



Agnieszka Lach

Modyfikacje testów zgodności w analizie
rozkładów stóp zwrotu z akcji

Modified goodness-of-fit tests in the analysis
of stock return distributions

Praca doktorska

Promotor: dr hab. Krzysztof Echaust, prof. UEP

Pracę przyjęto dnia:

Podpis Promotora

Poznań 2023

Spis treści

Wstęp.....	4
1. Testy zgodności i ich modyfikacje	13
1.1. Pojęcia podstawowe	13
1.2. Wybrane elementy teorii testowania hipotez statystycznych	18
1.3. Testy zgodności oparte na porównaniu dystrybuanty	23
1.3.1. Test zgodności Kołmogorowa.....	23
1.3.2. Statystyka Kuipera	24
1.3.3. Test Craméra-von Misesa	25
1.3.4. Test Andersona-Darlinga	26
1.4. Modyfikacje testów zgodności opartych na porównaniu dystrybuanty	27
1.4.1. Modyfikacje statystyki Andersona-Darlinga.....	28
1.4.2. Testy zgodności z lewostronnie uciętym rozkładem ciągłym	29
1.5. Bootstrapowe testy zgodności.....	32
1.6. Moc testów zgodności opartych na porównaniu dystrybuanty	34
2. Wybrane rozkłady stóp zwrotu	37
2.1. Wybrane jednowymiarowe bezwarunkowe rozkłady stóp zwrotu.....	37
2.1.1. Uzasadnienie wyboru rozkładów jednowymiarowych.....	37
2.1.2. Rozkład normalny (N)	39
2.1.3. Uogólniony rozkład hiperboliczny (GH).....	41
2.1.3.1. Rozkład hiperboliczny (Hyp).....	44
2.1.3.2. Normalny odwrotny rozkład gaussowski (NIG).....	46
2.1.4. Skośny uogólniony rozkład błędu ($SGED$).....	48
2.1.5. Skośny rozkład t -Studenta (ST).....	49
2.2. Kopule	52
2.2.1. Uzasadnienie doboru kopuli i wybrane elementy teorii	52
2.2.2. Kopule eliptyczne.....	54
2.2.2.1. Rozkłady eliptyczne	54
2.2.2.2. Kopula Gaussa	55
2.2.2.3. Kopula t -Studenta	56
2.2.3. Kopule archimedesowe	57
2.2.3.1. Wybrane elementy teorii kopul archimedesowych.....	57
2.2.3.2. Rodzina Franka	58

2.2.3.3.	Rodzina Claytona	59
2.2.3.4.	Rodzina Gumbela	61
2.2.3.5.	Rodzina Joe	62
3.	Testy zgodności z uciętymi rozkładami ciągłymi	64
3.1.	Statystyki dla obustronnie uciętych rozkładów ciągłych	64
3.2.	Testy zgodności z obustronnie uciętym rozkładem ciągłym	67
3.3.	Ocena jakości testów zgodności z obustronnie uciętym rozkładem ciągłym	70
3.4.	Charakterystyka danych	74
3.5.	Oszacowania rozmiarów testów zgodności	76
3.6.	Oszacowania mocy testów zgodności	81
4.	Badania empiryczne	102
4.1.	Opis badań	102
4.2.	Wpływ wyboru testu zgodności na oszacowania wartości zagrożonej na przykładzie rozkładów jednowymiarowych	104
4.3.	Wpływ wyboru testu zgodności na oszacowania wartości zagrożonej na przykładzie rozkładów dwuwymiarowych	112
4.4.	Wpływ wyboru testu zgodności na oszacowania wartości zagrożonej na przykładzie rozkładów jednowymiarowych – wyniki spoza próby	126
4.5.	Wpływ wyboru testu zgodności na jakość portfeli na przykładzie rozkładów dwuwymiarowych	135
	Zakończenie	141
	Bibliografia	145
	Spis tabel	155
	Spis rysunków	156
	Aneks	159

Wstęp

Poprawność doboru modelu jest oceniana w statystyce za pomocą testów statystycznych weryfikujących odpowiednio sformułowane hipotezy. Głównym przedmiotem rozprawy będą testy zgodności, których zadaniem jest „weryfikacja hipotezy dotyczącej zgodności (dopasowania) rozkładu zbioru wartości w próbie z rozkładem teoretycznym” (Magiera, 2002, s. 272). Z uwagi na fakt, że praca dotyczy rozkładów stóp zwrotu z instrumentów finansowych, będą to testy zgodności z wybranym rozkładem ciągłym.

W niektórych sytuacjach decydenta lub badacza nie interesuje dopasowanie rozkładu wartości z próby z całym rozkładem teoretycznym, tylko z jego wybranym fragmentem (Qin, 2019, s. 401–402). Funkcję gęstości rozkładu prawdopodobieństwa można podzielić na trzy główne części (Caers i Maes, 1998, s. 2): dolną część, część centralną i górną część. Dolna i górna część funkcji gęstości, czyli ogony rozkładów, są przedmiotami zainteresowania na przykład podczas szacowania wartości zagrożonej (Echaust i Just, 2020b; Jeziorski, 2014; Just, 2012) lub podczas pomiaru ryzyka operacyjnego, gdy wymagane jest określenie rozkładu wartości szkód (Chernobai, Burnecki, Rachev, Trück i Weron, 2006; Chernobai, Menn, Rachev i Trück, 2010; Matthias, Fischer i Jakob, 2016). Część centralna może natomiast przyciągać uwagę w przypadku pomiaru średniej lub mediany (Qin, 2019, s. 401). Istnieją również badania wychodzące poza wyżej wymieniony schemat podziału funkcji gęstości na ogony i część centralną, w których na przykład ogranicza się jedynie wpływ wartości ekstremalnych na modelowane zjawiska (Łuczak i Just, 2020a, 2020b). Badanie dopasowania z wybranymi fragmentami rozkładu prawdopodobieństwa wymaga doboru odpowiednich testów zgodności. W przypadku ogonów rozkładów takim testem może być test Andersona-Darlinga, nadający większą wagę ogonom rozkładu (Anderson i Darling, 1954). W przypadku części centralnej rozkładu, preferowanym testem może być test Kołmogorowa-Smirnowa, charakteryzujący się większą wrażliwością w okolicach mediany (Eicker, 1979, s. 117). Są to jednak testy zgodności weryfikujące dopasowanie z całym rozkładem. Bardziej odpowiednie mogłyby być testy weryfikujące zgodność z rozkładami uciętymi, które są przedmiotem niniejszej pracy. Taka weryfikacja pozwala na lepsze dopasowanie modelu do danych empirycznych w obszarze, którym zainteresowany jest decydent.

Jeśli decydent chce zweryfikować zgodność rozkładu wartości w próbie z zadaniem rozkładem teoretycznym, ale tylko w jego wybranym fragmencie, można rozważyć dwa

podejścia, gdy liczba uciętych danych jest znana i gdy liczba uciętych danych nie jest znana. Wpływa to na sposób zdefiniowania dystrybuant empirycznej i teoretycznej oraz statystyk testowych.

W pierwszym przypadku, gdy liczba uciętych danych jest znana, modelowany jest cały rozkład. Dystrybuanty empiryczna i teoretyczna, opisujące rozkład zmiennej losowej, uwzględniają wszystkie wartości z przestrzeni prób. W tym przypadku można zastosować statystyki dla rozkładów uciętych, gdy liczba uciętych danych jest znana. Takie modyfikacje były już w literaturze rozważane dla statystyki Kołmogorowa-Smirnowa (Barr i Davidson, 1973; Dufour i Maag, 1978) oraz statystyk Craméra-von Misesa, Andersona-Darlinga i Watsona (Pettitt i Stephens, 1976).

W drugim przypadku modelowany jest wybrany fragment rozkładu, a obserwacje przyjmują wartości tylko z pewnego przedziału. Liczba obserwacji z tego przedziału jest znana, natomiast liczby obserwacji z pozostałych przedziałów powstałych na skutek ucięcia nie są znane. W tym podejściu dystrybuanty empiryczna i teoretyczna opisują rozkład ucięty zmiennej losowej tylko w badanym przedziale. Tutaj próg ucięcia powinien być uwzględniony podczas estymacji parametrów rozkładu (Deluca i Corral, 2013; Szkutnik, 2016), gdyż wpływa on na kształt dystrybuanty teoretycznej, co z kolei powinno zostać uwzględnione w statystykach testowych. Chernobai, Rachev i Fabozzi (2015) wprowadzili istotne z punktu widzenia niniejszej pracy modyfikacje statystyk dla rozkładów lewostronnie uciętych, gdy liczba uciętych danych nie jest znana. W szczególności formuły obliczeniowe zostały wyznaczone dla statystyk Kołmogorowa-Smirnowa, Kuipera, Craméra-von Misesa oraz Andersona-Darlinga w wersjach supremum i kwadratowej. Dodatkowo we wspomnianej pracy przedstawiono formuły obliczeniowe dla statystyk Andersona-Darlinga z funkcją wagową $\psi(x) = 1/(1 - x)$, w wersjach supremum i kwadratowej, również dla rozkładów lewostronnie uciętych, gdy liczba uciętych danych nie jest znana. Statystyki te są przeznaczone do badania zgodności rozkładów w górnym ogonie.

Obydwa omówione powyżej podejścia, weryfikujące zgodność rozkładu wartości w próbie z zadaniem rozkładem teoretycznym w wybranym fragmencie, posiadają pewne ograniczenia. W pierwszym podejściu brakuje odpowiednich modyfikacji statystyk przeznaczonych do badania ogonów rozkładów. Główną wadą drugiego podejścia jest natomiast możliwość weryfikacji zgodności rozkładów tylko w ich górnym ogonie. Wspomniane ograniczenia oraz badania przeprowadzone na podstawie zmodyfikowanych statystyk dla rozkładów

lewostronnie uciętych (Echaust i Lach, 2017; Lach, 2017; Lach i Smaga, 2018), były inspiracją do powstania niniejszej pracy doktorskiej. Zamierzeniem autorki stało się stworzenie elastycznego narzędzia przeznaczonego do badania zgodności dowolnego fragmentu rozkładu bezwarunkowego z wybranym rozkładem teoretycznym oraz zbadanie jego walorów aplikacyjnych.

Bezwarunkowe rozkłady stóp zwrotu, chociaż są wypierane przez rozkłady warunkowe, nadal znajdują zastosowanie w finansach. Metody wyznaczania miar ryzyka takich jak wartość zagrożona (*Value at Risk, VaR*) czy warunkowa wartość zagrożona (*Conditional Value at Risk, CVaR*), w których wykorzystuje się rozkłady bezwarunkowe, mogą być wybierane, gdy ryzyko jest szacowane w długim horyzoncie czasowym lub w krótkich okresach czasu na potrzeby bieżącej działalności, w celu uniknięcia częstych zmian poziomów ryzyka (Danielsson i de Vries, 1998). Jest wiele modeli ekonometrycznych, w których rozkłady reszt są modelowane za pomocą rozkładów bezwarunkowych, zwłaszcza rozkładem normalnym oraz rozkładem *t*-Studenta (Boucher, Danielsson, Kouontchou, i Mailet, 2014; Danielsson, James, Valenzuela i Zer, 2016). Szczególnym zastosowaniem bezwarunkowych rozkładów stóp zwrotu jest ich użycie do konstruowania rozkładów wielowymiarowych za pomocą kopul. Kopule rozdzielają strukturę zależności od rozkładów brzegowych. Z jednej strony pozwala to na modelowanie zgodnych zmian cen w okresach wzrostów lub spadków, a z drugiej strony pozwala na kształtowanie rozkładów pojedynczych walorów i uwzględnienie takich cech jak grubość ogonów, skośność czy leptokurtyczność. Utworzone za pomocą kopul rozkłady wielowymiarowe mają wiele walorów aplikacyjnych, są stosowane między innymi w optymalizacji portfela (Lach, 2015; Patton, 2013) czy w zarządzaniu ryzykiem, przy czym wskazuje się, że poziom ryzyka zależy przede wszystkim od wyboru rozkładów brzegowych (Choi i Min, 2013). Z uwagi na tematykę proponowanej pracy doktorskiej, rozważony zostanie w pełni parametryczny model rozkładu wielowymiarowego, w którym określone zostają zarówno parametry gęstości kopuli, jak i parametry gęstości rozkładów brzegowych. Estymacji parametrów modelu w pełni parametrycznego dokonuje się zazwyczaj metodą największej wiarygodności (*maximum likelihood, ML*) lub metodą dwustopniową (*inference for margins, IFM*), przy czym z uwagi na prostsze obliczenia często preferowana jest metoda *IFM*. Obydwie metody nie są odporne na błędny wybór rozkładów brzegowych. Wskazanie niewłaściwych rozkładów brzegowych może mieć negatywny wpływ na oszacowania parametrów kopuli, powoduje również nieprzewidywalne działanie metody *IFM* (Kim, Silvapulle i Silvapulle, 2007).

Celem badawczym proponowanej rozprawy doktorskiej jest **opracowanie modyfikacji wybranych testów zgodności dla uciętych rozkładów ciągłych i ocena ich przydatności w analizie rozkładów stóp zwrotu z akcji.**

Do realizacji celu głównego przyczynią się następujące cele pomocnicze:

1. Modyfikacja statystyk dla rozkładów lewostronnie uciętych, przedstawionych w pracy Chernobai i in. (2015), umożliwiająca badanie rozkładów prawostronnie i obustronnie uciętych, gdy liczba uciętych danych nie jest znana. Wybrano następujące statystyki: Kołmogorowa-Smirnowa, Kuipera, Craméra-von Misesa, Andersona-Darlinga w wersjach supremum i kwadratowej oraz Andersona-Darlinga w wersjach supremum i kwadratowej z funkcją wagową dla ogona dolnego.
2. Opracowanie procedury wnioskowania statystycznego w oparciu o statystyki dla uciętych rozkładów ciągłych, która uwzględnia informację dotyczącą liczby uciętych danych. Cel ten zawiera opracowanie metody losowania danych z rozkładu, zapewniającej stałą liczbę losowanych danych z wybranego fragmentu dystrybuanty.
3. Ocena empirycznych rozmiarów i mocy testów zgodności dla rozkładów uciętych: lewostronnie, przedstawionych w pracy Chernobai i in. (2015); prawostronnie i obustronnie, opracowanych w niniejszej pracy. Jakość testów zostanie oceniona na podstawie wybranych przykładów, dla ogonów rozkładów oraz części dystrybuanty pozostającej po odrzuceniu ogonów.
4. Zastosowanie opracowanych w pracy zmodyfikowanych testów zgodności dla rozkładów lewostronnie i prawostronnie uciętych do weryfikacji zgodności rozkładu empirycznego z rozkładem teoretycznym w ogonach rozkładów stóp zwrotu z akcji. Ocena jakości oszacowań wartości zagrożonej w zależności od wyboru testu zgodności oraz opracowanie autorskiej procedury wyznaczania wartości zagrożonej w horyzoncie jednodniowym wykorzystującej testy zgodności.
5. Zastosowanie opracowanych w pracy zmodyfikowanych testów zgodności dla rozkładów obustronnie uciętych do weryfikacji zgodności rozkładu empirycznego z rozkładem teoretycznym w centralnej części rozkładów stóp zwrotu z akcji. Ocena jakości portfela papierów wartościowych, mierzona wybranymi charakterystykami rozkładu stopy zwrotu z portfela (odchyleniem standardowym, skośnością, kurtozą oraz wskaźnikiem Sharpe'a), w zależności od wyboru testu zgodności.

W rozprawie przyjęto następującą hipotezę główną:

H1. Zastosowanie opracowanych w pracy, zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych, pozwala na wybór nie gorszego modelu pomiaru ryzyka rynkowego niż model pomiaru ryzyka rynkowego wybrany na podstawie testów zgodności dla całych rozkładów.

Hipotezę główną uściślają dwie hipotezy szczegółowe:

H2. Zastosowanie opracowanych w pracy, zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych, w ogonach rozkładów, pozwala na wybór nie gorszego modelu wartości zagrożonej niż model wartości zagrożonej wybrany na podstawie testów zgodności dla całych rozkładów.

H3. W przypadku konstrukcji portfela o zadanej stopie zwrotu, zastosowanie opracowanych w pracy, zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych, w części centralnej rozkładów, pozwala na wybór nie gorszego modelu pomiaru ryzyka portfela papierów wartościowych niż model pomiaru ryzyka portfela papierów wartościowych wybrany na podstawie testów zgodności dla całych rozkładów.

Do zrealizowania pierwszego celu pomocniczego wyprowadzono formuły statystyk dla ciągłych rozkładów prawostronnie i obustronnie uciętych, gdy liczba uciętych danych nie jest znana, co jest autorskim wkładem naukowym w tym obszarze. Formuły wyznaczono dla wszystkich statystyk rozważanych przez Chernobai i in. (2015). Dodatkowo wyprowadzono formuły obliczeniowe dla statystyk Andersona-Darlinga w wersjach supremum i kwadratowej, gdy funkcja wagowa wynosi $\psi(x) = 1/x$, co pozwala weryfikować zgodność w dolnym ogonie rozkładu. Autorka nie znalazła w literaturze prac, w których przedstawiono podobne wyniki. Formuły są uogólnieniem formuł przedstawionych w pracy Chernobai i in. (2015). Gdy dystrybuanta nie jest ucinana, formuły obliczeniowe redukują się do klasycznych statystyk przeznaczonych do badania całych rozkładów. Gdy dystrybuanta jest ucinana tylko z lewej strony, otrzymywane są statystyki dla rozkładów lewostronnie uciętych, takie jak w pracy Chernobai i in. (2015). Gdy dystrybuanta jest ucinana tylko z prawej strony, otrzymywane są statystyki dla rozkładów prawostronnie uciętych. Wreszcie, gdy dystrybuanta jest ucinana z lewej i prawej strony jednocześnie, otrzymywane są statystyki dla rozkładów obustronnie uciętych.

Kolejny autorski wkład, związany z realizacją drugiego celu pomocniczego, to zaproponowanie procedury wnioskowania statystycznego, która uwzględnia informację dotyczącą liczby uciętych danych mimo, że w formułach statystyk nie ma tej informacji. Należy podkreślić, że po zastosowaniu tej procedury nie zostaną otrzymane wartości wspomnianych powyżej statystyk dla rozkładów obustronnie uciętych, gdy liczba uciętych danych jest znana (Barr i Davidson, 1973; Dufour i Maag, 1978; Pettitt i Stephens, 1976). Zasadnicza niezgodność między statystykami jest wyrażona w różnicy pomiędzy wartościami dystrybuanty empirycznej dla dwóch kolejnych punktów. Kolejna odmienność pomiędzy statystykami tkwi w wartościach przyjmowanych przez funkcję wagową. Funkcja wagowa w statystykach Andersona-Darlinga dla rozkładów uciętych, gdy liczba uciętych obserwacji jest znana, przyjmuje wartości dla argumentów będących pewnym podzbiorem przedziału $(0; 1)$. W statystykach zaproponowanych w niniejszej pracy, funkcja wagowa przyjmuje wartości dla argumentów z całego przedziału $(0; 1)$. Zaproponowane statystyki dla rozkładów obustronnie uciętych oraz procedura wnioskowania statystycznego uwzględniająca liczbę uciętych informacji, pozwalają na weryfikację dopasowania dowolnych fragmentów rozkładów bezwarunkowych, gdy liczba uciętych danych jest znana.

Następny wkład, służący realizacji trzeciego celu pomocniczego, to oszacowanie na podstawie wybranych przykładów empirycznych rozmiarów i mocy testów zgodności dla rozkładów uciętych, przedstawionych w pracy Chernobai i in. (2015) oraz zmodyfikowanych w niniejszej pracy. Empiryczne rozmiary i moc testów zgodności dla rozkładów lewostronnie uciętych były już badane przez Pavię (2015), jednak dla innych rozkładów lub innych parametrów rozkładów. W przypadku rozkładów prawostronnie i obustronnie uciętych, przedstawionych w niniejszej pracy, takich badań dotychczas nie przeprowadzono.

Ostatni autorski wkład, powiązany z realizacją czwartego celu pomocniczego, to opracowanie autorskiej procedury wyznaczania wartości zagrożonej w horyzoncie jednodniowym, która wykorzystuje testy dla rozkładów lewostronnie i prawostronnie uciętych do poprawy jakości oszacowań wartości zagrożonej.

Trzeci cel pomocniczy zakłada ocenę jakości testów zgodności, obejmującą ich rozmiar oraz moc. Realizacja tego celu wymaga bardzo czasochłonnych obliczeń, dlatego ocenę przeprowadzono dla czterech wybranych szeregów stóp zwrotu. Przedstawione w dysertacji badania nie są badaniami rynku akcji, ale badaniami nad testami zgodności i dotyczą ich rozmiaru i mocy oraz aplikacyjności w finansach. Wybrane przykłady szeregów czasowych są

obiektem, na których analizy są przeprowadzane. Obroty na akcjach dużych spółek notowanych na *GPW* w Warszawie znacznie przewyższają obroty na akcjach spółek małych i średnich, dlatego do dalszej analizy wybrano rozkłady stóp zwrotu z akcji dużych spółek. Jednak na rozkłady stóp zwrotu z akcji pojedynczych spółek, nawet tych dużych, mają wpływ liczne anomalie występujące na rynku (Zielonka, 2006, s. 35–46). Wybór konkretnej spółki zbyt mocno odzwierciedlałby ryzyko specyficzne, dlatego zdecydowano się wykorzystać indeksy akcji, które reprezentują portfele największych spółek. Jako reprezentację rynku akcji wybrano indeksy *WIG20* i *FTSE100*. Indeks *WIG20* został wybrany ze względu na chęć przeprowadzenia badania na rynku polskim. Wybór drugiego indeksu pozwoli na zbadanie zależności pomiędzy rynkiem krajowym, reprezentantem rynków środkowoeuropejskich, a indeksem reprezentującym rynki Europy Zachodniej.

Rozprawa składa się z czterech rozdziałów. Dwa pierwsze rozdziały stanowią część nieoryginalną, opracowaną na podstawie kwerendy literatury. Dwa kolejne rozdziały zawierają część oryginalną, opisującą opracowane metody oraz wyniki badań.

W rozdziale pierwszym zostały przedstawione wybrane zagadnienia dotyczące testów zgodności z rozkładem ciągłym, niezbędne do przedstawienia wyników badań. W pierwszej kolejności przedstawiono pojęcia podstawowe oraz wybrane elementy teorii dotyczące testowania hipotez statystycznych, w tym dotyczące również testów zgodności. Następnie przedstawiono testy zgodności oparte na porównaniu dystrybuanty, które będą podlegały modyfikacjom w rozdziale trzecim. Kolejne dwa podrozdziały zawierają istotne z punktu widzenia niniejszej pracy modyfikacje standardowych testów zgodności: zmiany funkcji wagowej w testach zgodności Andersona-Darlinga oraz opis testów zgodności dla rozkładów lewostronnie uciętych, opracowanych w pracy Chernobai i in. (2015). Następny podrozdział, w związku z wykorzystaniem metody bootstrapu do aproksymowania rozkładów statystyk testowych, zawiera krótkie omówienie bootstrapowych testów zgodności. Zaproponowane modyfikacje testów zgodności będą analizowane pod kątem ich mocy, stąd w ostatnim podrozdziale zawarto ogólne informacje dotyczące mocy testów zgodności opartych na porównaniu dystrybuanty.

W rozdziale drugim zawarto opis rozkładów prawdopodobieństwa zastosowanych w badaniach. Pierwszy podrozdział poświęcono jednowymiarowym rozkładom prawdopodobieństwa, na podstawie których badano empiryczny rozmiar i moc zmodyfikowanych testów statystycznych. Są to czteroparametryczne bezwarunkowe rozkłady

często stosowane do opisu rozkładów stóp zwrotu: rozkład hiperboliczny (*Hyp*), normalny odwrotny rozkład Gaussowski (*NIG*), skośny uogólniony rozkład błędu (*SGED*) oraz skośny rozkład *t*-Studenta (*ST*). Ze względu na historyczne znaczenie rozkładu normalnego (*N*), rozkład ten również został opisany i uwzględniony w badaniach w celach porównawczych. W drugim podrozdziale zawarto opis kopul, na podstawie których utworzono rozkłady dwuwymiarowe, zastosowane do badań ryzyka rynkowego oraz jakości portfela. Znalazły się tutaj dwie kopule eliptyczne, Gaussa i *t*-Studenta oraz wybrane kopule archimedesowe: Franka, Claytona, Gumbela i Joe.

W rozdziale trzecim przedstawiono oryginalnie opracowane metody służące ocenie dopasowania rozkładów. W pierwszym podrozdziale przedstawiono wzory obliczeniowe wybranych statystyk dla rozkładów obustronnie uciętych. W drugim podrozdziale przedstawiono zastosowane w badaniach empirycznych wnioski statystyczne, które umożliwiają badanie wybranych fragmentów całych rozkładów. W trzecim podrozdziale opisano metody, które zastosowano do oceny jakości zmodyfikowanych testów zgodności, w tym modyfikację sposobu losowania danych z rozkładu. W kolejnym podrozdziale opisano dane, na których oparto badanie empiryczne. Są to jednodniowe, procentowe, logarytmiczne stopy zwrotu z indeksów *WIG20* oraz *FTSE100* z lat 2019-2020, czyli z okresu przedpandemicznego i z okresu pandemii. W ostatnich dwóch podrozdziałach znajdują się, opracowane na podstawie wybranych przykładów, wyniki empirycznych oszacowań rozmiarów i mocy zmodyfikowanych testów zgodności.

W rozdziale czwartym przedstawiono możliwości zastosowania testów zgodności dla rozkładów uciętych w praktyce. W podrozdziałach drugim, trzecim i czwartym zbadano, w jaki sposób wybór testu zgodności wpływa na jakość oszacowań wartości zagrożonych. W podrozdziale czwartym znajduje się również autorska procedura wyznaczania wartości zagrożonej wykorzystująca testy zgodności. W ostatnim, piątym podrozdziale skonstruowano dwuskładnikowe portfele, rozwiązując wybrane zadanie optymalizacyjne. Zbadano, w jaki sposób wybór testu zgodności wpływa na jakość utworzonych portfeli. W podrozdziałach drugim i czwartym badania przeprowadzono na podstawie rozkładów jednowymiarowych, a w podrozdziałach trzecim i piątym na podstawie rozkładów dwuwymiarowych. Rozkłady dwuwymiarowe tworzą poprzez połączenie rozkładów brzegowych kopulami, do utworzenia rozkładów użyto metody w pełni parametrycznej. Badanie empiryczne w podrozdziałach drugim, trzecim i piątym, oparto na danych opisanych w rozdziale trzecim.

Badanie empiryczne w podrozdziale czwartym oparto na jednodniowych, procentowych, logarytmicznych stopach zwrotu z indeksów *WIG20* oraz *FTSE100* z lat 2017-2022.

Rozdziały zawierające wybrane zagadnienia teoretyczne powstały w oparciu o fizyczne zbiory zgromadzone w bibliotekach oraz zasoby elektroniczne zawarte głównie w bazach ScienceDirect i SpringerLink. Zarówno w przypadku zasobów fizycznych, jak i elektronicznych, analizie podlegała literatura polska i anglojęzyczna. Większość źródeł teoretycznych dotyczyła takich dziedzin jak statystyka matematyczna oraz ekonomia i finanse. Dane empiryczne do badań przeprowadzonych w rozdziałach trzecim i czwartym pobrano z serwisu <https://stooq.pl/>.

W przypadku dwóch pierwszych rozdziałów zastosowaną metodą badawczą były studia literaturowe. W dwóch kolejnych rozdziałach wykorzystano głównie metody statystyczne, w tym przede wszystkim wnioskowanie statystyczne, estymację parametrów i opisową analizę danych. Wszystkie obliczenia przeprowadzono w oparciu o autorskie programy napisane w języku R, wykorzystujące ogólnodostępne pakiety napisane dla tego języka.

1. Testy zgodności i ich modyfikacje

1.1. Pojęcia podstawowe

W tym podrozdziale zostaną przedstawione wybrane, podstawowe pojęcia związane z rachunkiem prawdopodobieństwa i statystyką matematyczną, które są niezbędne do zrozumienia pracy lub często pojawiają się w jej dalszej części. Wybór pojęć jest subiektywny i nie jest wyczerpujący. Selekcji dokonano w oparciu o przegląd wybranych pozycji z literatury (Domański, 2014; Domański i Pruska, 2000; Gerstenkorn i Śródka, 1976; Krysicki, 2004, 2012; Krzyśko, 2004; Magiera, 2002; Sobczyk, 2010; Trybuła, 2004). Jednakże, aby ujednoczyć opis, większość definicji związanych z rachunkiem prawdopodobieństwa zaczerpnięto z pracy Domańskiego i Pruski (2000), natomiast większość definicji związanych ze statystyką matematyczną pochodzi z pracy Krzyśko (2004).

W pierwszej części tego podrozdziału zostaną przedstawione podstawowe terminy związane z rachunkiem prawdopodobieństwa. Założono, że pojęcia zdarzenia elementarnego, przestrzeni zdarzeń elementarnych, zdarzenia losowego oraz σ -ciała zdarzeń losowych, są czytelnikowi znane (objaśnienia pojęć można znaleźć np. w pracy Gerstenkorna i Śródki, 1976, s. 57, 60, 73-74). Oznaczmy zdarzenie losowe przez A . Aksjomatyczna definicja prawdopodobieństwa według Kołmogorowa brzmi wówczas następująco:

Definicja 1.1.1 (Domański i Pruska, 2000, s. 16). Niech Ω będzie przestrzenią zdarzeń elementarnych, a \mathcal{F} - σ -ciałem zdarzeń losowych tej przestrzeni. Prawdopodobieństwem nazywamy funkcję P określoną na zbiorze \mathcal{F} taką, że:

1. $\forall_{A \in \mathcal{F}} P(A) \geq 0$,
2. $P(\Omega) = 1$,
3. dla każdego ciągu $(A_n)_{n \in \mathbb{N}}$ takiego, że $A_i \cap A_j = \emptyset$, gdy $i \neq j$, zachodzi $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$.

Trójka (Ω, \mathcal{F}, P) jest nazywana przestrzenią probabilistyczną i stanowi matematyczny opis pewnego doświadczenia losowego. Na tej podstawie można zdefiniować model matematyczny zjawisk obserwowanych w rzeczywistości, z czym związane jest pojęcie zmiennej losowej. Przyjmując, że ω oznacza zdarzenie elementarne, definicje jednowymiarowej i n -wymiarowej zmiennej losowej są następujące:

Definicja 1.1.2 (Domański i Pruska, 2000, s. 16). Niech dana będzie przestrzeń probabilistyczna (Ω, \mathcal{F}, P) . Jednowymiarową zmienną losową X o wartościach rzeczywistych nazywamy funkcją o wartościach rzeczywistych taką, że:

$$\forall_{r \in \mathbb{R}} \{\omega \in \Omega: X(\omega) < r\} \in \mathcal{F}. \quad (1.1.1)$$

Definicja 1.1.3 (Domański i Pruska, 2000, s. 17). Niech dana będzie przestrzeń probabilistyczna (Ω, \mathcal{F}, P) . Zmienną losową n -wymiarową o wartościach z przestrzeni \mathbb{R}^n lub wektorem losowym n -wymiarowym, gdzie $n \in \mathbb{N}$, nazywamy funkcję $\mathbf{X}: \Omega \rightarrow \mathbb{R}^n$ taką, że:

$$\forall_{B \in \mathbb{B}^n} \mathbf{X}^{-1}(B) \in \mathcal{F}, \quad (1.1.2)$$

gdzie \mathbb{B}^n jest σ -ciałem borelowskim przestrzeni \mathbb{R}^n , a $\mathbf{X}^{-1}(B)$ jest przeciwobrazem zbioru B .

Pojęcie σ -ciała borelowskiego przestrzeni \mathbb{R}^n zostało wyjaśnione w pracy Domańskiego i Pruskiej (2000, s. 17). Kolejnym istotnym pojęciem jest rozkład prawdopodobieństwa zmiennej losowej.

Definicja 1.1.4 (Domański i Pruska, 2000, s. 17). Rozkładem prawdopodobieństwa zmiennej losowej n -wymiarowej \mathbf{X} ($n \in \mathbb{N}$) nazywamy funkcję $P_{\mathbf{X}}$ postaci:

$$P_{\mathbf{X}}(B) = P(\mathbf{X}^{-1}(B)), \quad (1.1.3)$$

gdzie $B \in \mathbb{B}^n$, \mathbb{B}^n jest σ -ciałem borelowskim przestrzeni \mathbb{R}^n oraz $\mathbf{X}^{-1}(B)$ oznacza przeciwobraz zbioru B .

Rozkład zmiennej losowej można wyznaczyć poprzez dystrybuantę, która jednoznacznie wyznacza rozkład:

Definicja 1.1.5 (Domański i Pruska, 2000, s. 18). Dystrybuantą n -wymiarowej zmiennej losowej \mathbf{X} nazywamy funkcję postaci:

$$F(\mathbf{x}) = P(X_1 < x_1, \dots, X_n < x_n), \quad (1.1.4)$$

gdzie $\mathbf{X} = (X_1, \dots, X_n)$ oraz $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n, n \in \mathbb{N}$.

Zmienne losowe można podzielić na dyskretne i ciągłe. Stopy zwrotu z instrumentów finansowych są ciągłymi zmiennymi losowymi, dlatego zmienne losowe dyskretne zostaną pominięte w dalszych rozważaniach.

Definicja 1.1.6 (Domański i Pruska, 2000, s. 19). Zmienna losowa $\mathbf{X} = (X_1, \dots, X_n)$ jest typu ciągłego, jeżeli istnieje taka nieujemna funkcja f , że dystrybuanta F tej zmiennej losowej daje się wyrazić wzorem:

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(u_1, \dots, u_n) du_n, \dots, du_1 \quad (1.1.5)$$

dla $(x_1, \dots, x_n) \in \mathbb{R}^n$. Funkcję f nazywamy gęstością lub funkcją gęstości zmiennej losowej \mathbf{X} .

W dalszej części pracy wykorzystane zostaną również pewne charakterystyki liczbowe, które uzupełniają rozkład zmiennej losowej jednowymiarowej. Taką charakterystyką jest na przykład wartość oczekiwana, która w przypadku jednowymiarowej zmiennej losowej ciągłej o gęstości f jest określona wzorem:

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx. \quad (1.1.6)$$

Do określenia kolejnych charakterystyk liczbowych konieczne jest wprowadzenie pojęć momentów zwykłych i momentów centralnych. Momentem zwykłym rzędu $k > 0$ zmiennej losowej X nazywamy liczbę $m_k = E(X^k)$, a momentem centralnym rzędu $k > 0$ zmiennej losowej X nazywamy liczbę $M_k = E(X - E(X))^k$. Wielkościami opisującymi rozkłady zmiennych losowych, obok wartości oczekiwanej $E(X)$, są również: wariancja $Var(X)$, skośność $S(X)$ oraz kurtoza $K(X)$. Pomiedzy charakterystykami liczbowymi oraz momentami zachodzą zależności: $E(X) = m_1$; $Var(X) = M_2$; $S(X) = M_3/(M_2)^{\frac{3}{2}}$; $K(X) = M_4/(M_2)^2$. W pracy będzie również używane pojęcie ekscesu kurtozy, które jest równe $K(X) - 3$.

W drugiej części tego podrozdziału zostaną przedstawione terminy związane ze statystyką matematyczną. Wnioskowanie statystyczne opiera się na danych, które są wynikami losowego eksperymentu. Wyniki te interpretuje się jako wartości pewnej zmiennej losowej lub wektora losowego, a samą zmienną nazywa się obserwowalną zmienną losową. Niech \mathcal{X} oznacza zbiór wartości obserwowalnej zmiennej losowej X (przestrzeń zdarzeń elementarnych), a \mathcal{A} niech będzie wyróżnionym σ -ciałem podzbiorów (zdarzeń) zbioru \mathcal{X} . Rozkład prawdopodobieństwa obserwowalnej zmiennej losowej X nie jest zazwyczaj znany i wiadomo jedynie, że należy do pewnej rodziny rozkładów \mathcal{P} indeksowanej parametrem $\theta \in \Theta$, tzn. $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$. Zbiór Θ parametrów θ nazywa się przestrzenią parametrów. Rodzinę rozkładów

prawdopodobieństwa nazywa się parametryczną, jeśli Θ jest przestrzenią skończenie wymiarową oraz nieparametryczną, gdy Θ jest przestrzenią nieskończenie wymiarową.

Definicja 1.1.7 (Krzyśko, 2004, s. 11). Trójkę $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ nazywamy przestrzenią statystyczną indukowaną przez zmienną losową X lub modelem statystycznym.

Definicja 1.1.8 (Domański i Pruska, 2000, s. 91). Przestrzeń statystyczną $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ nazywamy produktem przestrzeni $(\mathcal{X}_i, \mathcal{A}_i, \mathcal{P}_i)$, $i = 1, 2, \dots, n$, jeżeli $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \dots \times \mathcal{X}_n$, $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_n$ (σ -ciało produktowe) oraz $\mathcal{P} = \{\mathcal{P}_1 \times \mathcal{P}_2 \times \dots \times \mathcal{P}_n\}$.

Jeżeli przestrzeń $(\mathcal{X}_i, \mathcal{A}_i, \mathcal{P}_i)$ jest indukowana przez zmienną losową X_i , to przestrzeń $(\mathcal{X}, \mathcal{A}, \mathcal{P})$ jest indukowana przez wektor n niezależnych zmiennych losowych $\mathbf{X} = (X_1, X_2, \dots, X_n)'$.

Jeśli zmienne losowe X_i , $i = 1, 2, \dots, n$, są niezależnymi kopiami tej samej zmiennej losowej X indukującej przestrzeń $(\mathcal{X}, \mathcal{A}, \mathcal{P})$, to przestrzeń mierzalną indukowaną przez wektor losowy $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ oznacza się przez $(\mathcal{X}, \mathcal{A})^n$, a odpowiednią przestrzeń statystyczną przez $(\mathcal{X}, \mathcal{A}, \mathcal{P})^n$.

Definicja 1.1.9 (Krzyśko, 2004, s. 12). Wektor losowy $\mathbf{X} = (X_1, X_2, \dots, X_n)'$, gdzie X_i dla $i = 1, 2, \dots, n$ są niezależnymi zmiennymi losowymi o jednakowym rozkładzie prawdopodobieństwa P_θ , $\theta \in \Theta$, nazywamy n -elementową próbą z rozkładu P_θ . (...)

Próbę z rozkładu P_θ zdefiniowaną powyżej nazywa się również próbą prostą (Domański i Pruska, 2000, s. 91).

Definicja 1.1.10 (Krzyśko, 2004, s. 12). Przestrzeń mierzalną $(\mathcal{X}, \mathcal{A})^n$, indukowaną przez próbę $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ nazywamy przestrzenią prób.

Zaobserwowane na elementach próby wartości zmiennych losowych X_i , $i = 1, \dots, n$ oznaczono w dalszej części pracy przez $\mathbf{x} = (x_1, x_2, \dots, x_n)'$. Pojęcie zmiennej losowej w rachunku prawdopodobieństwa ma swój odpowiednik w statystyce matematycznej, jest nim pojęcie statystyki (Krzyśko, 2004, s. 13).

Definicja 1.1.11 (Krzyśko, 2004, s. 12). Funkcje $\mathbf{T}(\mathbf{X}) = (T_1(\mathbf{X}), T_2(\mathbf{X}), \dots, T_k(\mathbf{X}))'$ próby $\mathbf{X} = (X_1, X_2, \dots, X_n)'$, mierzalne względem σ -ciała \mathcal{A} , nazywają się statystykami.

