00*	0,01*

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Kolejne zastosowane testy – D_{AD} i LR_B – pozwoliły rozszerzyć analizę dotyczącą zgodności kwantyli rozkładu na ocenę zgodności całych prognozowanych rozkładów stopy zwrotu z rozkładami teoretycznymi (tab. 4.4). Zgodność rozkładów nie została odrzucona jedynie w przypadku modelu homoskedastycznego z rozkładem t-Studenta, jednak przyjęcie nieco wyższego,

8% poziomu istotności skutkowałyby decyzją odwrotną. Uwzględniając dodatkowo odrzucenie pozostałych modeli z rozkładem t-Studenta, wyniki badania sugerowały raczej konieczność odrzucenia wszystkich rozważanych modeli w kontekście prognozowania rozkładu stopy zwrotu. Z drugiej strony wskazania testu LR_B , który sprawdza zgodność rozkładów jedynie poprzez dwa pierwsze momenty, nie wskazały podstaw do odrzucenia pozostałych modeli z rozkładem t-Studenta oraz dwóch modeli heteroskedastycznych z rozkładem normalnym, co pokazuje skuteczność tych modeli w przewidywaniu momentów rozkładu stopy zwrotu.

Tab. 4.4: P -wartości w testach VaR zgodności rozkładów dla indeksu WIG20, 1.01.1995–31.06.2013

Model	D_{AD}	LR_B
<i>Normal</i>	0,00*	0,00*
<i>GARCH(1, 1)-normal</i>	0,00*	0,35
<i>GJR-GARCH(1, 1)-normal</i>	0,00*	0,34
<i>T-Student</i>	0,08	0,51
<i>GARCH(1, 1)-t-Student</i>	0,02*	0,48
<i>GJR-GARCH(1, 1)-t-Student</i>	0,02*	0,48

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Ostatnia zastosowana grupa testów odnosiła się do otrzymanych oszacowań miary ES . Statystyka testu LB_B^{ES} nakazała odrzucenie wszystkich modeli, co pokazało niezgodność prognozowanych ogonów rozkładów z odpowiadającymi im obszarami rozważanych rozkładów teoretycznych (tab. 4.5).

Tab. 4.5: P -wartości w testach ES dla indeksu WIG20, 1.01.1995–31.06.2013

Model	LR_B^{ES}	S	U
<i>Normal</i>	0,00*	0,00*	0,00*
<i>GARCH(1, 1)-normal</i>	0,00*	0,00*	0,00*
<i>GJR-GARCH(1, 1)-normal</i>	0,00*	0,00*	0,00*
<i>T-Student</i>	0,00*	0,01*	0,13
<i>GARCH(1, 1)-t-Student</i>	0,00*	0,01*	0,54
<i>GJR-GARCH(1, 1)-t-Student</i>	0,00*	1,00	0,07
<i>POT</i>	0,00*	–	0,65
<i>GARCH(1, 1)-POT</i>	0,00*	–	0,65
<i>GJR-GARCH(1, 1)-POT</i>	0,00*	–	0,86

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

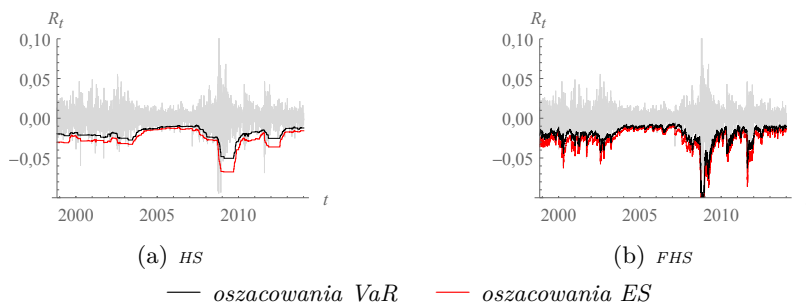
Wskazania statystyki S , sprawdzającej hipotezę, że rzeczywista wartość ES jest większa (co do wartości bezwzględnej) od prognozowanej, uzasadniły stosowanie rozkładu t -Studenta w szacowaniu ryzyka. P -wartość na poziomie 1 wskazała brak podstaw do odrzucenia modelu $GJR-GARCH(1,1)$ - t -Student. Wynik powyższy potwierdzony został za pomocą testu bootstrapowego U , również sprawdzającego hipotezę o wartości ES . Na podstawie testu U stwierdzono dodatkowo brak podstaw do odrzucenia innych modeli z rozkładem t -Studenta oraz modeli wywodzących się z teorii wartości ekstremalnych: POT , $GARCH(1,1)$ - POT i $GJR-GARCH(1,1)$ - POT .

Podsumowując wnioski płynące z zastosowania wszystkich grup wybranych testów – najlepszym modelem oceniającym ryzyko wynikające ze zmian wartości indeksu WIG20 okazał się model $GJR-GARCH(1,1)$ - t -Student, który jest modelem heteroskedastycznym z asymetrią zmienności, co oznacza, że opisuje on odmienne reakcje rynkowe na rosnące oraz spadające stopy zwrotu. Zgodnie z teorią ekonomii zjawisko asymetrii ma szczególne znaczenie w przypadku nowych rynków kapitałowych, lub rynków o stosunkowo niedługiej tradycji. W świetle przeprowadzonych testów, model $GJR-GARCH(1,1)$ - t -Student był w stanie dobrze szacować ryzyko zarówno mierzone miarą VaR jak i ES . Dzięki zastosowaniu tego modelu liczba przekroczeń VaR była zgodna z założonym poziomem tolerancji, przekroczenia te nie wykazywały tendencji do skupiania się w czasie a wielkość straty w przypadku przekroczenia VaR (t.j. wielkość ES) została poprawnie oszacowana. W odniesieniu do miary ES , poza modelem $GJR-GARCH(1,1)$ - t -Student, dobrze ocenione zostały testy POT oparte na teorii wartości ekstremalnych.

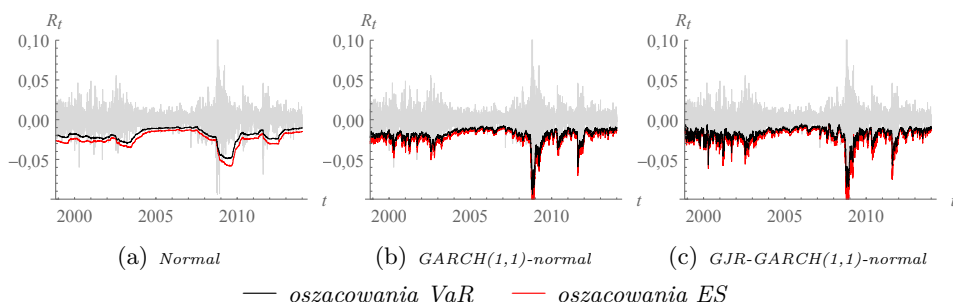
4.2.2. Ocena modeli VaR i ES dla indeksu S&P500

Graficzna analiza ocen miar ryzyka VaR i ES otrzymanych dla indeksu S&P500, podobnie jak dla indeksu WIG20, pokazała najmniejszą dynamikę oszacowań w przypadku modeli HS , $normal$, t -Student i POT , które nie uwzględniały zmian wartości parametrów rozkładu stopy zwrotu w czasie (rys. 4.12, 4.13, 4.14, 4.15). W szczególności uwidocznione to zostało na wykresie obrazującym oszacowania VaR i ES na podstawie modelu POT . Przy wykorzystaniu tego modelu w czasie kryzysu finansowego 2007/2008 nastąpił skokowy wzrost otrzymanych oszacowań poziomu ryzyka, które następnie utrzymywały się długi czas na podobnym poziomie. W przypadku zastosowania modeli ze zmienną wariancją – $GARCH(1,1)$ i $GJR-GARCH(1,1)$ z różnymi rozkładami oraz w połączeniu z metodą symulacji FHS – oszacowania miar ryzyka charakteryzowały się szybkim wzrostem (co do wartości

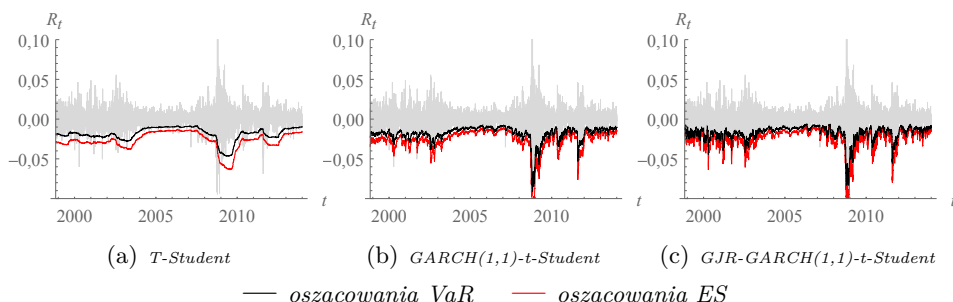
bezwzględnej) w okresach wzmożonej zmienności rynkowej. Dzięki temu oszacowania ryzyka odzwierciedliły gwałtowne zmiany wartości indeksu S&P500 związane z kryzysem finansowym.