Z punktu widzenia niniejszej pracy istotną statystyką jest dystrybuanta empiryczna. Zostanie jej poświęcona pozostała część tego podrozdziału.

Definicja 1.1.12 (Krzyśko, 2004, s.13). Dystrybuantą empiryczną lub dystrybuantą z próby $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ nazywamy funkcję $F_n: \mathbb{R} \times \mathbb{R}^n \rightarrow [0; 1]$ określoną wzorem (...):

$$F_n(x; \mathbf{X}) = \frac{\#\{1 \leq j \leq n: X_j \leq x\}}{n}, x \in \mathbb{R}, X_j \in \mathbb{R}^n, \quad (1.1.7)$$

gdzie # oznacza liczbę elementów zbioru.

Funkcja $F_n(x; \mathbf{X})$ posiada wszystkie własności charakteryzujące dystrybuantę: jej wartości są w zakresie $[0; 1]$, jest niemalejąca i prawostronnie ciągła względem argumentu x . Funkcja $F_n(x; \mathbf{X})$ dla każdej wartości argumentu x jest zmienną losową.

Dystrybuanta empiryczna jest lokalnie (punktowo) dobrym oszacowaniem nieznannej dystrybuanty F , o czym mówi następujące twierdzenie (Magiera, 2002, s. 102):

Twierdzenie 1.1.1. Dla każdego $x \in \mathbb{R}$ dystrybuanta empiryczna ma następujące własności:

1. $E_F(F_n(x; \mathbf{X})) = F(x)$ oraz $Var_F(F_n(x; \mathbf{X})) = F(x)[1 - F(x)]/n$,
2. $P\left(\lim_{n \rightarrow \infty} F_n(x; \mathbf{X}) = F(x)\right) = 1$,
3. $F_n(x; \mathbf{X}) \sim AN\left(F(x), \frac{F(x)[1-F(x)]}{n}\right)$, gdzie AN to rozkład asymptotycznie normalny.

Zgodnie z punktem pierwszym Twierdzenia 1.1.1, dystrybuanta empiryczna jest lokalnie estymatorem nieobciążonym dystrybuanty F . Punkt drugi mówi, że dystrybuanta empiryczna jest lokalnie mocno zgodnym estymatorem dystrybuanty F . Według punktu trzeciego natomiast, dystrybuanta empiryczna ma rozkład asymptotycznie normalny.

Dystrybuanta empiryczna dobrze oszacowuje również globalnie dystrybuantę F , o czym mówi twierdzenie Gliwienki-Cantellego, nazywane także podstawowym twierdzeniem statystyki matematycznej. Twierdzeniem tym zostanie zakończony ten podrozdział:

Twierdzenie 1.1.2 (Krzyśko, 2004, s.14). Niech

$$KS = \sup_{x \in \mathbb{R}} |F_n(x; \mathbf{X}) - F(x)| \quad (1.1.8)$$

będzie odległością Kołmogorowa dystrybuanty empirycznej F_n od dystrybuanty F . Jeżeli próba X_1, X_2, \dots, X_n pochodzi z rozkładu o dystrybuancie F , to

$$P\left(\lim_{n \rightarrow \infty} KS = 0\right) = 1. \quad (1.1.9)$$

1.2. Wybrane elementy teorii testowania hipotez statystycznych

W tym podrozdziale zostaną przedstawione podstawowe pojęcia związane z teorią testowania hipotez statystycznych, które będą używane w dalszej części pracy. Przedstawione informacje zawężono do rozkładów ciągłych, ponieważ stopy zwrotu z instrumentów finansowych są ciągłymi zmiennymi losowymi.

Definicja 1.2.1 (Fisz, 1969, s. 442). Hipotezą statystyczną nazywane jest każde przypuszczenie dotyczące nieznanego rozkładu zmiennej losowej.

Definicja 1.2.2 (Domański i Pruska, 2000, s. 150). Testem statystycznym nazywamy każdą taką regułę decyzyjną (funkcję decyzyjną), która każdej losowej próbie $x \in \mathcal{X}$ przyporządkowuje – z ustalonym prawdopodobieństwem – jedną z dwóch decyzji: odrzucić weryfikowaną hipotezę lub nie.

Nie można udowodnić z całą pewnością prawdziwości lub fałszywości hipotezy statystycznej, dokonuje się tego tylko z określonym prawdopodobieństwem (Domański i Pruska, 2000, s. 149). Z punktu widzenia niniejszej pracy, istotny jest podział hipotez statystycznych na parametryczne i nieparametryczne oraz proste i złożone. Hipoteza parametryczna określa wartości nieznanymi parametrów rozkładu, natomiast hipoteza nieparametryczna określa nieznaną postać funkcyjną rozkładu. Hipoteza prosta jednoznacznie określa rozkład danej populacji, a zatem zarówno postać funkcyjną rozkładu, jak i wartości jej parametrów, jest ona zbiorem jednoelementowym. Hipoteza złożona określa całą grupę rozkładów, jest to zbiór składający się z więcej niż jednego elementu.

Hipoteza statystyczna jest zazwyczaj formułowana w postaci układu dwóch hipotez, zerowej (H_0) i alternatywnej (H_1). Niech H_0 oznacza, że rozkład badanej zmiennej należy do klasy rozkładów \mathcal{P}_0 , zawartej w klasie rozkładów \mathcal{P}_θ i niech $\mathcal{P}_1 = \mathcal{P}_\theta \setminus \mathcal{P}_0$ (Domański i Pruska, 2000, s. 149). Ponadto, niech Θ_0 i Θ_1 oznaczają zbiory parametrów odpowiadające klasom \mathcal{P}_0 i \mathcal{P}_1 . Pomiędzy zbiorami parametrów zachodzą zależności: $\Theta_0 \cup \Theta_1 = \Theta$ oraz $\Theta_0 \cap \Theta_1 = \emptyset$. Układ hipotez w tym przypadku ma postać:

$$\begin{aligned} H_0: \theta &\in \Theta_0, \\ H_1: \theta &\in \Theta_1. \end{aligned} \tag{1.2.1}$$

W celu podjęcia decyzji o przyjęciu lub odrzuceniu hipotezy, przestrzeń prób \mathcal{X} jest dzielona na dwa rozłączne podzbiory.

Definicja 1.2.3 (Krzyśko, 2004, s. 165). Podzbiór W przestrzeni prób \mathcal{X} , dla którego hipoteza H_0 jest odrzucana, nazywa się obszarem odrzucenia lub obszarem krytycznym.

Podzbiór $W^c = \mathcal{X} \setminus W$ jest nazywany obszarem przyjęcia. Podjęcie decyzji następuje w oparciu o test hipotezy H_0 przeciwko hipotezie H_1 :

Definicja 1.2.4 (Krzyśko, 2004, s. 165). Testem hipotezy H_0 przeciwko hipotezie H_1 nazywamy funkcję $\varphi: \mathcal{X} \rightarrow \{0; 1\}$, zdefiniowaną następująco:

$$\varphi(\mathbf{x}) = \begin{cases} 1, & \text{jeżeli } \mathbf{x} \in W, \\ 0, & \text{jeżeli } \mathbf{x} \in W^c. \end{cases} \quad (1.2.2)$$

Test jest często wyrażony za pomocą funkcji próby $T(\mathbf{X}) = T(X_1, X_2, \dots, X_n)$, zwanej statystyką testową, wówczas obszar krytyczny również jest wyrażany za pomocą statystyki testowej. Testy weryfikujące hipotezy parametryczne są nazywane testami parametrycznymi, natomiast testy weryfikujące hipotezy nieparametryczne to testy nieparametryczne.

Podejmując decyzję można popełnić dwa rodzaje błędów. Błąd I rodzaju polega na odrzuceniu hipotezy H_0 wtedy, gdy jest ona prawdziwa. Błąd II rodzaju polega na przyjęciu hipotezy H_0 wtedy, gdy jest ona fałszywa. Błędy I i II rodzaju określają odpowiednio funkcje (Domański i Pruska, 2000, s. 153):

$$\alpha(\theta) = P\{\mathbf{x} \in W | H_0\} = P_\theta(W) \quad \text{dla } \theta \in \Theta_0, \quad (1.2.3)$$

$$\beta(\theta) = P\{\mathbf{x} \in (\mathcal{X} \setminus W) | H_1\} = P_\theta(W^c) = 1 - P_\theta(W) \quad \text{dla } \theta \in \Theta_1. \quad (1.2.4)$$

Z błędami I i II rodzaju są związane pojęcia mocy i rozmiaru testu, które będą podstawą do oceny testów w sekcji obliczeniowej.

Definicja 1.2.5 (Domański i Pruska, 2000, s. 153). Mocą testu nazywamy prawdopodobieństwo podjęcia prawidłowej decyzji polegającej na odrzuceniu weryfikowanej hipotezy na podstawie wyników próby wtedy, gdy jest on fałszywa.

Zgodnie z powyższą definicją, funkcja $1 - \beta(\theta)$, gdzie $\beta(\theta)$ określa prawdopodobieństwo błędu II rodzaju, jest funkcją mocy testu (Domański i Pruska, 2000, s. 153):

$$M(\theta) = P\{\mathbf{x} \in W | H_1\} = 1 - P\{\mathbf{x} \in (\mathcal{X} \setminus W) | H_1\} = 1 - \beta(\theta). \quad (1.2.5)$$

Prawdopodobieństwo popełnienia błędu II rodzaju zależy m. in. od wielkości próby, weryfikowanej hipotezy oraz poziomu istotności testu (Martinez i Martinez, 2008, s. 239).

Idealna funkcja mocy powinna być równa zero, gdy $\theta \in \Theta_0$ oraz jedności, gdy $\theta \in \Theta_1$ (Krzyśko, 2004, s. 167).

Definicja 1.2.6 (Domański i Pruska, 2000, s. 155). Kres górny funkcji opisującej prawdopodobieństwo błędu I rodzaju

$$\sup_{\theta \in \Theta_0} \alpha(\theta) \tag{1.2.6}$$

nazywamy rozmiarem testu.

Rozmiar testu powinien być bliski zadanemu poziomowi istotności. Nie można kontrolować obydwu błędów jednocześnie, istnieją natomiast metody konstrukcji testów, które pozwalają na kontrolę jednego z błędów. Jedną z tych metod jest konstrukcja testu najmocniejszego zaproponowanego przez Neymana i Pearsona.

Definicja 1.2.7 (Domański i Pruska, 2000, s. 153). Testem najmocniejszym nazywamy test, który jest oparty na takim obszarze krytycznym W_0 , dla którego przy danym z góry prawdopodobieństwie α błędu I rodzaju, prawdopodobieństwo błędu II rodzaju β jest najmniejsze, tzn.:

$$P\{x \in (\mathcal{X} \setminus W_0) | H_1\} = \min_W P\{x \in (\mathcal{X} \setminus W) | H_1\}. \tag{1.2.7}$$

Z powyższej definicji wynika, że błąd I rodzaju nie powinien przekraczać zadanego poziomu α , a spośród dostępnych testów powinno wybrać się ten, który maksymalizuje funkcję mocy. Z definicją testu najmocniejszego jest związana definicja testu jednostajnie najmocniejszego:

Definicja 1.2.8 (Domański i Pruska, 2000, s. 154). Test, którego funkcja mocy jest jednostajnie nie mniejsza od funkcji mocy jakiegokolwiek innego testu danej klasy, nazywa się testem jednostajnie najmocniejszym (*JNM*). Oznacza to, że dla każdego argumentu funkcji mocy wartość tej funkcji jest nie mniejsza od wartości funkcji mocy innego testu dla tego argumentu.

Oprócz mocy testu, innymi kryteriami wyboru testu mogą być następujące własności: nieobciążoność, zgodność oraz optymalność. Definicja testu optymalnego wymaga wprowadzenia definicji testu istotności i poziomu istotności.

Definicja 1.2.9 (Domański i Pruska, 2000, s. 154). Testem nieobciążonym dla hipotezy $H_0: \theta = \theta_0$ wobec $H_1: \theta = \theta_1$, gdzie $\theta_1 \in \Theta_1$, nazywamy taki test oparty na obszarze krytycznym W , dla którego zachodzi nierówność:

$$P\{x \in W_0 | H_1\} \geq P\{x \in W_0 | H_0\}. \quad (1.2.8)$$

Definicja 1.2.10 (Domański i Pruska, 2000, s. 154). Jeżeli dla danego testu zachodzi nierówność:

$$M(\theta) < \alpha, \text{ gdzie } \theta \in \Theta_1, \quad (1.2.9)$$

to taki test nazywamy testem obciążonym.

Zgodnie z definicją testu nieobciążonego, dla każdego $\theta \in \Theta_1$, moc takiego testu jest nie mniejsza niż prawdopodobieństwo α błędu I rodzaju. Testy obciążone natomiast częściej odrzucają hipotezę prawdziwą niż fałszywą.

Definicja 1.2.11 (Domański i Pruska, 2000, s. 155). Testem zgodnym dla hipotezy H_0 nazywamy taki test, dla którego zachodzi równość:

$$\lim_{n \rightarrow \infty} M_n(\theta) = 1, \quad (1.2.10)$$

gdzie $\theta \in \Theta_1$ oraz $M_n(\theta)$ oznacza funkcję mocy testu dla n -elementowej próby.

Definicja 1.2.12 (Domański i Pruska, 2000, s. 155). Testem istotności dla hipotezy H_0 nazywamy taki test, za pomocą którego odrzuca się hipotezę H_0 z przyjętym z góry prawdopodobieństwem błędu I rodzaju α , jeżeli $x \in W$, natomiast dla $x \notin W$ stwierdza się brak podstaw do odrzucenia H_0 .

Definicja 1.2.13 (Domański i Pruska, 2000, s. 155). Liczbę $0 \leq \alpha \leq 1$ nazywamy poziomem istotności pewnego testu, jeżeli dla tego testu zachodzi relacja:

$$\alpha(\theta) \leq \alpha, \text{ dla dowolnego } \theta \in \Theta_0. \quad (1.2.11)$$

Definicja 1.2.14 (Domański i Pruska, 2000, s. 155). Testem optymalnym nazywamy taki test, który jest testem jednostajnie najmocniejszym na danym poziomie istotności α .

Zbiór testów na poziomie istotności α zawiera zbiór testów rozmiaru α . Jeśli nie jest możliwe skonstruowanie testu o rozmiarze α , można posłużyć się testem na poziomie istotności α (Domański i Pruska, 2000, s. 155).

Kryteria testu jednostajnie najmocniejszego muszą być spełnione dla każdej wielkości próby i dla każdego rozkładu z danej klasy rozkładów, co często jest bardzo trudne do udowodnienia. Z tego powodu moc testów najczęściej bada się lokalnie, dla ściśle określonego przypadku, w którym hipoteza alternatywna jest bliska hipotezie zerowej (Thas, 2010, s. 44).

Dodatkową informacją zbieraną podczas przeprowadzania testów w części obliczeniowej pracy będzie wartość p , zwana również poziomem krytycznym testu (Krzyśko, 2004, s. 181):

$$p(\mathbf{x}) = \sup_{\theta \in \Theta_0} P_{\theta}(T(\mathbf{X}) \geq T(\mathbf{x})). \quad (1.2.12)$$

Jest to najmniejszy poziom istotności, przy którym należy weryfikowaną hipotezę odrzucić (Magiera, 2002, s. 205). Wartość p można też interpretować jako największy poziom istotności, przy którym hipoteza zerowa nie jest odrzucana lub jako poziom istotności testu, gdyby zaobserwowana wartość statystyki testowej była wartością krytyczną testu (Domański, 2014, s. 18).

W ramach testów nieparametrycznych wyróżnić można testy zgodności, których dotyczy niniejsza praca. Testom tym zostanie poświęcona pozostała część tego podrozdziału. Niech $F \in \mathcal{F}$ oznacza, że funkcja F należy do klasy dystrybuant \mathcal{F} .

Definicja 1.2.15 (Domański i Pruska, 2000, s. 167). Testem zgodności nazywamy test służący do weryfikacji hipotezy prostej albo złożonej, dotyczącej zgodności pomiędzy rozkładem zbioru wartości w próbie i rozkładem teoretycznym hipotezy postaci $H_0: F \in \mathcal{F}$.

Rozkłady stóp zwrotu nie są znane, stąd konieczne jest stosowanie testów zgodności. Testy zgodności, ze względu na sposób konstruowania statystyk testowych, można podzielić na oparte na (Domański i Pruska, 2000, s. 167): porównaniu gęstości, porównaniu dystrybuant, momentach oraz statystykach pozycyjnych. Do testów zgodności opartych na porównaniu dystrybuant, które są przedmiotem zainteresowania niniejszej pracy, należą testy bazujące na statystykach Kołmogorowa - Smirnowa oraz ich funkcjach (Domański i Pruska, 2000, s. 167).

Testy zgodności oparte na porównaniu dystrybuanty badają zgodność rozkładu empirycznego z pewnym rozkładem teoretycznym. Rozkład empiryczny jest wyrażony poprzez dystrybuantę empiryczną, a rozkład teoretyczny poprzez dystrybuantę teoretyczną. Ogólny opis przebiegu testu istotności można znaleźć w pracy Magiery (2002, s. 208). Wnioskowanie statystyczne w oparciu o testy zgodności przedstawione w niniejszej pracy będzie przebiegało według etapów opisanych w podrozdziale 3.2.

1.3. Testy zgodności oparte na porównaniu dystrybuanty

W tej sekcji zostaną zaprezentowane statystyki testowe dla całych rozkładów, które zostaną wykorzystane w części autorskiej pracy. Więcej informacji na temat statystyk opartych na porównaniu dystrybuanty można znaleźć na przykład w pracach Domańskiego (1979), Domańskiego i Pruski (2000) oraz Thasa (2010).

1.3.1. Test zgodności Kołmogorowa

Do przeprowadzenia testu zgodności Kołmogorowa korzysta się ze statystyki zwanej odległością Kołmogorowa, mierzącej odległość pomiędzy dystrybuantą empiryczną i dystrybuantą F . Zgodnie z hipotezą H_0 zakłada się, że $F = F_0$, statystyka testowa ma wówczas postać (Magiera, 2002, s. 255):

$$KS = \sup_{x \in \mathbb{R}} |F_n(x; \mathbf{X}) - F_0(x)|, \quad (1.3.1.1)$$

gdzie $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ jest próbą prostą pochodzącą z rozkładu o ciągłej dystrybuancie F .

Zgodnie z twierdzeniem Kołmogorowa, przy $n \rightarrow \infty$ prawdziwe jest zdanie: $\sqrt{n}KS \xrightarrow{d} X$, gdzie zmienna losowa X ma rozkład Kołmogorowa, a symbol \xrightarrow{d} oznacza zbieżność według rozkładu. Zmienna losowa X ma rozkład Kołmogorowa, gdy jej dystrybuanta ma postać:

$$F(x) = 1 - 2 \sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2 x^2}, \quad x > 0. \quad (1.3.1.2)$$

Przy dostatecznie dużej liczbie obserwacji n , obszar krytyczny testu ma postać:

$$W = \{x: \sqrt{n}KS \geq \lambda(1 - \alpha)\}, \quad (1.3.1.3)$$

gdzie $\lambda(1 - \alpha)$ oznacza kwantyl rzędu $(1 - \alpha)$ z rozkładu Kołmogorowa, a α jest poziomem istotności. Przy małej liczbie obserwacji ($n \leq 100$), obszar krytyczny ma postać:

$$W = \{x: KS \geq d_n(1 - \alpha)\}, \quad (1.3.1.4)$$

gdzie wartości d_n podane są w tablicach. Tablice można znaleźć na przykład w pracy Magiery (2002, s. 373).

W praktyce do wyliczenia statystyki KS korzysta się z następujących wzorów:

$$KS^+ = \max_{1 \leq i \leq n} \left(\frac{i}{n} - F_0(X_{(i)}) \right), \quad (1.3.1.5)$$

$$KS^- = \max_{1 \leq i \leq n} \left(F_0(X_{(i)}) - \frac{i-1}{n} \right), \quad (1.3.1.6)$$

$$KS = \max\{KS^+, KS^-\}, \quad (1.3.1.7)$$

gdzie $X_{(i)}$ jest i -tą statystyką pozycyjną, $i = 1, 2, \dots, n$. Korzystanie z powyższych wzorów jest uzasadnione tym, że supremum statystyki KS realizuje się w jednym z punktów skoku dystrybuanty empirycznej $F_n(x; \mathbf{X})$.

Powyższy test Kołmogorowa może być stosowany do weryfikacji hipotez prostych. Wówczas, jeśli F_0 jest dystrybuantą ciągłą, rozkład prawdopodobieństwa statystyki KS nie zależy od F_0 , zależy jedynie od n (Magiera, 2002, s. 256). Jeśli jednak dystrybuanta teoretyczna zależy od parametrów, które są szacowane na podstawie próby, wówczas rozkład statystyki KS zależy od dystrybuanty F_0 . Co więcej, w ogólnym przypadku rozkład statystyki KS zależy również od parametrów dystrybuanty. W tej sytuacji nie można zastosować testów z obszarami krytycznymi wyznaczonymi według powyższych wzorów. Istnieją szczególne przypadki, dla których wyznaczono rozkład statystyki testowej KS , gdy parametry są szacowane z próby. Lilliefors podał kwantyle dokładnego rozkładu statystyki testowej KS , gdy F_0 jest dystrybuantą rozkładu normalnego $N(\mu; \sigma^2)$ lub wykładniczego $Ex(\lambda)$, a parametry rozkładu są szacowane na podstawie próby (Magiera, 2002, s. 257). W ogólnym przypadku, gdy parametry dystrybuanty teoretycznej są szacowane z próby, stosuje się metody Monte Carlo w celu wyznaczenia rozkładu statystyki testowej KS .

Testy zgodności oparte na statystyce Kołmogorowa-Smirnowa są bardziej czułe na rozbieżności pomiędzy dystrybuantą empiryczną i teoretyczną w części centralnej rozkładu niż w ogonach rozkładu (Eicker, 1979, s. 117).

1.3.2. Statystyka Kuipera

Z odległością Kołmogorowa jest związana statystyka Kuipera:

$$V = \sqrt{n} \left(\sup_{x \in \mathbb{R}} \{F_n(x) - F_0(x)\} + \sup_{x \in \mathbb{R}} \{F_0(x) - F_n(x)\} \right). \quad (1.3.2.1)$$

Kuiper wykazał, że dla dużej liczby obserwacji n , rozkład statystyki $\sqrt{n}V$ można opisać następująco (Stephens, 1965, s. 309):

$$P(\sqrt{n}V \geq z) = \sum_{i=1}^{\infty} 2(4i^2z^2 - 1)e^{-2i^2z^2} - \frac{8z}{3\sqrt{n}} \sum_{i=1}^{\infty} i^2(4i^2z^2 - 3)e^{-2i^2z^2} + o\left(\frac{1}{n}\right), \quad (1.3.2.2)$$

gdzie $O(1/n)$ oznacza resztę co najwyżej rzędu $1/n$. Tablice kwantyli dokładnego rozkładu statystyki V , dla wybranych liczebności prób i poziomów istotności, można znaleźć w pracy Stephensa (1965).

Obszar krytyczny tego testu jest prawostronny. W praktyce do wyliczenia statystyki V korzysta się ze wzoru:

$$V = KS^+ + KS^-, \quad (1.3.2.1)$$

gdzie KS^+ i KS^- zostały określone odpowiednio wzorami (1.3.1.5) oraz (1.3.1.6).

Rozkład statystyki V , pod warunkiem prawdziwości hipotezy zerowej, nie zależy od F_0 (Stephens, 1965, s. 309). Istnieją modyfikacje testu Kuipera, umożliwiające testowanie złożonej hipotezy zerowej, np. dla rozkładu normalnego, gdy nie są znane parametry populacji generalnej oraz dla rozkładu wykładniczego, gdy nieznana jest wartość oczekiwana (Domański, 1979, s. 104).

Test Kuipera jest niezmienniczy ze względu na cykliczne transformacje. Oznacza to, że nie ma znaczenia, który punkt zostanie wybrany jako startowy. Test ten ma szczególne znaczenie w przypadku testowania danych cyklicznych (Górecki, 2011, s. 231), może jednak być stosowany do danych na prostej rzeczywistej (Pewsey, 2018, s. 154).

1.3.3. Test Craméra-von Misesa

Test Craméra-von Misesa został zaproponowany przez Craméra w 1928 r. i niezależnie przez von Misesa w 1931 r. Statystyka Craméra-von Misesa wyraża się wzorem:

$$Q_n = \int_{-\infty}^{\infty} [F_n(x; \mathbf{X}) - F_0(x)]^2 dK(x), \quad (1.3.3.1)$$

gdzie K jest pewną niemalejącą funkcją. Dla $K = nF_0$ statystyka testowa Craméra-von Misesa przyjmuje postać:

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x; \mathbf{X}) - F_0(x)]^2 dF_0(x). \quad (1.3.3.2)$$

Ze względu na złożoność dokładnego rozkładu statystyki W_n^2 , często wykorzystywany jest jej rozkład graniczny (Magiera, 2002, s. 247):

$$W_n^2 \xrightarrow{d} G = \sum_{i=1}^{\infty} i^{-2} \pi^{-2} \chi_{1,i}^2, \quad (1.3.3.3)$$

gdzie $\chi_{1,i}^2$ dla $i = 1, 2, \dots$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie χ^2 z jednym stopniem swobody.

Obszar krytyczny ma postać:

$$W = \{\mathbf{x}: w_n^2 > g(1 - \alpha)\}, \quad (1.3.3.4)$$

gdzie $g(1 - \alpha)$ oznacza kwantyl rzędu $(1 - \alpha)$ z rozkładu G . Wartości kwantyli można odczytać z tablic (Zieliński i Zieliński, 1990, s. 312).

Do obliczeń wykorzystuje się wzór (Magiera, 2002, s. 247):

$$w_n^2 = \sum_{i=1}^n \left(F_0(x_{i:n}) - \frac{i - 0.5}{n} \right)^2 + \frac{1}{12n}. \quad (1.3.3.5)$$

Gdy hipoteza H_0 jest prawdziwa, dokładny rozkład statystyki Craméra-von Misesa nie zależy od dystrybuanty teoretycznej F_0 . Jeśli jednak dystrybuanta teoretyczna zależy od parametrów, które są szacowane na podstawie próby, wówczas rozkład statystyki W_n^2 zależy od dystrybuanty F_0 oraz od parametrów dystrybuanty (Thas, 2010, s. 140). Więcej informacji na temat rozkładu statystyki Craméra-von Misesa można znaleźć w pracy Thasa (2010, s. 137–138).

1.3.4. Test Andersona-Darlinga

Anderson i Darling (1952) zaproponowali modyfikacje statystyk Kołmogorowa oraz Craméra-von Misesa polegające na wprowadzeniu funkcji wagowej. Funkcje wagowe zostały dobrane w ten sposób, aby nadać większe znaczenie ogonom rozkładu. W ogólnej postaci statystyki te są następujące:

$$AD = \sup_{x \in \mathbb{R}} \sqrt{n} |F_n(x; \mathbf{X}) - F_0(x)| \sqrt{\Psi[F_0(x)]}, \quad (1.3.4.1)$$

$$AD^2 = n \int_{-\infty}^{\infty} [F_n(x; \mathbf{X}) - F_0(x)]^2 \Psi[F_0(x)] dF_0(x), \quad (1.3.4.2)$$

gdzie $\Psi(x)$ jest funkcją wagową taką, że $\Psi(x) \geq 0$ dla $0 \leq x \leq 1$.

Anderson i Darling rozważyli dwie funkcje wagowe: $\Psi(x) = 1$ oraz $\Psi(x) = 1/[x(1 - x)]$. Dla funkcji wagowej $\Psi(x) = 1$, wzory (1.3.4.1) oraz (1.3.4.2) redukują się odpowiednio do statystyk Kołmogorowa oraz Craméra-von Misesa. Gdy $\Psi(x) = 1/[x(1 - x)]$, statystyki są następujące:

$$AD = \sup_{x \in \mathbb{R}} \sqrt{n} \left| \frac{F_n(x; \mathbf{X}) - F_0(x)}{\sqrt{F_0(x)(1 - F_0(x))}} \right|, \quad (1.3.4.3)$$

$$AD^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x; \mathbf{X}) - F_0(x))^2}{F_0(x)(1 - F_0(x))} dF_0(x). \quad (1.3.4.4)$$

Statystyki te przypisują większe znaczenie ogonom rozkładu niż wartościom w okolicach mediany, gdyż funkcja Ψ osiąga największe wartości w otoczeniu 0 i 1. Funkcja wagowa jest w tym przypadku odwrotnością granicznej wariancji ciągu zmiennych losowych $F_n(x; \mathbf{X})$.

Ze względu na złożoność dokładnego rozkładu statystyki AD^2 , często wykorzystywany jest jej rozkład graniczny (Thas, 2010, s. 139):

$$AD^2 \xrightarrow{d} G = \sum_{i=1}^{\infty} \frac{1}{j(j+1)} \chi_{1,i}^2, \quad (1.3.4.5)$$

gdzie $\chi_{1,i}^2$ dla $i = 1, 2, \dots$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie χ^2 z jednym stopniem swobody.

Obszar krytyczny dla statystyki AD^2 ma postać:

$$W = \{x: ad^2 > g(1 - \alpha)\}, \quad (1.3.4.6)$$

gdzie $g(1 - \alpha)$ oznacza kwantyl rzędu $(1 - \alpha)$ z rozkładu G .

W praktyce do wyliczenia statystyk można skorzystać ze wzorów:

$$AD = \sqrt{n} \max \left\{ \sup_i \left\{ \frac{\frac{i}{n} - F_0(X_{(i)})}{\sqrt{F_0(X_{(i)})(1 - F_0(X_{(i)})}} \right\}, \sup_i \left\{ \frac{F_0(X_{(i)}) - \frac{i-1}{n}}{\sqrt{F_0(X_{(i)})(1 - F_0(X_{(i)})}} \right\} \right\}, \quad (1.3.4.7)$$

$$AD^2 = -n - \sum_{i=1}^n \frac{2i-1}{n} \left(\ln F_0(X_{(i)}) + \ln (1 - F_0(X_{(n+1-i)})) \right). \quad (1.3.4.8)$$

Gdy hipoteza H_0 jest prawdziwa, rozkład statystyki Andersona-Darlinga nie zależy od dystrybuanty teoretycznej F_0 . Jeśli jednak dystrybuanta teoretyczna zależy od parametrów, które są szacowane na podstawie próby, wówczas rozkład statystyki AD^2 zależy od dystrybuanty F_0 oraz od parametrów dystrybuanty (Thas, 2010, s. 140). Więcej informacji na temat rozkładu statystyki AD^2 można znaleźć u Thasa (2010, s. 138), a na temat rozkładu statystyki AD u Chapmana (1958, s. 665–666).

1.4. Modyfikacje testów zgodności opartych na porównaniu dystrybuanty

W tym podrozdziale zostaną przedstawione modyfikacje testów zgodności istotne z punktu widzenia niniejszej pracy: modyfikacje funkcji wagowych w statystyce Andersona-Darlinga oraz testy dla rozkładów uciętych.

1.4.1. Modyfikacje statystyki Andersona-Darlinga

Statystyki Andersona-Darlinga dają możliwość wstawienia dowolnej funkcji wagowej, najczęściej stosowane są jednak tylko dwie: $\psi(x) = 1$ oraz $\psi(x) = 1/[x(1-x)]$. Obydwie funkcje są symetryczne względem prostej $x = 1/2$, a druga z tych funkcji nadaje większe znaczenie ogonom rozkładu. W tym podrozdziale zostaną zaprezentowane inne funkcje wagowe zaproponowane dotychczas w literaturze, funkcje będą rozpatrywane dla $x \in (0, 1)$.

Ahmad, Sinclair i Spurr (1988) zaproponowali użycie funkcji wagowej $\psi_1(x) = 1/x$, natomiast Sinclair, Spurr i Ahmad (1987) użycie funkcji wagowej $\psi_2(x) = 1/(1-x)$. Funkcja ψ_1 nadaje większe znaczenie dolnemu ogonowi rozkładu, a funkcja ψ_2 górnemu ogonowi rozkładu. Zmodyfikowane statystyki AD^2 , odpowiednio dla funkcji wagowych ψ_1 i ψ_2 , mają postać:

$$A_1^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x; \mathbf{X}) - F_0(x))^2}{F_0(x)} dF_0(x), \quad (1.4.1.1)$$

$$A_2^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x; \mathbf{X}) - F_0(x))^2}{1 - F_0(x)} dF_0(x), \quad (1.4.1.2)$$

Zachodzi następująca zależność pomiędzy statystyką Andersona-Darlinga w wersji kwadratowej a zmodyfikowanymi statystykami:

$$AD^2 = A_1^2 + A_2^2. \quad (1.4.1.3)$$

Sinclair i in. (1987) wyznaczyli formułę obliczeniową dla statystyki A_2^2 . Do obliczenia statystyki A_1^2 można skorzystać z zależności (1.4.1.3).

$$A_2^2 = \frac{n}{2} - 2 \sum_{i=1}^n F_0(X_{(i)}) - \sum_{i=1}^n \left(2 - \frac{2i-1}{n}\right) \ln(1 - F_0(X_{(i)})). \quad (1.4.1.4)$$

Dla rozkładu EV1, Sinclair i in. (1987) wyznaczyli przybliżone rozkłady statystyki A_2^2 , zarówno gdy parametry rozkładu są określone, jak i szacowane na podstawie próby.

Rodriguez i Viollaz (1995) zaproponowali funkcję wagową $\psi_3(x) = [x(2-x)]^{-1}$, przypisującą większe wagi dolnemu ogonowi rozkładu oraz funkcję wagową $\psi_4(x) = (1-x^2)^{-1}$, nadającą większe wagi górnemu ogonowi rozkładu. Funkcje ψ_3 oraz ψ_4 są symetryczne względem prostej $x = 1/2$. Zmodyfikowana statystyka Andersona-Darlinga w wersji kwadratowej ma wówczas formę, odpowiednio dla funkcji ψ_3 oraz ψ_4 :

$$V_n^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x; \mathbf{X}) - F_0(x))^2}{F_0(x)(2 - F_0(x))} dF_0(x), \quad (1.4.1.5)$$

$$V_n^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x; \mathbf{X}) - F_0(x))^2}{(1 - F_0^2(x))} dF_0(x). \quad (1.4.1.6)$$

Rozkłady statystyk określonych wzorami (1.4.1.5) i (1.4.1.6), w przypadku prawdziwości prostej hipotezy zerowej, są takie same i dążą do następującego rozkładu granicznego:

$$V^2 \xrightarrow{d} G = \sum_{i=1}^{\infty} \frac{1}{2j(2j+1)} \chi_{1,i}^2, \quad (1.4.1.7)$$

gdzie $\chi_{1,i}^2$ dla $i = 1, 2, \dots$ są niezależnymi zmiennymi losowymi o tym samym rozkładzie χ^2 z jednym stopniem swobody.

Rodriguez i Viollaz (1995) zbadali moc testów dla statystyki z funkcją wagową $\psi_4(x) = (1 - x^2)^{-1}$, na 5% poziomie istotności, dla wybranych hipotez prostych. Otrzymane wyniki były zgodne z oczekiwaniami, funkcja ψ_4 przypisywała większe znaczenie górnemu ogonowi rozkładu. Dla przypadku, gdy rozkład ma grubszy prawy ogon, moc testu z funkcją wagową ψ_4 była większa niż moc testu przy zastosowaniu statystyki Craméra-von Misesa oraz przy zastosowaniu statystyki Andersona-Darlinga. Dla przypadku, gdy grubszy jest lewy ogon rozkładu, moc testu z funkcją wagową ψ_4 była mniejsza w porównaniu do testów opartych na statystykach Craméra-von Misesa oraz AD^2 .

W kolejnym artykule Rodriguez i Viollaz (1999) zaproponowali użycie funkcji wagowych $\psi_1(x) = 1/x$ oraz $\psi_2(x) = 1/(1 - x)$, które były rozważane w literaturze już wcześniej (Sinclair i in., 1987). Dla każdej statystyki wyznaczono asymptotyczne rozkłady statystyk, pod warunkiem prawdziwości prostej hipotezy zerowej.

W części autorskiej pracy zostaną wykorzystane modyfikacje statystyk testowych AD i AD^2 , polegające na wprowadzeniu funkcji wagowych $\psi_1(x) = 1/x$ oraz $\psi_2(x) = 1/(1 - x)$.

1.4.2. Testy zgodności z lewostronnie uciętym rozkładem ciągłym

W tym podrozdziale zostaną opisane testy zgodności z lewostronnie uciętym rozkładem ciągłym, przedstawione w pracy Chernobai i in. (2015). Testy te zostaną zmodyfikowane w autorskiej części pracy.

Rozkłady ucięte występują, gdy wartości badanej cechy mieszczą się w przedziale $[H1; H2]$, natomiast wartości zmiennej losowej należą do przedziału $(a; b)$

i $[H1; H2] \subset (a; b)$ (Domański i Pruska, 2000, s. 34). W tym podrozdziale rozważania dotyczące rozkładu uciętego zostaną zawężone do rozkładu lewostronnie uciętego.

Definicja 1.4.2.1 (Domański i Pruska, 2000,s.34). Niech X będzie zmienną losową o rozkładzie P_X przyjmującą wartości z przedziału $(a; b)$ i niech $[H1; b) \subset (a; b)$, gdzie $a, b, H1 \in \mathbb{R}$. Rozkład zmiennej losowej $Y = X|H1 \leq X < b$ nazywamy lewostronnie uciętym rozkładem P_X .

W celu sformułowania hipotezy statystycznej dla rozkładu lewostronnie uciętego, należy zdefiniować dystrybuantę teoretyczną i empiryczną tego rozkładu. Najpierw zostanie określona dystrybuanta teoretyczna dla rozkładu lewostronnie uciętego.

Niech F_0 oznacza dystrybuantę teoretyczną, a z_H wartość dystrybuanty teoretycznej w punkcie ucięcia H , tzn. $z_H = F_0(H)$. Jeśli rozkład ucięty w punkcie H jest zgodny z dystrybuantą teoretyczną, wówczas $F_0(X) \sim U[F_0(H); 1]$. Dystrybuanta teoretyczna rozkładu lewostronnie uciętego ma postać:

$$F_0^*(x) = \begin{cases} \frac{F_0(x) - F_0(H)}{1 - F_0(H)}, & \text{dla } x \geq H, \\ 0, & \text{dla } x < H. \end{cases} \quad (1.4.2.1)$$

Jeśli rozkład próby jest zgodny z rozkładem teoretycznym, wówczas $F_0^*(x) \sim U[0; 1]$.

Zdefiniujmy teraz dystrybuantę empiryczną rozkładu lewostronnie uciętego. Rozważmy w tym celu n_c -elementową próbę z pewnego rozkładu. Uporządkowana próba $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n_c)}$ ma dystrybuantę empiryczną określoną wzorem (1.1.7). Różnica pomiędzy wartościami dystrybuanty empirycznej dla dwóch kolejnych punktów wynosi $1/n_c$. W przypadku rozkładu lewostronnie uciętego znane są jedynie obserwacje nie mniejsze niż próg ucięcia H . Załóżmy, że ich liczba wynosi n . Niech liczba nieznanymi obserwacji, mniejszych niż próg ucięcia H , wynosi m . Całkowita liczba obserwacji jest równa $n_c = m + n$. Dystrybuanta empiryczna rozkładu lewostronnie uciętego jest także określona wzorem (1.1.7), ale tu różnica pomiędzy jej wartościami dla dwóch kolejnych punktów wynosi $1/n$.

Dystrybuanta empiryczna dla zaobserwowanej części całej populacji wynosi:

$$F_n(x) * (1 - F_0(H)) + F_0(H) = \begin{cases} F_0(H) & x < x_{(1)}, \\ \frac{j}{n} (1 - F_0(H)) + F_0(H) & x_{(j)} \leq x < x_{(j+1)}, j = 1, 2, \dots, n - 1, \\ 1 & x \geq x_{(n)}. \end{cases} \quad (1.4.2.2)$$

Układ hipotez dla rozkładu lewostronnie uciętego jest następujący:

$$\begin{aligned} H_0: F &= F_0^*, \\ H_1: F &\neq F_0^*. \end{aligned} \quad (1.4.2.3)$$

Jeśli parametry dystrybuant teoretycznych nie są znane i są szacowane na podstawie próby, dystrybuanta teoretyczna zostanie oznaczona daszkiem i układ hipotez jest następujący:

$$\begin{aligned} H_0: F &= \hat{F}_0^*, \\ H_1: F &\neq \hat{F}_0^*. \end{aligned} \quad (1.4.2.4)$$

Do weryfikacji testów statystycznych dla rozkładów uciętych stosuje się statystyki dla rozkładów uciętych. W tabeli 1.4.1 znajdują się wzory statystyk testowych dla ciągłych rozkładów lewostronnie uciętych wraz z ich formułami obliczeniowymi, przedstawione w pracy Chernobai i in. (2015). Są one modyfikacjami statystyk dla całych rozkładów omówionych w rozdziale 1.3 i w podrozdziale 1.4.1. Statystyki dla rozkładów lewostronnie uciętych zostaną w niniejszej pracy oznaczone gwiazdką (*). Zastosowano następujące oznaczenia: $F_n(x)$ oznacza dystrybuantę empiryczną określoną wzorem (1.1.7), $\hat{F}_0^*(x)$ oznacza dystrybuantę teoretyczną rozkładu lewostronnie uciętego z parametrami oszacowanymi na podstawie próby określoną wzorem (1.4.2.1), $z_H = \hat{F}_0(H)$ oznacza wartość dystrybuanty teoretycznej w punkcie ucięcia H , $\hat{F}_0(x)$ oznacza dystrybuantę dla całego rozkładu z parametrami oszacowanymi na podstawie próby, $z_j = \hat{F}_0(x_{(j)})$ oznacza wartość dystrybuanty teoretycznej w punkcie $x_{(j)}$, gdzie $x_{(j)}$ oznacza j -tą statystykę pozycyjną z próby, $j = 1, 2, \dots, n$, natomiast n jest wielkością próby z rozkładu lewostronnie uciętego.

Tabela 1.4.1. Statystyki dla rozkładów lewostronnie uciętych

Statystyka	Wzór statystyki i formuła obliczeniowa
KS^*	$KS^* = \sqrt{n} \sup_x F_n(x) - \hat{F}_0^*(x) $ <p>Formuła obliczeniowa:</p> $KS^* = \frac{\sqrt{n}}{1 - z_H} \max \left\{ \sup_j \left\{ z_H + \frac{j}{n}(1 - z_H) - z_j \right\}, \sup_j \left\{ z_j - \left(z_H + \frac{j-1}{n}(1 - z_H) \right) \right\} \right\}$
V^*	$V^* = \sqrt{n} \left(\sup_x \{F_n(x) - \hat{F}_0^*(x)\} + \sup_x \{\hat{F}_0^*(x) - F_n(x)\} \right)$ <p>Formuła obliczeniowa:</p> $V^* = \frac{\sqrt{n}}{1 - z_H} \max \left\{ \sup_j \left\{ z_H + \frac{j}{n}(1 - z_H) - z_j \right\} + \sup_j \left\{ z_j - \left(z_H + \frac{j-1}{n}(1 - z_H) \right) \right\} \right\}$
AD^*	$AD^* = \sqrt{n} \sup_x \left \frac{F_n(x) - \hat{F}_0^*(x)}{\sqrt{\hat{F}_0^*(x)(1 - \hat{F}_0^*(x))}} \right $

Statystyka	Wzór statystyki i formuła obliczeniowa
AD^*	Formuła obliczeniowa: $AD^* = \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_H + \frac{j}{n}(1 - z_H) - z_j}{\sqrt{(z_j - z_H)(1 - z_j)}} \right\}, \sup_j \left\{ \frac{z_j - \left(z_H + \frac{j-1}{n}(1 - z_H) \right)}{\sqrt{(z_j - z_H)(1 - z_j)}} \right\} \right\}$
AD_{up}^*	$AD_{up}^* = \sqrt{n} \sup_x \left \frac{F_n(x) - \hat{F}_0^*(x)}{1 - \hat{F}_0^*(x)} \right $ Formuła obliczeniowa: $AD_{up}^* = \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_H + \frac{j}{n}(1 - z_H) - z_j}{1 - z_j} \right\}, \sup_j \left\{ \frac{z_j - \left(z_H + \frac{j-1}{n}(1 - z_H) \right)}{1 - z_j} \right\} \right\}$
AD^{2*}	$AD^{2*} = n \int_H^\infty \frac{(F_n(x) - \hat{F}_0^*(x))^2}{\hat{F}_0^*(x)(1 - \hat{F}_0^*(x))} d\hat{F}_0^*(x)$ Formuła obliczeniowa: $AD^{2*} = -n + 2n \log(1 - z_H) - \frac{1}{n} \sum_{j=1}^n (1 + 2(n - j)) \log(1 - z_j) + \frac{1}{n} \sum_{j=1}^n (1 - 2j) \log(z_j - z_H)$
W^{2*}	$W^{2*} = n \int_H^\infty (F_n(x) - \hat{F}_0^*(x))^2 d\hat{F}_0^*(x)$ Formuła obliczeniowa: $W^2 = \frac{n}{3} + \frac{nz_H}{1 - z_H} + \frac{1}{n(1 - z_H)} \sum_{j=1}^n (1 - 2j)z_j + \frac{1}{(1 - z_H)^2} \sum_{j=1}^n (z_j - z_H)^2$
AD_{up}^{2*}	$AD_{up}^{2*} = n \int_H^\infty \frac{(F_n(x) - \hat{F}_0^*(x))^2}{(1 - \hat{F}_0^*(x))^2} d\hat{F}_0^*(x)$ Formuła obliczeniowa: $AD_{up}^{2*} = -2n \log(1 - z_H) + 2 \sum_{j=1}^n \log(1 - z_j) + \frac{1 - z_H}{n} \sum_{j=1}^n (1 + 2(n - j)) \frac{1}{1 - z_j}$

Źródło: Chernobai i in. (2015)

1.5. Bootstrapowe testy zgodności

Wyznaczenie rozkładów zmodyfikowanych statystyk testowych metodami analitycznymi często jest bardzo trudne. Jeśli dodatkowo weryfikowana hipoteza jest złożona, rozkład statystyk może zależeć od rozkładu teoretycznego, estymowanych parametrów, metody estymacji tych parametrów oraz wielkości próby (Pewsey, 2018, s. 154). W takich sytuacjach można aproksymować rozkłady statystyk testowych, a więc wyznaczyć ich wartości krytyczne i oszacować wartości p , rozkładem asymptotycznym lub zastosować metody bootstrapowe.

Aproksymowanie rozkładu statystyki rozkładem asymptotycznym, gdy próba jest mała, może nie być właściwe z kilku powodów (Orzeszko, 2014, s. 339): tempo zbieżności rozkładu małopróbkowego do rozkładu asymptotycznego może nie być znane; mogą być wymagane założenia dotyczące rozkładu populacji, które w praktyce mogą nie być spełnione; mogą wystąpić trudności z wyznaczeniem kwantyli dla rozkładu asymptotycznego. W przypadku ogonów rozkładów stóp zwrotu często liczba obserwacji jest mała, z tego powodu nie rozważano aproksymowania rozkładów statystyk testowych za pomocą rozkładów asymptotycznych. Ponadto metoda bootstrapu jest metodą rekomendowaną w przypadku weryfikowania hipotez złożonych (Thas, 2010, s. 159). Więcej na temat wnioskowania małopróbkowego można znaleźć w pracy Kęłowski (2013) lub Orzeszko (2014).

Testy statystyczne wykorzystujące do aproksymacji rozkładów statystyk testowych metody bootstrapowe są nazywane testami bootstrapowymi. Stosowanie testów bootstrapowych nie wymaga założeń dotyczących rozkładu zmiennej losowej i liczebności próby (Domański, 2014, s. 119), co jest istotne z punktu widzenia przeprowadzonych badań. Rozkłady stóp zwrotu nie są znane, ponadto liczebność prób pochodzących z ogonów rozkładów często jest niewielka.

Metody bootstrapowe zaproponował Efron (1979). Opis metod bootstrapowych i ich zastosowanie do weryfikacji hipotez statystycznych można znaleźć w pracach Domańskiego (2014) oraz Domańskiego i Pruski (2000). W związku z pojawiającymi się w literaturze niezgodnościami dotyczącymi nazewnictwa metod bootstrapowych i Monte Carlo (Martinez i Martinez, 2008, s. 256), poniżej zostanie wyjaśnione, w jakim znaczeniu terminy te należy rozumieć w niniejszej pracy.

Oznaczmy przez $\mathbf{X} = (X_1, \dots, X_n)$ próbę z rozkładu określonego przez dystrybuantę $F(x)$. Jeśli dystrybuanta jest znana, wówczas można użyć metody *Monte Carlo*. Metoda ta bazuje na niezależnych pseudolosowych próbach $\mathbf{X}^{(i)} = (X_1^{(i)}, \dots, X_n^{(i)})$, gdzie $i = 1, 2, \dots, k$, wygenerowanych z rozkładu określonego przez $F(x)$ (Magiera, 2002, s. 163). Jeśli natomiast dystrybuanta nie jest znana, wtedy można skorzystać z procedury próbkowania wtórnego nazywanej bootstrapem. Bootstrap można podzielić na nieparametryczny i parametryczny. W obydwu przypadkach konieczne jest wyznaczenie estymatorów rzeczywistej dystrybuanty. W bootstrapie nieparametrycznym tym estymatorem jest dystrybuanta empiryczna $F_n(x)$. W bootstrapie parametrycznym estymatorem dystrybuanty rzeczywistej jest pewna wybrana dystrybuanta teoretyczna z parametrami oszacowanymi na podstawie próby $\hat{F}(x)$. W metodzie bootstrapu nieparametrycznego, próby $\mathbf{X}^* = (X_1^*, \dots, X_n^*)$ tworzone są przez

losowy wybór ze zwracaniem n elementów z próby X . Wybór każdego elementu do próby X^* jest jednakowo prawdopodobny i wynosi $1/n$. Zmienne X_1^*, \dots, X_n^* są warunkowo niezależnymi zmiennymi losowymi dla próby X . W metodzie bootstrapu parametrycznego, niezależne próby $X^{*(i)} = (X_1^{*(i)}, \dots, X_n^{*(i)})$, gdzie $i = 1, 2, \dots, k$, są wygenerowane z rozkładu określonego przez $\hat{F}(x)$ (Magiera, 2002, s. 160). Metoda bootstrapu parametrycznego została zastosowana w części autorskiej pracy do aproksymowania rozkładu statystyk testowych.

W idealnych warunkach wartość p ma rozkład jednostajny na przedziale $[0; 1]$ (Davison i Hinkley, 1997, s. 175). W rzeczywistości rozkład wartości p otrzymany metodą bootstrapu parametrycznego może znacznie odbiegać od rozkładu jednostajnego, dlatego w celu otrzymania rozkładu wartości p bardziej zbliżonego do rozkładu jednostajnego, stosowana jest metoda podwójnego bootstrapu parametrycznego (Davison i Hinkley, 1997, s. 177). Z uwagi na czasochłonność tej procedury (Davison i Hinkley, 1997, s. 437), w badaniach empirycznych, do oszacowania rozkładów statystyk testowych, wykorzystana zostanie metoda pojedynczego bootstrapu parametrycznego.

Testy zgodności oparte na porównaniu dystrybuanty, korzystające z metody bootstrapu parametrycznego, mają asymptotycznie właściwy rozmiar i są zgodne (Babu i Rao, 2004; Stute, Manteiga i Quindimil, 1993). Pewsey (2018, s. 156) zwraca jednak uwagę, że w niektórych przypadkach rozmiar testów opartych na bootstrapie parametrycznym może odbiegać od poziomu oczekiwanego. Ma to miejsce w przypadku rozkładów opisanych więcej niż dwoma parametrami, gdy próby są małej lub średniej wielkości lub parametry rozkładu nie posiadają estymatorów największej wiarygodności w formie analitycznej.

1.6. Moc testów zgodności opartych na porównaniu dystrybuanty

Zagadnienie mocy testów nieparametrycznych jest znacznie mniej zbadane, niż zagadnienie mocy testów parametrycznych. Wpływają na to z jednej strony trudności obliczeniowe, a z drugiej trudności z wyborem hipotezy alternatywnej spośród wszystkich możliwych hipotez alternatywnych (Fisz, 1969, s. 588; Lehmann i Romano, 2005, s. 583).

Większość badań nad testami zgodności wskazuje, że są one zgodne, a więc ich moc dąży do jedności wraz ze zwiększaniem wielkości próby (Vilijandas, Kruopis i Nikulin, 2011, s. 13). Brakuje jednak ogólnej teorii wskazującej, które testy są optymalne w sensie asymptotycznym,

istnieją jedynie odosobnione przykłady opisujące takie testy (Lehmann i Romano, 2005, s. 583). W przypadku wielu możliwych hipotez alternatywnych jest mało prawdopodobne, aby jeden, wybrany test, miał podobną moc w rozpoznawaniu wszystkich hipotez. Z pewnością natomiast nie istnieje w takich przypadkach test jednostajnie najmocniejszy (Lehmann i Romano, 2005, s. 599).

W przypadku testów zgodności wprowadza się pojęcie hipotez lokalnych, to znaczy takich, w których rozkład znajdujący się w hipotezie alternatywnej jest bardzo bliski rozkładowi z hipotezy zerowej (Thas, 2010, s. 44). Testy zgodności mogą charakteryzować się dobrą mocą tylko lokalnie, w odniesieniu do wybranych hipotez (Janssen, 2000; Lehmann i Romano, 2005, s. 622). Istnieją takie hipotezy lokalne, dla których dany test będzie charakteryzował się słabą mocą nawet wówczas, gdy zwiększana jest liczebność próby (Lehmann i Romano, 2005, s. 622). Nie jest to w sprzeczności ze stwierdzeniem, że moc testów zgodności dąży do jedności wraz ze zwiększaniem wielkości próby, stwierdzenie to dotyczy bowiem hipotez nielokalnych.

W związku z pojawiającymi się problemami związanymi z określeniem mocy testów zgodności, ich analiza często ma na celu jedynie przedstawienie słabych i mocnych stron testów, a wybór konkretnego testu zależy od testowanej hipotezy i możliwych hipotez alternatywnych. Liczne badania symulacyjne wskazują jednak na dwa testy, Andersona-Darlinga i Craméra-von Misesa, jako te mające ogólnie dobrą moc przy rozpatrywaniu rozmaitych hipotez i z punktu widzenia mocy testu, są one często preferowane ponad innymi testami (Thas, 2010, s. 159). Poniżej zostaną zaprezentowane wyniki wybranych badań nad mocami testów zgodności.

Pearson (1963) porównał moc testów opartych na statystykach KS , W^2 , V i Watsona (U^2) do weryfikacji hipotezy prostej, zakładającej, że próba ma charakter losowy. Sprowadza się to do sprawdzenia hipotezy, że próba jest prosta (Domański, 1979, s. 60). Testy oparte na minimalnej i maksymalnej odległości pomiędzy dystrybuantami empiryczną i teoretyczną, Kołmogorowa-Smirnowa i Kuipera, w większości badanych przykładów miały większą moc niż testy oparte na normie L_2 , Craméra-von Misesa i Watsona. Ponadto, testy oparte na statystykach V i U^2 były bardziej wrażliwe na odchylenia od losowości niż testy oparte na statystykach KS i W^2 . W tym badaniu Pearson wykorzystał rangowanie testów na podstawie ich wartości p .

Shapiro, Wilk i Chen (1968) badali moc testów zgodności opartych na dziewięciu statystykach, w tym na statystykach KS , W^2 i AD^2 . W testach badających normalność

rozkładu rozważano 45 rozkładów alternatywnych pochodzących z 12 rodzin rozkładów prawdopodobieństwa oraz 5 wielkości prób. Testy oparte na odległości Kołmogorowa okazały się zazwyczaj niewrażliwe na niewielkie zmiany parametrów.

Stephens (1974) badał moc testów opartych na statystykach KS , W^2 , AD^2 , V i U^2 . W badaniach weryfikowano hipotezę prostą zakładającą zgodność rozkładu z rozkładem jednostajnym oraz hipotezy złożone badające zgodność z rozkładem normalnym. Testy oparte na statystykach W^2 i AD^2 zazwyczaj charakteryzowały się większą mocą niż testy oparte na pozostałych statystykach.

Szczegółowe omówienie mocy testów można znaleźć np. w pracach Fisz (1969), Janssen (2000), Lehmann i Romano (2005) oraz Thasa (2010). Więcej przykładowych badań można znaleźć w pracach Lehmann i Romano (2005, s. 616) oraz Thasa (2010).

2. Wybrane rozkłady stóp zwrotu

Przedstawione w rozdziale trzecim modyfikacje wybranych testów statystycznych zostaną zastosowane do analizy przykładowych empirycznych rozkładów stóp zwrotu w rozdziale czwartym. Modyfikacje dotyczą rozkładów jednowymiarowych, jednak rozkłady stóp zwrotu często są analizowane w kontekście wielowymiarowym. Szczególnym zastosowaniem bezwarunkowych, jednowymiarowych rozkładów stóp zwrotu, jest ich użycie do konstruowania rozkładów wielowymiarowych za pomocą kopul. Kopule rozdzielają strukturę zależności od rozkładów brzegowych. Z jednej strony pozwala to na modelowanie zgodnych zmian cen w okresach wzrostów lub spadków, a z drugiej strony pozwala na kształtowanie rozkładów pojedynczych walorów i uwzględnienie takich cech jak skośność czy leptokurtyczność. Z uwagi na tematykę pracy doktorskiej, rozważony zostanie specyficzny sposób tworzenia takiego rozkładu wielowymiarowego, model w pełni parametryczny, w którym określone zostają parametry gęstości kopuli i parametry gęstości brzegowych. Jeśli inwestor zdecyduje się na użycie modelu w pełni parametrycznego, rozkłady brzegowe powinny zostać prawidłowo dobrane. Wskazanie niewłaściwych rozkładów brzegowych, między innymi, może mieć negatywny wpływ na oszacowania parametrów kopuli (Kim i in., 2007). W tym rozdziale zostaną przedstawione wybrane do badań jednowymiarowe bezwarunkowe rozkłady stóp zwrotu oraz kopule, które wybrano do utworzenia rozkładów dwuwymiarowych.

2.1. Wybrane jednowymiarowe bezwarunkowe rozkłady stóp zwrotu

2.1.1. Uzasadnienie wyboru rozkładów jednowymiarowych

Początki badań nad rozkładami stóp zwrotu sięgają 1900 roku. Louis Bachelier zauważył wówczas, że zgodnie z Centralnym Twierdzeniem Granicznym rozkłady stóp zwrotu w długim okresie czasu powinny mieć rozkład normalny (Haas i Pigorsch, 2009, s. 3406). Koncepcja zakładająca normalność rozkładów stóp zwrotu przeważała do roku 1963, kiedy to Mandelbrot zaobserwował grube ogony rozkładów logarytmów cen bawełny (Mandelbrot, 1963). Obecnie uważa się, że rozkłady stóp zwrotu z okresów krótszych niż miesiąc, powinny charakteryzować się grubymi ogonami, leptokurtycznością oraz niezerową skośnością (asymetrią) (Doman, M. i Doman, R., 2009, s. 24-27). Rozkłady pozwalające na uchwycenie tych cech powinny mieć co

najmniej cztery parametry (Cont, 2001, s. 226): parametr określający położenie, parametr określający skalę, parametr określający grubość ogonów oraz parametr pozwalający na sterowanie asymetrią, gdy grubość lewego i prawego ogona różni się. Większą elastyczność w określaniu rozkładu zapewniają rozkłady z pięcioma parametrami, takie jak uogólniony rozkład hiperboliczny (*GH*), które posiadają dodatkowy parametr sterujący kształtem rozkładu. Rozkłady z pięcioma parametrami wymagają jednak zbyt czasochłonnych obliczeń, stąd badania zostały zawężone do wybranych rozkładów z czterema parametrami, które często są stosowane do opisu rozkładu stóp zwrotu: rozkładu hiperbolicznego (*Hyp*), normalnego odwrotnego rozkładu Gaussowskiego (*NIG*), skośnego uogólnionego rozkładu błędu (*SGED*) oraz skośnego rozkładu *t*-Studenta (*ST*). Ze względu na historyczne znaczenie rozkładu normalnego (*N*), rozkład ten również został uwzględniony w badaniach w celach porównawczych. W tabeli 2.1.1. przedstawiono wybrane przykłady zastosowania wymienionych wyżej rozkładów w badaniach empirycznych. W tabeli uwzględniono również rozkład *GH*, ponieważ rozkłady *Hyp* oraz *NIG* są jego szczególnymi przypadkami.

Tabela 2.1.1. Przykłady zastosowania rozkładów *GH*, *Hyp*, *N*, *NIG*, *SGED* oraz *ST* w badaniach empirycznych

Autor	Rozkład	Cel zastosowania rozkładu	Typ instrumentu finansowego	Okres	Wielkość próby*
Küchler, Neumann, Sørensen i Streller (1999)	<i>Hyp</i>	Opis rozkładów stóp zwrotu	Akcje (30 akcji wchodzących w skład indeksu DAX), indeksy (DAX, FAZ)	03.01.1974-31.03.1992	4568
Prause (1999)	<i>GH</i> <i>Hyp</i> <i>NIG</i>	Opis rozkładów stóp zwrotu	Akcje (6 akcji notowanych na rynku niemieckim)	01.01.1988-24.05.1994	1598
			Indeksy (NYSE Industrial Index, NYSE Finance Index, NYSE Composite Index, NYSE Transport Index, NYSE Utility Index)	02.01.1990-29.11.1996	1746
Eberlein i Prause (2002)	<i>GH</i> <i>Hyp</i> <i>NIG</i>	Opis rozkładów stóp zwrotu	Indeks (DAX)	15.12.1993-26.11.1997	992
Komunjer (2007)	<i>SGED</i>	Opis rozkładów innowacji w modelach GARCH	Akcje (Microsoft), Indeksy (S&P500, NASDAQ), kursy walut (BP/USD)	02.01.1990-31.12.2002	-

Autor	Rozkład	Cel zastosowania rozkładu	Typ instrumentu finansowego	Okres	Wielkość próby*
Theodossiou (2008)	<i>ST</i>	Opis rozkładów stóp zwrotu	Indeksy (S&P500, TSE300, Topix), kursy walut (CAD/USD, JY/USD)	Wybrane podokresy lat 1973-1995	-
Cizek, Härdle i Weron (2011)	<i>Hyp NIG</i>	Opis rozkładów stóp zwrotu	Indeksy (DJIA, WIG20)	03.01.2000-31.12.2009	2494
Cizek, Härdle i Weron (2011)	<i>NIG ST</i>	Szacowanie ryzyka	Indeks (DAX)	28.09.1998-24.10.2008	2630
			Indeks (DAX)	12.11.2007-24.10.2008	250
Czapkiewicz i Basiura (2014)	<i>ST</i>	Opis rozkładów innowacji w modelach GARCH	Indeksy (WIG, WIG20, ATX, BEL20, SOFIX, TSX, HSI, PX, HEX, CAC40, DAX, BUX, NIKKEI, OSE, BET, RTS, SAX, KOSPI, IBEX, SMI, AEX, FTM, DJIA, ISE)	06.2003-03.2012 (podzielono na cztery podokresy)	-
Broda, Krause i Paoletta (2018)	<i>N</i>	Opis rozkładów innowacji w modelach GARCH	Akcje (29 akcji wchodzących w skład DJIA w lipcu 2015 r., z badania wyłączono Visa Inc.)	03.01.2005-31.12.2014	1000 (przesuwane przedziały estymacji)
Guegan i Hassani (2018)	<i>GH</i>	Szacowanie ryzyka	Dane reprezentujące ryzyko operacyjne	2009-2014	98 082
Pasieczna i Szydłowska (2021)	<i>N</i>	Szacowanie ryzyka	Akcje (BNP Paribas, Credit Agricole, Commerzbank, Deutsche Bank)	2002-2019	125, 250, 500 (przesuwane przedziały estymacji)
Duda, Gurgul i Syrek (2022)	<i>ST</i>	Opis rozkładów innowacji w modelach GARCH	Indeksy (CAC40, DAX30, DJIA, FTSE250, WIG20)	2006-2017, 2018-2021	-

* Wielkość próby podano tylko w tych przypadkach, gdy została ona podana w tekście opracowania.

Źródło: opracowanie własne

W kolejnych podrozdziałach zostaną opisane wybrane do badań rozkłady prawdopodobieństwa. Rozkłady zostaną przedstawione w parametryzacjach, które były stosowane podczas obliczeń.

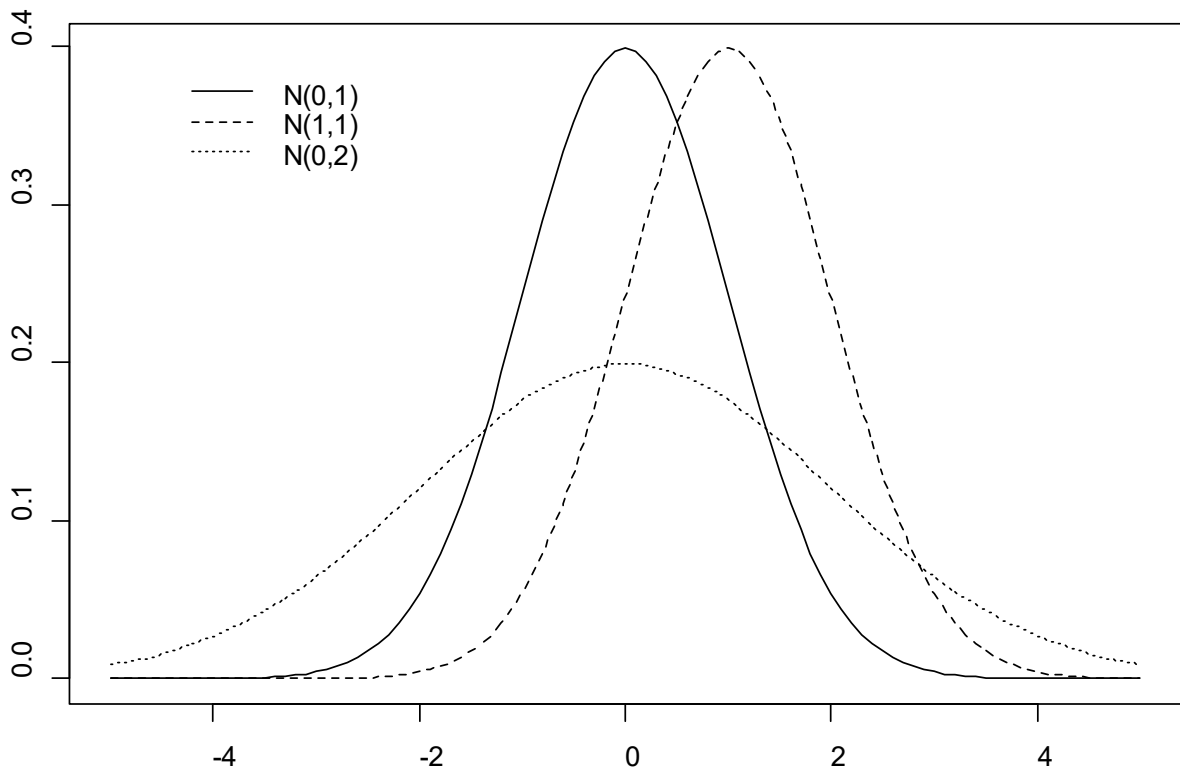
2.1.2. Rozkład normalny (*N*)

Rozkład normalny został odkryty w 1733 r. przez francuskiego uczonego Abrahama de Moivre'a, w roku 1774 został wspomniany przez kolejnego francuskiego matematyka Pierre'a Simon'a de Laplace'a, ostatecznie jednak często przypisuje się jego odkrycie Carlowi Friedrichowi Gaussowi, który opisał rozkład normalny w 1809 roku.

Funkcja gęstości rozkładu normalnego ma postać:

$$f_N(x; \mu; \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (2.1.2.1)$$

Parametry określające rozkład normalny, $\mu \in \mathbb{R}$ i $\sigma \in \mathbb{R}^+$, są parametrami określającymi odpowiednio położenie oraz skalę. Przykładowe wykresy funkcji gęstości rozkładu normalnego przedstawiono na rysunku 2.1.1.



Rysunek 2.1.1. Funkcja gęstości rozkładu normalnego $N(\mu; \sigma^2)$

Źródło: opracowanie własne

Jeśli zmienna losowa $X \sim N(\mu; \sigma^2)$, to zmienna losowa $Y = (X - \mu)/\sigma$ ma rozkład $N(0; 1)$ nazywany rozkładem normalnym standaryzowanym. Jeśli zmienna losowa $X \sim N(\mu; \sigma^2)$, to zmienna losowa $Y = e^X$ ma rozkład logarymiczno-normalny $LN(\mu; \sigma^2)$, a zmienna $\ln Y$ ma rozkład $N(\mu; \sigma^2)$. Zmienna o rozkładzie logarymiczno-normalnym przyjmuje wartości nieujemne i często służy do opisywania zmian kursów akcji.

Rozkład normalny posiada wszystkie momenty centralne. Wartość oczekiwana, wariancja, skośność i kurtoza są określone następująco:

$$E(X) = \mu, \quad (2.1.2.2)$$

$$Var(X) = \sigma^2, \quad (2.1.2.3)$$

$$S(X) = 0, \quad (2.1.2.4)$$

$$K(X) = 3. \quad (2.1.2.5)$$

Jest to rozkład symetryczny o cienkich ogonach, które zanikają szybciej niż w tempie wykładniczym. W badaniach estymatory wartości oczekiwanej oraz odchylenia standardowego zostały wyznaczone metodą największej wiarygodności, za pośrednictwem funkcji *nFit* z pakietu *fBasics* w programie R.

2.1.3. Uogólniony rozkład hiperboliczny (GH)

Rozkład hiperboliczny i normalny odwrotny rozkład Gaussowski są przypadkami szczególnymi uogólnionego rozkładu hiperbolicznego. Według Cruza, Petersa i Shevchenki (2015, s. 350-351), rozkład hiperboliczny jako pierwszy opisał Etienne Halphen (Halphen 1941, Halphen 1953). Cizek, Härdle i Weron (2005, s. 36) również twierdzą, że początki rozkładu hiperbolicznego sięgają lat 40-tych XX wieku, jednak nie podają, kto jako pierwszy opisał ten rozkład. W obydwu pozycjach książkowych zaznaczono jednak, że rozkład hiperboliczny został ponownie opisany w latach 70-tych XX wieku przez Barndorffa-Nielsen (1977), który w tej samej pracy jako pierwszy opisał uogólniony rozkład hiperboliczny. Pierwsze zastosowania uogólnionego rozkładu hiperbolicznego do analizy danych finansowych znalazły się w pracy Eberleina i Kellera z roku 1995. Uogólniony rozkład hiperboliczny pozwala na modelowanie średnio-ciężkich ogonów rozkładów i jest często stosowany w finansach (Haas i Pigorsch, 2009, s. 3408).

Jeśli zmienna losowa ma uogólniony rozkład hiperboliczny, co oznacza się poprzez zapis $X \sim GH(\mu; \delta; \alpha; \beta; \lambda)$, to jej funkcja gęstości ma postać (M. Fischer, 2011, s. 589–590):

$$f_{GH}(x; \mu; \delta; \alpha; \beta; \lambda) = \frac{\left[(\sqrt{\alpha^2 - \beta^2})^\lambda (\sqrt{\delta^2 + (x - \mu)^2})^{\lambda - \frac{1}{2}} K_{\lambda - \frac{1}{2}}(\alpha \sqrt{\delta^2 + (x - \mu)^2}) \right]}{\left[\sqrt{2\pi} \alpha^{\lambda - \frac{1}{2}} \delta^\lambda K_\lambda(\delta \sqrt{\alpha^2 - \beta^2}) e^{-\beta(x - \mu)} \right]}, \quad (2.1.3.1)$$

gdzie K_λ oznacza zmodyfikowaną funkcję Bessela trzeciego rodzaju:

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty t^{\lambda - 1} e^{-\frac{1}{2}x(t + t^{-1})} dt. \quad (2.1.3.2)$$

Parametry uogólnionego rozkładu hiperbolicznego, $\mu \in \mathbb{R}$, $\delta \in \mathbb{R}_0^+$, $\alpha \in \mathbb{R}^+$, $\beta \in (-\alpha, \alpha)$ oraz $\lambda \in \mathbb{R}$ oznaczają odpowiednio parametr położenia, parametr skali, parametr spłaszczenia, parametr skośności oraz parametr, który decydując o postaci całej funkcji gęstości, wyznacza

podklasy rozkładu. Pomiędzy parametrami rozkładu zachodzą następujące zależności (Haas i Pigorsch, 2009, s. 3417):

$$\begin{aligned} \delta &\geq 0 \text{ oraz } |\beta| < \alpha, \text{ gdy } \lambda > 0, \\ \delta &> 0 \text{ oraz } |\beta| < \alpha, \text{ gdy } \lambda = 0, \\ \delta &> 0 \text{ oraz } |\beta| \leq \alpha, \text{ gdy } \lambda < 0. \end{aligned} \tag{2.1.3.3}$$

Rozkład jest lewostronnie skośny dla $\beta < 0$, prawostronnie skośny dla $\beta > 0$, a dla $\beta = 0$ rozkład jest symetryczny. Do najbardziej popularnych podklas uogólnionego rozkładu hiperbolicznego należą rozkład hiperboliczny i normalny odwrotny rozkład gaussowski (Barndorff-Nielsen i Stelzer, 2005, s. 619). Dla $\lambda = 1$ otrzymujemy rozkład hiperboliczny, a dla $\lambda = -1/2$ normalny odwrotny rozkład gaussowski. Rozkłady te zostaną krótko omówione w dwóch kolejnych podrozdziałach. Do wyznaczenia rozkładów granicznych uogólnionego rozkładu hiperbolicznego zazwyczaj stosuje się alternatywną parametryzację (Haas i Pigorsch, 2009, s. 3418), w której parametry α i β zostają zastąpione parametrami χ oraz ξ , określonymi następująco:

$$\chi = \frac{\beta}{\alpha \sqrt{1 + \delta \sqrt{\alpha^2 - \beta^2}}}, \tag{2.1.3.4}$$

$$\xi = \frac{1}{\sqrt{1 + \delta \sqrt{\alpha^2 - \beta^2}}}. \tag{2.1.3.5}$$

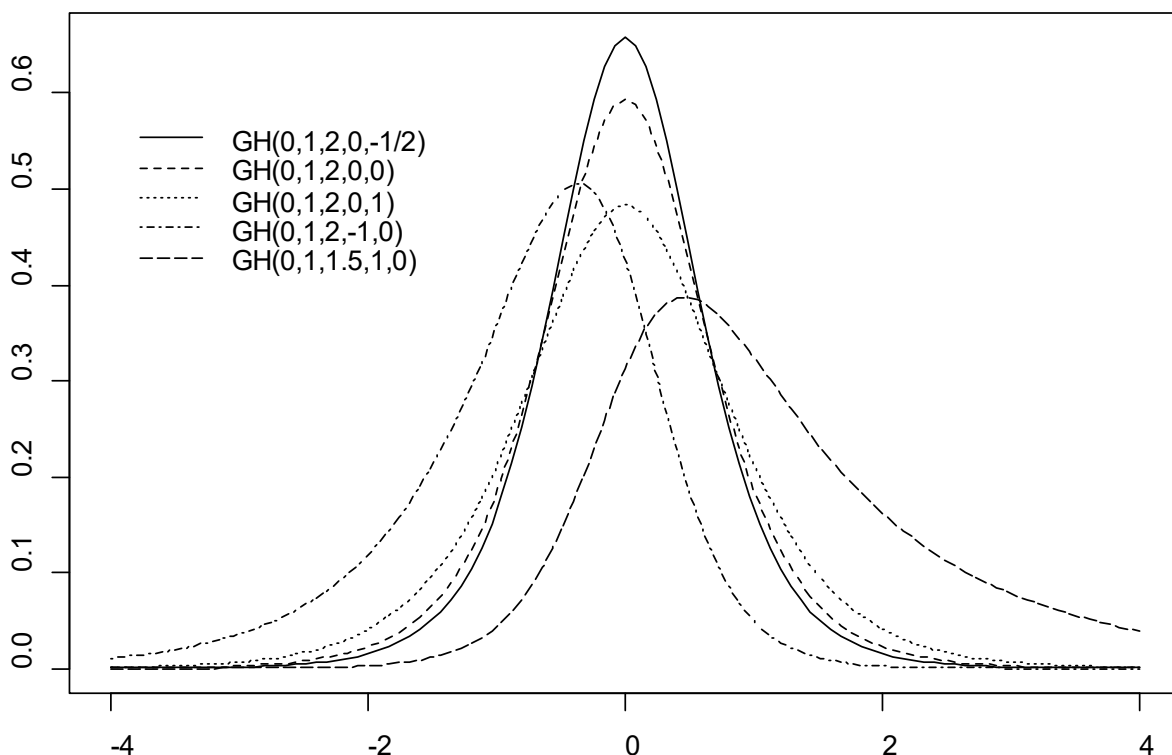
Parametry te spełniają zależność $0 < |\chi| < \xi < 1$, stąd wartości, które może przyjmować uporządkowana para (χ, ξ) , leżą wewnątrz trójkąta ograniczonego punktami $(-1, 1)$, $(1, 1)$ i $(0, 0)$ (Eberlein i Hammerstein, 2004, s. 223). Rozkłady graniczne uogólnionego rozkładu hiperbolicznego występują wtedy, gdy uporządkowana para (χ, ξ) przyjmuje wartości znajdujące się na brzegach trójkąta. Zestawienie rozkładów granicznych, opracowane na podstawie pracy Eberleina i Hammersteina (2004), przedstawiono w tabeli 2.1.2. Rozkład Laplace'a jest przypadkiem szczególnym, a rozkład gamma przypadkiem granicznym rozkładu Variance-Gamma (Gaunt, 2014, s. 3); rozkład Cauchy'ego jest przypadkiem szczególnym rozkładu t-Studenta (Magiera, 2002, s. 76); rozkład wykładniczy jest przypadkiem szczególnym rozkładu gamma (Magiera, 2002, s. 47); rozkłady gamma oraz odwrotny gamma są przypadkami szczególnymi uogólnionego odwrotnego rozkładu gaussowskiego (Scott i in., 2011, s. 462).