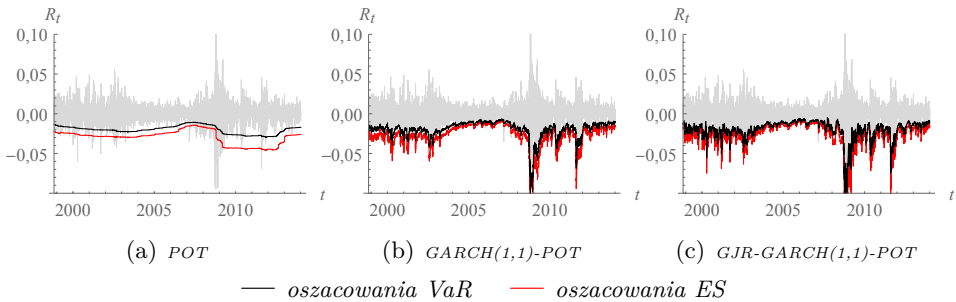
Rys. 4.12: Oszacowania *VaR* i *ES* w modelach symulacji historycznej dla indeksu S&P500, 1.01.1995–31.06.2013



Rys. 4.13: Oszacowania *VaR* i *ES* w modelach z rozkładem normalnym dla indeksu S&P500, 1.01.1995–31.06.2013



Rys. 4.14: Oszacowania *VaR* i *ES* w modelach z rozkładem t-Studenta dla indeksu S&P500, 1.01.1995–31.06.2013



Rys. 4.15: Oszacowania *VaR* i *ES* w modelach *POT* dla indeksu S&P500, 1.01.1995–31.06.2013

Przeprowadzenie testów T_1 , Z_{uc} , LM_{uc} , weryfikujących zgodność udziału przekroczeń *VaR* z przyjętym poziomem 5%, dla indeksu S&P500 pokazało odmienne wnioski niż dla indeksu WIG20 (tab. 4.6). Niemalże wszystkie testy, na poziomie istotności 5% nakazały odrzucenie modeli z rozkładem t-Studenta, ze względu na zbyt duży udział przekroczeń *VaR*, podczas gdy modele te były uznane za dobre w świetle testów dla indeksu WIG20.

Tab. 4.6: *P*-wartości w testach rozkładu bezwarunkowego przekroczeń *VaR* dla indeksu S&P500, 1.01.1995–31.06.2013

Model	T_1	Z_{uc}	LM_{uc}
<i>HS</i>	0,52	0,50	1,00
<i>FHS</i>	0,18	0,19	0,38
<i>Normal</i>	0,06	0,06	0,13
<i>GARCH(1,1)-normal</i>	0,10	0,11	0,22
<i>GJR-GARCH(1,1)-normal</i>	0,00*	0,00*	0,01*
<i>T-Student</i>	0,00*	0,00*	0,00*
<i>GARCH(1,1)-t-Student</i>	0,00*	0,00*	0,00*
<i>GJR-GARCH(1,1)-t-Student</i>	0,00*	0,00*	0,00*
<i>POT</i>	0,01*	0,01*	0,03*
<i>GARCH(1,1)-POT</i>	0,34	0,36	0,72
<i>GJR-GARCH(1,1)-POT</i>	0,51	0,53	0,94

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Porównanie *p*-wartości w grupie testów autokorelacji – *DQ*, *LB*, LR_{Weib} i SD_{AD} – pokazało, że autokorelację przekroczeń *VaR* dla indeksu S&P500 wyeliminowano w przypadku wszystkich modeli heteroskedastycznej zmienności, wykorzystujących proces *GARCH*: *FHS*, *GARCH(1,1)-normal*, *GJR-GARCH(1,1)-normal*, *GARCH(1,1)-t-Student*, *GJR-GARCH(1,1)-t-Student*, *GARCH(1,1)-POT* i *GJR-GARCH(1,1)-POT* (tab. 4.7). Za-

tem w przeciwieństwie do indeksu WIG20, w przypadku S&P500, w celu prawidłowego szacowania VaR nie było niezbędne stosowanie modeli z asymetrią zmienności.

Na podstawie statystyki wielowymiarowego testu LR_H nie stwierdzono podstaw do odrzucenia wymienionych wyżej modeli parametrycznych z heteroskedastyczną zmiennością. Potwierdziło to skuteczność tych modeli zarówno w przypadku szacowania VaR na poziomie tolerancji 5% jak i na poziomie tolerancji 1%.

Tab. 4.7: P -wartości w testach rozkładu warunkowego przekroczeń VaR dla indeksu S&P500, 1.01.1995–31.06.2013

Model	DQ	LB	LR_{Weib}	SD_{AD}	LR_H
<i>HS</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>FHS</i>	0,05*	0,04*	0,94	0,04*	0,02*
<i>Normal</i>	0,00*	0,00*	0,00*	0,00*	0,04*
<i>GARCH(1,1)-normal</i>	0,02*	0,03*	0,93	0,20	0,98
<i>GJR-GARCH(1,1)-normal</i>	0,01*	0,15	0,04*	0,52	0,62
<i>T-Student</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>GARCH(1,1)-t-Student</i>	0,01*	0,14	0,98	0,27	0,95
<i>GJR-GARCH(1,1)-t-Student</i>	0,00*	0,09	0,07	0,12	0,67
<i>POT</i>	0,00*	0,00*	0,00*	0,00*	0,00*
<i>GARCH(1,1)-POT</i>	0,12	0,08	0,75	0,17	0,87
<i>GJR-GARCH(1,1)-POT</i>	0,38	0,25	0,11	0,36	0,94

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Obydwa zastosowane testy zgodności – D_{AD} i LR_B – nakazały odrzucenie wszystkich rozważanych modeli (tab. 4.8). Wyniki testów zgodności świadczą więc o braku możliwości prognozowania rozkładu stopy zwrotu za pomocą zastosowanych modeli.

Tab. 4.8: P -wartości w testach VaR zgodności rozkładów dla indeksu S&P500, 1.01.1995–31.06.2013

Model	D_{AD}	LR_B
<i>Normal</i>	0,00*	0,00*
<i>GARCH(1,1)-normal</i>	0,00*	0,00*
<i>GJR-GARCH(1,1)-normal</i>	0,00*	0,00*
<i>T-Student</i>	0,00*	0,00*
<i>GARCH(1,1)-t-Student</i>	0,00*	0,01*
<i>GJR-GARCH(1,1)-t-Student</i>	0,00*	0,00*

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Ostatnią zastosowaną grupą testów były testy weryfikujące hipotezę dotyczącą otrzymanych oszacowań ES . Statystyka LB_B^{ES} , podobnie jak dla indeksu WIG20, nakazała odrzucenie wszystkich modeli (tab. 4.9). Pokazało to niezgodność prognozowanych ogonów rozkładów z przyjmowanymi rozkładami teoretycznymi. Jedynym modelem dopuszczonym w świetle wyników dwóch testów S i U , sprawdzających hipotezę, że rzeczywista wartość ES jest większa (co do wartości bezwzględnej) od prognozowanej, był model $GARCH(1,1)$ - t -Student. Wyniki testu bootstrapowego U dla modeli opartych na teorii wartości ekstremalnych pokazały dodatkowo brak podstaw do odrzucenia któregośkolwiek z modeli POT .