Tabela 2.1.2. Zestawienie rozkładów granicznych dla rozkładu GH

Rozkład graniczny	Parametry		
	$ \chi $	ξ	λ
Variance-Gamma	$[0, 1)$	1	$(0, \infty)$
Laplace'a	$[0, 1)$	1	1
Skośny rozkład t -Studenta	$[0, 1)$	1	$(-\infty, 0)$
Cauchy'ego	$[0, 1)$	1	$-1/2$
Uogólniony odwrotny rozkład gaussowski	$ \chi = \xi$	$(0, 1)$	\mathbb{R}
Gamma	1	1	$(0, \infty)$
Wykładniczy	1	1	1
Odwrotny gamma	1	1	$(-\infty, 0)$
Normalny	0	0	\mathbb{R}

Źródło: opracowanie własne na podstawie pracy Eberleina i Hammerstein'a (2004)

Wykresy funkcji gęstości rozkładu GH , dla różnych wartości parametrów α, β, λ , przedstawiono na rysunku 2.1.2.



Rysunek 2.1.2. Funkcja gęstości rozkładu $GH(\mu = 0; \delta = 1; \alpha; \beta; \lambda)$ dla różnych wartości parametrów α, β, λ

Źródło: opracowanie własne

Uogólniony rozkład hiperboliczny posiada wszystkie momenty centralne. Wartość oczekiwana i wariancja wynoszą (Scott i in., 2011, s. 467):

$$E(X) = \mu + \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}} \cdot \frac{K_{\lambda+1}(\delta\sqrt{\alpha^2 - \beta^2})}{K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})}, \quad (2.1.3.6)$$

$$\begin{aligned} Var(X) = \sigma^4\beta^2 \left(\delta\sqrt{\alpha^2 - \beta^2}\right)^{-2} & \left[\frac{K_{\lambda+2}(\delta\sqrt{\alpha^2 - \beta^2})}{K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})} - \left(\frac{K_{\lambda+1}(\delta\sqrt{\alpha^2 - \beta^2})}{K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})} \right)^2 \right] \\ & + \sigma^2 \left(\delta\sqrt{\alpha^2 - \beta^2}\right)^{-1} \frac{K_{\lambda+1}(\delta\sqrt{\alpha^2 - \beta^2})}{K_{\lambda}(\delta\sqrt{\alpha^2 - \beta^2})}. \end{aligned} \quad (2.1.3.7)$$

Autorka nie odnalazła w literaturze pracy, w której podano wzory na skośność i kurtozę rozkładu GH . Wzory na trzeci i czwarty moment centralny, które pozwalają na wyznaczenie skośności i kurtozy, można odnaleźć np. w pracy Scotta i in. (2011, s. 467), przy czym sami autorzy zaznaczają, że wzory są skomplikowane i w związku z tym rzadko są przytaczane w pracach naukowych. Z tych powodów w niniejszej pracy wzory na momenty centralne nie zostaną podane.

Ogony uogólnionego rozkładu hiperbolicznego są cięższe od ogonów rozkładu normalnego, zmieniają się w tempie potęgowo-wykładniczym i są określane jako średnio ciężkie. Kształt ogonów rozkładu GH można opisać następująco (Cruz i in. 2015, s. 343):

$$f_X(x; \mu = 0; \delta; \alpha; \beta; \lambda) \sim C|x|^{\lambda-1}e^{[(\beta-\alpha)x]}, x \rightarrow \infty, \quad (2.1.3.8)$$

gdzie C oznacza pewną stałą.

2.1.3.1. Rozkład hiperboliczny (*Hyp*)

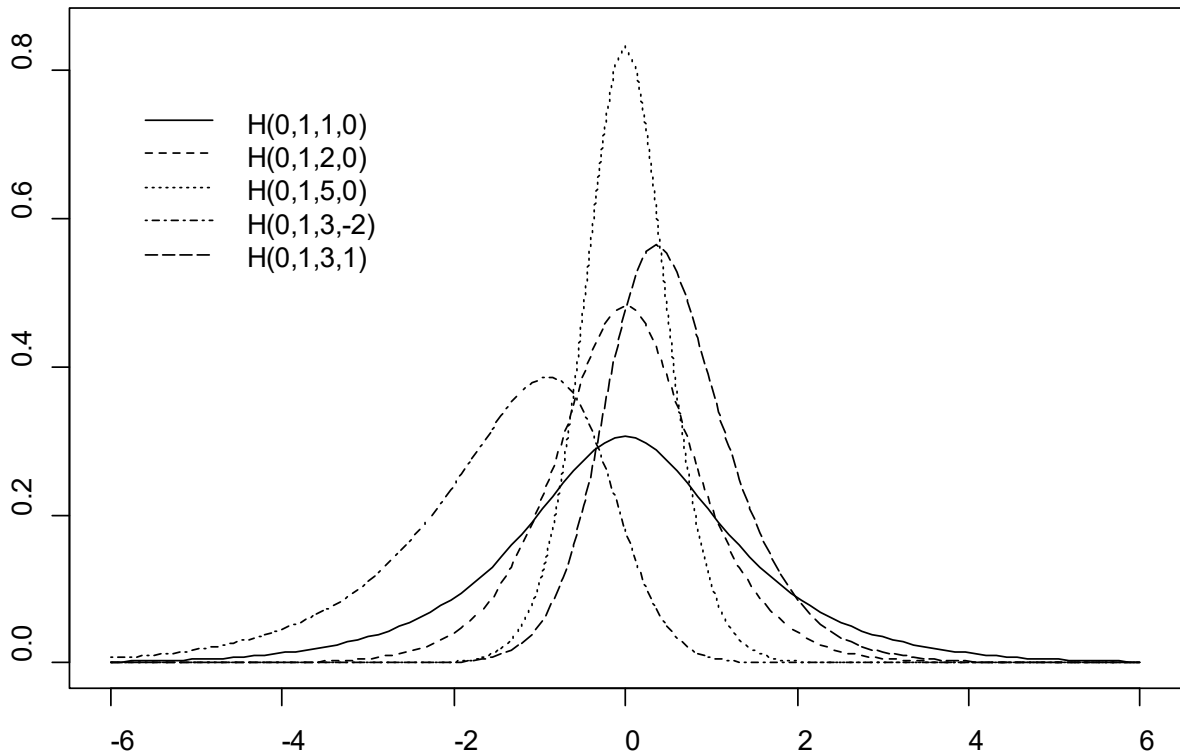
Rozkład hiperboliczny (*Hyp*) jest szczególnym przypadkiem uogólnionego rozkładu hiperbolicznego dla parametru $\lambda = 1$, tzn. gdy $X \sim GH(\mu; \delta; \alpha; \beta; 1)$.

Jeśli zmienna losowa ma rozkład hiperboliczny, $X \sim Hyp(\mu; \delta; \alpha; \beta)$, to jej funkcja gęstości ma postać (Prause, 1999, s. 3):

$$f_{Hyp}(x; \mu; \delta; \alpha; \beta) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\delta\alpha K_1(\delta\sqrt{\alpha^2 - \beta^2})} e^{(-\alpha\sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu))}. \quad (2.1.3.1.1)$$

Parametry rozkładu *Hyp* mają to samo znaczenie, co w rozkładzie GH , natomiast K_1 oznacza zmodyfikowaną funkcję Bessela trzeciego rodzaju stopnia pierwszego. Rozkłady graniczne rozkładu hiperbolicznego można odczytać z tabeli 2.1.2, przyjmując, że $\lambda = 1$. W zależności od

przyjętych wartości parametrów χ oraz ξ , rozkładami granicznymi mogą być: rozkład Laplace'a, uogólniony odwrotny rozkład gaussowski, rozkład wykładniczy oraz rozkład normalny. Wykres logarytmu funkcji gęstości rozkładu hiperbolicznego ma kształt hiperboli. Wykresy funkcji gęstości rozkładu hiperbolicznego dla różnych wartości parametrów α i β przedstawiono na rysunku 2.1.3.



Rysunek 2.1.3. Funkcja gęstości rozkładu $Hyp(\mu = 0; \delta = 1; \alpha; \beta)$ dla różnych wartości parametrów α i β

Źródło: opracowanie własne

Aby uzyskać wzory na momenty rozkładu hiperbolicznego, należy podstawić $\lambda = 1$ we wzorach na momenty uogólnionego rozkładu hiperbolicznego (Scott i in., 2011, s. 469). W związku z tym, aby uzyskać wzory na wartość oczekiwaną i wariancję rozkładu hiperbolicznego, wystarczy podstawić $\lambda = 1$ do wzorów na wartość oczekiwaną i wariancję uogólnionego rozkładu hiperbolicznego (odpowiednio wzory 2.1.3.6 oraz 2.1.3.7). Wartość oczekiwana i wariancja rozkładu hiperbolicznego wynoszą:

$$E(X) = \mu + \frac{\beta\delta}{\sqrt{\alpha^2 - \beta^2}} \cdot \frac{K_2(\delta\sqrt{\alpha^2 - \beta^2})}{K_1(\delta\sqrt{\alpha^2 - \beta^2})}, \quad (2.1.3.1.2)$$

$$\begin{aligned}
\text{Var}(X) = \sigma^4 \beta^2 \left(\delta \sqrt{\alpha^2 - \beta^2} \right)^{-2} & \left[\frac{K_3(\delta \sqrt{\alpha^2 - \beta^2})}{K_1(\delta \sqrt{\alpha^2 - \beta^2})} - \left(\frac{K_2(\delta \sqrt{\alpha^2 - \beta^2})}{K_1(\delta \sqrt{\alpha^2 - \beta^2})} \right)^2 \right] \\
& + \sigma^2 \left(\delta \sqrt{\alpha^2 - \beta^2} \right)^{-1} \frac{K_2(\delta \sqrt{\alpha^2 - \beta^2})}{K_1(\delta \sqrt{\alpha^2 - \beta^2})}.
\end{aligned} \tag{2.1.3.1.3}$$

Ogony rozkładu hiperbolicznego zanikają w tempie wykładniczym i są określane jako cienie. W przeprowadzonych badaniach parametry rozkładu *Hyp* zostały wyznaczone metodą największej wiarygodności, zaimplementowaną w funkcji *hyperbFit* pakietu *GeneralizedHyperbolic* w programie R. Domyślną metodą optymalizacji w tej funkcji jest metoda Nelder-Mead. Gdy działanie funkcji kończyło się błędem, rozwiązania poszukiwano metodą *nlm*, oznaczającą nieliniową metodę optymalizacji Newtona.

2.1.3.2. Normalny odwrotny rozkład gaussowski (*NIG*)

Normalny odwrotny rozkład gaussowski jest szczególnym przypadkiem uogólnionego rozkładu hiperbolicznego dla parametru $\lambda = -1/2$, tzn. gdy $X \sim GH(\mu; \delta; \alpha; \beta; -1/2)$.

Jeśli zmienna losowa ma normalny odwrotny rozkład gaussowski, $X \sim NIG(\mu; \delta; \alpha; \beta)$, to jej funkcja gęstości ma postać (Prause, 1999, s. 3):

$$f_{NIG}(x; \mu; \delta; \alpha; \beta) = \frac{\alpha \delta K_1(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{\pi \sqrt{\delta^2 + (x - \mu)^2}} e^{(\delta \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu))}. \tag{2.1.3.2.1}$$

Parametry rozkładu *NIG* mają to samo znaczenie, co w rozkładzie *GH*, a K_1 oznacza zmodyfikowaną funkcję Bessela trzeciego rodzaju stopnia pierwszego. Rozkłady graniczne rozkładu *NIG* można odczytać z tabeli 2.1.2, przyjmując, że $\lambda = -1/2$. W zależności od przyjętych wartości parametrów χ oraz ξ , rozkładami granicznymi mogą być: rozkład Cauchy'ego, uogólniony odwrotny rozkład gaussowski, rozkład odwrotny gamma oraz rozkład normalny. Wykresy funkcji gęstości rozkładu *NIG* dla różnych wartości parametrów α i β przedstawione zostały na rysunku 2.1.4.

Wartość oczekiwana, wariancja, skośność i kurtoza rozkładu *NIG* wynoszą (Brunlid, 2007, s. 3):

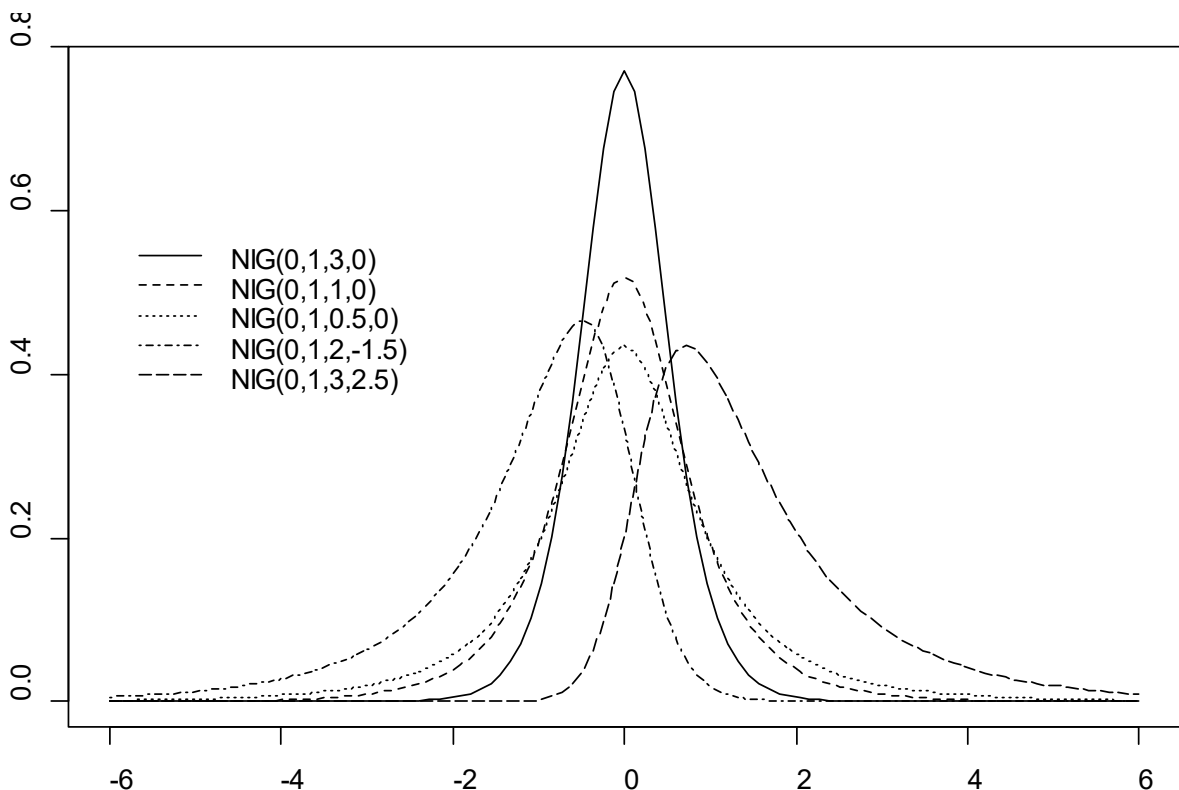
$$E(X) = \mu + \frac{\beta \delta}{\sqrt{\alpha^2 - \beta^2}}, \tag{2.1.3.2.2}$$

$$\text{Var}(X) = \frac{\delta\alpha^2}{(\alpha^2 - \beta^2)^{3/2}}, \quad (2.1.3.2.3)$$

$$S(X) = \frac{3\beta}{\alpha\sqrt{\delta}(\alpha^2 - \beta^2)^{1/4}}, \quad (2.1.3.2.4)$$

$$K(X) = \frac{3\left(1 + 4\left(\frac{\beta}{\alpha}\right)^2\right)}{\delta\sqrt{\alpha^2 - \beta^2}}. \quad (2.1.3.2.5)$$

Ogony rozkładu *NIG* są określane jako średnio-ciężkie, są one grubsze od rozkładu normalnego. W przeprowadzonych badaniach parametry rozkładu *NIG* zostały wyznaczone metodą największej wiarygodności, zaimplementowaną w funkcji *nigFit* pakietu *GeneralizedHyperbolic* w programie R. Podobnie jak w przypadku funkcji *hyperbFit*, domyślną metodą optymalizacji funkcji *nigFit* jest metoda Nelder-Meada i gdy działanie funkcji kończyło się błędem, rozwiązania poszukiwano metodą *nlm*, oznaczającą nieliniową metodę optymalizacji Newtona.



Rysunek 2.1.4. Funkcja gęstości rozkładu $NIG(\mu = 0; \delta = 1; \alpha; \beta)$ dla różnych wartości parametrów α i β

Źródło: opracowanie własne

2.1.4. Skośny uogólniony rozkład błędu (SGED)

W niniejszej pracy zostanie przedstawiony skośny uogólniony rozkład błędu wprowadzony przez Fernández, Osiewalskiego i Steela (1995). Odniesienia do skośnego uogólnionego rozkładu błędu w innych parametryzacjach można odnaleźć w artykule Zhu i Zinde-Walsh (2009).

Zmienna losowa X typu ciągłego ma skośny uogólniony rozkład błędu z parametrami położenia $\mu \in \mathbb{R}$, skali $\sigma \in \mathbb{R}_0^+$, kształtu $\alpha \in \mathbb{R}_0^+$ oraz skośności $\kappa \in \mathbb{R}_0^+$, gdy jej funkcję gęstości opisuje wzór (Ayebo i Kozubowski, 2004):

$$f_{SGED}(x; \mu; \sigma; \alpha; \kappa) = \frac{\alpha}{\sigma \Gamma\left(\frac{1}{\alpha}\right)} \frac{\kappa}{1 + \kappa^2} e^{\left(\frac{-\kappa^\alpha}{\sigma^\alpha} [(x-\mu)^+]^\alpha - \frac{1}{\sigma^\alpha \kappa^\alpha} [(x-\mu)^-]^\alpha\right)}, \quad (2.1.4.1)$$

gdzie:

$$u^+ = \begin{cases} u & \text{gdy } u \geq 0, \\ 0 & \text{gdy } u < 0, \end{cases} \quad \text{i} \quad u^- = \begin{cases} -u & \text{gdy } u \leq 0, \\ 0 & \text{gdy } u > 0 \end{cases} \quad (2.1.4.2)$$

oraz Γ oznacza niekompletną funkcję gamma:

$$\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt, \quad a > 0, x > 0. \quad (2.1.4.3)$$

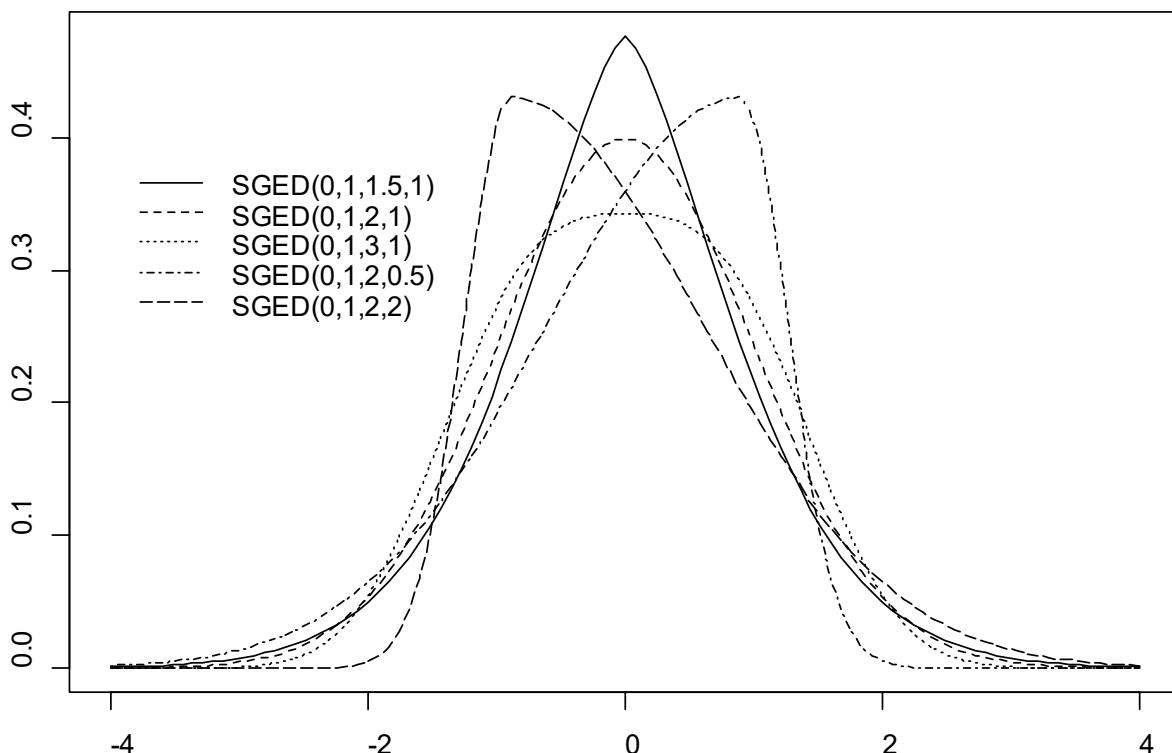
Gdy $\kappa = 1$, otrzymywany jest symetryczny rozkład potęgowo-wykładniczy. Gdy $\kappa \neq 1$ oraz $\alpha = 1$, otrzymywany jest rozkład Laplace'a. Dla $\alpha = 2$ otrzymywany jest skośny rozkład normalny. Wykresy funkcji gęstości rozkładu *SGED* dla różnych wartości parametrów κ oraz α przedstawiono na rysunku 2.1.5.

Wartość oczekiwana i wariancja skośnego uogólnionego rozkładu błędu wynoszą (Ayebo i Kozubowski, 2004):

$$E(X) = \mu + \sigma \left(\frac{1}{\kappa} - \kappa\right) \frac{\Gamma\left(\frac{2}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)}, \quad (2.1.4.4)$$

$$Var(X) = \sigma^2 \frac{\Gamma\left(\frac{3}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)} \frac{1 + \kappa^6}{\kappa^2(1 + \kappa^2)} - \sigma^2 \frac{\Gamma^2\left(\frac{2}{\alpha}\right)}{\Gamma^2\left(\frac{1}{\alpha}\right)} \frac{(1 - \kappa^2)^2}{\kappa^2}. \quad (2.1.4.5)$$

Wzory na skośność i kurtozę są skomplikowane i nie zostaną podane w niniejszej pracy, można je odnaleźć w pracy Ayebo i Kozubowskiego (2004).



Rysunek 2.1.5. Funkcja gęstości rozkładu $SGED(\mu = 0; \delta = 1; \alpha; \kappa)$ dla różnych wartości parametrów α i κ

Źródło: opracowanie własne

Tempo zanikania ogona uogólnionego rozkładu błędu zależy od parametru α i rośnie wraz ze spadkiem wartości tego parametru. Dla $\alpha < 2$ rozkład posiada ogony grubsze, a dla $\alpha > 2$ cieńsze niż rozkład normalny (Marín i Sucarrat, 2012).

W przeprowadzonych badaniach parametry rozkładu $SGED$ zostały wyznaczone metodą największej wiarygodności, poprzez funkcję *sgedFit* z pakietu *fGarch* w programie R. W związku z licznymi błędami numerycznymi, wartości parametrów zostały ograniczone do przedziałów: $\mu \in [-99; 99]$, $\sigma \in [0,05; 10]$, $\alpha \in [0,1; 3]$, $\kappa \in [0,6; 5]$, podobnie jak w pracy Marína i Sucarrata (2012). W związku z wprowadzonymi ograniczeniami, przyjętą metodą optymalizacyjną była metoda *L-BFGS-B*, z rodziny metod quasi-Newtona.

2.1.5. Skośny rozkład *t*-Studenta (*ST*)

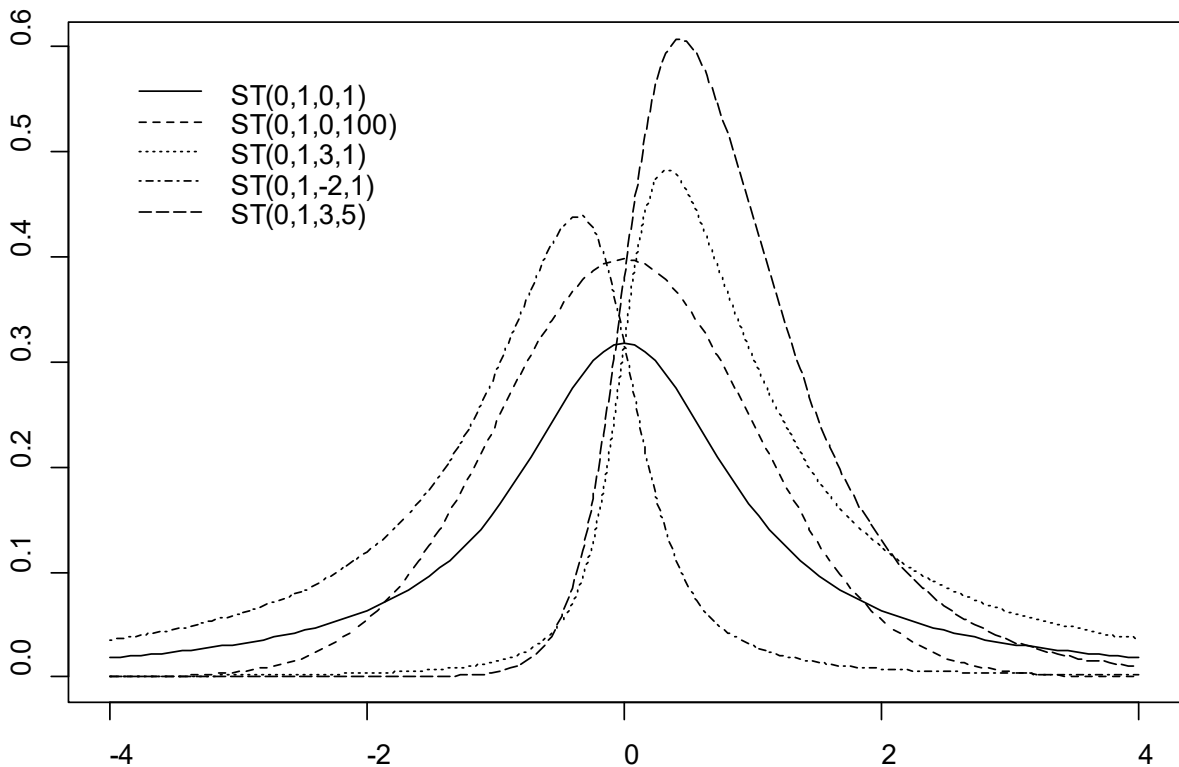
W literaturze zaproponowano kilka rozkładów o nazwie skośny rozkład *t*-Studenta (Haas i Pigorsch, 2009, s. 3421). W niniejszej pracy badaniom będzie podlegał skośny rozkład *t*-Studenta zaproponowany w pracy Azzalini i Capitanio (2003).

Jeśli zmienna losowa X typu ciągłego ma skośny rozkład t -Studenta, $X \sim ST(\xi; \omega^2; \alpha; \nu)$, to jej funkcja gęstości ma postać (Azzalini i Arellano-Valle, 2013, s. 429):

$$f_{ST}(x; \xi; \omega^2; \alpha; \nu) = \frac{2}{\omega} t\left(\frac{x - \xi}{\omega}; \nu\right) T\left(\alpha \frac{x - \xi}{\omega} \sqrt{\frac{\nu + 1}{\nu + Q_x}}; \nu + 1\right), \quad (2.1.5.1)$$

gdzie $\xi \in \mathbb{R}$ jest parametrem położenia, $\omega \in \mathbb{R}^+$ parametrem skali, $\alpha \in \mathbb{R}$ parametrem określającym skośność, natomiast $\nu \in \mathbb{R}^+$ oznaczają stopnie swobody. Ponadto $t(\cdot; \nu)$ oznacza rozkład t -Studenta z ν stopniami swobody, $T(\cdot; \nu + 1)$ oznacza dystrybuantę rozkładu t -Studenta z $\nu + 1$ stopniami swobody oraz $Q_x = \omega^{-2}(x - \xi)^2$.

Dla $\alpha = 0$ otrzymywany jest rozkład t -Studenta, dla $\nu \rightarrow \infty$ rozkład redukuje się do skośnego rozkładu normalnego, a dla $\nu = 1$ otrzymywany jest skośny rozkład Cauchy'ego. Wykresy funkcji gęstości rozkładu ST dla różnych wartości parametrów α oraz ω przedstawiono na rysunku 2.1.6.



Rysunek 2.1.6. Funkcja gęstości rozkładu $ST(\xi = 0; \omega = 1; \alpha; \nu)$ dla różnych wartości parametrów α i ν

Źródło: opracowanie własne

Przyjmijmy teraz, że $z = (x - \xi)/\omega$. Po zastosowaniu następujących podstawień:

$$\delta = \frac{\alpha}{\sqrt{1 + \alpha^2}}, \quad (2.1.5.2)$$

$$b_\nu = \sqrt{\frac{\nu \Gamma((\nu - 1)/2)}{\pi \Gamma(\frac{\nu}{2})}}, \quad (2.1.5.3)$$

$$\sigma_z^2 = \text{Var}(z) = \frac{\nu}{\nu - 2} - (b_\nu \delta)^2, \quad (2.1.5.4)$$

wartość oczekiwaną, wariancję, skośność i kurtozę rozkładu ST określają następujące wzory:

$$E(X) = \xi + \omega b_\nu \delta, \quad (2.1.5.5)$$

$$\text{Var}(X) = \omega^2 \sigma_z^2, \text{ dla } \nu > 2, \quad (2.1.5.6)$$

$$S(X) = \frac{b_\nu \delta}{\sigma_z^{3/2}} \left[\frac{\nu(3 - \delta^2)}{\nu - 3} - \frac{3\nu}{\nu - 2} + 2(b_\nu \delta)^2 \right], \text{ dla } \nu > 3, \quad (2.1.5.7)$$

$$K(X) = \frac{1}{\sigma_z^4} \left[\frac{3\nu^2}{(\nu - 2)(\nu - 4)} - \frac{4(b_\nu \delta)^2 \nu(3 - \delta^2)}{\nu - 3} + \frac{6\nu(b_\nu \delta)^2}{\nu - 2} - 3(b_\nu \delta)^4 \right], \text{ dla } \nu > 4. \quad (2.1.5.8)$$

Obydwa ogony rozkładu zaproponowanego przez Azzalini i Capitanio (2003) można opisać przez wielomiany. Oznacza to, że rozkład ten będzie dobrze opisywał rozkłady z obydwoma grubymi ogonami. Rozkład ten może nie być odpowiedni do opisu rozkładów ze znaczną skośnością, gdy jeden z ogonów nie jest określany jako gruby (Aas i Haff, 2008, s. 2).

Zastosowanie metody największej wiarygodności do estymacji parametru określającego skośność, w przypadku rozkładu ST , powoduje problemy ze zbieżnością estymatora. W związku z tym zastosowano następującą funkcję do wyznaczania estymatorów parametrów rozkładu ST (Azzalini i Arellano-Valle, 2013, s. 422):

$$\ell_p(\theta) = \ell(\theta) - Q, \quad (2.1.5.9)$$

gdzie $\ell(\theta)$ oznacza funkcję wiarygodności dla parametru θ , a Q funkcję kary. Wartość $\hat{\theta}$, która maksymalizuje funkcję $\ell_p(\theta)$, jest estymatorem parametru θ . Funkcja kary w ogólnej postaci jest następująca:

$$Q = c_1 \log(1 + c_2 \alpha^2), \quad (2.1.5.10)$$

gdzie c_1 i c_2 oznaczają pewne stałe. Szczegóły dotyczące metody można znaleźć w pracy Azzalini i Arellano-Valle (2013). W przeprowadzonych badaniach parametry rozkładu ST zostały wyznaczone opisaną w tym paragrafie metodą największej wiarygodności z funkcją kary, zaimplementowaną w funkcji *st.mple* z pakietu *sn* w programie R. Domyślnie stosowano metodę optymalizacji wdrożoną w funkcji *nlminb*, a w przypadku błędu metodę Nelder-Meada.

2.2. Kopule

2.2.1. Uzasadnienie doboru kopuli i wybrane elementy teorii

Powszechnie stosowaną w finansach miarą zależności jest współczynnik korelacji liniowej, jednak powinien on być stosowany jedynie wówczas, gdy łączny rozkład danych jest eliptyczny. Gdy tak nie jest, do opisu zależności można stosować kopule. Podstawowe zalety stosowania kopuli to właśnie możliwość modelowania zależności nieliniowych oraz oddzielenie struktury zależności od rozkładów brzegowych. Kopule zastosowane w badaniach w niniejszej pracy wybrano na podstawie przeglądu następujących pozycji literatury: Cizek i in. (2011), Cruz i in. (2015), Doman (2011), Echaust (2014), Fang, Fang i Kotz (2002), Heilpern (2007), Lee i Lee (2015), Nelsen (1999) oraz Patton (2013). Kopule dobrano tak, aby z jednej strony reprezentowały najważniejsze rodziny kopuli stosowane w badaniach dotyczących struktury zależności na rynkach finansowych, a z drugiej strony pozwalały w różny sposób modelować zależności w ogonach rozkładów. Część wybranych kopuli należy do rodziny kopuli eliptycznych (Gausa i t -Studenta), a część do rodziny kopuli archimedesowych (tu znajdują się kopule: Franka, Clayтона, Gumbela i Joe). Kopule Gausa i Franka nie mają zależności w ogonach, kopula Clayтона ma zależność tylko w dolnym ogonie, kopule Gumbela i Joe mają zależności tylko w górnym ogonie, natomiast kopula t -Studenta ma zależności w obydwu ogonach. Główną wadą kopuli eliptycznych jest to, że modelują symetryczne struktury zależności. Główną wadą wielowymiarowych kopul archimedesowych jest natomiast to, że wszystkie pary powiązanych zmiennych mają taką samą strukturę zależności. Ponadto, generatory kopul archimedesowych mają małą liczbę parametrów. Te dwa ograniczenia powodują, że są one stosowane przede wszystkim do modelowania zależności pomiędzy dwiema zmiennymi (Doman, 2011, s. 25). W dalszej części podrozdziału zostaną przedstawione wybrane elementy teorii dotyczące kopuli: twierdzenie Sklára, pojęcie kopuli, pojęcie rozkładu dwuwymiarowego utworzonego za pomocą kopuli oraz współczynniki zależności w ogonach rozkładów dwuwymiarowych utworzonych za pomocą kopuli.

Kopule, poprzez opisanie struktury zależności pomiędzy rozkładami brzegowymi, umożliwiają ich łączenie w rozkłady wielowymiarowe. Rozważania dotyczące kopul zostaną zawężone do przypadku dwuwymiarowego, ponieważ tylko takie będą rozpatrywane w części empirycznej pracy. Z pojęciem kopuli związane jest twierdzenie Sklára:

Twierdzenie 2.2.1.1 (Heilpern, 2007, s. 16).

1. Niech F będzie dwuwymiarową dystrybuantą z dystrybuantami brzegowymi F_1, F_2 ; wtedy istnieje funkcja łącząca C taka, że dla każdego x, y

$$F(x; y) = C(F_1(x), F_2(y)). \quad (2.2.1.1)$$

Jeśli dystrybuanty F_1, F_2 są ciągłe, to funkcja C jest jedyna. W przeciwnym wypadku jest ona jednoznacznie określona na iloczynie kartezjańskim zbiorów wartości dystrybuant brzegowych.

2. Jeśli C jest dwuwymiarową funkcją łączącą i F_1, F_2 są jednowymiarowymi dystrybuantami, to funkcja F , określona wzorem (2.2.1.1), jest dwuwymiarową dystrybuantą, a F_1, F_2 są jej dystrybuantami brzegowymi.

Zamienną nazwą dla funkcji łączącej występującej w powyższym twierdzeniu jest „kopuła” i ta nazwa będzie stosowana w dalszej części pracy. Kopulę ze wzoru (2.2.1.1) można wyodrębnić następująco:

$$C(u_1; u_2) = F(F_1^{(-1)}(u_1), F_2^{(-1)}(u_2)), \quad (2.2.1.2)$$

gdzie $F_i^{(-1)}$ oznacza funkcję kwantylową dystrybuanty F_i . Wzór (2.2.1.2) pokazuje, jak wyodrębnić wyłącznie strukturę zależności z rozkładu wielowymiarowego. Ze wzoru tego wynika również, że kopule nie zależą od rozkładów brzegowych, tylko od zmiennych u_i , które posiadają rozkład jednostajny na przedziale $[0; 1]$.

Funkcja gęstości dla dwuwymiarowego rozkładu łącznego utworzonego za pomocą kopuli wynosi (Heilpern, 2014, s. 86):

$$f(x_1; x_2) = c(F_1(x_1), F_2(x_2))f_1(x_1)f_2(x_2), \quad (2.2.1.3)$$

gdzie c oznacza gęstość funkcji łączącej C , natomiast f_i gęstość zmiennej losowej X_i .

Zależności w ogonach rozkładów dwuwymiarowych utworzonych za pomocą kopul można zdefiniować za pomocą współczynników λ^L dla ogona dolnego i λ^U dla ogona górnego (Doman, 2011, s. 18; Echaust, 2014, s. 175):

$$\lambda^L = \lim_{u \rightarrow 0^+} P(X_1 \leq F_1^{(-1)}(u) | X_2 \leq F_2^{(-1)}(u)) = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}, \quad (2.2.1.4)$$

$$\lambda^U = \lim_{u \rightarrow 1^-} P(X_1 > F_1^{(-1)}(u) | X_2 > F_2^{(-1)}(u)) = \lim_{u \rightarrow 1^-} \frac{1 + 2u + C(u, u)}{1 - u}. \quad (2.2.1.5)$$

Wartości równe zero wskazują na brak zależności pomiędzy ogonami, a wartości bliskie jedności oznaczają silną zależność między zmiennymi.

W następnych podrozdziałach zostaną zaprezentowane wybrane kopule eliptyczne i archimedesowe, które zastosowano w części empirycznej pracy.

2.2.2. Kopule eliptyczne

2.2.2.1. Rozkłady eliptyczne

Podanie definicji rozkładu eliptycznego wymaga wprowadzenia pojęcia rozkładu sferycznego.

Definicja 2.2.2.1.1 (Heilpern, 2007, s. 59). Wektor losowy $\mathbf{X} = (X_1, \dots, X_n)^T$ ma rozkład sferyczny, jeśli dla każdego przekształcenia ortogonalnego opisanego macierzą $\mathbf{\Gamma} \in \mathbb{R}^{n \times n}$ zachodzi

$$\mathbf{\Gamma}\mathbf{X} =_d \mathbf{X}. \quad (2.2.2.1.1)$$

W powyższej definicji zapis „ $=_d$ ” oznacza równość rozkładów. Macierz przekształcenia ortogonalnego spełnia natomiast warunek $\mathbf{\Gamma}\mathbf{\Gamma}^T = \mathbf{I}_n$, gdzie \mathbf{I}_n oznacza macierz jednostkową o wymiarach $n \times n$.

Definicja 2.2.2.1.2 (Heilpern, 2007, s. 62). Losowy, n -wymiarowy wektor \mathbf{Y} ma rozkład eliptyczny, jeśli istnieje wektor $\boldsymbol{\mu} \in \mathbb{R}^n$, macierz $\boldsymbol{\Sigma} \in \mathbb{R}^{n \times n}$ postaci $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^T$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ oraz sferyczny wektor losowy \mathbf{X} taki, że

$$\mathbf{Y} =_d \boldsymbol{\mu} + \mathbf{A}\mathbf{X}. \quad (2.2.2.1.2)$$

Wektor $\boldsymbol{\mu}$ oznacza wektor położenia, a $\boldsymbol{\Sigma}$ jest symetryczną, nieujemnie określoną macierzą dyspersji. Funkcja charakterystyczna wektora eliptycznego jest postaci (Heilpern, 2007, s. 63):

$$\psi(\mathbf{t}) = E(e^{it^T\mathbf{Y}}) = e^{(it^T\boldsymbol{\mu})} \varphi(\mathbf{t}^T\boldsymbol{\Sigma}\mathbf{t}), \quad (2.2.2.1.3)$$

gdzie $\mathbf{t} = (t_1, \dots, t_n)$, a φ jest funkcją nazywaną generatorem charakterystycznym. Wektor \mathbf{Y} o rozkładzie eliptycznym oznacza się przez $\mathbf{Y} \sim E_n(\boldsymbol{\mu}; \boldsymbol{\Sigma}; \varphi)$. Gęstość wektora eliptycznego, jeśli istnieje, ma postać (Heilpern, 2007, s. 63):

$$f(\mathbf{x}) = \frac{1}{\sqrt{|\boldsymbol{\Sigma}|}} g((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})), \quad (2.2.2.1.4)$$

gdzie g oznacza generator gęstości. Rozkłady brzegowe wektora eliptycznego również są eliptyczne i mają ten sam generator charakterystyczny.