Tab. 4.9: P -wartości w testach ES dla indeksu S&P500, 1.01.1995–31.06.2013

Model	LR_B^{ES}	S	U
<i>Normal</i>	0,00*	0,00*	0,00*
<i>GARCH(1,1)-normal</i>	0,00*	0,00*	0,00*
<i>GJR-GARCH(1,1)-normal</i>	0,00*	0,35	0,00*
<i>T-Student</i>	0,00*	0,05*	0,58
<i>GARCH(1,1)-t-Student</i>	0,01*	0,00*	0,06
<i>GJR-GARCH(1,1)-t-Student</i>	0,00*	1,00	0,12
<i>POT</i>	0,00*	–	0,08
<i>GARCH(1,1)-POT</i>	0,00*	–	0,49
<i>GJR-GARCH(1,1)-POT</i>	0,00*	–	0,33

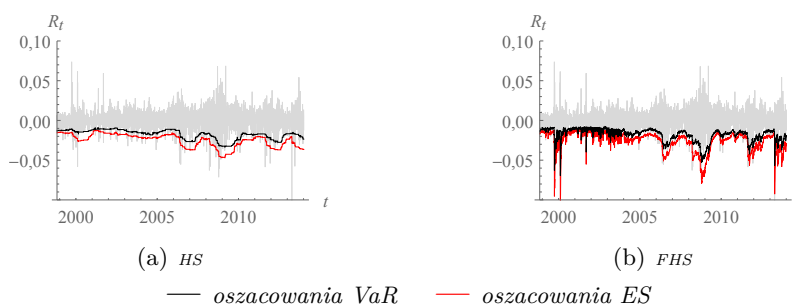
Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Podsumowując wyniki wszystkich testów, dobrze ocenionymi modelami, w odniesieniu do szacowania miary VaR , były testy heteroskedastycznej zmienności, przy czym zastosowanie rozkładu t -Studenta w połączeniu z tymi modelami skutkowało przeszacowaniem liczby przekroczeń VaR . Wniosek ten pokazuje odmiennosc rynków finansowych o różnej długości funkcjonowania, przejawiającą się tym, że w przypadku rynku o krótszej tradycji szczególnie zalecane jest stosowanie rozkładu t -Studenta – w celu uwzględniania obserwacji odstających – oraz modeli z asymetrią zmienności – w celu uwzględniania asymetrycznych reakcji inwestorów w sytuacjach wzrostów i spadków. Nie jest to konieczne by prawidłowo szacować ryzyko na rynku o dłuższej tradycji. Z kolei w kontekście prawidłowego szacowania wielkości straty w przypadku przekroczenia VaR , t.j. prognozowania ES , konsekwentnie najlepsze są modele oparte na teorii wartości ekstremalnych.

4.3. Wyniki badania empirycznego dla rynku towarowego

4.3.1. Ocena modeli VaR i ES dla indeksu *Gold Bullion LBM*

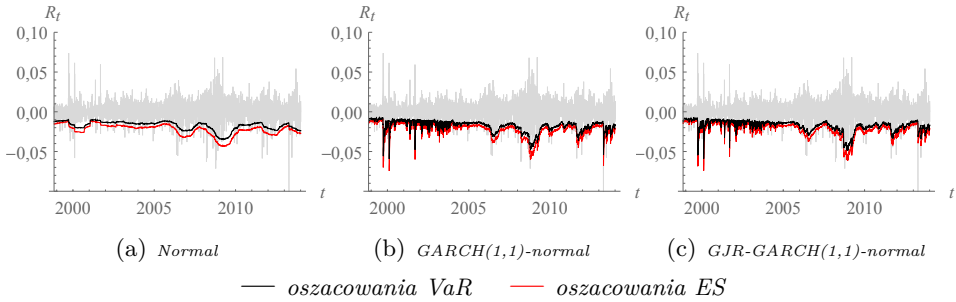
W empirycznym badaniu ryzyka przeprowadzonym na przykładzie indeksu *Gold Bullion LBM*, podobnie jak dla indeksów rynku kapitałowego, w pierwszym etapie przeprowadzono estymację VaR i ES na podstawie modeli symulacji historycznej (rys. 4.16) a następnie na podstawie modeli parametrycznych z rozkładami normalnym (rys. 4.17) i t-Studenta (rys. 4.18) oraz modelu POT wykorzystującego teorię wartości ekstremalnych (rys. 4.19).



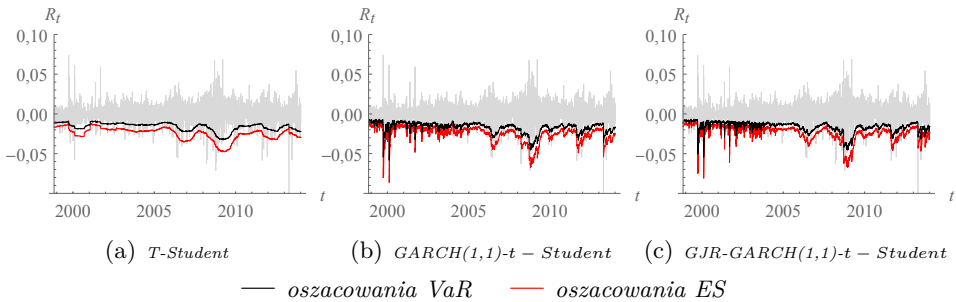
Rys. 4.16: Oszacowania VaR i ES w modelach symulacji historycznej dla indeksu *Gold Bullion LBM*, 1.01.1995–31.06.2013

Graficzna analiza otrzymanych oszacowań pokazała, że dla indeksu obliczającego ceny złota otrzymano niższe (co do wartości bezwzględnej) oceny ryzyka niż dla indeksów WIG20 i S&P500, pochodzących z rynków kapitałowych. Wyniki te odzwierciedliły stosunkowo niską przeciętną zmienność indeksu *Gold Bullion LBM*, świadczącą o mniejszym ryzyku inwestycji.

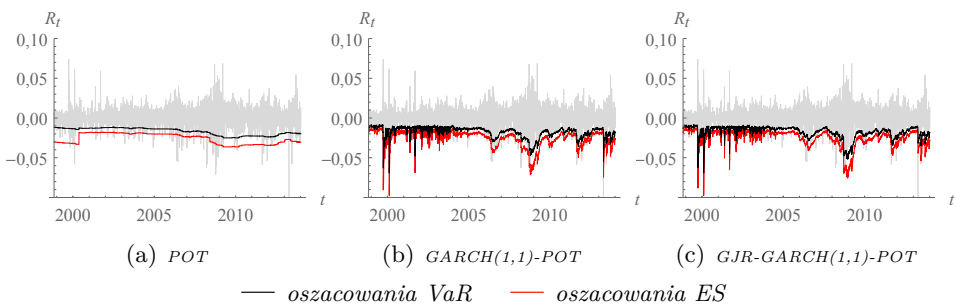
Mniejsza zmienność rozważanego indeksu przełożyła się na wolne zmiany otrzymanych ocen ryzyka w czasie, zwłaszcza przy zastosowaniu modeli zakładających stałość rozkładu stopy zwrotu. Oszacowania otrzymane na podstawie modelu POT dały obraz w postaci krzywej o bardzo niewielkiej zmienności. Dla rynku złota, podobnie jak dla rozważanych wcześniej indeksów rynku kapitałowego, potwierdzona została zwiększona dynamika oszacowań miar VaR i ES otrzymana dzięki zastosowaniu modeli heteroskedastycznych $GARCH(1,1)$ i $GJR-GARCH(1,1)$.



Rys. 4.17: Oszacowania *VaR* i *ES* w modelach z rozkładem normalnym dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013



Rys. 4.18: Oszacowania *VaR* i *ES* w modelach z rozkładem t-Studenta dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013



Rys. 4.19: Oszacowania *VaR* i *ES* w modelach *POT* dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013

Wyniki oceny modeli *VaR* i *ES* za pomocą testów statystycznych T_1 , Z_{uc} i LM_{uc} , sprawdzających udział przekroczeń *VaR*, nakazały odrzucenie

modeli parametrycznych, które nie uwzględniają zmienności wariancji rozkładu stopy zwrotu w czasie (tab. 4.10). Odrzucono zarówno modele z rozkładem normalnym, t-Studenta jak i homoskedastyczne modele *POT*. Wyniki badania pokazały zatem, że w celu prawidłowego szacowania *VaR* konieczne było uwzględnienie warunkowej heteroskedastyczności rozkładu stopy zwrotu. Dodatkowo wskazania wszystkich statystyk testowych, badających udział przekroczeń *VaR*, wskazały na konieczność odrzucenia modelu *GJR-GARCH(1,1)-t-Student*, typowego dla danych pochodzących z rynków kapitałowych.

Tab. 4.10: *P*-wartości w testach rozkładu bezwarunkowego przekroczeń *VaR* dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013

Model	T_1	Z_{uc}	LM_{uc}
<i>HS</i>	0,93	0,93	0,15
<i>FHS</i>	0,40	0,41	0,83
<i>Normal</i>	0,01*	0,01*	0,02*
<i>GARCH(1,1)-normal</i>	0,40	0,41	0,83
<i>GJR-GARCH(1,1)-normal</i>	0,27	0,28	0,56
<i>T-Student</i>	0,00*	0,00*	0,00*
<i>GARCH(1,1)-t-Student</i>	0,32	0,33	0,66
<i>GJR-GARCH(1,1)-t-Student</i>	0,00*	0,00*	0,00*
<i>POT</i>	0,00*	0,00*	0,00*
<i>GARCH(1,1)-POT</i>	0,51	0,53	0,94
<i>GJR-GARCH(1,1)-POT</i>	0,42	0,44	0,89

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

W przypadku indeksu Gold Bullion LBM ocena *p*-wartości w testach autokorelacji, odnoszących się do warunkowego rozkładu wyjątków *VaR*, była dla większości modeli zgodna z wskazaniami, omówionych wyżej, testów udziału przekroczeń (tab. 4.11). Na podstawie wszystkich testów autokorelacji odrzucono modele homoskedastyczne z rozkładami normalnym i t-Studenta oraz homoskedastyczne modele *POT*, natomiast modele heteroskedastyczne prawidłowo oceniały *VaR*, w tym sensie, że przekroczenia *VaR* nie wykazywały tendencji do grupowania się w czasie.

Na podstawie statystyki wielowymiarowego testu LR_H , uwzględniającej jednocześnie oszacowania *VaR* na poziomie tolerancji 1% i 5%, odrzucono wymienione wyżej modele homoskedastyczne oraz dodatkowo model *GJR-GARCH(1,1)-normal*.

Podobnie jak w przypadku poprzednich indeksów, obydwa zastosowane testy zgodności D_{AD} i LR_B , pokazały niezgodność prognozowanego rozkładu

Tab. 4.11: P -wartości w testach rozkładu warunkowego przekroczeń VaR dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013

Model	DQ	LB	LR_{Weib}	SD_{AD}	LR_H
HS	0,00*	0,00*	0,01*	0,00*	0,02*
FHS	0,75	0,66	0,00*	0,19	0,09
$Normal$	0,00*	0,00*	0,02*	0,03*	0,13
$GARCH(1,1)$ -normal	0,95	0,90	0,01*	0,89	0,99
GJR - $GARCH(1,1)$ -normal	0,62	0,56	0,56	0,12	0,03*
T -Student	0,00*	0,00*	0,00*	0,00*	0,00*
$GARCH(1,1)$ - t -Student	0,90	0,85	0,00*	0,50	0,60
GJR - $GARCH(1,1)$ - t -Student	0,03*	0,51	0,52	0,57	0,91
POT	0,00*	0,00*	0,00*	0,00*	0,02*
$GARCH(1,1)$ - POT	0,90	0,82	0,01*	0,90	0,90
GJR - $GARCH(1,1)$ - POT	0,31	0,22	0,97	0,05*	0,27

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

stopy zwrotu z rozważanymi rozkładami teoretycznymi (tab. 4.12). Jedynym wyjątkiem było wskazanie statystyki testu LR_B dla modelu homoskedastycznego z rozkładem t -Studenta, jednak niska p -wartość oraz wskazania drugiego testu zgodności, sugerowały ostrożność przy akceptacji tego modelu w kontekście prognozowania stopy zwrotu.

Tab. 4.12: P -wartości w testach VaR zgodności rozkładów dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013

Model	D_{AD}	LR_B
$Normal$	0,00*	0,00*
$GARCH(1,1)$ -normal	0,00*	0,00*
GJR - $GARCH(1,1)$ -normal	0,00*	0,00*
T -Student	0,00*	0,11
$GARCH(1,1)$ - t -Student	0,00*	0,00*
GJR - $GARCH(1,1)$ - t -Student	0,00*	0,00*

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Analizując wyniki testów ES dla indeksu Gold Bullion LBM, tak samo jak dla obydwóch indeksów rynku kapitałowego, na podstawie statystyki LB_B^{ES} odrzucono wszystkie modele, co pokazało niezgodność prognozowanych ogonów rozkładów z rozważanymi rozkładami teoretycznymi (tab. 4.13).

Wskazania dwóch statystyk S i U , sprawdzających hipotezę, że rzeczywista wartość ES jest większa (co do wartości bezwzględnej) od prognozowanej, uzasadniły stosowanie modeli z rozkładem t -Studenta. Ponadto wyniki te-

stu U pokazały brak podstaw do odrzucenia modeli POT wywodzących się z teorii wartości ekstremalnych.

Tab. 4.13: P -wartości w testach ES dla indeksu Gold Bullion LBM, 1.01.1995–31.06.2013

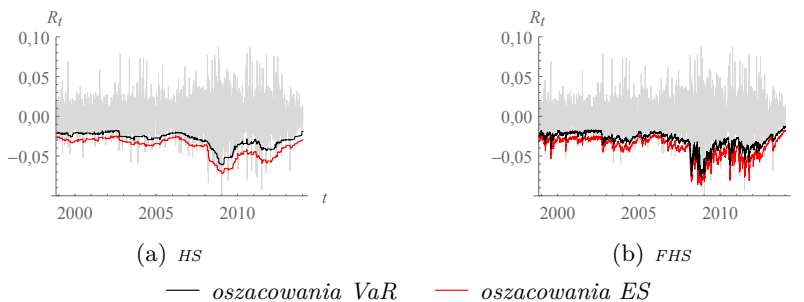
Model	LR_B^{ES}	S	U
<i>Normal</i>	0,00*	0,00*	0,00*
<i>GARCH(1, 1)-normal</i>	0,00*	0,00*	0,00*
<i>GJR-GARCH(1, 1)-normal</i>	0,00*	0,00*	0,00*
<i>T-Student</i>	0,00*	0,02*	0,92
<i>GARCH(1, 1)-t-Student</i>	0,02*	0,04*	0,65
<i>GJR-GARCH(1, 1)-t-Student</i>	0,00*	0,86	0,53
<i>POT</i>	0,00*	–	0,11
<i>GARCH(1, 1)-POT</i>	0,00*	–	0,50
<i>GJR-GARCH(1, 1)-POT</i>	0,00*	–	0,45

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Podsumowując wyniki testów dla indeksu Gold Bullion LBM, w odniesieniu do szacowania wartości VaR dobrymi modelami były wszystkie modele heteroskedastyczne – uwzględniające zmienność rozkładu stopy zwrotu w czasie. W kontekście szacowania ES potwierdzona została przewaga modeli opartych na teorii wartości ekstremalnych oraz modelu $GJR-GARCH(1, 1)-t-Student$, jednak ten ostatni model przeszacowywał udział przekroczeń VaR .

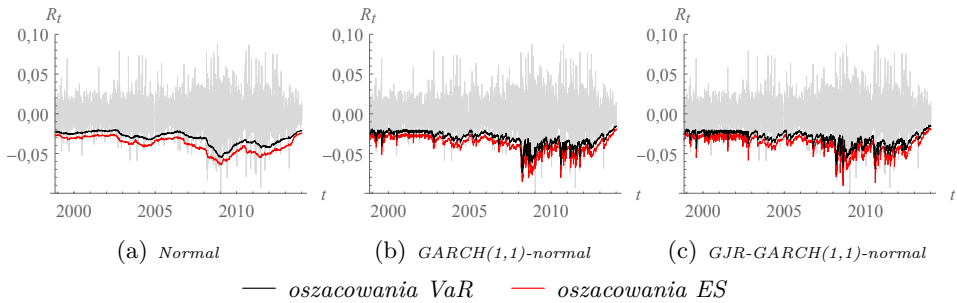
4.3.2. Ocena modeli VaR i ES dla indeksu S&P GSCI Wheat

Oceny miar ryzyka VaR i ES otrzymane dla indeksu S&P GSCI Wheat charakteryzowały się wyższym poziomem (co do wartości bezwzględnej) niż dla innych rozważanych indeksów, co odzwierciedlało dużą zmienność cen na rynku pszenicy (rys. 4.20 – 4.23). W szczególności były one znacznie większe niż dla indeksu cen złota, również pochodzącego z rynku towarowego.