Do najbardziej znanych rozkładów eliptycznych należą wielowymiarowy rozkład normalny oraz wielowymiarowy rozkład t -Studenta. Kopule eliptyczne można wyodrębnić z wielowymiarowych rozkładów eliptycznych na podstawie wzoru (2.2.1.2). Poniżej zostaną przedstawione dwie kopule zastosowane w części empirycznej pracy: kopula Gaussa i t -Studenta.

Rozkłady brzegowe wektora eliptycznego również są rozkładami eliptycznymi tego samego typu, o czym już powyżej wspomniano. Wielowymiarowych rozkładów eliptycznych nie powinno się więc stosować, gdy zmienne mają różne rozkłady brzegowe. Problem ten można rozwiązać poprzez wyznaczenie dystrybuant dowolnych rozkładów brzegowych oraz powiązanie ich kopulami eliptycznymi. Utworzony w ten sposób rozkład wielowymiarowy nosi nazwę metaeliptycznego.

2.2.2.2. Kopula Gaussa

Kopula Gaussa generuje wielowymiarowy standaryzowany rozkład normalny tylko wtedy, gdy rozkłady brzegowe są również standaryzowanymi rozkładami normalnymi. Dwuwymiarowa kopula Gaussa, z macierzą korelacji \mathbf{R} , gdzie $\mathbf{R}_{ij} = \Sigma_{ij}/\sqrt{\Sigma_{ii}\Sigma_{jj}}$, ma postać:

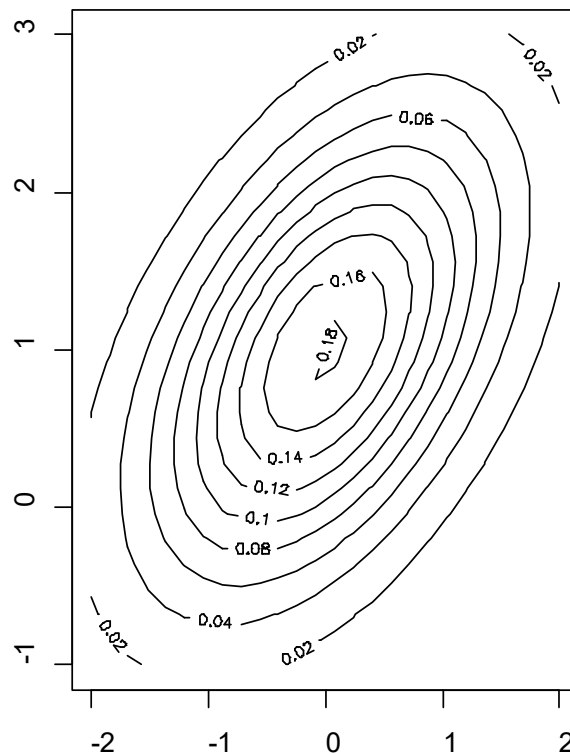
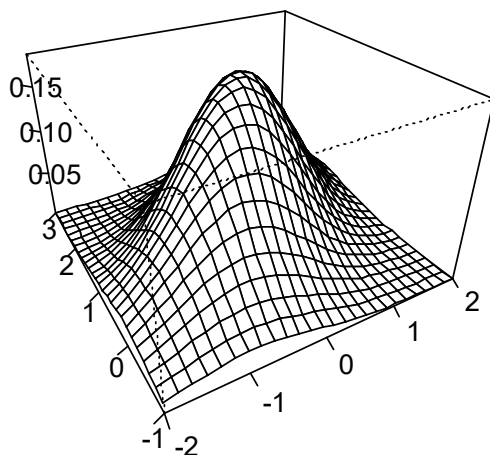
$$C_{\mathbf{R}}^{Ga}(u_1; u_2) = \Phi_{\mathbf{R}}^2(\Phi^{-1}(u_1), \Phi^{-1}(u_2)), \quad (2.2.2.2.1)$$

gdzie $\Phi_{\mathbf{R}}^2$ jest dystrybuantą dwuwymiarowego standaryzowanego rozkładu normalnego, a Φ^{-1} jest odwrotnością dystrybuanty jednowymiarowego standardowego rozkładu normalnego.

W przypadku dwuwymiarowym, gdy $r = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}} = \Sigma_{21}/\sqrt{\Sigma_{22}\Sigma_{11}}$, wzór (2.2.2.2.1) można przedstawić następująco:

$$C_{\mathbf{R}}^{Ga}(u; v) = \frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} e^{\left(\frac{-(s^2-2rs+t^2)}{2(1-r^2)}\right)} ds dt. \quad (2.2.2.2.2)$$

Funkcja (2.2.2.2.2) ma jeden parametr, współczynnik korelacji r . Kopula Gaussa nie posiada zależności w ogonach rozkładu. Wykresy funkcji gęstości dla dwuwymiarowego rozkładu, w którym rozkłady brzegowe połączono kopulą Gaussa, pokazano na rysunku 2.2.1.



Rysunek 2.2.1. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopułą Gaussa ($r = 0,5$)

Źródło: opracowanie własne

2.2.2.3. Kopuła t -Studenta

W przypadku dwuwymiarowego rozkładu t -Studenta z ν stopniami swobody rozkłady brzegowe również są rozkładami t -Studenta z tą samą liczbą stopni swobody. Kopuła t -Studenta, z macierzą korelacji \mathbf{R} oraz liczbą stopni swobody ν ma postać:

$$C_{\nu, \mathbf{R}}^t(u_1; u_2) = t_{\nu, \mathbf{R}}^2(t_{\nu}^{-1}(u_1), t_{\nu}^{-1}(u_2)), \quad (2.2.2.3.1)$$

gdzie $t_{\nu, \mathbf{R}}^2$ jest dystrybuantą dwuwymiarowego rozkładu t -Studenta, a t_{ν}^{-1} jest odwrotnością dystrybuanty jednowymiarowego rozkładu t -Studenta.

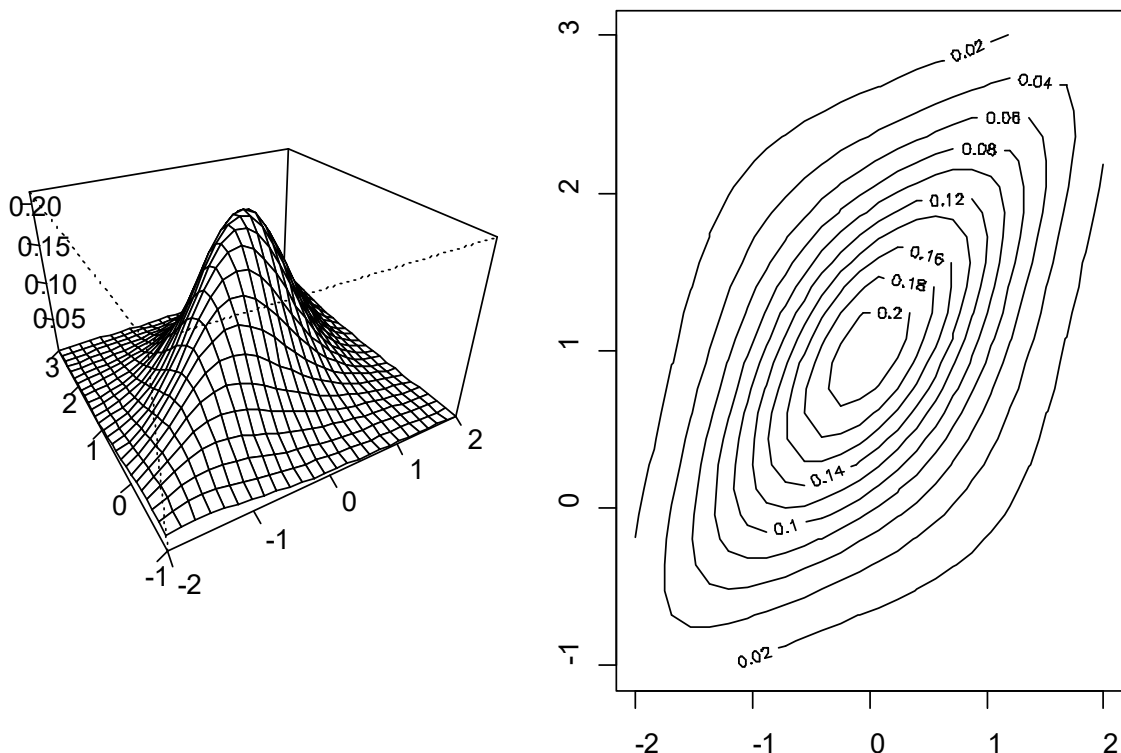
W przypadku dwuwymiarowym, gdzie $r = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}} = \Sigma_{21}/\sqrt{\Sigma_{22}\Sigma_{11}}$, wzór (2.2.2.3.1) można przedstawić następująco:

$$C_{\nu, \mathbf{R}}^t(u; v) = \frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{t_{\nu}^{-1}(u)} \int_{-\infty}^{t_{\nu}^{-1}(v)} \left(1 + \frac{(s^2 - 2rst + t^2)}{\nu(1-r^2)}\right)^{-(\nu+2)/2} ds dt. \quad (2.2.2.3.2)$$

Funkcja (2.2.2.3.2) ma dwa parametry, współczynnik korelacji r oraz liczbę stopni swobody ν . Kopuła t -Studenta posiada symetryczną zależność w dolnym i górnym ogonie rozkładu, gdy $r > -1$. W przypadku dwuwymiarowym, zależność wyraża się wzorem:

$$\lambda^L = \lambda^U = 2t_{\nu+1} \left(-\sqrt{\frac{(\nu+1)(1-r)}{1+r}} \right). \quad (2.2.2.3.3)$$

Zależność w ogonach rozkładu maleje wraz ze wzrostem liczby stopni swobody oraz rośnie wraz ze wzrostem współczynnika korelacji r . Wykresy funkcji gęstości dla rozkładu, w którym rozkłady brzegowe połączono kopułą t -Studenta, pokazano na rysunku 2.2.2.



Rysunek 2.2.2. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopułą t -Studenta ($r = 0,5, \nu = 3$)

Źródło: opracowanie własne

2.2.3. Kopule archimedesowe

2.2.3.1. Wybrane elementy teorii kopuł archimedesowych

Kopule archimedesowe wyznacza się za pomocą funkcji nazywanych generatorami. Generator może być addytywny lub multiplikatywny. W przypadku generatora addytywnego, kopule archimedesowe spełniają warunek (Heilpern, 2007, s. 30–31):

$$C_\varphi(u; v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)), \quad (2.2.3.1.1)$$

gdzie generator $\varphi: [0; 1] \rightarrow [0; \infty]$ jest ściśle malejącą funkcją, spełniającą warunek $\varphi(1) = 0$ oraz $\varphi^{[-1]}: [0; \infty] \rightarrow [0; 1]$ oznacza funkcję pseudoodwrotną generatora zdefiniowaną następująco:

$$\varphi^{[-1]}(u) = \begin{cases} \varphi^{-1}(u) & 0 \leq u \leq \varphi(0) \\ 0 & \varphi(0) \leq u \leq \infty. \end{cases} \quad (2.2.3.1.2)$$

Kopule archimedesowe można również przedstawić za pomocą generatora multiplikatywnego, który ma postać $g(u) = e^{-\varphi(u)}$. Zachodzi wówczas zależność (Heilpern, 2007, s. 34):

$$C(u; v) = g^{[-1]}(g(u)g(v)), \quad (2.2.3.1.3)$$

gdzie g jest rosnąca i spełnia warunki $g(0) = e^{-\varphi(0)}$ i $g(1) = 1$, natomiast funkcja pseudoodwrotna $g^{[-1]}: [0; 1] \rightarrow [0; 1]$ jest zdefiniowana następująco:

$$g^{[-1]}(u) = \begin{cases} 0 & 0 \leq u \leq g(0) \\ g^{-1}(u) & g(0) \leq u \leq 1. \end{cases} \quad (2.2.3.1.4)$$

Kopule archimedesowe umożliwiają opis asymetrycznych zależności w ogonach rozkładów, ale ich użycie jest ograniczone z dwóch powodów – mają zazwyczaj niewiele parametrów (jeden lub dwa), a wszystkie pary zmiennych charakteryzuje ten sam typ zależności (Doman, 2011, s. 25).

2.2.3.2. Rodzina Franka

Rodzina Franka została opisana na podstawie pracy Heilperna (2007, s. 41–43). Rodzinę Franka tworzą kopule indeksowane parametrem α , opisane wzorem:

$$C_\alpha^{Fr}(u; v) = \frac{1}{\alpha} \ln \left(1 + \frac{(e^{\alpha u} - 1)(e^{\alpha v} - 1)}{e^\alpha - 1} \right), \quad (2.2.3.2.1)$$

gdzie parametr $\alpha \in \mathbb{R} \setminus \{0\}$. Generatory addytywny i multiplikatywny są odpowiednio równe:

$$\varphi_\alpha(u) = -\ln \left(\frac{e^{\alpha u} - 1}{e^\alpha - 1} \right), \quad (2.2.3.2.2)$$

$$g(u) = \frac{e^{\alpha u} - 1}{e^\alpha - 1}. \quad (2.2.3.2.3)$$

Gęstość kopuli Franka wynosi:

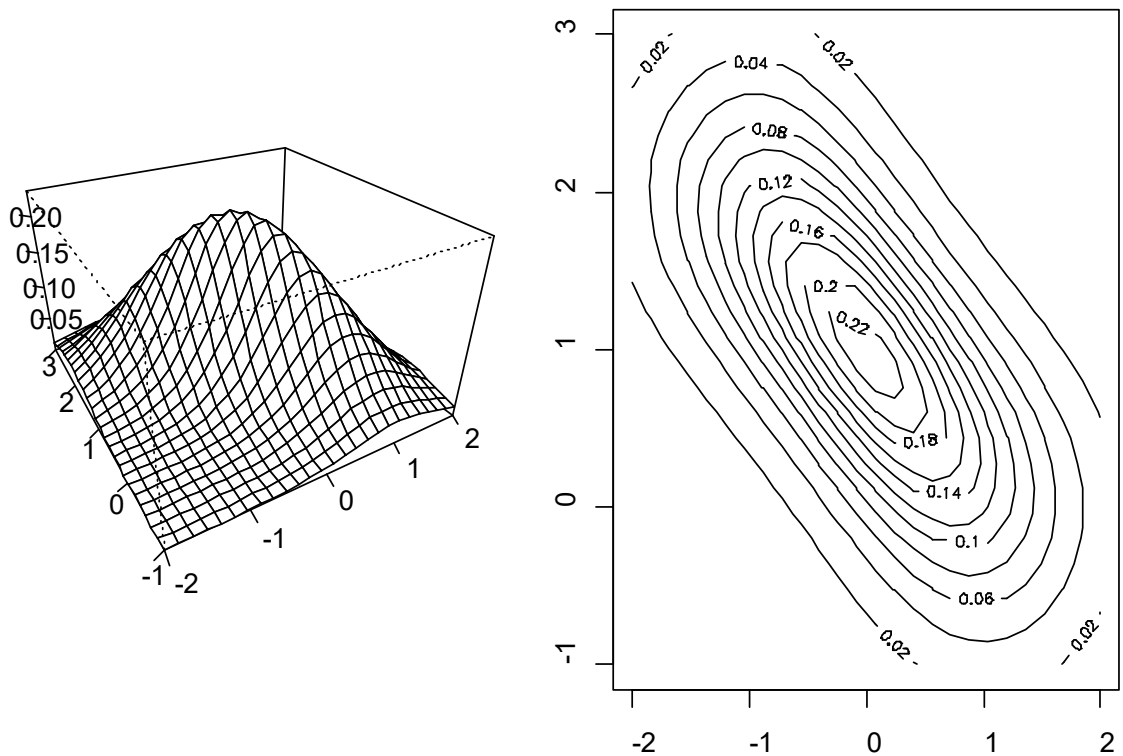
$$c_\alpha^{Fr}(u; v) = \frac{\alpha(e^\alpha - 1)e^{\alpha(u+v)}}{(e^\alpha - e^{\alpha u} - e^{\alpha v} + e^{\alpha(u+v)})^2} \quad (2.2.3.2.4)$$

i jest symetryczna względem prostych $v = u$ oraz $v = 1 - u$. Największe wartości, wynoszące $\alpha/(e^\alpha - 1)$, funkcja gęstości przyjmuje w punktach $(0, 0)$ oraz $(1, 1)$.

Rodzina Franka jest malejąca ze względu na parametr α , gdyż dla każdego $u, v \in [0; 1]$ prawdziwa jest zależność:

$$\alpha \leq \beta \Rightarrow C_\alpha^F(u; v) \geq C_\beta^F(u; v). \quad (2.2.3.2.5)$$

Kopule Franka opisują zależności symetryczne, w ogonach ich rozkładów nie ma zależności. Wykresy przykładowych funkcji gęstości dla rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych kopulą Franka znajdują się na rysunku 2.2.3.



Rysunek 2.2.3. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Franka ($\alpha = -5$)

Źródło: opracowanie własne

2.2.3.3. Rodzina Claytona

Rodzina Claytona została opisana na podstawie pracy Heilperna (2007, s. 43–44). Wykresy przykładowych funkcji gęstości dla rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych kopulą Claytona znajdują się na rysunku 2.2.4. Rodzinę Claytona tworzą kopule indeksowane parametrem α , opisane wzorem:

$$C_{\alpha}^{Cl}(u; v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}, \quad (2.2.3.3.1)$$

gdzie parametr $\alpha > 0$. Generatory addytywny i multiplikatywny są odpowiednio równe:

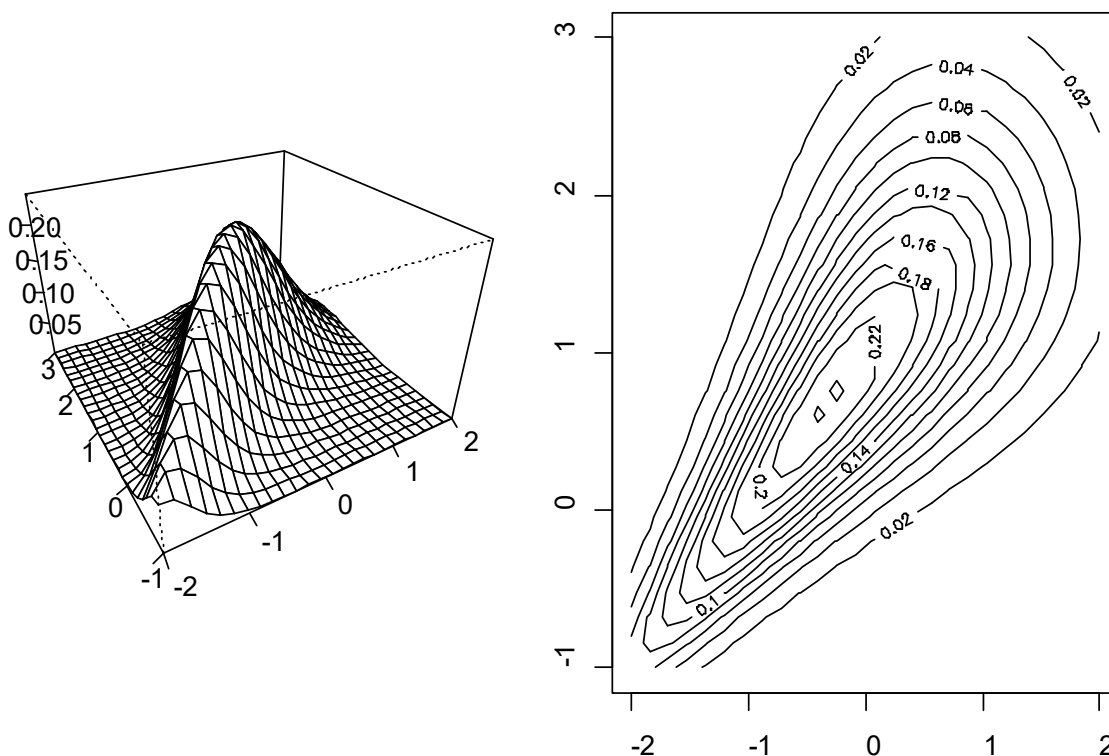
$$\varphi_{\alpha}(u) = u^{-\alpha} - 1, \quad (2.2.3.3.2)$$

$$g(u) = e^{1-u^{-\alpha}}. \quad (2.2.3.3.3)$$

Gęstość kopuli Claytona wynosi:

$$c_{\alpha}^{Cl}(u; v) = (1 + \alpha)(uv)^{-\alpha-1}(u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha-2} \quad (2.2.3.3.4)$$

i jest symetryczna względem prostej $v = u$, a w punkcie $(0, 0)$ dąży do nieskończoności.



Rysunek 2.2.4. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Claytona ($\alpha = 2$)

Źródło: opracowanie własne

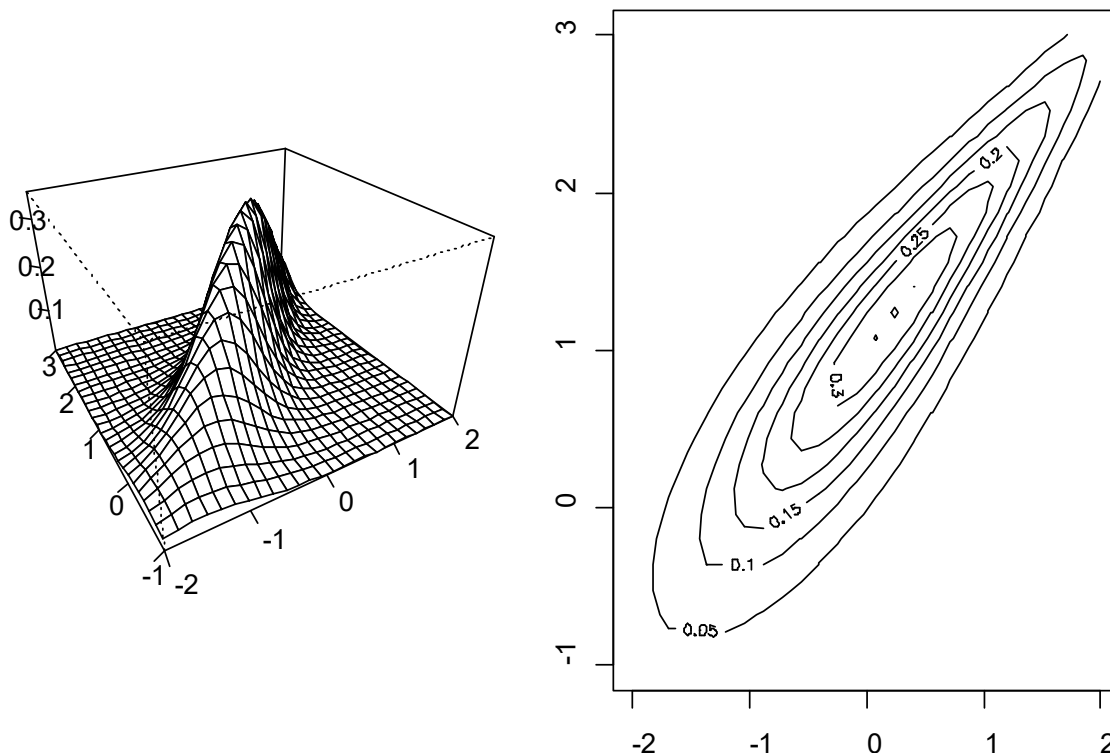
Rodzina Claytona jest rosnąca ze względu na parametr α , gdyż dla każdego $u, v \in [0; 1]$ prawdziwa jest zależność:

$$\alpha \leq \beta \Rightarrow C_{\alpha}^{Cl}(u; v) \leq C_{\beta}^{Cl}(u; v). \quad (2.2.3.3.5)$$

Rodzinę Claytona można rozszerzyć, przyjmując, że parametr $\alpha \in [-1; 0) \cup (0; \infty)$. Kopuła Claytona nie ma zależności w górnym ogonie. Zależność w dolnym ogonie występuje, gdy $\alpha > 0$ i wynosi $\lambda^L = 2^{-1/\alpha}$ (Doman, 2011, s. 27).

2.2.3.4. Rodzina Gumbela

Rodzina Gumbela również została opisana na podstawie pracy Heilperna (2007, s. 45). Wykresy przykładowych funkcji gęstości dla rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych kopulą Gumbela znajdują się na rysunku 2.2.5.



Rysunek 2.2.5. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Gumbela ($\alpha = 3$)

Źródło: opracowanie własne

Rodzinę Gumbela tworzą kopule indeksowane parametrem α , opisane wzorem:

$$C_{\alpha}^{Gu}(u; v) = e^{-[(-\ln u)^{\alpha} + (-\ln v)^{\alpha}]^{1/\alpha}}, \quad (2.2.3.4.1)$$

gdzie parametr $\alpha > 1$. Generatory addytywny i multiplikatywny są odpowiednio równe:

$$\varphi_{\alpha}(u) = (-\ln u)^{\alpha}, \quad (2.2.3.4.2)$$

$$g(u) = e^{-(-\ln u)^{\alpha}}. \quad (2.2.3.4.3)$$

Gęstość kopuli Gumbela wynosi:

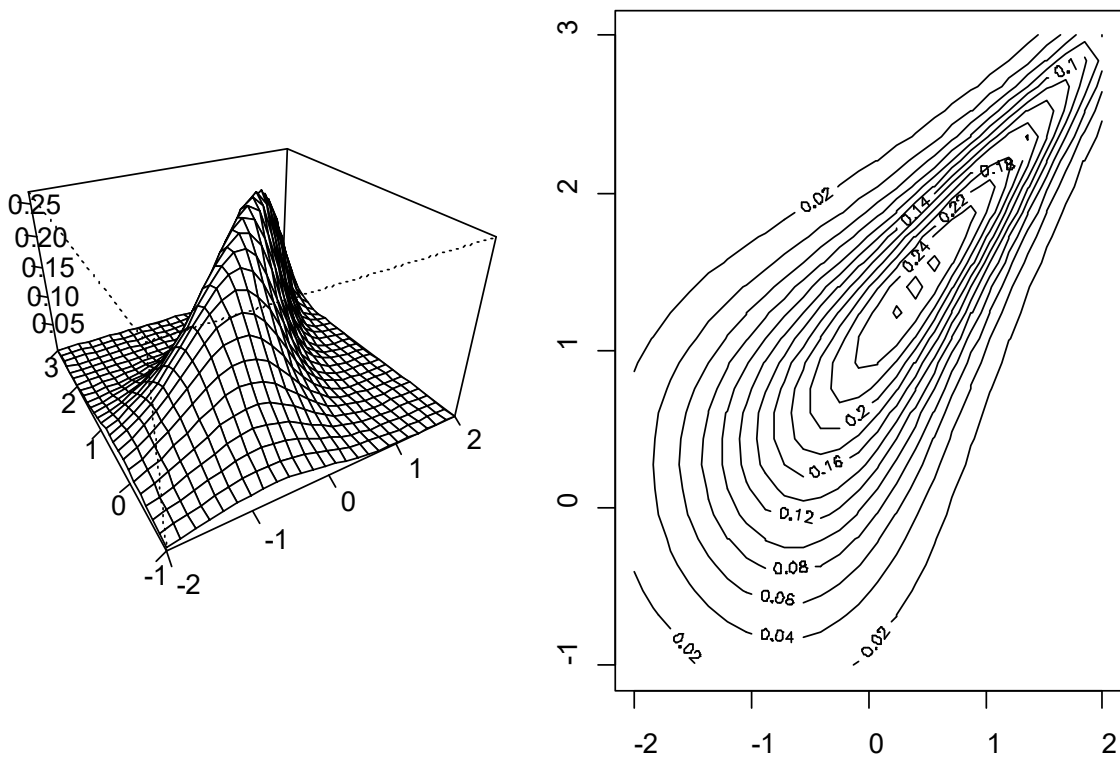
$$c_{\alpha}^{Gu}(u; v) = C_{\alpha}^{Gu}(u, v) \frac{(\tilde{u}\tilde{v})^{\alpha-1}}{uv(\tilde{u}^{\alpha} + \tilde{v}^{\alpha})^{2-1/\alpha}} ((\tilde{u}^{\alpha} + \tilde{v}^{\alpha})^{1/\alpha} + \alpha - 1), \quad (2.2.3.4.4)$$

gdzie $\tilde{u} = -\ln u$ i $\tilde{v} = -\ln v$.

Rodzina Gumbela jest rosnąca ze względu na parametr α . Kopula Gumbela nie ma zależności w dolnym ogonie, a zależność w górnym ogonie wynosi $\lambda^U = 2 - 2^{1/\alpha}$ (Doman, 2011, s. 28).

2.2.3.5. Rodzina Joe

Rodzina Joe została opisana na podstawie pracy Nelsena (1991, s. 112-117). Wykresy przykładowych funkcji gęstości dla rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych kopulą Joe znajdują się na rysunku 2.2.6.



Rysunek 2.2.6. Wykres perspektywny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Joe ($\alpha = 3$)

Źródło: opracowanie własne

Rodzinę Joe tworzą kopule indeksowane parametrem α , opisane wzorem:

$$C_{\alpha}^{Jo}(u; v) = 1 - [(1 - u)^{\alpha} + (1 - v)^{\alpha} - (1 - u)^{\alpha}(1 - v)^{\alpha}]^{\frac{1}{\alpha}}, \quad (2.2.3.5.1)$$

gdzie parametr $\alpha \geq 1$. Generatory addytywny i multiplikatywny są odpowiednio równe:

$$\varphi_{\alpha}(u) = -\ln(1 - (1 - u)^{\alpha}), \quad (2.2.3.5.2)$$

$$g(u) = 1 - (1 - u)^{\frac{1}{\alpha}}. \quad (2.2.3.5.3)$$

Gęstość kopuli Joe wynosi (Joe, 1997, s. 142):

$$c_{\alpha}^{Jo}(u; v) = (\bar{u}^{\alpha} + \bar{v}^{\alpha} - \bar{u}^{\alpha}\bar{v}^{\alpha})^{(-2+\frac{1}{\alpha})} \bar{u}^{(\alpha-1)} \bar{v}^{(\alpha-1)} (\alpha - 1 + \bar{u}^{\alpha} + \bar{v}^{\alpha} - \bar{u}^{\alpha}\bar{v}^{\alpha}), \quad (2.2.3.5.4)$$

gdzie $\bar{u} = 1 - u$ i $\bar{v} = 1 - v$. Kopula ta została oznaczona przez Joe jako B5.

Kopula Joe nie ma zależności w dolnym ogonie, a zależność w górnym ogonie wynosi $\lambda^U = 2 - 2^{1/\alpha}$ (Nelsen, 1999, s. 215).

3. Testy zgodności z uciętymi rozkładami ciągłymi

3.1. Statystyki dla obustronnie uciętych rozkładów ciągłych

Modyfikacje statystyk dla rozkładów lewostronnie uciętych, przedstawione w podrozdziale 1.4.2 oraz przeprowadzone na nich badania (Echaust i Lach, 2017; Lach, 2017; Lach i Smaga, 2018), stały się inspiracją do ich dalszej modyfikacji. W tym podrozdziale testy zgodności z lewostronnie uciętymi rozkładami ciągłymi zostaną uogólnione do testów zgodności z obustronnie uciętymi rozkładami ciągłymi. Dodatkowo dwie statystyki przeznaczone do badania zgodności w dolnym ogonie rozkładu, zaprezentowane w podrozdziale 1.4.1, również zostaną uogólnione do testów zgodności z obustronnie uciętymi rozkładami ciągłymi.

Dostosowana do potrzeb niniejszej pracy definicja rozkładu obustronnie uciętego jest następująca:

Definicja 3.1.1 (Domański i Pruska, 2000, s. 34). Niech X będzie zmienną losową o rozkładzie P_X przyjmującą wartości z przedziału $(a; b)$ i niech $[H1; H2] \subset (a; b)$, gdzie $H1, H2, a, b \in \mathbb{R}$. Rozkład zmiennej losowej $Y = X|H1 \leq X \leq H2$, czyli zmiennej losowej, której zbiór wartości stanowią wartości zmiennej losowej X należące do przedziału $[H1; H2]$ nazywamy uciętym rozkładem P_X albo obustronnie uciętym rozkładem P_X .

Sformułowanie hipotezy dla rozkładu obustronnie uciętego, tak jak w przypadku rozkładu lewostronnie uciętego, wymaga zdefiniowania dystrybuanty teoretycznej i empirycznej tego rozkładu.

Niech F_0 oznacza dystrybuantę teoretyczną, a z_{H1} wartość dystrybuanty teoretycznej w punkcie ucięcia $H1$ oraz z_{H2} wartość dystrybuanty teoretycznej w punkcie ucięcia $H2$, tzn. $z_{H1} = F_0(H1)$ i $z_{H2} = F_0(H2)$, ponadto niech $H1 < H2$. Jeśli rozkład ucięty w punktach $H1$ oraz $H2$ jest zgodny z dystrybuantą teoretyczną, wówczas $F_0(X) \sim U[F_0(H1); F_0(H2)]$. Dystrybuanta teoretyczna ciągłego rozkładu obustronnie uciętego ma postać:

$$F_0^{**}(x) = \begin{cases} 0 & \text{dla } x \leq H1, \\ \frac{F_0(x) - F_0(H1)}{F_0(H2) - F_0(H1)} & \text{dla } H1 < x \leq H2, \\ 1 & \text{dla } x > H2. \end{cases} \quad (3.1.1)$$

Jeśli rozkład próby jest zgodny z rozkładem teoretycznym, wówczas $F_0^{**}(x) \sim U[0; 1]$.

Dystrybuanta empiryczna dla rozkładu obustronnie uciętego jest taka sama, jak dystrybuanta empiryczna dla całego rozkładu oraz dystrybuanta empiryczna dla rozkładu lewostronnie uciętego i wyraża się wzorem:

$$F_n(x) = \begin{cases} 0 & x < x_{(1)}, \\ \frac{j}{n} & x_{(j)} \leq x < x_{(j+1)}, \quad j = 1, 2, \dots, n-1, \\ 1 & x \geq x_{(n)}. \end{cases} \quad (3.1.2)$$

Dystrybuanta empiryczna dla zaobserwowanej części całej populacji wynosi natomiast:

$$F_n(x) * (F_0(H2) - F_0(H1)) + F_0(H1) = \begin{cases} F_0(H1) & x < x_{(1)}, \\ \frac{j}{n} (F_0(H2) - F_0(H1)) + F_0(H1) & x_{(j)} \leq x < x_{(j+1)}, \quad j = 1, 2, \dots, n-1, \\ F_0(H2) & x \geq x_{(n)}. \end{cases} \quad (3.1.3)$$

W oparciu o wzory (3.1.1)–(3.1.3) wyprowadzono formuły obliczeniowe statystyk dla obustronnie uciętych rozkładów ciągłych. Wzory statystyk testowych oraz ich formuły obliczeniowe znajdują się w tabeli 3.1.1. Wyprowadzenie wzorów wraz z opisem przyjętych oznaczeń znajduje się w załączniku 1 umieszczonym w aneksie. Gdy $z_{H1} = 0$ oraz $z_{H2} = 1$, wzory redukują się do wzorów dla całych rozkładów. Gdy $z_{H1} > 0$ oraz $z_{H2} = 1$, otrzymuje się wzory dla rozkładów lewostronnie uciętych, gdy $z_{H1} = 0$ oraz $z_{H2} < 1$, wzory dla rozkładów prawostronnie uciętych, a gdy $z_{H1} > 0$ oraz $z_{H2} < 1$, wzory dla rozkładów obustronnie uciętych.

Tabela 3.1.1. Wzory oraz formuły obliczeniowe statystyk dla obustronnie uciętych rozkładów ciągłych

Statystyka	Wzór statystyki i formuła obliczeniowa
KS^{**}	$KS^{**} = \sqrt{n} \sup_x F_n(x) - \hat{F}_0^{**}(x) $ <p>Formuła obliczeniowa:</p> $KS^{**} = \frac{\sqrt{n}}{z_{H2} - z_{H1}} \max \left\{ \sup_j \left\{ z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right\}, \sup_j \left\{ z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right\} \right\}$

Statystyka	Wzór statystyki i formuła obliczeniowa
V^{**}	$V^{**} = \sqrt{n} \left(\sup_x \{F_n(x) - \hat{F}_0^{**}(x)\} + \sup_x \{\hat{F}_0^{**}(x) - F_n(x)\} \right)$ <p>Formuła obliczeniowa:</p> $V^{**} = \frac{\sqrt{n}}{z_{H2} - z_{H1}} \left(\sup_j \left\{ z_{H1} + \frac{j}{n}(z_{H2} - z_{H1}) - z_j \right\} + \sup_j \left\{ z_j - \left(z_{H1} + \frac{j-1}{n}(z_{H2} - z_{H1}) \right) \right\} \right)$
AD^{**}	$AD^{**} = \sqrt{n} \sup_x \left \frac{F_n(x) - \hat{F}_0^{**}(x)}{\sqrt{\hat{F}_0^{**}(x)(1 - \hat{F}_0^{**}(x))}} \right $ <p>Formuła obliczeniowa:</p> $AD^{**} = \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_{H1} + \frac{j}{n}(z_{H2} - z_{H1}) - z_j}{\sqrt{(z_j - z_{H1})(z_{H2} - z_j)}} \right\}, \sup_j \left\{ \frac{z_j - \left(z_{H1} + \frac{j-1}{n}(z_{H2} - z_{H1}) \right)}{\sqrt{(z_j - z_{H1})(z_{H2} - z_j)}} \right\} \right\}$
AD_{up}^{**}	$AD_{up}^{**} = \sqrt{n} \sup_x \left \frac{F_n(x) - \hat{F}_0^{**}(x)}{1 - \hat{F}_0^{**}(x)} \right $ <p>Formuła obliczeniowa:</p> $AD_{up}^{**} = \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_{H1} + \frac{j}{n}(z_{H2} - z_{H1}) - z_j}{z_{H2} - z_j} \right\}, \sup_j \left\{ \frac{z_j - \left(z_{H1} + \frac{j-1}{n}(z_{H2} - z_{H1}) \right)}{z_{H2} - z_j} \right\} \right\}$
AD_{down}^{**}	$AD_{down}^{**} = \sqrt{n} \sup_x \left \frac{F_n(x) - \hat{F}_0^{**}(x)}{\hat{F}_0^{**}(x)} \right $ <p>Formuła obliczeniowa:</p> $AD_{down}^{**} = \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_{H1} + \frac{j}{n}(z_{H2} - z_{H1}) - z_j}{z_j - z_{H1}} \right\}, \sup_j \left\{ \frac{z_j - \left(z_{H1} + \frac{j-1}{n}(z_{H2} - z_{H1}) \right)}{z_j - z_{H1}} \right\} \right\}$
AD^{2**}	$AD^{2**} = n \int_{H1}^{H2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{\hat{F}_0^{**}(x)(1 - \hat{F}_0^{**}(x))} d\hat{F}_0^{**}(x)$ <p>Formuła obliczeniowa:</p> $AD^{2**} = -n + 2n \ln(z_{H2} - z_{H1}) + \frac{1}{n} \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) - \frac{1}{n} \sum_{j=1}^n (1 + 2(n - j)) \ln(z_{H2} - z_j)$

Statystyka	Wzór statystyki i formuła obliczeniowa
W^{2**}	$W^{2**} = n \int_{H_1}^{H_2} (F_n(x) - \hat{F}_0^{**}(x))^2 d\hat{F}_0^{**}(x)$ <p>Formuła obliczeniowa:</p> $W^{2**} = \frac{n}{3} + \frac{nz_{H1}}{(z_{H2} - z_{H1})} + \frac{1}{n(z_{H2} - z_{H1})} \sum_{j=1}^n (1 - 2j)z_j + \frac{1}{(z_{H2} - z_{H1})^2} \left(\sum_{j=1}^n (z_j - z_{H1})^2 \right)$
AD_{up}^{2**}	$AD_{up}^{2**} = n \int_{H_1}^{H_2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{(1 - \hat{F}_0^{**}(x))^2} d\hat{F}_0^{**}(x)$ <p>Formuła obliczeniowa:</p> $AD_{up}^{2**} = -2n \ln(z_{H2} - z_{H1}) + 2 \sum_{j=1}^n \ln(z_{H2} - z_j) + \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n (1 + 2(n - j)) \frac{1}{z_{H2} - z_j}$
AD_{down}^{2**}	$AD_{down}^{2**} = n \int_{H_1}^{H_2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{(\hat{F}_0^{**}(x))^2} d\hat{F}_0^{**}(x)$ <p>Formuła obliczeniowa:</p> $AD_{down}^{2**} = -2n \ln(z_{H2} - z_{H1}) + 2 \sum_{j=1}^n \ln(z_j - z_{H1}) + \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n (1 - 2j) \frac{1}{z_{H1} - z_j}$

Źródło: obliczenia własne

3.2. Testy zgodności z obustronnie uciętym rozkładem ciągłym

W tym podrozdziale zostanie opisane zastosowane w badaniach empirycznych wnioskowanie statystyczne, które umożliwia badanie wybranych fragmentów całych rozkładów. Ogólny przebieg wnioskowania statystycznego w oparciu o testy zgodności przedstawiony w niniejszej pracy będzie przebiegał według następujących etapów:

1. Sformułowanie układu hipotez statystycznych.

W przypadku testów zgodności z obustronnie uciętym rozkładem ciągłym układ hipotez jest następujący:

$$H_0: F = F_0^{**},$$

$$H_1: F \neq F_0^{**}, \tag{3.2.1}$$

gdzie F_0^{**} jest pewną ustaloną i ciągłą dystrybuantą rozkładu obustronnie uciętego.