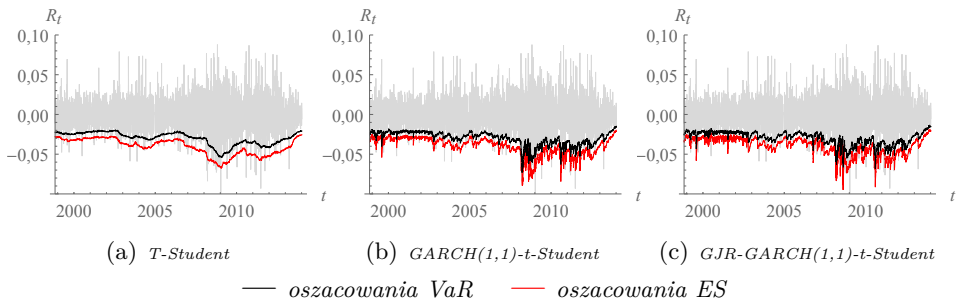


Rys. 4.20: Oszacowania VaR i ES w modelach symulacji historycznej dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013

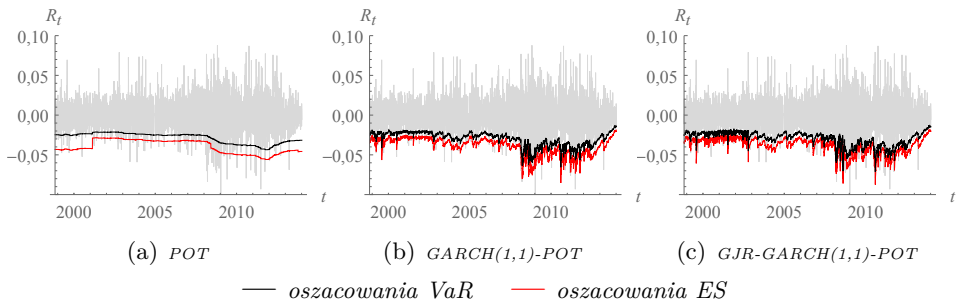
Analiza graficzna, podobnie jak dla pozostałych rozważanych indeksów, pokazała najmniejszą dynamikę oszacowań w przypadku modeli *HS*, *normal*, *t-Student* i *POT*, które nie uwzględniały zmian rozkładu stopy zwrotu w czasie.



Rys. 4.21: Oszacowania *VaR* i *ES* w modelach z rozkładem normalnym dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013



Rys. 4.22: Oszacowania *VaR* i *ES* w modelach z rozkładem t-Studenta dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013



Rys. 4.23: Oszacowania *VaR* i *ES* w modelach *POT* dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013

W ostatnim kroku badania empirycznego wykorzystano wybrane statystyki testowe do przeprowadzenia oceny modeli VaR i ES dla indeksu S&P GSCI Wheat. W przypadku wszystkich rozważanych modeli, każda z zastosowanych statystyk testowych, badających udział przekroczeń $VaR - T_1$, Z_{uc} i LM_{uc} – wskazała na brak podstaw do odrzucenia modelu (tab. 4.14). Wyniki badania pokazały zatem, że zarówno modele homoskedastyczne jak i heteroskedastyczne były skuteczne w kontekście zgodności udziału przekroczeń z założonym poziomem tolerancji. Wniosek powyższy był niezależny od wyboru rozkładu przyjętego w modelach parametrycznych.

Tab. 4.14: P -wartości w testach rozkładu bezwarunkowego przekroczeń VaR dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013

Model	T_1	Z_{uc}	LM_{uc}
<i>HS</i>	0,80	0,79	0,42
<i>FHS</i>	0,54	0,56	0,88
<i>Normal</i>	0,91	0,92	0,16
<i>GARCH(1,1)-normal</i>	0,98	0,98	0,03*
<i>GJR-GARCH(1,1)-normal</i>	0,97	0,97	0,06
<i>T-Student</i>	0,76	0,77	0,46
<i>GARCH(1,1)-t-Student</i>	0,87	0,88	0,24
<i>GJR-GARCH(1,1)-t-Student</i>	0,80	0,81	0,38
<i>POT</i>	0,03*	0,03*	0,06
<i>GARCH(1,1)-POT</i>	0,65	0,67	0,66
<i>GJR-GARCH(1,1)-POT</i>	0,48	0,50	1,00

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Zgodnie z analizą p -wartości w testach autokorelacji – DQ , LB , LR_{Weib} i SD_{AD} – wśród rozważanych modeli większość okazała się skuteczna w eliminowaniu autokorelacji przekroczeń VaR dla indeksu S&P GSCI Wheat (tab. 4.15). Jedynym modelem odrzuconym przez wszystkie statystyki testowe z tej grupy był homoskedastyczny model POT , który charakteryzował się najmniejszą dynamiką dostosowania ocen ryzyka do zmienności stopy zwrotu.

Dzięki zastosowaniu wielowymiarowego testu LR_H , weryfikującego modele VaR jednocześnie dla dwóch poziomów tolerancji, uzyskano dodatkowy wniosek o konieczności odrzucenia modeli nie uwzględniających zmienności rozkładu stopy zwrotu w czasie. Odrzucono modele HS oraz modele parametryczne z rozkładem normalnym i t-Studenta. Ponadto spośród modeli heteroskedastycznych odrzucono $GARCH(1,1)-POT$.

Zgodność prognozowanych rozkładów stóp zwrotu z rozkładem rzeczywistym została potwierdzona, na podstawie obydwóch zastosowanych testów

Tab. 4.15: P -wartości w testach rozkładu warunkowego przekroczeń VaR dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013

Model	DQ	LB	LR_{Weib}	SD_{AD}	LR_H
HS	0,17	0,16	0,47	0,13	0,01*
FHS	0,75	0,64	0,15	0,70	0,98
$Normal$	0,33	0,35	0,70	0,63	0,05*
$GARCH(1,1)$ -normal	0,32	0,71	0,41	0,71	0,97
GJR - $GARCH(1,1)$ -normal	0,48	0,81	0,84	0,72	0,67
T -Student	0,24	0,17	0,57	0,42	0,00*
$GARCH(1,1)$ - t -Student	0,72	0,77	0,56	0,73	0,73
GJR - $GARCH(1,1)$ - t -Student	0,83	0,82	0,87	0,74	0,43
POT	0,00*	0,00*	0,01*	0,00*	0,00*
$GARCH(1,1)$ - POT	0,75	0,64	0,74	0,22	0,00*
GJR - $GARCH(1,1)$ - POT	0,58	0,46	0,96	0,17	0,20

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

zgodności D_{AD} i LR_B w przypadku modeli z rozkładem t -Studenta (tab. 4.16). Na podstawie testu LR_B , w którym weryfikacja zgodności przeprowadzana jest z wykorzystaniem jedynie dwóch pierwszych momentów, nie stwierdzono również podstaw do odrzucenia modeli z rozkładem normalnym.

Tab. 4.16: P -wartości w testach VaR zgodności rozkładów dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013

Model	D_{AD}	LR_B
$Normal$	0,00*	0,84
$GARCH(1,1)$ -normal	0,00*	0,78
GJR - $GARCH(1,1)$ -normal	0,00*	0,92
T -Student	0,17	0,84
$GARCH(1,1)$ - t -Student	0,11	0,80
GJR - $GARCH(1,1)$ - t -Student	0,14	0,90

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Ocena modeli ES , przeprowadzona na podstawie statystyki testu LB_B^{ES} nakazała odrzucenie wszystkich modeli poza modelami z rozkładem t -Studenta (tab. 4.17). Indeks S&P GSCI Wheat był jedynym rozważanym indeksem, dla którego test LB_B^{ES} dopuścił zgodność rozkładu ogona z jednym z przyjętych rozkładów teoretycznych.

Analizując wskazania statystyk S i U , sprawdzających hipotezę, że rzeczywista wartość ES jest większa (co do wartości bezwzględnej) od prognozowanej, stwierdzono, że dobre oszacowania wielkości straty w sytuacjach

ekstremalnych otrzymywano na podstawie modeli z warunkową heteroskedastycznością z rozkładem t-Studenta lub heteroskedastycznych modeli *POT*.

Tab. 4.17: *P*-wartości w testach *ES* dla indeksu S&P GSCI Wheat, 1.01.1995–31.06.2013

Model	LR_B^{ES}	<i>S</i>	<i>U</i>
<i>Normal</i>	0,00*	0,00*	0,15
<i>GARCH(1,1)-normal</i>	0,00*	0,00*	0,01*
<i>GJR-GARCH(1,1)-normal</i>	0,00*	0,01*	0,01*
<i>T-Student</i>	0,65	0,43	0,06
<i>GARCH(1,1)-t-Student</i>	0,29	0,95	0,42
<i>GJR-GARCH(1,1)-t-Student</i>	0,22	0,99	0,51
<i>POT</i>	0,00*	–	0,01*
<i>GARCH(1,1)-POT</i>	0,00*	–	0,45
<i>GJR-GARCH(1,1)-POT</i>	0,00*	–	0,27

Uwaga: Wyniki istotne na poziomie 5% oznaczono *.

Podsumowując, testy oceniające otrzymane oszacowania miar ryzyka dla indeksu S&P GSCI Wheat dopuściły możliwość zastosowania najszerszej klasy modeli. Jedynym modelem odrzuconym przez większość testów odnoszących się do wartości miary *VaR* był homoskedastyczny model *POT*. Indeks S&P GSCI Wheat był jedynym z rozważanych indeksów, dla którego stwierdzono zgodność prognozowanego rozkładu stopy zwrotu z jednym z rozkładów teoretycznych i był to rozkład t-Studenta. W kontekście szacowania *ES* dopuszczono zarówno heteroskedastyczne modele wywodzące się z teorii wartości ekstremalnych jak i heteroskedastyczne modele z rozkładem t-Studenta.

Podsumowanie

Praca poświęcona została statystycznej ocenie modeli ryzyka rynkowego. Nawiązując do aktualnych kierunków badań, skupiono się na weryfikacji hipotez związanych z modelami zagrożenia VaR i ES . Jako podstawa badania określone zostały ekonomiczne kryteria oceny modeli ryzyka. Kryteria te zostały sformułowane jako hipotezy statystyczne, stanowiące punkt wyjścia do konstrukcji testów. Przedstawiono istniejące w literaturze propozycje testów, wskazano możliwe modyfikacje oraz zaproponowano rozszerzenie analizy modeli opartych na miarach VaR i ES z wykorzystaniem nowych testów. Uwzględniając najnowsze zalecenia międzynarodowego nadzoru bankowego (tzw. *Basel III*) analizę testów VaR uzupełniono propozycją testu wielowymiarowego, co pozwoliło na jednoczesną weryfikację modelu dla wielu poziomów tolerancji z zachowaniem kontroli błędu pierwszego rodzaju.

Kluczową częścią pracy było przeprowadzenie badania rozmiaru i mocy testów stosowanych w ocenie modeli ryzyka rynkowego. Badanie oparto na obszernym eksperymencie symulacyjnym, który został zaprojektowany w taki sposób, by zapewnić reprezentację zjawiska gromadzenia zmienności, uznawanego za główny problem związany z kontrolą ryzyka rynkowego. Poza procesem *GARCH*, typowym dla danych finansowych, wykorzystano w nim procesy *BGAR*, *BGMA* i proces Markowa. Praca stanowi zatem rozwinięcie dotychczasowych analiz mocy testów VaR i ES , obejmujące szerszą klasę procesów ekonomicznych.

Wyniki uzyskane na podstawie przeprowadzonych badań pozwoliły wyłonić testy statystyczne gwarantujące najwyższy stopień kontroli nad błędem pierwszego rodzaju oraz jednocześnie charakteryzujące się najwyższą skutecznością wykrywania błędnych modeli ryzyka. Ponadto wskazano przypadki, w których, dzięki pewnym modyfikacjom można uzyskać dokładniejsze zachowanie rozmiaru lub poprawę mocy.

Przeprowadzone porównanie rozmiaru i mocy testów *VaR* i *ES*, w szczególności zaproponowanych modyfikacji, pozwoliło na wybór testów, które zastosowano następnie w badaniu empirycznym. W oparciu o rzeczywiste szeregi czasowe, pochodzące z rynku finansowego i towarowego, przedstawiono przykład zastosowania wybranych metod statystycznych do oceny konkurencyjnych modeli ryzyka. Przeprowadzono obszerną procedurę testową, która podkreśliła między innymi specyfikę polskiego rynku kapitałowego, polegającą na silnie asymetrycznych zachowaniach inwestorów, oraz wykazała potrzebę stosowania odmiennych metod analizy ryzyka dla rynków o różnej tradycji giełdowej, w szczególności rynku polskiego oraz amerykańskiego.

Wyniki uzyskane w pracy mogą znaleźć zastosowanie w praktyce gospodarczej. Wybór odpowiednich modeli ryzyka oraz ich właściwa ocena ma wartość praktyczną dla przedsiębiorstw operujących na rynkach finansowych lub towarowych. Przedstawione badania i ich wyniki mogą być również wykorzystane przez organy krajowe i międzynarodowe sprawujące nadzór nad systemem finansowym, w tym systemem bankowym, do opracowania wytycznych dotyczących weryfikacji modeli ryzyka w instytucjach finansowych. Z punktu widzenia praktyki gospodarczej, przedstawiona analiza porównawcza umożliwia przedsiębiorstwom wprowadzenie, uzasadnionych na gruncie teorii statystyki, naukowych metod oceny wewnętrznych systemów zarządzania ryzykiem. W szczególności analiza eksperymentów symulacyjnych uwzględniających różne struktury autokorelacji może być pomocna w wyborze modelu odpowiedniego dla procesów ekonomicznych generujących ryzyko przy określonym rodzaju działalności gospodarczej. Część empiryczna niniejszej pracy może stanowić również wskazanie dotyczące wyboru modeli ryzyka dla poszczególnych rynków finansowych i towarowych.

Z doświadczeń uzyskanych na podstawie analiz prezentowanych w pracy wynika, że rozważane problemy badawcze wymagają kontynuacji. W szczególności na uwagę zasługuje testowanie modeli opartych na mierze *ES* i następujące zagadnienia z tym związane:

- budowa parametrycznych testów *ES*, w szczególności konstrukcja testu opartego na technice punktu siodłowego z wykorzystaniem funkcji generujących momenty dla rozkładów innych niż rozkład normalny,
- konstrukcja wielowymiarowych testów *ES*, pozwalających oceniać modele oparte na tej mierze z jednoczesnym uwzględnieniem wielu poziomów tolerancji.

Bibliografia

- Acerbi C., Spectral measures of risk: a coherent representation of subjective risk aversion, *Journal of Banking & Finance* 26(7), [2002], s. 1505–1518.
- Acerbi C., Nordio C., Sirtori C., *Expected Shortfall as a Tool for Financial Risk Management*, Working paper, Italian Association for Financial Risk Management, [2001].
- Acerbi C., Tasche D., On the coherence of Expected Shortfall, *Journal of Banking & Finance* 26(7), [2002], s. 1487–1503.
- Al-Subaihi A.A., Simulating Correlated Multivariate Pseudorandom Numbers, *Journal of Statistical Software* 9(4), [2004], s. 1–20.
- Anderson R.L., Distribution of the serial correlation coefficient, *The Annals of Mathematical Statistics* 13, [1942], s. 1–13.
- Anderson T.W., Walker A.M., On the Asymptotic Distribution of the Autocorrelations of a Sample From a Linear Stochastic Process, *Annals of Mathematical Statistics* 35, [1964], s. 1296–1303.
- Artzner P., Delbaen F., Eber J.M., Heath D., Coherent Measures of Risk, *Mathematical Finance* 9(3), [1999], s. 203–228.
- Artzner P., Delbaen F., Eber J.M., Heath D., Thinking coherently, *Risk* 10, [1997], s. 68–71.
- Bałamut T., Metody estymacji Value at Risk, *Materiały i studia NBP* 147, Warszawa, [2002].
- Barbanchan J.S.F., Ornelas J.R.H., Farias A.R., Goodness-of-fit Tests Focus on Value-at-Risk Estimation, *Brazilian Review of Econometrics* 26(2), [2006], s. 309–326.
- Barone-Adesi G., Giannopoulos K., Vosper L., VaR without correlations for non-linear portfolios. *Journal of Futures Markets* 19, [1999], s. 583–602.
- Basel Committee on Banking Supervision, *Amendment to the capital accord to incorporate market risks* [1996, online], dostępne w Internecie: <http://www.bis.org/publ/bcbs24.pdf> [dostęp: 20 grudnia 2013].
- Basel Committee on Banking Supervision, *Amendment to the capital accord to incorporate market risks* [2005, online], dostępne w Internecie: www.bis.org/publ/bcbs119.pdf [dostęp: 20 grudnia 2013].
- Basel Committee on Banking Supervision, *Basel II: International Convergence of Capital Measurement and Capital Standards: a Revised Framework* [2005, online], dostępne w Internecie: <http://www.bis.org/publ/bcbs118.htm> [dostęp: 20 grudnia 2013].
- Basel Committee on Banking Supervision, *Fundamental Review of the Trading Book* [2012, online], dostępne w Internecie: <http://www.bis.org/publ/bcbs219.pdf> [dostęp: 6 czerwca 2016].
- Basel Committee on Banking Supervision, *Fundamental Review of the Trading Book: a Revised Market Risk Framework* [2013, online], dostępne w Internecie: <http://www.bis.org/publ/bcbs265.pdf> [dostęp: 6 czerwca 2016].
- Basel Committee on Banking Supervision, *Fundamental Review of the Trading Book: Outstanding Issues* [2014, online], dostępne w Internecie: <http://www.bis.org/bcbs/publ/d305.pdf> [dostęp: 6 czerwca 2016].