W badaniach empirycznych parametry dystrybuant teoretycznych nie będą znane i będą szacowane na podstawie próby. Nieznajomość parametrów oznacza, że

weryfikowane hipotezy statystyczne są złożone. Aby podkreślić, że parametry są szacowane, dystrybuanta teoretyczna zostanie oznaczona daszkiem. Ostatecznie układ hipotez jest następujący:

$$\begin{aligned} H_0: F &= \hat{F}_0^{**}, \\ H_1: F &\neq \hat{F}_0^{**}. \end{aligned} \tag{3.2.2}$$

Badania empiryczne zostały przeprowadzone w oparciu o rozkłady opisane w rozdziale drugim.

2. Wybór statystyki testowej.

Statystykami testowymi będą statystyki dla rozkładów obustronnie uciętych zawarte w tabeli 3.1.1. Rozkłady statystyk testowych biorących udział w badaniu nie są znane i będą aproksymowane zmodyfikowaną metodą bootstrapu parametrycznego. Opis podstawowej metody bootstrapu i uzasadnienie wyboru tej metody znajduje się w podrozdziale 1.5. Opis modyfikacji metody bootstrapu znajduje się w podrozdziale 3.3.

3. Wyznaczenie wartości p .

Pojęcie wartości p zostało opisane w podrozdziale 1.2. Rozważana hipoteza zerowa jest prawostronna, stąd przeciwko niej będą świadczyć duże wartości statystyk testowych. W tym przypadku wartość p wyraża się następująco: $p = P_{H_0}(T \geq T_0)$.

4. Podjęcie decyzji.

Na tym etapie zostanie podjęta jedna z dwóch decyzji: hipoteza zerowa zostanie odrzucona lub zostanie stwierdzony brak podstaw do jej odrzucenia. Decyzje będą podejmowane w oparciu o wartość p . Hipoteza zerowa H_0 jest odrzucana, gdy wartość p jest mniejsza od zadanego poziomu istotności α .

Szczegółowa procedura, weryfikująca układ hipotez (3.2.2) i jednocześnie umożliwiająca badanie wybranych fragmentów całych rozkładów, została opracowana na podstawie pracy Pewsey'a (2018, s. 155) oraz Chernobai i in. (2015, s. 579):

Procedura I – weryfikacja układu hipotez:

1. Dla danej n -elementowej próby $\mathbf{X} = (X_1, \dots, X_n)$:

- a) oszacuj parametry $\hat{\theta}$ rozkładu $F_0(x)$, gdzie F_0 oznacza dystrybuantę całego rozkładu,

- b) oblicz wartość statystyki testowej T_0 (w oparciu o estymowane wartości parametrów).
2. Pobierz z rozkładu $\hat{F}_0(x)$ n -elementową próbę losową $\mathbf{X}^{*(i)} = (X_1^{*(i)}, \dots, X_n^{*(i)})$.
Dla danej n -elementowej próby $\mathbf{X}^{*(i)}$:
- a) oszacuj parametry $\hat{\theta}$ rozkładu $F_0(x)$,
- b) oblicz wartość statystyki testowej T_i (w oparciu o estymowane wartości parametrów).
3. Powtórz B -krotnie punkt 2. Na podstawie wartości T_1, \dots, T_B zostanie oszacowany rozkład statystyki testowej T pod warunkiem prawdziwości hipotezy zerowej.
4. Oblicz wartość p według wzoru:
- $$\hat{p} = \frac{\#\{1 \leq i \leq B: T_i > T_0\} + 1}{B + 1}. \quad (3.2.3)$$
5. Wyznacz poziom istotności α .
6. Odrzuć hipotezę zerową, jeśli \hat{p} jest mniejsza od zadanego poziomu istotności α .

W porównaniu do procedury weryfikującej zgodność z lewostronnie uciętym rozkładem ciągłym, zaproponowanej w pracy Chernobai i in. (2015, s. 579), procedura I zawiera kilka modyfikacji. Ich wprowadzenie było możliwe dzięki założeniu, że próba pochodzi z całego rozkładu, a nie tylko z jego fragmentu. W stosunku do oryginalnej procedury zaproponowanej w pracy Chernobai i in. (2015, s. 579), zmiany są następujące:

1. W punktach 1 i 2 procedury I próby pochodzą z całego rozkładu, a nie z rozkładu lewostronnie uciętego w punkcie H .
2. W punktach 1a i 2a procedury I parametry $\hat{\theta}$ rozkładu $F_0(x)$ są estymowane na podstawie całego rozkładu, a nie rozkładu lewostronnie uciętego. W przypadku zastosowania metody największej wiarygodności, parametry $\hat{\theta}$ estymuje się na podstawie całego rozkładu według wzoru:

$$\hat{\theta}_{MLE} = \arg \max_{\theta} \log \left(\prod_{k=1}^n f_{\theta}(x_k) \right), \quad (3.2.4)$$

gdzie f_{θ} jest gęstością całego rozkładu, $\mathbf{x} = (x_1, \dots, x_n)'$ jest realizacją próby \mathbf{X} . Dla rozkładu lewostronnie uciętego z kolei parametry $\hat{\theta}$ estymuje się według wzoru:

$$\hat{\theta}_{MLE} = \arg \max_{\theta} \log \left(\prod_{k=1}^n \frac{f_{\theta}(x_k)}{1 - F_{\theta}(H)} \right). \quad (3.2.5)$$

3.3. Ocena jakości testów zgodności z obustronnie uciętym rozkładem ciągłym

W tym podrozdziale zostaną opisane metody, które zastosowano do oceny jakości zmodyfikowanych testów zgodności. Ocenę jakości testów przeprowadzono w środowisku R w oparciu o autorskie programy, wykorzystujące ogólnodostępne pakiety.

Jakość testów określa się poprzez zbadanie ich zachowania się przy założeniu prawdziwości lub nieprawdziwości hipotezy zerowej. Liczba odrzuconych prawdziwych hipotez zerowych określa rozmiar testu α (prawdopodobieństwo popełnienia błędu I rodzaju). Liczba odrzuconych nieprawdziwych hipotez zerowych określa moc testu. Jeśli prawdopodobieństwo popełnienia błędu II rodzaju wynosi β , wówczas moc testu wynosi $(1 - \beta)$. Rozmiar testu powinien być bliski założonemu poziomowi istotności, moc testu powinna być jak największa.

Testy statystyczne wybrane do badań empirycznych są testami bootstrapowymi. Ich własności zostaną natomiast ocenione metodą Monte Carlo. Takie podejście do oceny własności testów bootstrapowych zastosowano m. in. w pracy Domańskiego (2014, s. 135).

Procedury szacowania błędów I i II rodzaju metodą Monte Carlo (oznaczone jako procedury II i III), zostały opracowane na podstawie Martinez i Martinez (2008, s. 252–253) oraz Stute i in. (1993, s. 5).

Procedura II – szacowanie błędu I rodzaju metodą Monte Carlo:

1. Określ rzeczywisty rozkład populacji F oraz parametry tego rozkładu. W przypadku badania rozmiaru testu, dystrybuenta teoretyczna F_0 jest taka sama, jak F . Niech $i = 0$, gdzie i oznacza numer symulacji.
2. Wygeneruj n -elementową próbkę losową (replikację) z rozkładu F , $i = i + 1$.
3. Zweryfikuj hipotezę zerową używając wartości p (procedura I).
4. Jeśli w wyniku przeprowadzonego testu statystycznego odrzucono hipotezę H_0 , został popełniony błąd I rodzaju. Przypisz wartość zmiennej I_i :

$$I_i = \begin{cases} 1 & \text{popełniono błąd I rodzaju,} \\ 0 & \text{nie popełniono błędu I rodzaju.} \end{cases} \quad (3.3.1)$$

5. Powtórz M -krotnie punkty 2-4.
6. Oszacowanie prawdopodobieństwa popełnienia błędu I rodzaju wynosi:

$$\hat{\alpha} = \frac{1}{M} \sum_{i=1}^M I_i. \quad (3.3.2)$$

Procedura III – szacowanie błędu II rodzaju metodą Monte Carlo:

1. Określ rzeczywisty rozkład populacji F oraz parametry tego rozkładu. W przypadku badania mocy testu, dystrybuanta teoretyczna F_0 różni się od F . Niech $i = 0$, gdzie i oznacza numer symulacji.
2. Wygeneruj n -elementową próbę losową (replikację) z rozkładu F , $i = i + 1$.
3. Zweryfikuj hipotezę zerową używając wartości p (procedura I).
4. Jeśli w wyniku przeprowadzonego testu statystycznego nie było podstaw do odrzucenia hipotezy zerowej, został popełniony błąd II rodzaju. Przypisz wartość zmiennej I_i :

$$I_i = \begin{cases} 1 & \text{popełniono błąd II rodzaju,} \\ 0 & \text{nie popełniono błędu II rodzaju.} \end{cases} \quad (3.3.3)$$

5. Powtórz M -krotnie punkty 2-4.
6. Oszacowanie prawdopodobieństwa popełnienia błędu II rodzaju wynosi:

$$\hat{\beta} = \frac{1}{M} \sum_{i=1}^M I_i. \quad (3.3.4)$$

Prawdopodobieństwo popełnienia błędu II rodzaju zależy m. in. od wielkości próby, weryfikowanej hipotezy oraz poziomu istotności testu, dlatego jakość testów postanowiono zbadać dla następujących przypadków:

1. Sześciu wielkości prób: 100, 200, 300, 400, 500, 1000.
2. Dwóch zestawów poziomów ucięcia: 0,05 i 0,95 oraz 0,10 i 0,90.
3. Pięciu rozkładów: normalnego (N), normalnego odwrotnego rozkładu Gaussowskiego (NIG), hiperbolicznego (Hyp), skośnego uogólnionego rozkładu błędu ($SGED$) i skośnego rozkładu t-Studenta (ST).

Wielkości prób zostały ustalone na podstawie długości szeregów, które zostaną poddane badaniom w części empirycznej oraz na podstawie dostępnych możliwości obliczeniowych.

Każdy rozkład można podzielić na trzy części: dolny ogon, część centralną i górny ogon (Caers i Maes, 1998, s. 2). Ustalenie poziomów ucięcia wymaga odpowiedzi na pytanie, gdzie zaczynają się ogony rozkładu. Wyznaczaniem początku ogona zajmowali się np. Caeiro i Gomes (2016), Caers i Maes (1998), Danielsson i in. (2016), Echaust i Just (2020b). Ustalenie progów ucięcia ma duże znaczenie dla rozkładów warunkowych stosowanych do opisów ogonów (np. *generalized Pareto distribution, GPD*). Zbyt duża liczba elementów w ogonie powoduje

obciążenie estymatorów parametrów rozkładów, natomiast zbyt mała powoduje wzrost ich wariancji. W zakresie ustalenia progu ucięcia zostaną wykorzystane w niniejszej pracy sugestie innych autorów. Jeśli przez k oznaczymy liczbę elementów w ogonie rozkładu, a przez n liczbę wszystkich elementów, wówczas dla $50 < n < 500$ stosunek k/n powinien wynosić 0,2, a dla $n > 500$ natomiast 0,1 (Boos, 1984, s. 5). Według Hasofera (1996, za: Caers i Maes, 1998, s. 8), liczba elementów w ogonie powinna wynosić $k \approx 1,5\sqrt{n}$. Biorąc pod uwagę powyższe kryteria, dla przyjętych w badaniu wielkości prób od 100 do 500 stosunek k/n mieści się w przedziale (0,067; 0,15), a dla prób wielkości 500 lub 1000 w przedziale (0,047; 0,067). W badaniu przyjęto dwa zestawy poziomów ucięcia, 0,05 i 0,95 oraz 0,10 i 0,90.

W procedurach II i III należy określić rozkład rzeczywisty, a więc wybrać jego typ i określić parametry. Wybór typów rozkładów prawdopodobieństwa został uzasadniony w rozdziale drugim. Szczegóły dotyczące wyboru parametrów rozkładów znajdują się w podrozdziale 3.4 zatytułowanym „Charakterystyka danych”.

W badaniu przyjęto poziom istotności $\alpha = 0,05$. Podział rozkładów na trzy części pozwolił na przetestowanie wszystkich skrajnych przypadków dla statystyk zaprezentowanych w podrozdziale 3.1. Każda próba była badana przez 24 statystyki: siedem dla dolnego ogona, siedem dla górnego ogona, pięć dla środkowej części dystrybuanty i pięć dla całego rozkładu. W przypadku części dystrybuanty $F \in [z_{H1}; z_{H2}]$ oraz całego rozkładu nie zastosowano statystyk zaprojektowanych do badania zgodności w ogonach rozkładów.

W celu oceny jakości testów statystycznych należy jeszcze ustalić: liczbę replikacji M występujących w procedurze II i III, liczbę prób bootstrapowych B występujących w procedurze I, generator liczb pseudolosowych oraz metodę losowania.

Brakuje jednoznacznych wskazań, jaką liczbę symulacji M należy przyjąć w metodzie Monte Carlo (Martinez i Martinez, 2008, s. 256). Do określenia rozmiaru i mocy bootstrapowych testów zgodności, Stute i in. (1993, s. 5) przyjęli $M = 1000$ i $B = 500$ dla prób o wielkościach $n = 20, 50, 100$ oraz rozkładów o dwóch parametrach. Pomimo znacznego wzrostu możliwości obliczeniowych, Pewsey (2018) przyjął $M = 1000$ i $B = 199$, dla prób o wielkościach $n = 20, 50, 100$ oraz rozkładów określonych czterema parametrami. W przeprowadzonych badaniach, ze względu na dostępne możliwości obliczeniowe, zdecydowano się przyjąć $M = 1000$ i $B = 100$. Należy podkreślić, że badaniu poddano sześć prób o liczebnościach od 100 do 1000, ponadto cztery z pięciu badanych rozkładów są określone przez cztery parametry, co znacznie wydłuża obliczenia.

Badaniu zostaną poddane szeregi liczb pseudolosowych. W obliczeniach wykorzystano obliczenia równoległe, dlatego wybrano generator liczb pseudolosowych *L'Ecuyer-CMRG (MRG32k3a)* (L'Ecuyer, 1999), zalecany przy tego typu obliczeniach (Karl, Eubank, Milovanovic, Reiser i Young, 2014).

W pozostałej części podrozdziału zostanie opisana modyfikacja metody bootstrapu, którą zastosowano w badaniach. Modyfikacja ta polega na zmianie sposobu losowania danych. W zmodyfikowanej metodzie bootstrapu losowanie n -elementowej próby (punkt 2 procedur I, II i III) będzie podzielone na trzy części. Dla dolnego ogona zostanie wylosowanych k elementów, dla części środkowej $n - 2k$ elementów i dla górnego ogona zostanie wylosowanych również k elementów. Takie losowanie zapewnia taką samą liczbę obserwacji dla każdego z wyróżnionych fragmentów dystrybuanty podczas liczenia statystyk testowych oraz podczas szacowania empirycznych rozmiarów i mocy testów. W podstawowej metodzie bootstrapu losowanie n -elementowej próby nie jest dzielone na części. Dwie metody bootstrapu można ze sobą porównać poprzez zestawienie wariacji estymowanych tymi metodami parametrów (Davison i Hinkley, 1997, s. 440), w przypadku niniejszej pracy tymi parametrami są wartości p . Modyfikacje powinny przyczynić się do redukcji wariacji estymatorów parametrów. W celu porównania podstawowej metody bootstrapu z metodą zmodyfikowaną przeprowadzono dwie symulacje:

1. W punkcie 2 procedury II i w punkcie 2 procedury I losowano 1000-elementową próbę.
2. W punkcie 2 procedury II i w punkcie 2 procedury I losowano trzy próby: 100-elementowe dla ogona górnego i dolnego oraz 800-elementową dla części dystrybuanty $F \in [z_{H1}; z_{H2}]$. Tak wylosowane dane łączono ze sobą w jedną próbę.

Dane w pierwszym punkcie powyższej procedury losowano więc podstawową metodą bootstrapu, a dane w drugim punkcie pozyskano zmodyfikowaną metodą bootstrapu. Eksperyment przeprowadzono dla rozkładu normalnego standardowego, $M = 1000$, $B = 100$, $z_{H1} = 0,1$ oraz $z_{H2} = 0,9$. Oszacowano wartości p testów zgodności dla całych rozkładów, opartych na statystykach: KS^{**} , V^{**} , AD^{**} , W^{2**} , AD^{2**} . Wariacje oszacowanych wartości p znajdują się w tabeli 3.3.1. Nowa metoda przyczyniła się w analizowanym przypadku do redukcji wariacji w porównaniu do podstawowej metody bootstrapu. Najmniejsza redukcja wariacji estymatora wartości p nastąpiła dla statystyki AD^{**} , w pozostałych przypadkach wielkość redukcji kształtowała się podobnie i wyniosła 9–10%.

Tabela 3.3.1. Porównanie wariancji estymatora wartości p dla podstawowej i zmodyfikowanej metody bootstrapu

L.p.	Statystyka	KS**	V**	AD**	W ^{2**}	AD ^{2**}
1	Wariancja estymatora wartości p - podstawowa metoda	0,0879	0,0852	0,0842	0,0881	0,0868
2	Wariancja estymatora wartości p - zmodyfikowana metoda	0,0805	0,0773	0,0822	0,0806	0,0794
3	L.p. 1 / L.p. 2	1,0918	1,1023	1,0236	1,0927	1,0935

Źródło: obliczenia własne

3.4. Charakterystyka danych

Badaniom empirycznym będą podlegały szeregi dziennych, logarytmicznych, procentowych stóp zwrotu. Stopy te są obliczane zgodnie ze wzorem:

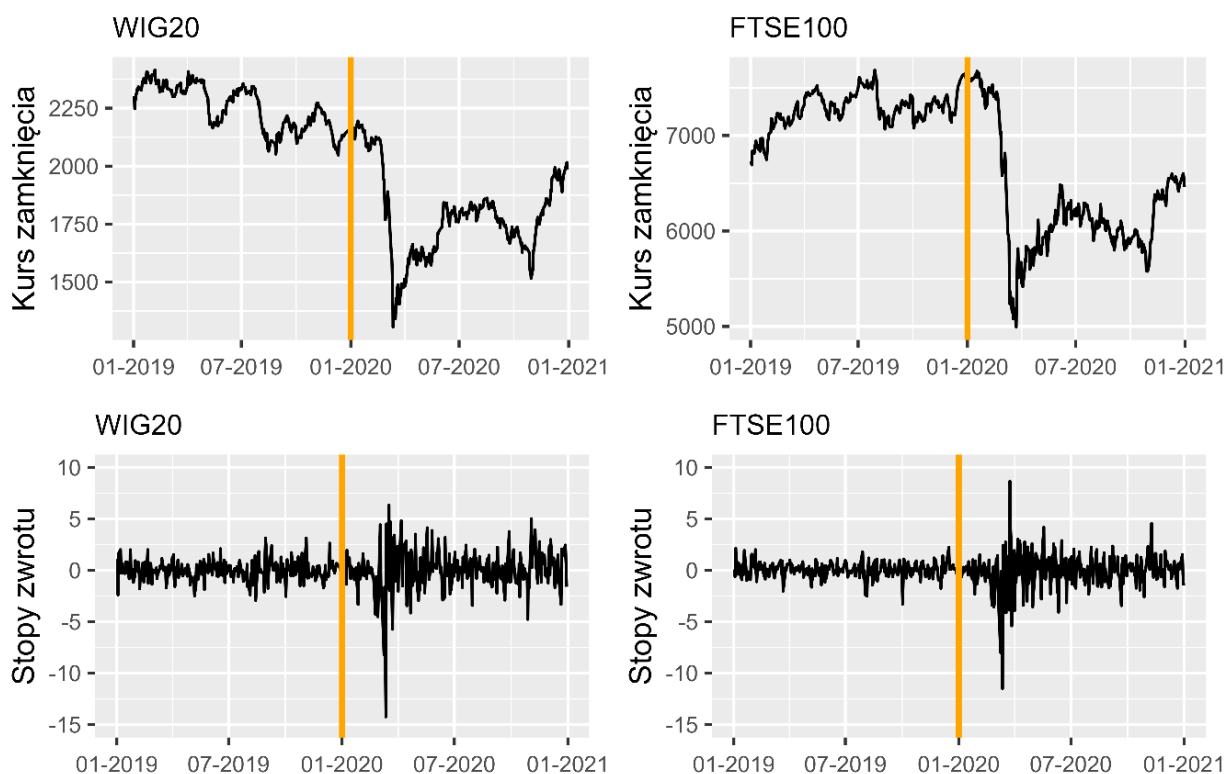
$$r_t = 100 \cdot (\ln P_t - \ln P_{t-1}), \quad (3.4.1)$$

gdzie r_t oznacza dzienną procentową logarytmiczną stopę zwrotu, a P_t i P_{t-1} oznaczają odpowiednio ceny instrumentu finansowego w okresie t oraz $t - 1$. Przyjęto, że cenami instrumentów finansowych są kursy zamknięcia dla indeksów *WIG20* oraz *FTSE100*, pochodzące z lat 2019-2020, przy czym okresy te będą w pracy rozpatrywane osobno. Dane empiryczne do badań pobrano z serwisu <https://stooq.pl/>. Dane te w podanym serwisie uzyskuje się poprzez wybranie symbolu „WIG20” dla indeksu *WIG20* oraz symbolu „^UKX” dla indeksu *FTSE100*. Statystyki opisowe badanych szeregów z uwzględnieniem podziału na okresy znajdują się w tabeli 3.4.1. Ceny zamknięcia z lat 2019-2020 dla wybranych indeksów oraz wyliczone na ich podstawie logarytmiczne, procentowe stopy zwrotu znajdują się na rysunku 3.4.1. Rok 2019 to okres przed pandemią, rok 2020 natomiast to rok pandemii COVID-19. Na podstawie wykresu i tabeli można stwierdzić, że rok 2019 jest okresem o małej zmienności obu indeksów, a rok 2020 jest okresem o dużej zmienności, również obu indeksów.

Tabela 3.4.1. Statystyki opisowe badanych szeregów

Indeks	Rok	Liczba obserwacji	Min	Max	Średnia arytmetyczna	Odchylenie standardowe	Skośność	Eksces Kurtozy
<i>WIG20</i>	2019	248	-2,9301	3,1388	-0,0231	0,9501	0,1102	1,0665
<i>WIG20</i>	2020	252	-14,2456	6,3353	-0,0319	2,0877	-1,3726	9,0687
<i>FTSE100</i>	2019	253	-3,2839	2,2272	0,0452	0,7397	-0,4402	2,1702
<i>FTSE100</i>	2020	254	-11,5124	8,6668	-0,0610	1,8584	-0,9965	7,8597

Źródło: obliczenia własne



Rysunek 3.4.1. Notowania i stopy zwrotu dla indeksów WIG20 oraz FTSE100 w latach 2019-2020

Źródło: opracowanie własne

W poprzednim podrozdziale stwierdzono, że w procedurach II (szacowanie błędu I rodzaju) i III (szacowanie błędu II rodzaju), należy określić rozkład rzeczywisty, a więc wybrać jego typ i określić parametry. Wybór typów rozkładów prawdopodobieństwa został uzasadniony w rozdziale drugim. Parametry rozkładów ustalono natomiast na podstawie dziennych logarytmicznych procentowych stóp zwrotu wyliczonych dla kursów zamknięcia indeksów WIG20 i FTSE100 w latach 2019-2020. Zestawienie typów rozkładów wraz z oszacowanymi dla nich parametrami znajduje się w tabeli 3.4.2. Parametryzacja rozkładów jest zgodna z parametryzacją zastosowaną w użytych pakietach programu R. Są to pakiety: *fBasics* i *stats* (rozkład normalny), *GeneralizedHyperbolic* (rozkłady *NIG* i *Hyp*), *fGarch* (*SGED*) oraz *sn* (*ST*). Ustalony typy rozkładów oraz ich parametry staną się wzorcami dla rozkładów rzeczywistych w eksperymentach szacujących rozmiar i moc testów w kolejnych dwóch podrozdziałach.

Tabela 3.4.2. Wzorce rozkładów rzeczywistych biorących udział w badaniu rozmiarów i mocy testów

Rok	WIG20	FTSE100
2019	$N(\mu = -0,02; \sigma = 0,95)$	$N(\mu = 0,05; \sigma = 0,74)$
	$NIG(\mu = -0,05; \delta = 1,14; \alpha = 1,25; \beta = 0,03)$	$NIG(\mu = 0,16; \delta = 1,00; \alpha = 1,90; \beta = -0,22)$
	$Hyp(\mu = -0,06; \delta = 0,60; \alpha = 1,74; \beta = 0,04)$	$Hyp(\mu = 0,17; \delta = 0,70; \alpha = 2,57; \beta = -0,24)$
	$SGED(\mu = -0,02; \sigma = 0,95; \alpha = 1,30; \kappa = 1,03)$	$SGED(\mu = 0,04; \sigma = 0,73; \alpha = 1,34; \kappa = 0,94)$
	$ST(\xi = -0,06; \omega = 0,77; \alpha = 0,05; \nu = 5,39)$	$ST(\xi = 0,21; \omega = 0,61; \alpha = -0,30; \nu = 5,94)$
2020	$N(\mu = -0,03; \sigma = 2,08)$	$N(\mu = -0,06; \sigma = 1,85)$
	$NIG(\mu = 0,01; \delta = 1,37; \alpha = 0,32; \beta = -0,01)$	$NIG(\mu = 0,24; \delta = 1,36; \alpha = 0,45; \beta = -0,10)$
	$Hyp(\mu = 0,05; \delta = 0,27; \alpha = 0,74; \beta = -0,02)$	$Hyp(\mu = 0,21; \delta = 0,30; \alpha = 0,84; \beta = -0,09)$
	$SGED(\mu = -0,03; \sigma = 2,01; \alpha = 0,94; \kappa = 0,96)$	$SGED(\mu = -0,06; \sigma = 1,79; \alpha = 0,94; \kappa = 0,95)$
	$ST(\xi = -0,06; \omega = 1,24; \alpha = 0,06; \nu = 2,75)$	$ST(\xi = 0,54; \omega = 1,24; \alpha = -0,50; \nu = 3,24)$

Źródło: obliczenia własne

3.5. Oszacowania rozmiarów testów zgodności

Empiryczne rozmiary testów α_{emp} zostały oszacowane na podstawie procedury II. W eksperymentach oceniających rozmiar testu, typ rozkładu teoretycznego F_0 jest taki sam, jak rozkładu rzeczywistego F . Jeśli więc przyjmuje się, że rozkładem rzeczywistym jest rozkład normalny, to w hipotezie zerowej zakładamy, że rozkładem teoretycznym F_0 jest również rozkład normalny. Żeby wygenerować próbę z rozkładu rzeczywistego, oprócz typu rozkładu, należy jeszcze określić jego parametry. W przeprowadzonych badaniach dla każdego typu rozkładu wybrano po cztery zestawy parametrów, określono je na podstawie roku 2019 i roku 2020 dla indeksów WIG20 i FTSE100. Wzorce dla rozkładów rzeczywistych F zostały przedstawione w tabeli 3.4.2. Rozmiar testów był szacowany dla każdego typu rozkładu dla czterech wybranych zestawów parametrów, jednak ze względu na podobieństwo wyników, zostaną przedstawione jedynie rezultaty otrzymane dla rozkładów z parametrami oszacowanymi na podstawie roku 2020 dla indeksu FTSE100. Wyniki oszacowań rozmiarów analizowanych testów zgodności znajdują się w załączniku 2 umieszczonym w aneksie.

Dolny ogon był testowany przez siedem procedur statystycznych, w oparciu o statystyki $KS^{**}, V^{**}, AD^{**}, AD_{down}^{**}, AD^{2**}, W^{2**}, AD_{down}^{2**}$. Dla poziomu ucięcia $z_{H1} = 0,1$, analizie poddano próby o wielkościach 10, 20, 30, 40, 50, 100, a dla poziomu ucięcia $z_{H1} = 0,05$, próby o wielkościach 5, 10, 15, 20, 25, 50. Stąd w przypadku dolnego ogona oszacowano

rozmiar 420 testów statystycznych (7 statystyk testowych x 2 poziomy ucięcia x 6 wielkości prób x 5 rozkładów, opisanych w rozdziale drugim). Ogon górny był testowany również przez 7 procedur statystycznych, w oparciu o statystyki $KS^{**}, V^{**}, AD^{**}, AD_{up}^{**}, AD^{2**}, W^{2**}, AD_{up}^{2**}$, dla tych samych wielkości prób, co ogon dolny, dla przyjętych poziomów ucięcia $z_{H2} = 0,95$ oraz $z_{H2} = 0,9$. Zatem w przypadku ogona górnego także oszacowano rozmiar 420 testów statystycznych. Część dystrybuanty $F \in [z_{H1}; z_{H2}]$ była testowana przez 5 procedur statystycznych, w oparciu o statystyki $KS^{**}, V^{**}, AD^{**}, AD^{2**}, W^{2**}$. Dla poziomów ucięcia $z_{H1} = 0,1$ i $z_{H2} = 0,9$, analizie poddano próby o wielkościach 80, 160, 240, 320, 400, 800, a dla poziomów ucięcia $z_{H1} = 0,05$ i $z_{H2} = 0,95$, próby o wielkościach 90, 180, 270, 360, 450, 900. W przypadku części dystrybuanty $F \in [z_{H1}; z_{H2}]$ oszacowano rozmiar 300 testów statystycznych (5 statystyk testowych x 2 zestawy poziomów ucięcia x 6 wielkości prób x 5 rozkładów). Cała dystrybuanta była testowana przez 5 procedur statystycznych, w oparciu o statystyki testowe $KS^{**}, V^{**}, AD^{**}, AD^{2**}, W^{2**}$. Analizie poddano próby o wielkościach 100, 200, 300, 400, 500, 1000. Stąd dla całej dystrybuanty oszacowano rozmiar 150 testów statystycznych (5 statystyk testowych x 6 wielkości prób x 5 rozkładów). Łącznie oszacowano rozmiar 1290 testów statystycznych. Dla każdego testu generowano $M = 1000$ replikacji oraz $B = 100$ próbek bootstrapowych.

W celu zweryfikowania istotności różnicy pomiędzy empirycznym rozmiarem testu oraz przyjętym poziomem istotności zastosowano test dla wskaźnika struktury (Orzeszko, 2014, s. 342), w którym weryfikuje się następujący układ hipotez:

$$\begin{aligned} H_0: \alpha_{emp} &= \alpha, \\ H_1: \alpha_{emp} &\neq \alpha. \end{aligned} \tag{3.5.1}$$

Zakładając prawdziwość hipotezy zerowej, empiryczny rozmiar testu α_{emp} ma asymptotyczny rozkład normalny $N(\alpha, \frac{\alpha(1-\alpha)}{M})$. Stąd, α_{emp} należy do przedziału $[3,9; 6,1]$ z prawdopodobieństwem równym 90%, do przedziału $[3,6; 6,3]$ z prawdopodobieństwem równym 95% oraz do przedziału $[3,2; 6,8]$ z prawdopodobieństwem równym 99%. Końce przedziałów zostały oszacowane na podstawie wzoru (Pewsey, 2018, s. 156):

$$\alpha \mp Z_{0,5+u/2} * \sqrt{\frac{\alpha(1-\alpha)}{M}}, \tag{3.5.2}$$

gdzie α oznacza przyjęty poziom istotności, u przyjęte prawdopodobieństwo dla przedziału ufności, natomiast M oznacza liczbę replikacji. W oszacowaniu końców przedziałów przyjęto: $\alpha = 0,05$, $u \in \{0,90; 0,95; 0,99\}$, $M = 1000$.

W załączniku 2 pogrubiono te wartości rozmiarów empirycznych, dla których odrzucono hipotezę H_0 testu dla wskaźnika struktury. Wartości znajdujące się poza 90%, 95% i 99% przedziałem ufności zostały oznaczone odpowiednio przez *, **, ***. Gdy rozmiary testów leżą poniżej dolnej granicy wyznaczonego przedziału ufności, test taki określa się mianem konserwatywnego. Gdy leżą powyżej górnej granicy tego przedziału, test jest zbyt liberalny. W tabeli 3.5.1 przedstawiono procentowe zestawienie liczby odrzuconych hipotez H_0 testu dla wskaźnika struktury dla $\alpha = 0,05$. W tych przypadkach różnica pomiędzy empirycznym rozmiarem testu a przyjętym poziomem istotności $\alpha = 0,05$ była zbyt duża.

Tabela 3.5.1. Udziały odrzuconych hipotez zerowych testu dla wskaźnika struktury, dla $\alpha = 0,05$ (w %)

90% przedział ufności [3,9; 6,1]		95% przedział ufności [3,6; 6,3]		99% przedział ufności [3,2; 6,8]	
Przekroczenia dolnej granicy ($\alpha_{emp} < 3,9$)	Przekroczenia górnej granicy ($\alpha_{emp} > 6,1$)	Przekroczenia dolnej granicy ($\alpha_{emp} < 3,6$)	Przekroczenia górnej granicy ($\alpha_{emp} > 6,3$)	Przekroczenia dolnej granicy ($\alpha_{emp} < 3,2$)	Przekroczenia górnej granicy ($\alpha_{emp} > 6,8$)
23,3	1,7	13,3	0,7	7,3	0,1

Źródło: obliczenia własne

Badane testy są testami konserwatywnymi. Oznacza to, że w wyniku przeprowadzenia testu rzadziej odrzucano prawdziwą hipotezę H_0 zakładającą zgodność rozkładu rzeczywistego z rozkładem hipotetycznym, niż wynikało to z przyjętego poziomu istotności. Dalsze wnioski zostaną przedstawione dla 95% przedziału ufności. Z punktu widzenia przeprowadzonych badań kluczowe znaczenie ma podział dystrybuanty na fragmenty, dlatego w dalszej części podrozdziału analiza zostanie przeprowadzona z uwzględnieniem tego podziału.

W tabeli 3.5.2 zawarto procentowe zestawienie liczby odrzuconych hipotez H_0 testu dla wskaźnika struktury w podziale na badane statystyki testowe. W przypadku górnego i dolnego ogona rozkładu oszacowania empirycznych rozmiarów testów w większości przypadków były właściwe. Dla części dystrybuanty $[z_{H1}; z_{H2}]$, oszacowania empirycznych rozmiarów były w większości przypadków prawidłowe dla testu opartego na statystyce AD^{**} . Testy dla tej części dystrybuanty oparte na pozostałych statystykach często okazywały się zbyt

konserwatywne. Oszacowania empirycznych rozmiarów testów były znacząco mniejsze od założonego poziomu istotności w 60% przeprowadzonych testów opartych na statystykach W^{2**} i AD^{2**} . W przypadku całej dystrybuanty testy również często okazywały się konserwatywne, jednak tutaj udział odrzuconych hipotez zerowych testu dla wskaźnika struktury z powodu przekroczenia dolnych granic przedziałów ufności nie przekraczał 30%.

Tabela 3.5.2 Udziały odrzuconych hipotez zerowych testu dla wskaźnika struktury, w podziale na przekroczenia dolnej lub górnej granicy 95% przedziału ufności, statystyki testowe oraz fragment dystrybuanty, dla $\alpha = 0,05$ (w %)

Fragment dystrybuanty	[0; z_{H1}]		[z_{H2} ; 1]		[z_{H1} ; z_{H2}]		[0; 1]		Średnia	
	< 3,6	> 6,3	< 3,6	> 6,3	< 3,6	> 6,3	< 3,6	> 6,3	< 3,6	> 6,3
Statystyka:										
KS^{**}	5,0	0,0	1,7	0,0	41,7	0,0	16,7	0,0	16,2	0,0
V^{**}	0,0	0,0	1,7	5,0	26,7	0,0	23,3	0,0	11,4	1,4
AD^{**}	3,3	3,3	1,7	0,0	0,0	3,3	0,0	3,3	1,4	2,4
$AD_{down}^{**}, AD_{up}^{**}$	6,7	0,0	1,7	0,0	X	X	X	X	4,2	0,0
W^{2**}	1,7	0,0	5,0	0,0	60,0	0,0	23,3	0,0	22,4	0,0
AD^{2**}	1,7	0,0	3,3	1,7	60,0	0,0	30,0	0,0	22,9	0,5
$AD_{down}^{2**}, AD_{up}^{2**}$	11,7	0,0	6,7	0,0	X	X	X	X	9,2	0,0

Źródło: obliczenia własne

W tabeli 3.5.3 zawarto procentowe zestawienie liczby odrzuconych hipotez H_0 testu dla wskaźnika struktury z uwzględnieniem podziału na statystyki testowe oraz rozkłady prawdopodobieństwa. W takim przekroju, w przypadku ogonów oraz części środkowej rozkładu, znajduje się jedynie po 12 obserwacji (6 prób różnej wielkości x 2 poziomy ucięcia), a w przypadku całego rozkładu po 6 obserwacji (6 prób różnej wielkości). Testy statystyczne dla części dystrybuanty [z_{H1} ; z_{H2}] często okazywały się zbyt konserwatywne, zwłaszcza gdy zastosowano statystyki W^{2**} i AD^{2**} . Dla tej części dystrybuanty testy tylko w dwóch przypadkach okazały się zbyt liberalne, gdy w testach zastosowano statystykę testową AD^{**} , a dane generowano z rozkładów N lub NIG . W przypadku testów statystycznych dla całej dystrybuanty, empiryczny rozmiar testu był znacząco większy od założonego poziomu istotności tylko w jednym przypadku, gdy w testach zastosowano statystykę testową AD^{**} , a dane generowano z rozkładu ST . W przypadku ogonów rozkładu większość oszacowań empirycznych rozmiarów testów była prawidłowa. Dla ogona dolnego testy były zbyt liberalne, gdy użyto statystyki testowej AD^{**} , a dane symulowano z rozkładów $SGED$ lub ST .