- Basel Committee on Banking Supervision, *Fundamental review of the trading book - interim impact analysis* [2015, online], dostępne w Internecie: <http://www.bis.org/bcbs/publ/d346.pdf> [dostęp: 6 czerwca 2016].
- Basel Committee on Banking Supervision, *Minimum capital requirements for market risk* [2016, online], dostępne w Internecie: <http://www.bis.org/bcbs/publ/d352.pdf> [dostęp: 6 czerwca 2016].
- Berkowitz J., Testing Density Forecasts with Applications to Risk Management, *Journal of Business and Economic Statistics* 19, [2001], s. 465–474.
- Berkowitz J., Christoffersen P., Pelletier D., Evaluating Value-at-Risk Models with Desk-Level Data, *Management Science* 12(57), [2011], s. 2213–2227.
- Best P., *Wartość narażona na ryzyko*, Dom wydawniczy ABC, Kraków, [2000].
- Bollerslev T., Generalised Autoregressive Heteroscedasticity, *Journal of Econometrics* 31, [1986], s. 307–327.
- Box G.E.P., Pierce D.A., Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, *Journal of the American Statistical Association* 65, [1970], s. 1509–1526.
- Butler R.W., Wood A.T.A., Saddlepoint Approximation for Moment Generating Functions of Truncated Random Variables, *The Annals of Statistics* 32(6), [2004], s. 2712–2730.
- Campbell S.D., A Review of Backtesting and Backtesting Procedures, *Finance and Economics Discussion Series Divisions of Research and Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.*, [2005].
- Carver L., Mooted VAR substitute cannot be backtested, says top quant, *Risk* [2013, online], dostępne w Internecie: <http://www.risk.net/risk-magazine/news/2253463/mooted-var-substitute-cannot-be-backtested-says-top-quant> [dostęp: 1 lipca 2016].
- Carver L., Backtesting expected shortfall: mission possible? *Risk* [2014, online], dostępne w Internecie: <http://www.risk.net/risk-magazine/feature/2375204/back-testing-expected-shortfall-mission-possible> [dostęp: 1 lipca 2016].
- Chernick M.R., *Bootstrap Methods: a Guide for Practitioners and Researchers*, John Wiley & Sons, Hoboken, New Jersey, [2008].
- Christoffersen P.F., *Elements of Financial Risk Management*, Elsevier Inc., Oxford, [2012].
- Christoffersen P.F., Evaluating Interval Forecasts, *International Economic Review* 39, [1998], s. 841–862.
- Christoffersen P.F., Pelletier D., Backtesting Value-at-Risk: a Duration-Based Approach, *Journal of Financial Econometrics* 1(2), [2004], s. 84–108.
- Crnkovic C., Drachman J., *A Universal Tool to Discriminate Among Risk Measurement Techniques*, Working paper, J. P. Morgan, [1995].
- Crnkovic C., Drachman J., Quality Control, *Risk* 9, [1996], s. 139–143.
- Delbaen F., *Coherent risk measures on general probability spaces*, Working paper, ETH Zurich, [1998].
- Deo R.S., Spectral tests of the martingale hypothesis under conditional heteroscedasticity, *Journal of Econometrics* 99, [2000], s. 291–315.

- Diebold F.X., Gunther T.A., Tay A.S., Evaluating Density Forecasts, *International Economic Review* 39(4), [1998], s. 863–883.
- Domański Cz. (red.), *Nieklasyczne metody oceny efektywności i ryzyka. Otwarte fundusze emerytalne*, PWE, Warszawa, [2011].
- Domański Cz., *Testy statystyczne*, PWE, Warszawa, [1990].
- Domański Cz., Pruska K., *Nieklasyczne metody statystyczne*, PWE, Warszawa, [2000].
- Domański Cz., Pekasiewicz D., Baszczyńska A., Witaszczyk A., *Testy statystyczne w procesie podejmowania decyzji*, Wyd. Uniwersytetu Łódzkiego, Łódź, [2014].
- Dufour J.M., Monte Carlo Tests with Nuisance Parameters: a General Approach to Finite-Sample Inference and Nonstandard Asymptotics, *Journal of Econometrics* 133(2), [2006], s. 443–477.
- Durlauf S.N., Spectral Based Testing of the Martingale Hypothesis, *Journal of Econometrics* 50(3), [1991], s. 355–376.
- Dwass M., Modified randomization tests for nonparametric hypotheses, *Annals of Mathematical Statistics* 28(1), [1957], s. 181–187.
- Efron B., Tibshirani R.J., *An Introduction to the Bootstrap*, Chapman & Hall, New York, [1993].
- Embrechts P., Kaufmann R., Patie P., Strategic long-term financial risks: Single risk factors, *Computational Optimization and Applications* 32(1–2), [2005], s. 61–90.
- Embrechts P., Kluppelberg C., Mikosch T., *Modelling Extremal Events*, Springer, New York, [1997].
- Embrechts P., McNeil A.J., Straumann D., Correlation: pitfalls and alternatives, *Risk Magazine* 12(5), [1999], s. 69–71.
- Engle R. F., Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation, *Econometrica* 50, [1982], s. 987–1008.
- Engle R.F., Wald, Likelihood Ratio and Lagrange Multiplier Tests in Econometrics w: Z. Griliches, M.D. Introlligator (red.) *Handbook of Econometrics* 2, North-Holland, Amsterdam, [1984].
- Engle R.F., Manganelli A., CAViAR: Conditional Autoregressive Value-at-Risk by Regression Quantiles, *Journal of Business and Economic Statistics* 22, [2004], s. 367–381.
- Engle R.F., Patton J., What good is a volatility model? *Quantitative Finance* 1(2), [2001], s. 237–245.
- Engle R.F., Russel J.R., Autoregressive Conditional Duration: a New Model for Irregularly Spaced Transaction Data, *Econometrica* 66(5), [1998], s. 1127–62.
- Fiszeder P., *Modele klasy GARCH w empirycznych badaniach finansowych*, UMK, Toruń, [2009].
- Follmer H., Schied A., Convex Measures of Risk and Trading Constrains, *Finance and Stochastics* 6, [2002], s. 429–447.
- Franses P. H., Dijk D., *Non-Linear Time Series Models in Empirical Finance*, 6th ed., Cambridge University Press, Cambridge, [2008].
- Greene W.H., *Econometric Analysis*, Prentice Hall, New Jersey, [2011].

- Haas M., *New methods in backtesting*, Working Paper, Financial Engineering Research Center Caesar, Bonn, [2001].
- Hamidi B., Hurlin Ch., Kouontchou P., Maillet B., *Towards a Well-diversified Risk Measure: a DARE Approach*, [2010, online], dostępne w Internecie: www.researchgate.net/publication/228712607_Towards_a_Well-diversified_Risk_Measure_A_DARE_Approach [dostęp: 29 stycznia 2013].
- Hull J., *Risk Management and Financial Institutions*, John Wiley & Sons, Hoboken, [2015].
- Hull J., White A., Hull and White on the pros and cons of expected shortfall, *Risk*, [2014, online], dostępne w Internecie: <http://www.risk.net/risk-magazine/opinion/2375185/hull-and-white-on-the-pros-and-cons-of-expected-shortfall> [dostęp: 1 lipca 2016].
- Hurlin Ch., Tokpavi S., Backtesting value-at-risk accuracy: a simple new test, *The Journal of Risk* 9(2), [2007], s. 19–37.
- Jajuga K., Miary ryzyka rynkowego – część trzecia, *Rynek Terminowy* 8, [2000], s. 112–117.
- Jajuga K., Value-at-Risk, *Rynek Terminowy* 9, [2000], s. 18–20.
- Jajuga K., Jajuga T., *Inwestycje: instrumenty finansowe, aktywa niefinansowe, ryzyko finansowe, inżynieria finansowa*, PWN, Warszawa, [2011].
- Jędrzejczak A., Estimation of Standard Errors of Selected Income Concentration Measures on the Basis of Polish HBS, *International Advances in Economic Research* 18(3), [2012], s. 287–297.
- Jorion P., *Value at Risk. The New Benchmark for Managing Financial Risk*, McGraw-Hill, Singapore, [2007].
- JP Morgan, *Riskmetrics*, Technical document, Morgan Guaranty Trust Company, New York, [1994].
- Kjellson B., *Forecasting Expected Shortfall. An Extreme Value Approach*, Centre for Mathematical Sciences, Lund University, Lund, [2013].
- Kończak G., Wykorzystanie metod Monte Carlo w badaniach statystycznych, w: Z. Zieliński (red.), *Rola informatyki w naukach ekonomicznych i społecznych. Innowacje i implikacje interdyscyplinarne*, wyd. Wyższej Szkoły Handlowej, Kielce, [2011].
- Krzyżko M., *Podstawy wielowymiarowego wnioskowania statystycznego*, Wydawnictwo Naukowe UAM, Poznań, [2009].
- Kupiec P., Techniques for Verifying the Accuracy of Risk Measurement Models, *Journal of Derivatives* 3, [1995], s. 73–84.
- Kusuoka S., On law invariant coherent risk measures, *Advances in Mathematical Economics* 3, [2001], s. 83–95.
- Lewis P.A.W., McKenzie E., Hugus D.K., Gamma processes, *Communications in Statistics. Stochastic Models* 5(1), [1989], s. 1–30.
- Li W.K., McLeod A.I., Distribution of the Residual Autocorrelations in Multivariate ARMA Time Series Models, *Journal of the Royal Statistical Society serie B* 43, [1981], s. 231–239.
- Ljung G.M., Box G.E.P., On a Measure of Lack of Fit in Time Series Models, *Biometrika* 65, [1978], s. 297–303.

- Lopez J., Methods for Evaluating Value-at-Risk Estimates, *FRBSF Economic Review* 2, [1999], s. 3–17.
- Lopez J.A., *Regulatory Evaluation of Value-at-Risk Models*, Working Paper, Wharton Financial Institutions Center, [1996].
- Lugannani R., Rice S.O., Saddlepoint approximation for the distribution of the sum of independent random variables, *Advanced Applied Probability* 12, [1980], s. 475–490.
- Maddala G.S., *Ekonometria*, PWN, Warszawa, [2008].
- Małecka M., Prognozowanie zmienności indeksów giełdowych przy wykorzystaniu modelu klasy GARCH, *Ekonomista* 6, [2011], s. 843–860.
- Małecka M., Experimental design in evaluating VaR forecasts, *Acta Universitatis Lodzianensis. Folia Oeconomica* 286, [2013], s. 277–290.
- Małecka M., Ocena efektywności prognoz VaR w kontekście wybranych okresów historii notowań WIG20 w: J. Lisowski, P. Manikowski (red.) *Problemy współczesnego rynku ubezpieczeń*, wyd. Uniwersytetu Ekonomicznego w Poznaniu, Poznań, [2014].
- Małecka M., Spectral density tests in VaR failure correlation analysis, *Research Papers of Wrocław University of Economics* 381, [2015], s. 235–249.
- Małecka M., Pekasiewicz D., Application of probability weighted method of moments to estimation of financial return distribution tail w: H Vojáčková (red.) *Proceedings of the 31st International Conference Mathematical Methods in Economics 2013*, College of Polytechnics, Jihlava, [2013].
- McNeil A.J., Frey R., Embrechts P., *Quantitative Risk Management*, Princeton University Press, New Jersey, [2005].
- McNeil A.J., Frey R., Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach, *Journal of Empirical Finance* 7(3–4), [2000], s. 271–300.
- Mentel G., *Value at Risk w warunkach polskiego rynku kapitałowego*, CeDuWu.pl Wydawnictwa fachowe, Warszawa, [2011].
- Neyman J., Pearson E.S., On the use of interpretation of certain test criteria for purpose of statistical inference, *Biometrika* 20, [1928], s. 175–240.
- Osińska M., *Ekonometria finansowa*, Polskie Wydawnictwo Ekonomiczne, Warszawa, [2006].
- Pekasiewicz D., Bootstrapowa weryfikacja hipotez o wartości oczekiwanej populacji o rozkładzie asymetrycznym, *Acta Universitatis Lodzianensis. Folia Oeconomica* 271, [2012], s. 151–159.
- Pflug G., Some remarks on the value-at-risk and the conditional value-at-risk w: S. Uryasev (red.), *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, Dordrecht, [2000].
- Piontek K., *Analiza mocy wybranych testów metod szacowania Value-at-Risk. Badania symulacyjne*, The 7th Professor Aleksander Zelias International Conference on Modelling and Forecasting of Socio-Economic Phenomena, Zakopane, [2013].
- Pipień M., *Wnioskowanie bayesowskie w ekonometrii finansowej*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków, [2006].

- Rao C.R., Large Sample Tests of Statistical Hypotheses Concerning Several Parameters with Application to Problems of Estimation, *Proceedings of the Cambridge Philosophical Society* 44, [1948], s. 50–57.
- Rencher A.C., *Methods of Multivariate Analysis*, John Wiley & Sons, New York, [2002].
- Riehl H., *Zarządzanie ryzykiem na rynku pieniężnym, walutowym i instrumentów pochodnych*, Warszawski Instytut Bankowości, Warszawa, [2001].
- Righi M.B., Ceretta P.S., Individual and flexible expected shortfall backtesting, *Journal of Risk Model Validation* 7(3), [2013], s. 3–20.
- Rockafellar R.T., Uryasev S., Optimization of conditional value-at-risk, *Journal of Risk* 2, [2000], s. 21–41.
- Rockafellar R.T., Uryasev S., Conditional value-at-risk for general loss distributions, *Journal of Banking & Finance* 26(7), [2002], s. 1443–1471.
- Rockafellar R.T., Uryasev S., Zabarankin M., Generalized deviations in risk analysis, *Finance and Stochastics* 10(1), [2006], s. 51–74.
- Rosenblatt M., Remarks on a multivariate transformation, *Annals of Mathematical Statistics* 23, [1952], s. 470–472.
- Silvey S.D., The Lagrange multiplier test, *Annals of Mathematical Statistics* 30, [1959], s. 389–407.
- Skoglund J., Chen W., *References in Financial Risk Management: Applications in Market, Credit, Asset and Liability Management and Firmwide Risk*, John Wiley & Sons, Hoboken, [2015].
- Szego G. (red.), *Risk Measures for the 21st Century*, John Wiley & Sons, New York, [2004].
- Szego G., Measures of risk, *Journal of Banking & Finance* 26(12), [2002], s. 1253–1272.
- Tarczyński W., Mojsiewicz M., *Zarządzanie ryzykiem. Podstawowe zagadnienia*, PWE, Warszawa, [2001].
- Tasche D., Expected shortfall and beyond, *Journal of Banking & Finance* 26(7), [2002], s. 1519–1533.
- Trzpiot G., *Wielowymiarowe metody statystyczne w analizie ryzyka inwestycyjnego*, PWE, Warszawa, [2010].
- Utkin J., *Statyczne miary ryzyka i straty w skończonych modelach struktury terminowej*, Oficyna Wydawnicza SGH, Warszawa, [2010].
- Wald A., Tests of statistical hypothesis concerning several parameters when the number of observations is large, *Transaction of American Mathematical Society* 54, [1943], s. 426–482.
- Wald A., Wolfowitz J., On a Test Whether Two Samples are from the Same Population, *Annals of Mathematical Statistics* 11(2), [1940], s. 147–162.
- Wang C.J., *Risk Measures and Capital Allocation*, North Carolina State University, Raleigh, North Carolina, [2010].
- Wong W., Backtesting Trading Risk of Commercial Banks using Expected Shortfall, *Journal of Banking & Finance* 32(7), [2008], s. 1404–1415.