Tabela 3.5.3. Udziały odrzuconych hipotez zerowych testu dla wskaźnika struktury, w podziale na przekroczenia dolnej lub górnej granicy 95% przedziału ufności, statystyki testowe oraz rozkłady prawdopodobieństwa, dla $\alpha = 0,05$ (w %)

Fragment dystrybuanty		[0; z_{H1}]		[z_{H2} ; 1]		[z_{H1} ; z_{H2}]		[0; 1]		Średnia	
α_{emp}		< 3,6	> 6,3	< 3,6	> 6,3	< 3,6	> 6,3	< 3,6	> 6,3	< 3,6	> 6,3
Rozkład*	Statystyka										
<i>N</i>	<i>KS</i> **	8,3	0,0	0,0	0,0	66,7	0,0	16,7	0,0	23,8	0,0
	<i>V</i> **	0,0	0,0	0,0	0,0	8,3	0,0	0,0	0,0	2,4	0,0
	<i>AD</i> **	0,0	0,0	0,0	0,0	0,0	8,3	0,0	0,0	0,0	2,4
	<i>AD</i> ** _{down} , <i>AD</i> ** _{up}	0,0	0,0	0,0	0,0	×	×	×	×	0,0	0,0
	<i>W</i> ^{2**}	0,0	0,0	0,0	0,0	100,0	0,0	0,0	0,0	28,6	0,0
	<i>AD</i> ^{2**}	0,0	0,0	0,0	0,0	91,7	0,0	16,7	0,0	28,6	0,0
	<i>AD</i> ^{2**} _{down} , <i>AD</i> ^{2**} _{up}	0,0	0,0	0,0	0,0	×	×	×	×	0,0	0,0
<i>NIG</i>	<i>KS</i> **	0,0	0,0	0,0	0,0	25,0	0,0	0,0	0,0	7,1	0,0
	<i>V</i> **	0,0	0,0	0,0	0,0	8,3	0,0	0,0	0,0	2,4	0,0
	<i>AD</i> **	0,0	0,0	0,0	0,0	0,0	8,3	0,0	0,0	0,0	2,4
	<i>AD</i> ** _{down} , <i>AD</i> ** _{up}	33,3	0,0	0,0	0,0	×	×	×	×	16,7	0,0
	<i>W</i> ^{2**}	0,0	0,0	0,0	0,0	50,0	0,0	16,7	0,0	16,7	0,0
	<i>AD</i> ^{2**}	0,0	0,0	0,0	0,0	50,0	0,0	16,7	0,0	16,7	0,0
	<i>AD</i> ^{2**} _{down} , <i>AD</i> ^{2**} _{up}	41,7	0,0	0,0	0,0	×	×	×	×	20,8	0,0
<i>Hyp</i>	<i>KS</i> **	8,3	0,0	8,3	0,0	66,7	0,0	66,7	0,0	33,3	0,0
	<i>V</i> **	0,0	0,0	0,0	8,3	83,3	0,0	100,0	0,0	38,1	2,4
	<i>AD</i> **	0,0	0,0	8,3	0,0	0,0	0,0	0,0	0,0	2,4	0,0
	<i>AD</i> ** _{down} , <i>AD</i> ** _{up}	0,0	0,0	0,0	0,0	×	×	×	×	0,0	0,0
	<i>W</i> ^{2**}	0,0	0,0	8,3	0,0	75,0	0,0	66,7	0,0	33,3	0,0
	<i>AD</i> ^{2**}	0,0	0,0	8,3	0,0	75,0	0,0	66,7	0,0	33,3	0,0
	<i>AD</i> ^{2**} _{down} , <i>AD</i> ^{2**} _{up}	0,0	0,0	8,3	0,0	×	×	×	×	4,2	0,0
<i>SGED</i>	<i>KS</i> **	0,0	0,0	0,0	0,0	33,3	0,0	0,0	0,0	9,5	0,0
	<i>V</i> **	0,0	0,0	0,0	8,3	16,7	0,0	0,0	0,0	4,8	2,4
	<i>AD</i> **	8,3	8,3	0,0	0,0	0,0	0,0	0,0	0,0	2,4	2,4
	<i>AD</i> ** _{down} , <i>AD</i> ** _{up}	0,0	0,0	8,3	0,0	×	×	×	×	4,2	0,0
	<i>W</i> ^{2**}	8,3	0,0	8,3	0,0	50,0	0,0	16,7	0,0	21,4	0,0
	<i>AD</i> ^{2**}	0,0	0,0	8,3	8,3	58,3	0,0	33,3	0,0	23,8	2,4
	<i>AD</i> ^{2**} _{down} , <i>AD</i> ^{2**} _{up}	16,7	0,0	8,3	0,0	×	×	×	×	12,5	0,0
<i>ST</i>	<i>KS</i> **	8,3	0,0	0,0	0,0	16,7	0,0	0,0	0,0	7,1	0,0
	<i>V</i> **	0,0	0,0	8,3	8,3	16,7	0,0	16,7	0,0	9,5	2,4
	<i>AD</i> **	8,3	8,3	0,0	0,0	0,0	0,0	0,0	16,7	2,4	4,8
	<i>AD</i> ** _{down} , <i>AD</i> ** _{up}	0,0	0,0	0,0	0,0	×	×	×	×	0,0	0,0
	<i>W</i> ^{2**}	0,0	0,0	8,3	0,0	25,0	0,0	16,7	0,0	11,9	0,0
	<i>AD</i> ^{2**}	8,3	0,0	0,0	0,0	25,0	0,0	16,7	0,0	11,9	0,0
	<i>AD</i> ^{2**} _{down} , <i>AD</i> ^{2**} _{up}	0,0	0,0	16,7	0,0	×	×	×	×	8,3	0,0

* Oszacowania parametrów rozkładów dla procentowych logarytmicznych dziennych stóp zwrotu z indeksu *FTSE100* w roku 2020 są następujące: $N(\mu = -0,06; \sigma = 1,85)$, $NIG(\mu = 0,24; \delta = 1,36; \alpha = 0,45; \beta = -0,10)$, $Hyp(\mu = 0,21; \delta = 0,30; \alpha = 0,84; \beta = -0,09)$, $SGED(\mu = -0,06; \sigma = 1,79; \alpha = 0,94; \kappa = 0,95)$, $ST(\xi = 0,54; \omega = 1,24; \alpha = -0,50; \nu = 3,24)$.

Źródło: obliczenia własne

Dla ogona górnego testy były zbyt liberalne w czterech przypadkach, gdy zastosowano statystykę testową V^{**} i dane generowano z rozkładów *Hyp*, *SGED* lub *ST* oraz gdy zastosowano statystykę testową AD^{2**} i dane symulowano z rozkładu *SGED*. Najbardziej konserwatywne w przypadku ogonów okazały się testy oparte na statystyce AD_{down}^{2**} , gdy dane generowano z rozkładu *NIG*.

Procentowe zestawienie liczby odrzuconych hipotez H_0 testu dla wskaźnika struktury z uwzględnieniem podziału na wielkość próby lub poziom ucięcia nie będą interpretowane z uwagi na niewielką liczbę obserwacji w każdej grupie oraz brak wyraźnych prawidłowości.

Podsumowując ten podrozdział można stwierdzić, że większość analizowanych testów zgodności dla ogona górnego i dolnego charakteryzowała się prawidłowymi oszacowaniami rozmiarów empirycznych. Natomiast testy zgodności dla centralnej części dystrybuanty oraz dla całej dystrybuanty często okazywały się zbyt konserwatywne, co oznacza, że błąd I rodzaju był tu popełniany rzadziej, niż zakładał przyjęty poziom istotności. Jednakże, według Tollenaara i Mooijaarta (2003, s. 282), jeśli konserwatywne testy zgodności charakteryzują się rozsądnymi oszacowaniami mocy, takie testy nadal można stosować.

3.6. Oszacowania mocy testów zgodności

Empiryczne moce testów zostały oszacowane na podstawie procedury III. W badaniach mocy testów rozkład rzeczywisty F różni się od rozkładu teoretycznego F_0 i oczekuje się, że moc będzie jak największa. Oszacowanie mocy testów zgodności podzielono w niniejszej pracy na cztery części. W każdej części znajdowały się rozkłady z parametrami szacowanymi na podstawie tylko jednego okresu dla wybranego indeksu (tabela 3.4.2). Załóżmy, że tym okresem jest rok 2020 dla indeksu *FTSE100*. Oszacowania parametrów rozkładów są wówczas następujące: $N(-0,06; 1,85)$, $NIG(0,24; 1,36; 0,45; -0,10)$, $Hyp(0,21; 0,30; 0,84; -0,09)$, $SGED(-0,06; 1,79; 0,94; 0,95)$, $ST(0,54; 1,24; -0,50; 3,24)$. W pierwszym eksperymencie przyjęto, że rozkładem rzeczywistym F jest rozkład $N(-0,06; 1,85)$, a rozkładem teoretycznym F_0 jest rozkład *NIG*. W eksperymentach drugim, trzecim i czwartym rozkład rzeczywisty nie zmienia się, nadal jest to rozkład $N(-0,06; 1,85)$, natomiast przyjmuje się, że rozkładami teoretycznymi są kolejno rozkłady: *Hyp*, *SGED* i *ST*. Następnie przyjmijmy, że rozkładem rzeczywistym jest rozkład $NIG(0,24; 1,36; 0,45; -0,10)$, a rozkładami teoretycznymi są kolejno rozkłady: *N*, *Hyp*, *SGED* i *ST*, itd. Analogiczne eksperymenty przeprowadzono dla rozkładów z parametrami szacowanymi na podstawie: roku 2019 dla

indeksu *FTSE100*, roku 2019 dla indeksu *WIG20* oraz roku 2020 dla indeksu *WIG20*. Ograniczenie się do jednego okresu i jednego indeksu jest uzasadnione tym, że badane rozkłady powinny znajdować się blisko siebie. W przypadku mocno oddalonych rozkładów, moc testu powinna dążyć do jedności. Ostatecznie rozkład rzeczywisty F został określony na 20 sposobów. Dla każdego rozkładu rzeczywistego rozkład teoretyczny F_0 wybierano na 4 sposoby. Łącznie testowano więc 80 układów hipotez statystycznych. Wyniki oszacowań mocy analizowanych testów zgodności znajdują się w załączniku 3 umieszczonym w aneksie.

Zgodność rozkładów empirycznych z teoretycznymi w ich dolnych i górnych ogonach sprawdzano stosując procedury statystyczne zbudowane w oparciu o: 7 statystyk testowych, 2 poziomy ucięcia, 6 prób różnej wielkości (łącznie 84 procedury dla każdego ogona). Zgodność rozkładów empirycznych z teoretycznymi w części centralnej dystrybuanty, tzn. gdy $F \in [z_{H1}, z_{H2}]$, sprawdzano stosując procedury statystyczne zbudowane w oparciu o: 5 statystyk testowych, 2 poziomy ucięcia, 6 prób różnej wielkości (łącznie 60 procedur). Zgodność całych rozkładów empirycznych z całymi rozkładami teoretycznymi sprawdzano stosując procedury statystyczne zbudowane w oparciu o 5 statystyk testowych i 6 prób różnej wielkości (łącznie 30 procedur). Razem w badaniach wykorzystano zatem 258 procedur statystycznych, które przemnożone przez liczbę układów hipotez statystycznych, dają łącznie 20 640 wykonanych testów statystycznych. W ramach każdego testu generowano $M = 1000$ replikacji oraz $B = 100$ próbek bootstrapowych.

W następnych paragrafach zostaną omówione oszacowania mocy testów dla ustalonych rozkładów teoretycznych, rozkłady występują w następującej kolejności: N , NIG , Hyp , $SGED$ i ST . Wszystkie rysunki zostały umieszczone na końcu tego podrozdziału. Rozpatrzmy najpierw przypadki, gdy rozkładem teoretycznym jest rozkład normalny. Niezależnie od rozkładu rzeczywistego, można zauważyć kilka cech charakteryzujących oszacowania mocy poszczególnych testów. Zostanie to zobrazowane na przykładzie, gdy rozkładem rzeczywistym jest rozkład ST , co przedstawiono na rysunkach 3.6.1.–3.6.3. W przypadku dolnego i górnego ogona rozkładu, testy statystyczne można podzielić na trzy grupy o różnych oszacowaniach mocy. Największe oszacowania mocy uzyskano dla trzech testów opartych na statystykach AD_{down}^{2**} , AD_{down}^{**} i AD^{**} dla ogona dolnego i odpowiednio AD_{up}^{2**} , AD_{up}^{**} oraz AD^{**} dla ogona górnego. Najmniejsze oszacowania mocy uzyskano dla testów opartych na statystykach KS^{**} , V^{**} i W^{2**} . Do trzeciej grupy przyporządkowano jeden test, oparty na statystyce AD^{2**} , którego oszacowania mocy, w zależności od rozkładu rzeczywistego, są albo zbliżone do tych

uzyskanych w grupie testów o najmniejszych oszacowaniach mocy, albo znajdują się pomiędzy wyróżnionymi powyżej grupami. Oszacowania mocy testów w większości przypadków wzrastają wraz ze wzrostem punktu ucięcia dla ogona górnego oraz zmniejszaniem punktu ucięcia dla ogona dolnego, co jest równoznaczne ze zmniejszaniem wielkości prób. W przypadku centralnej części dystrybuanty najniższe oszacowania mocy uzyskano dla testów opartych na statystyce AD^{**} . Oszacowania mocy testów dla tej części dystrybuanty wyraźnie zmieniają się wraz ze zmianą punktów ucięcia. Im mniej uciętych danych, tym są one większe. Bez względu na badany fragment dystrybuanty, oszacowania mocy przeważającej części testów są większe, gdy parametry rozkładów rzeczywistych były szacowane na podstawie roku 2020, niż wtedy, gdy parametry rozkładów rzeczywistych były szacowane na podstawie roku 2019, zarówno w przypadku stóp zwrotu z indeksu $FTSE100$, jak i $WIG20$. Jest to szczególnie widoczne w przypadku testów przeznaczonych do badania całego rozkładu. Gdy parametry rozkładów rzeczywistych były szacowane na podstawie roku 2020, oszacowania mocy testów szybko wzrastają, by już przy próbach o liczebności 200 osiągnąć wartość niemal 100%. Gdy parametry rozkładów były szacowane na podstawie roku 2019, oszacowania mocy wszystkich testów wzrastają prawie liniowo dla prób o liczebności od 100 do 500. Wartości zbliżone do 100% są osiągnięte dla prób o liczebności 1000.

Teraz założmy, że rozkładem teoretycznym jest rozkład NIG . Gdy rozkładem rzeczywistym jest rozkład normalny, oszacowania mocy wszystkich testów są zbliżone do oszacowań ich rozmiarów. Jest to zgodne z oczekiwaniami, ponieważ jednym z przypadków granicznych rozkładu NIG jest rozkład normalny. Gdy rozkładami rzeczywistymi były rozkłady Hyp lub $SGED$, oszacowania mocy testów kształtowały się podobnie. Gdy rozkładem rzeczywistym był rozkład ST , zaobserwowano większe oszacowania mocy testów dla ogonów rozkładów oraz mniejsze oszacowania mocy dla części centralnej dystrybuanty. Oszacowania mocy zostaną opisane na przykładzie, gdy rozkładem rzeczywistym jest rozkład $SGED$, co przedstawiono na rysunkach 3.6.4–3.6.6. W tym przypadku oszacowania mocy testów dla ogonów dolnych i górnych nie przekraczają 11,6%. Oszacowania mocy testów znacznie wzrastają, gdy weryfikowana jest zgodność rozkładu empirycznego z teoretycznym w części centralnej dystrybuanty oraz całej dystrybuanty. Oszacowania te osiągają maksymalne wartości bliskie 70%, ale tylko gdy rozkłady rzeczywiste były wzorowane na danych z roku 2020. Gdy rozkłady rzeczywiste były wzorowane na danych z roku 2019, oszacowania mocy nie przekraczają 30%. Różnica pomiędzy wartościami oszacowań może wynikać z większej zmienności stóp zwrotu

w roku 2020 niż w roku 2019. Wśród testów dedykowanych dla części dystrybuanty $[z_{H1}; z_{H2}]$, największymi oszacowaniami mocy charakteryzują się testy oparte na statystykach V^{**} i W^{2**} , a wśród testów przeznaczonych dla całej dystrybuanty, testy oparte na statystykach AD^{2**} i W^{2**} . Na podstawie wykresów można wnioskować, że główne różnice pomiędzy rozkładami występują w części dystrybuanty po odrzuceniu ogonów i są one widoczne przede wszystkim dla rozkładów wzorowanych na podstawie roku 2020.

Gdy rozkładem teoretycznym jest rozkład hiperboliczny, a rozkładem rzeczywistym rozkład normalny lub rozkład *SGED*, oszacowania mocy testów są zbliżone do poziomu istotności testu. W pozostałych przypadkach, gdy rozkładem rzeczywistym jest rozkład *NIG* lub *ST*, wnioski płynące z oszacowań mocy testów kształtują się podobnie, z tym, że oszacowania mocy, gdy rozkładem rzeczywistym jest rozkład *ST*, są większe. Analiza zostanie przeprowadzona na przykładzie, gdy rozkładem rzeczywistym jest rozkład *NIG* (rysunki 3.6.7–3.6.9). Oszacowania mocy wszystkich testów, gdy rozkłady rzeczywiste były wzorowane na danych z roku 2019, są zbliżone do oszacowań rozmiarów tych testów. Natomiast gdy rozkłady rzeczywiste były wzorowane na danych z roku 2020, podobnie jak podczas weryfikowania hipotez, gdy rozkładem teoretycznym był rozkład normalny, testy statystyczne dla ogonów rozkładów można podzielić na trzy grupy o różnych oszacowaniach mocy. Największe oszacowania mocy uzyskano dla trzech testów opartych na statystykach AD_{down}^{2**} , AD_{down}^{**} i AD^{**} dla ogona dolnego i odpowiednio AD_{up}^{2**} , AD_{up}^{**} oraz AD^{**} dla ogona górnego. Najmniejsze oszacowania mocy uzyskano dla testów opartych na statystykach KS^{**} , V^{**} i W^{2**} , w trzeciej grupie znalazł się natomiast jeden test, oparty na statystyce AD^{2**} , którego oszacowania mocy znajdują się najczęściej pomiędzy oszacowaniami mocy uzyskanymi dla dwóch wyróżnionych powyżej grup. W przypadku centralnej części dystrybuanty, oszacowania mocy większości testów są zbliżone do poziomu istotności testów. Testy weryfikujące zgodność całych rozkładów empirycznych z rozkładami teoretycznymi charakteryzują się podobnymi oszacowaniami mocy i nie przekraczają 70% dla prób o liczebności 1000. Na podstawie wykresów można wnioskować, że gdy rozkłady rzeczywiste były wzorowane na podstawie roku 2020, różnice pomiędzy rozkładami występują głównie w ogonach rozkładów.

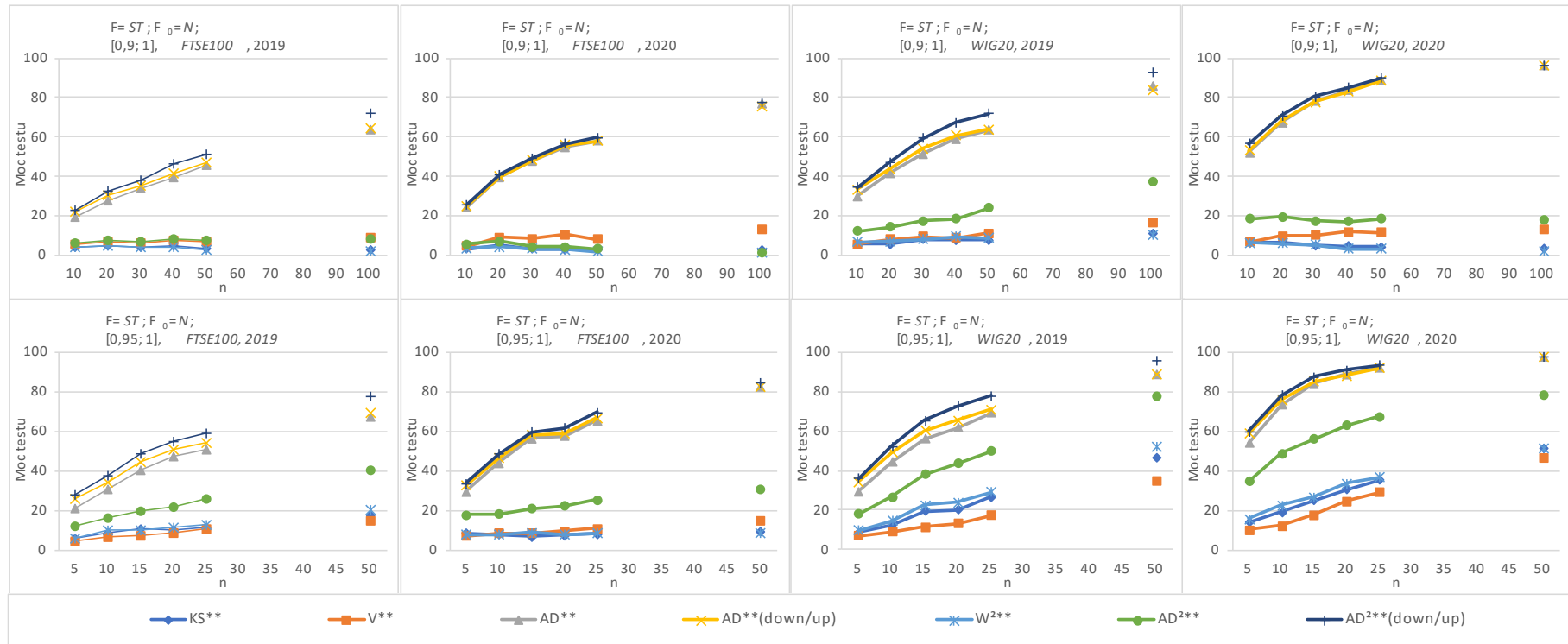
Gdy rozkładem teoretycznym jest rozkład *SGED*, a rozkładem rzeczywistym rozkład normalny, oszacowania mocy testów są zbliżone do poziomu istotności testu, ponieważ jednym z przypadków granicznych rozkładu *SGED* jest rozkład normalny. Gdy rozkładem rzeczywistym jest rozkład hiperboliczny, oszacowania mocy testów nie przekraczają 16,7%.

Gdy rozkładem rzeczywistym jest rozkład *NIG* lub *ST*, oszacowania mocy testów są podobne. W dalszej części paragrafu uwaga zostanie skupiona na przypadku, gdy rozkładem rzeczywistym jest rozkład *ST* (rysunki 3.6.10-3.6.12). Oszacowania mocy testów dla ogonów rozkładów są tutaj podobne do oszacowań uzyskanych w przypadku weryfikowania hipotez badających zgodność rozkładu empirycznego z rozkładem normalnym lub z rozkładem *NIG*. Testy te również można podzielić na te same trzy grupy. Ponadto oszacowania mocy testów dla ogonów rozkładów, gdy rozkłady rzeczywiste były wzorowane na danych z roku 2020, są znacznie większe niż oszacowania mocy testów dla ogonów rozkładów, gdy rozkłady rzeczywiste były wzorowane na danych z roku 2019. Testy przeznaczone do badania zgodności rozkładów w części centralnej dystrybuanty również można podzielić na trzy grupy. Największe oszacowania mocy uzyskano dla testu opartego na statystyce V^{**} , najmniejsze natomiast dla testu opartego na statystyce AD^{**} . Pomiędzy nimi znalazły się trzy pozostałe testy, oparte na statystykach: AD^{2**} , W^{2**} , KS^{**} . Wśród testów dla całej dystrybuanty największe oszacowania mocy uzyskano dla testów opartych na statystyce Andersona-Darlinga, zarówno w wersji supremum, jak i kwadratowej. Oszacowania mocy testów weryfikujących zgodność całych rozkładów lub części centralnej rozkładów są większe, gdy rozkłady rzeczywiste były wzorowane na danych z roku 2020. Na podstawie wykresów można wnioskować, że w roku 2020 różnice pomiędzy rozkładami występują zarówno w ogonach rozkładów, jak i w części centralnej dystrybuanty.

Rozpatrzmy ostatni przypadek, gdy rozkładem teoretycznym jest rozkład *ST*. Ponownie, gdy rozkładem rzeczywistym jest rozkład normalny, oszacowania mocy testów są zbliżone do poziomu istotności testu, ponieważ jednym z przypadków granicznych rozkładu *ST* jest rozkład normalny. Gdy rozkładem rzeczywistym jest rozkład *NIG*, oszacowania mocy testów nie przekraczają 22,7% dla prób o liczebności 1000. Gdy rozkładem rzeczywistym jest rozkład *Hyp* lub *SGED*, oszacowania mocy testów są podobne. Na rysunkach 3.6.13–3.6.15 znajdują się wykresy z oszacowaniami mocy, gdy rozkładem rzeczywistym jest rozkład *SGED*. Testy dla ogonów rozkładów, AD_{down}^{2**} i AD_{down}^{**} dla ogona dolnego oraz odpowiednio AD_{up}^{2**} i AD_{up}^{**} dla ogona górnego, w przeciwieństwie do omówionych powyżej przykładów, charakteryzują się bardzo niskimi oszacowaniami mocy, ponieważ różnice pomiędzy rozkładem empirycznym a rozkładem teoretycznym w tej części rozkładu są niewielkie. Wszystkie rozpatrywane wzorce rozkładów rzeczywistych *SGED* przyjęły parametr kształtu $\alpha < 2$, co oznacza, że rozkład ten posiada grube ogony. Rozkład teoretyczny *ST* natomiast dobrze opisuje rozkłady z grubymi

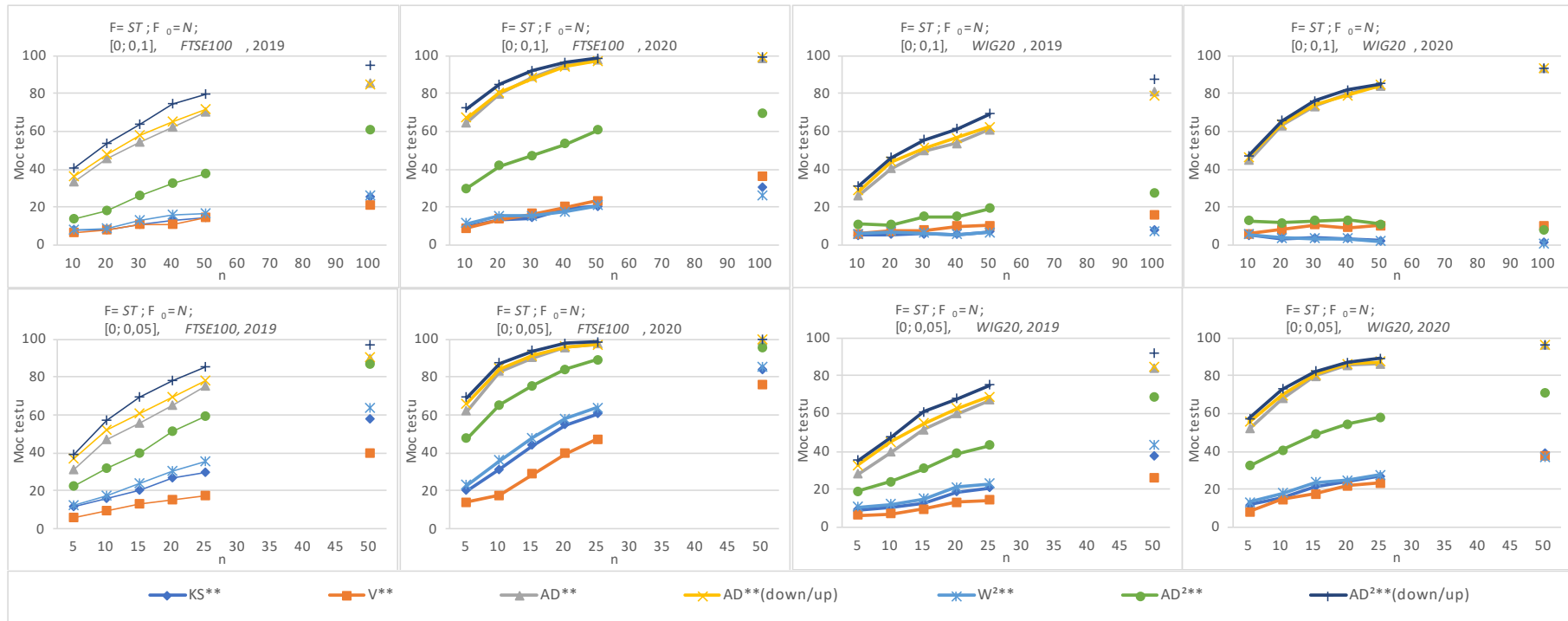
ogonami. Oszacowania mocy testów weryfikujących zgodność całych rozkładów oraz części centralnej rozkładów różnią się w zależności od okresu będącego podstawą do wyznaczenia rozkładu rzeczywistego. Oszacowania te są znacznie większe, gdy rozkłady rzeczywiste były wzorowane na danych z roku 2020. Wyjątek stanowią tu testy oparte na statystyce AD^{**} , dla których uzyskano wyraźnie najmniejsze oszacowania mocy. Na podstawie wykresów można wnioskować, że w roku 2020 różnice pomiędzy rozkładami występują głównie w części centralnej dystrybuanty.

Jak już wspomniano w podrozdziale 1.2, prawdopodobieństwo popełnienia błędu II rodzaju zależy m. in. od wielkości próby oraz weryfikowanej hipotezy (Martinez i Martinez, 2008, s. 239). Uzyskane wyniki potwierdzają istnienie zależności pomiędzy oszacowaniami mocy testów zgodności oraz tymi wielkościami. Z punktu widzenia niniejszej pracy istotniejsze jest jednak spostrzeżenie, że oszacowania mocy testów zgodności zależą również od badanego fragmentu dystrybuanty oraz w mniejszym stopniu również od zastosowanych poziomów ucięcia.



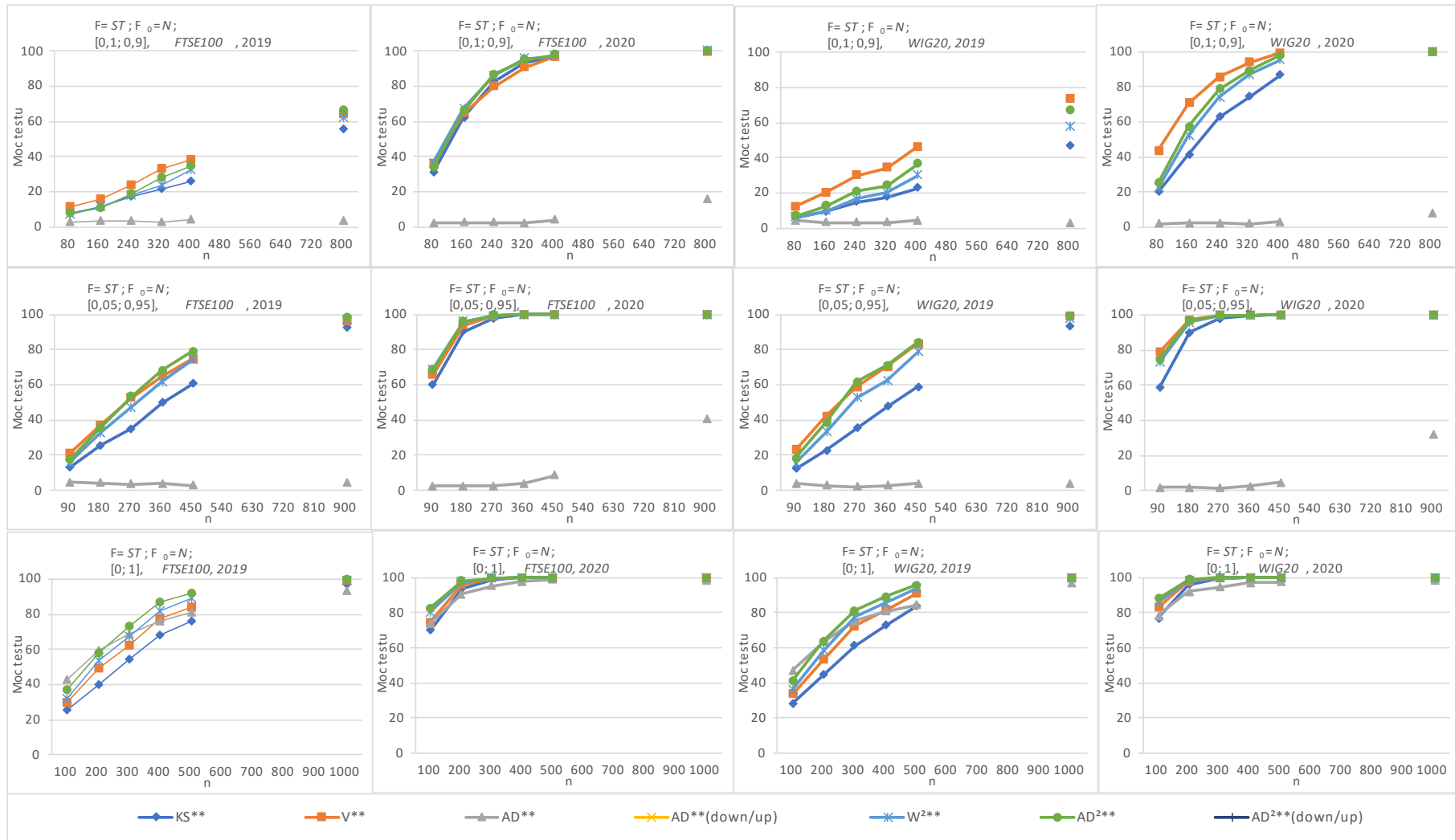
Rysunek 3.6.1. Oszacowania mocy testów zgodności dla $H_0: F = N$ (ogon górny)

Źródło: obliczenia własne



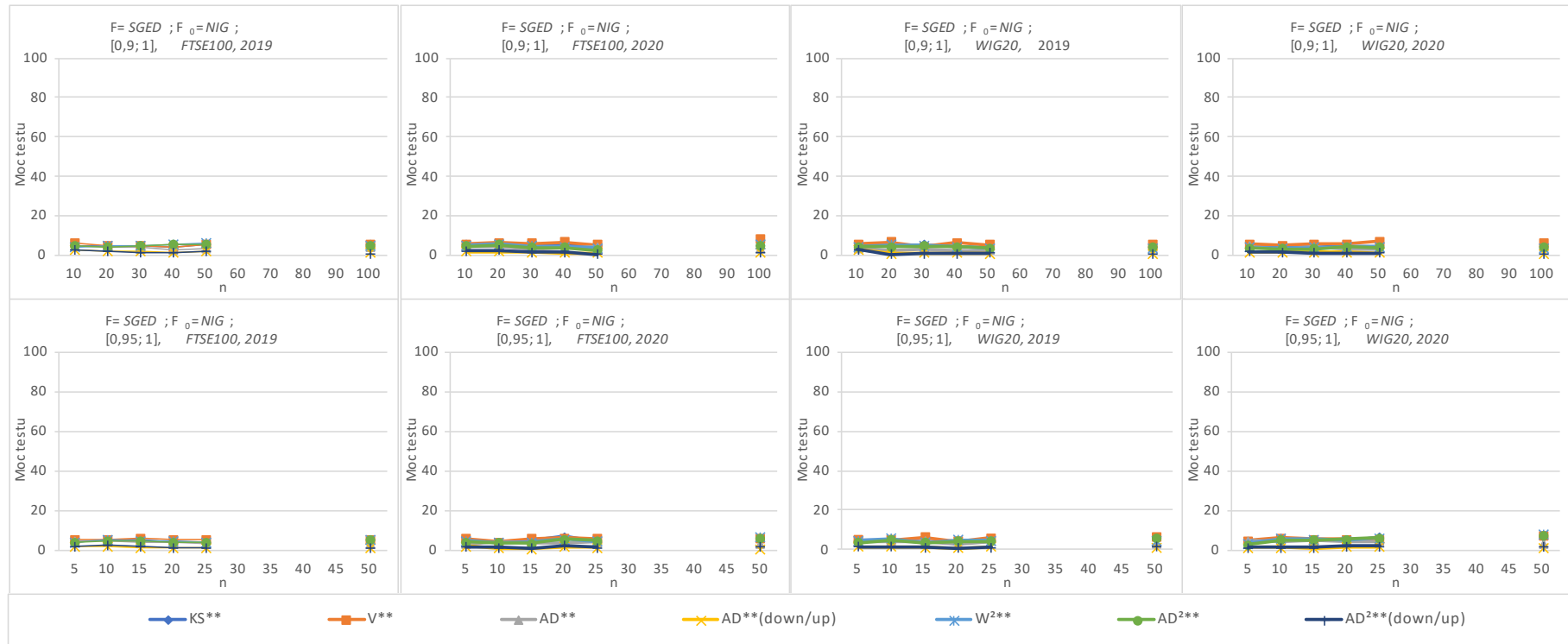
Rysunek 3.6.2. Oszacowania mocy testów zgodności dla $H_0: F = N$ (ogon dolny)

Źródło: obliczenia własne



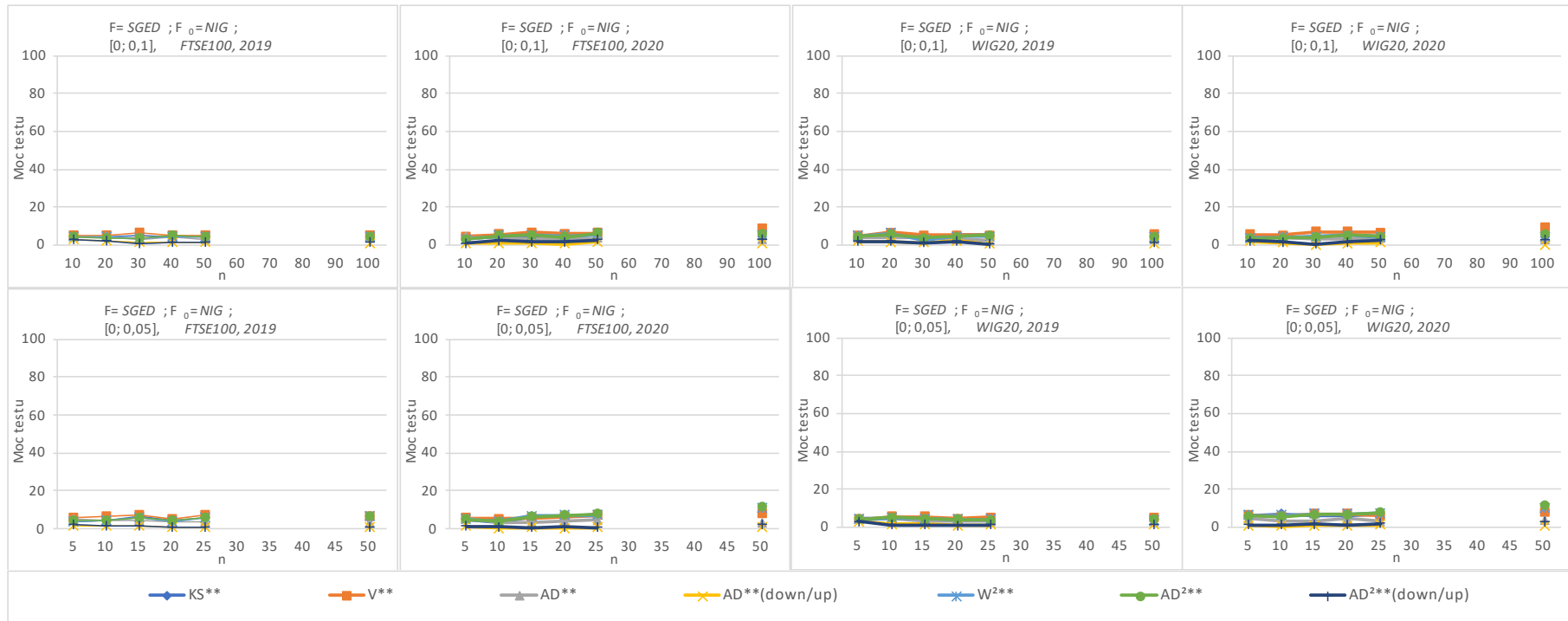
Rysunek 3.6.3. Oszacowania mocy testów zgodności dla $H_0: F = N$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład)

Źródło: obliczenia własne



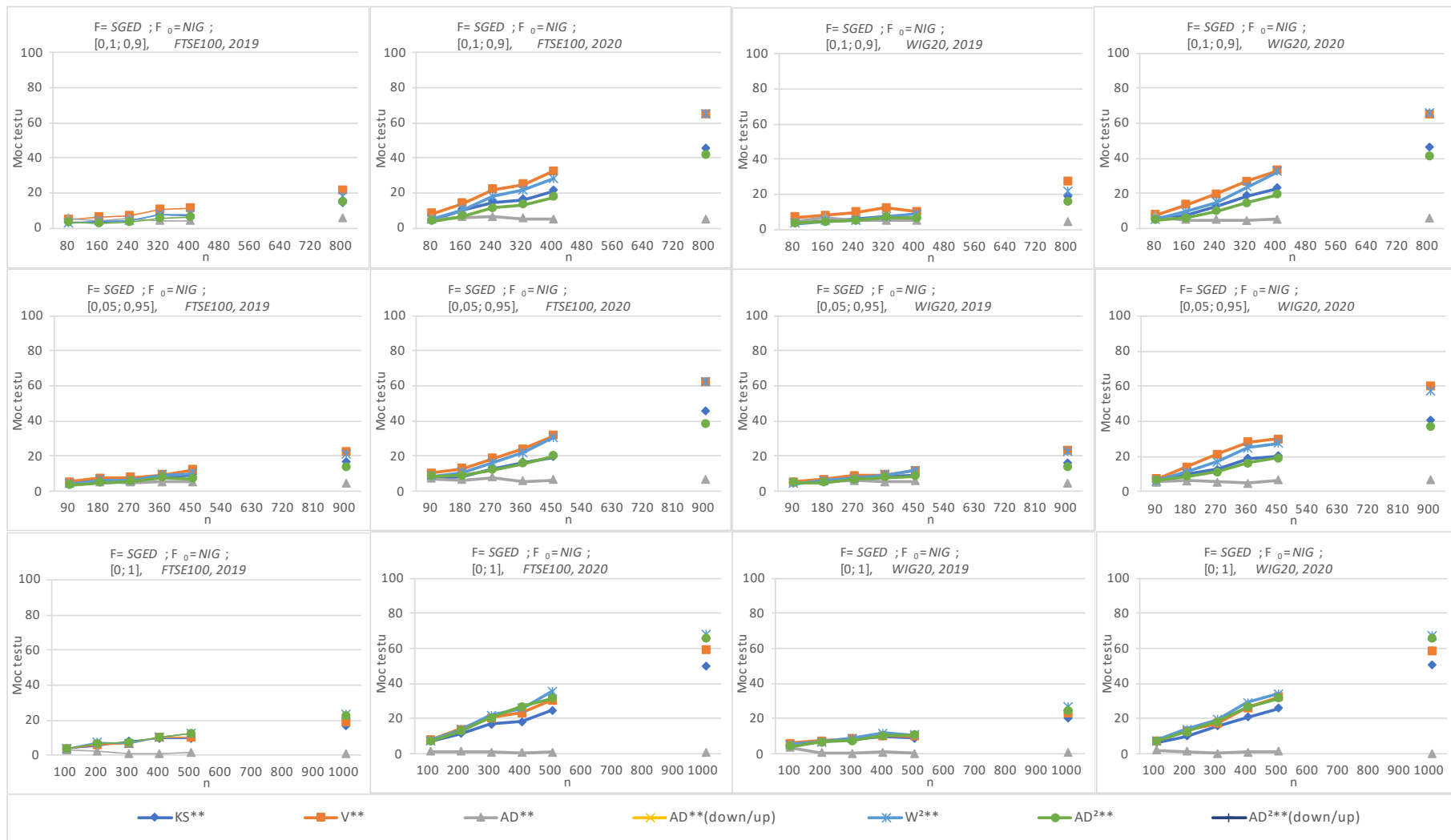
Rysunek 3.6.4. Oszacowania mocy testów zgodności dla $H_0: F = NIG$ (ogon górny)

Źródło: obliczenia własne



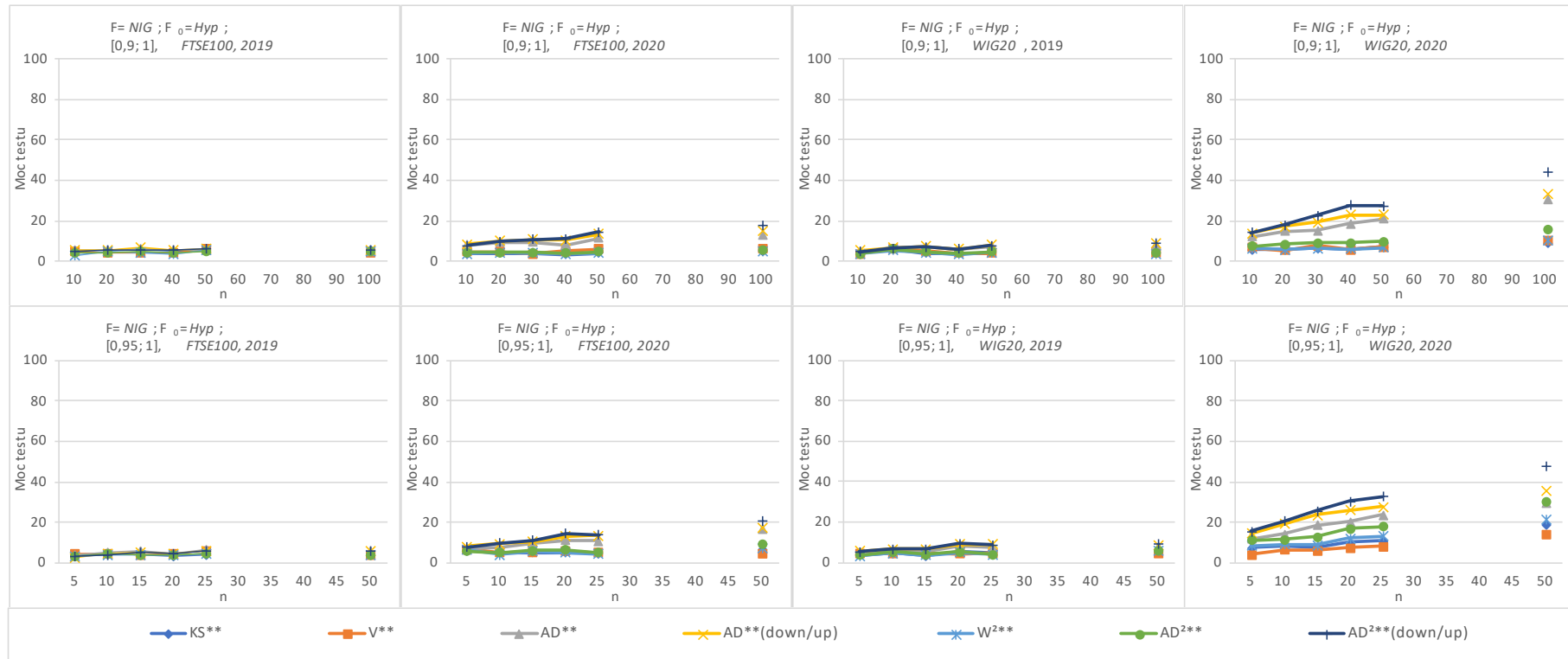
Rysunek 3.6.5. Oszacowania mocy testów zgodności dla $H_0: F = NIG$ (ogon dolny)

Źródło: obliczenia własne



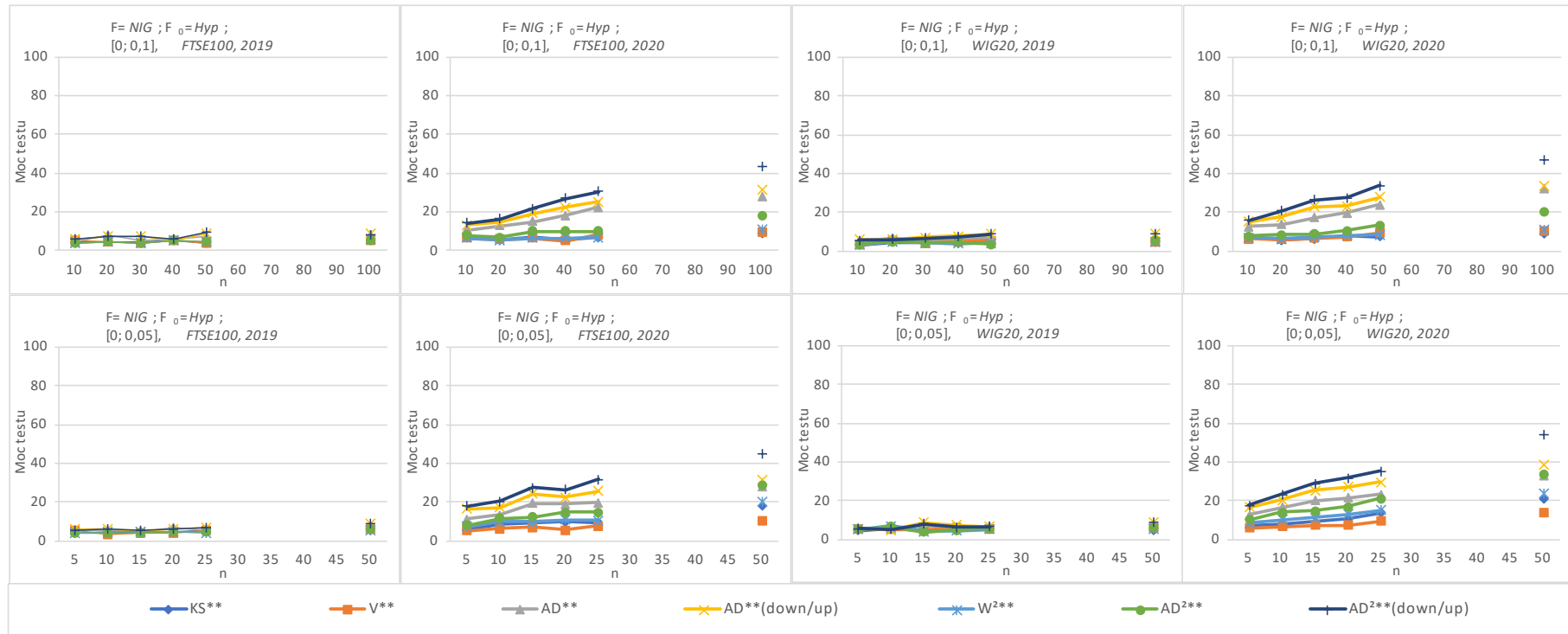
Rysunek 3.6.6. Oszacowania mocy testów zgodności dla $H_0: F = NIG$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład)

Źródło: obliczenia własne



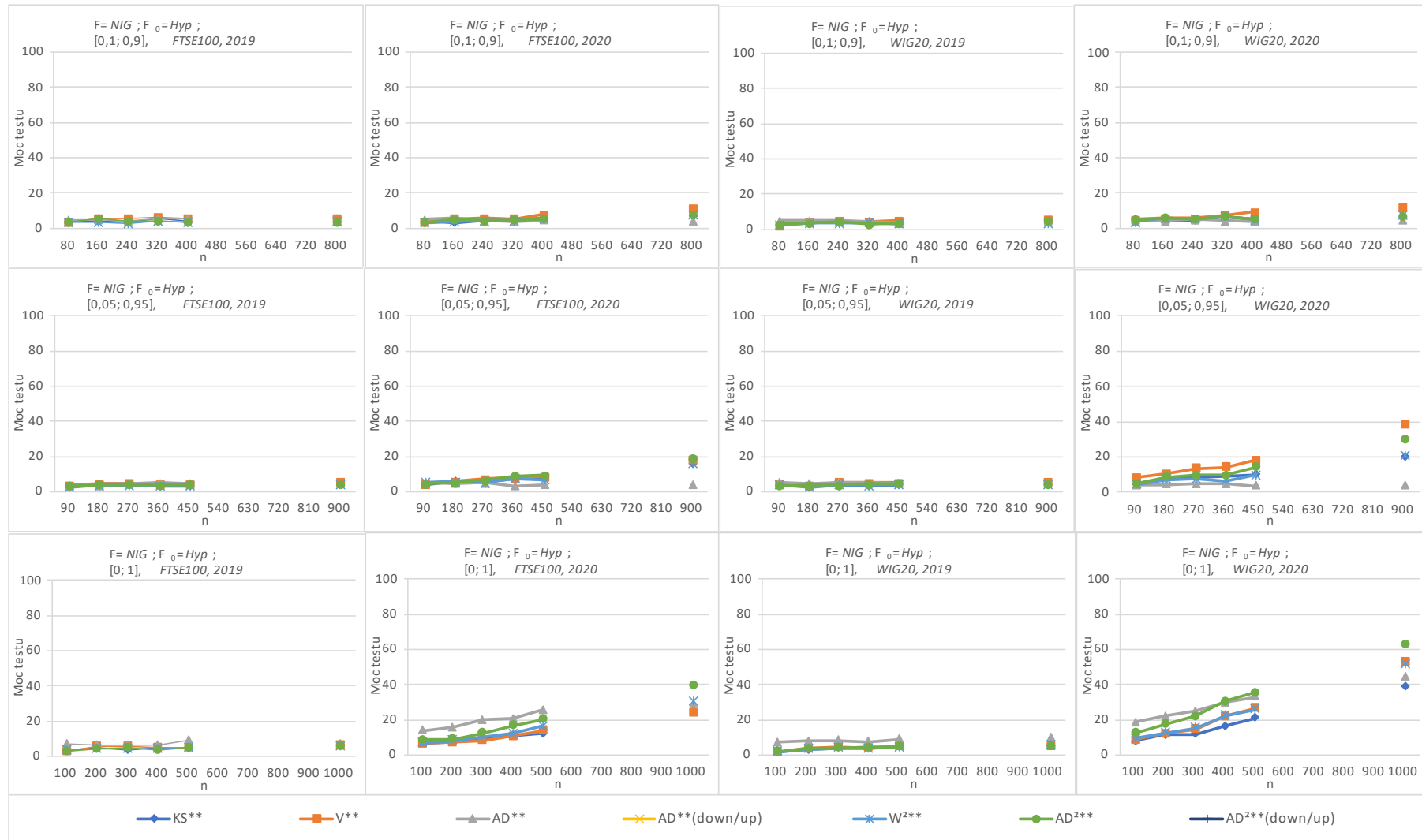
Rysunek 3.6.7. Oszacowania mocy testów zgodności dla $H_0: F = Hyp$ (ogon górny)

Źródło: obliczenia własne



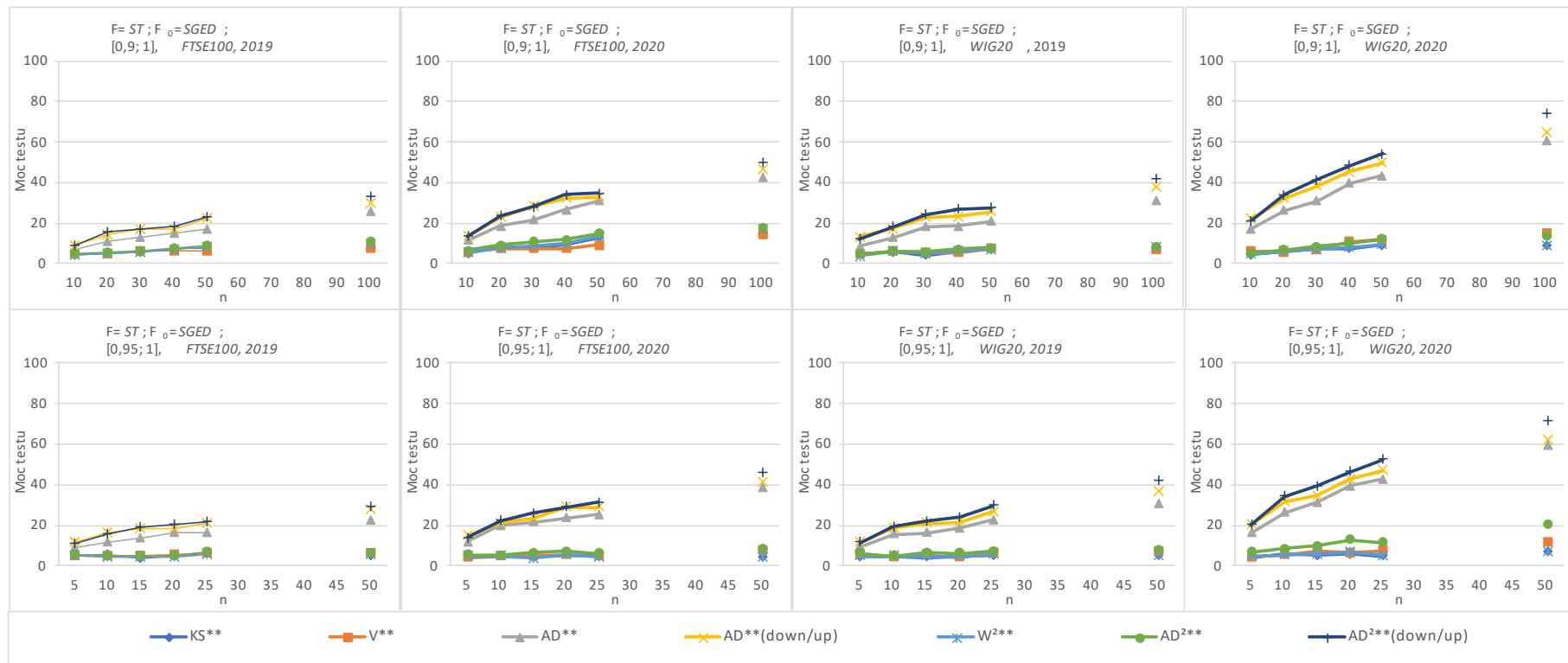
Rysunek 3.6.8. Oszacowania mocy testów zgodności dla $H_0: F = Hyp$ (ogon dolny)

Źródło: obliczenia własne



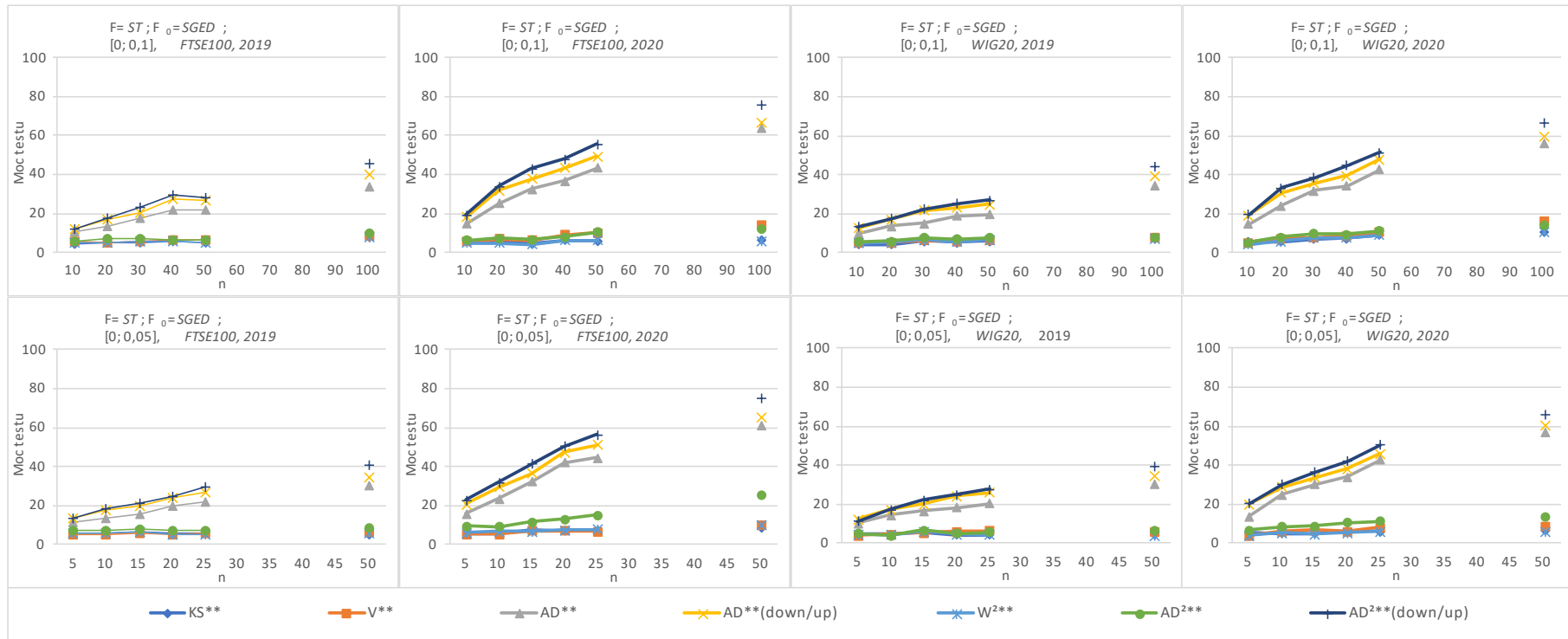
Rysunek 3.6.9. Oszacowania mocy testów zgodności dla $H_0: F = Hyp$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład)

Źródło: obliczenia własne



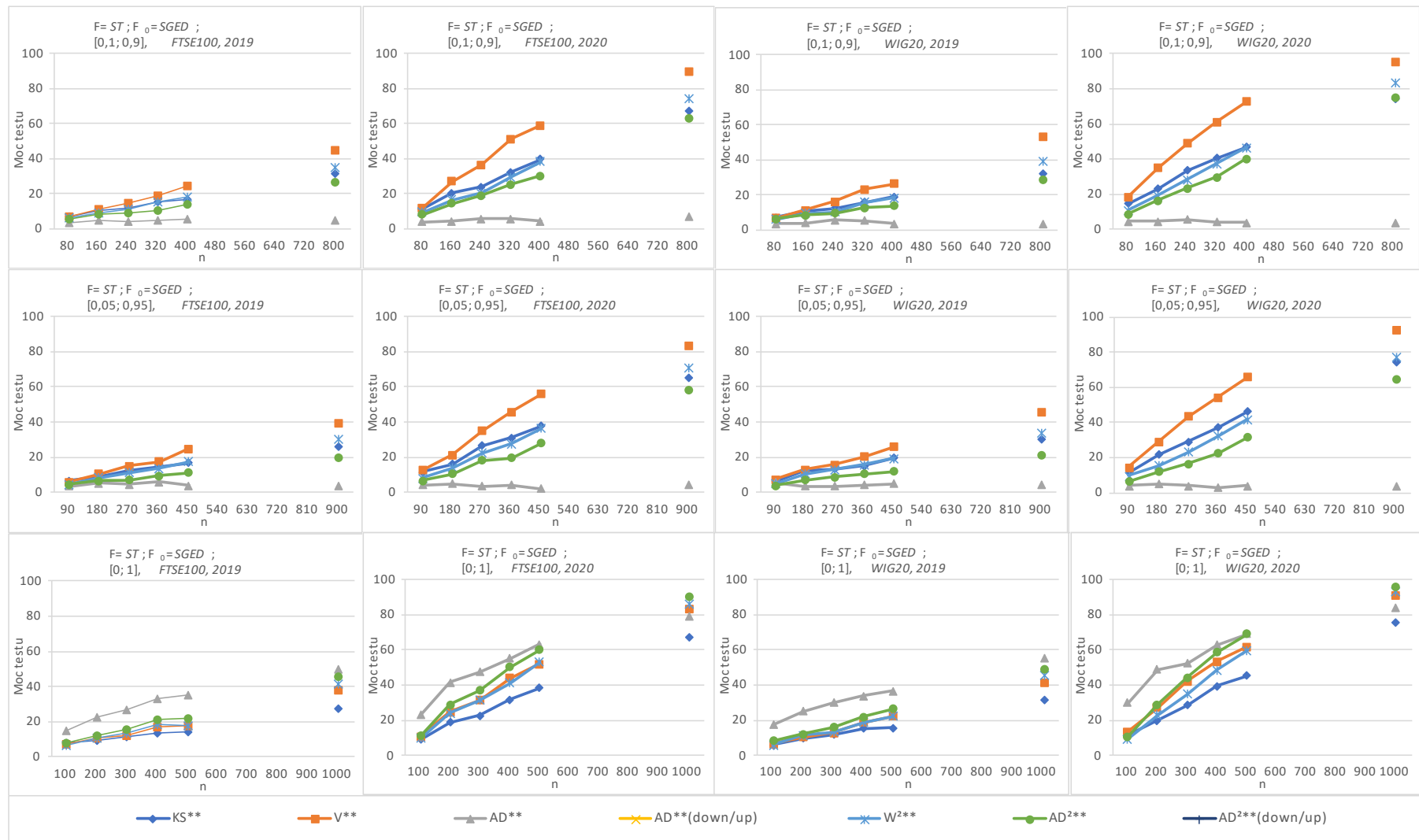
Rysunek 3.6.10. Oszacowania mocy testów zgodności dla $H_0: F = SGED$ (ogon górny)

Źródło: obliczenia własne



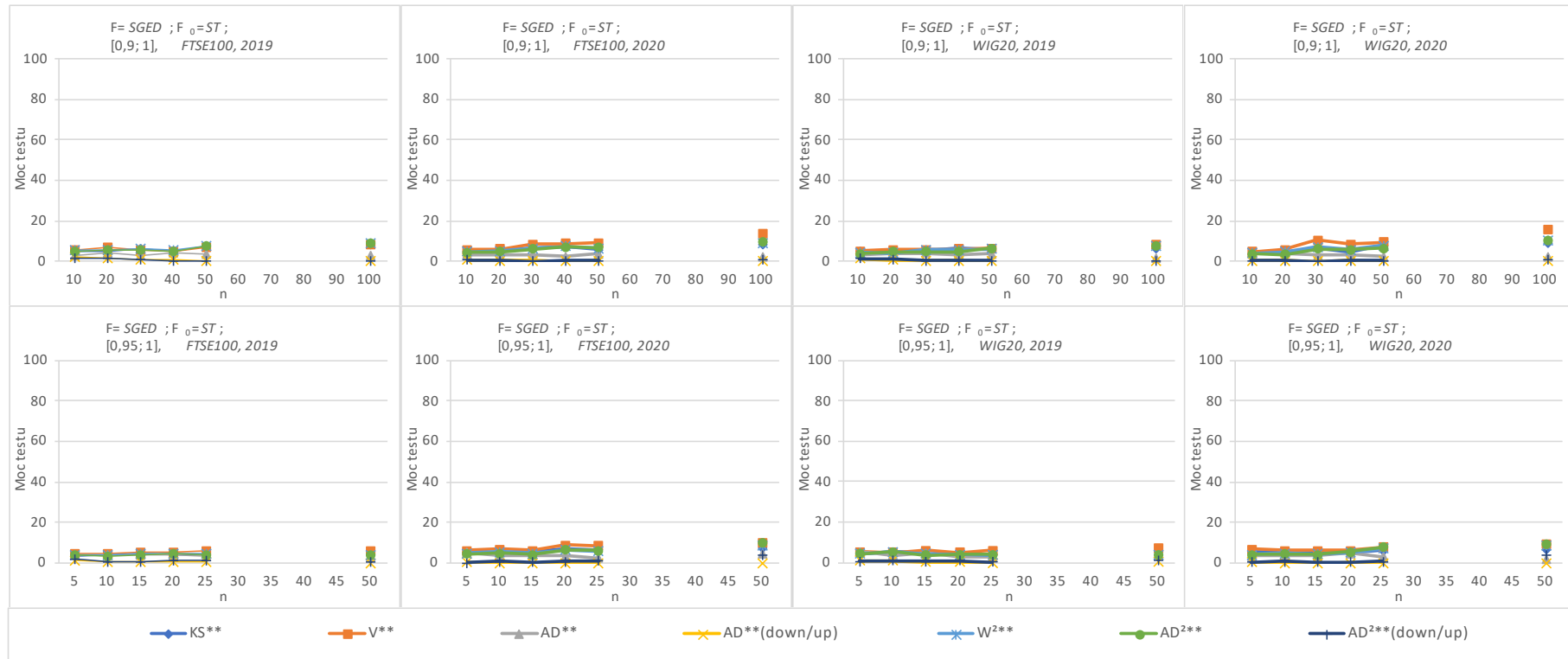
Rysunek 3.6.11. Oszacowania mocy testów zgodności dla $H_0: F = SGED$ (ogon dolny)

Źródło: obliczenia własne



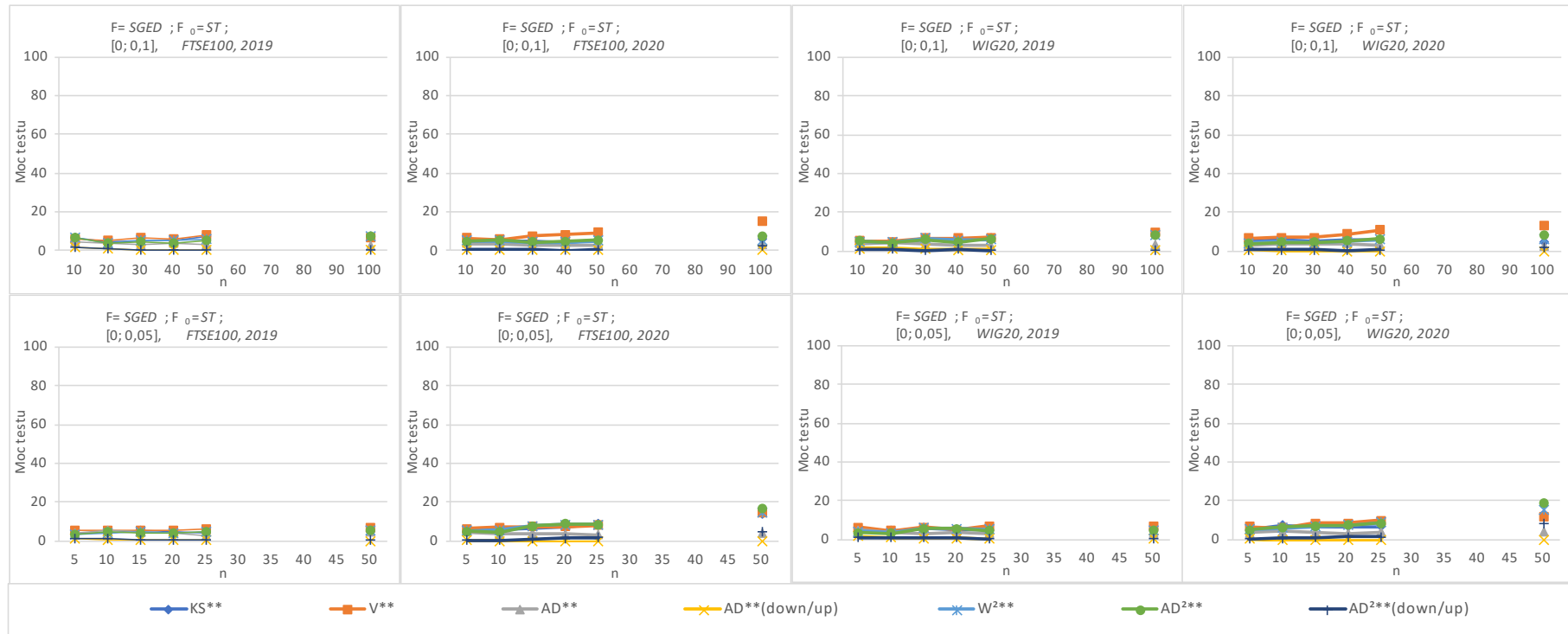
Rysunek 3.6.12. Oszacowania mocy testów zgodności dla $H_0: F = SGED$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład)

Źródło: obliczenia własne



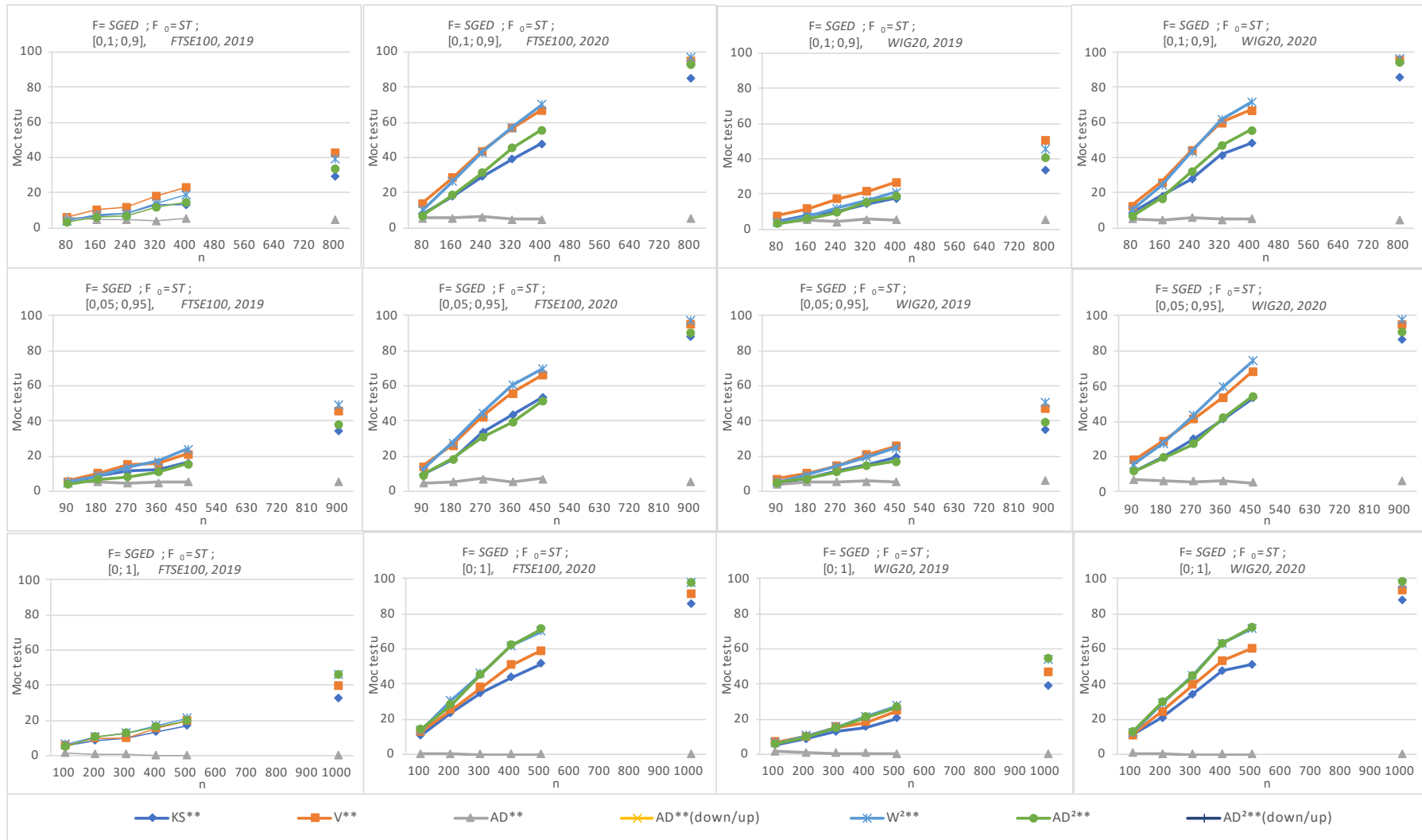
Rysunek 3.6.13. Oszacowania mocy testów zgodności dla $H_0: F = ST$ (ogon górny)

Źródło: obliczenia własne



Rysunek 3.6.14. Oszacowania mocy testów zgodności dla $H_0: F = ST$ (ogon dolny)

Źródło: obliczenia własne



Rysunek 3.6.15. Oszacowania mocy testów zgodności dla $H_0: F = ST$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład)

Źródło: obliczenia własne

4. Badania empiryczne

4.1. Opis badań

Na podstawie wyników badań empirycznych przeprowadzonych w podrozdziałach drugim, trzecim i czwartym, zostanie udzielona odpowiedź na pytanie, czy zastosowanie zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych w ogonach rozkładów pozwala na wybór nie gorszego modelu wartości zagrożonej niż model wartości zagrożonej wybrany na podstawie testów zgodności dla całych rozkładów.

W podrozdziałach drugim i trzecim przekroczenia wartości zagrożonej będą badane na podstawie tej samej próby, która była podstawą do szacowania wartości zagrożonej. W tych podrozdziałach rozkłady jednowymiarowe oraz rozkłady brzegowe w przypadku rozkładów dwuwymiarowych będą wybierane na podstawie testów zgodności dla całych rozkładów oraz na podstawie testów zgodności do badania ogonów rozkładów. Testy statystyczne do badania ogonów rozkładów są tu rozumiane jako szczególny przypadek testów do badania rozkładów obustronnie uciętych, z progami ucięcia $z_{H_1} \neq 0$ i $z_{H_2} = 1$ oraz $z_{H_1} = 0$ i $z_{H_2} \neq 1$, są to więc testy do badania rozkładów lewostronnie lub prawostronnie uciętych. Jakość oszacowań wartości zagrożonych zostanie oceniona w zależności od wyboru testu zgodności. Wartość zagrożona będzie badana na standardowo przyjmowanych poziomach tolerancji 0,01 i 0,05 oraz w celach poznawczych na poziomie tolerancji 0,10. Ten ostatni poziom tolerancji nie występuje w zastosowaniach praktycznych, jednak został przyjęty ze względu na ustalone poziomy ucięcia dystrybuanty, jeden z nich wynosi właśnie 0,10. W podrozdziale czwartym przekroczenia wartości zagrożonej będą weryfikowane poza próbą. W tym podrozdziale przedstawione zostaną trzy modele oceny wartości zagrożonej, na podstawie których będą wyznaczane jednodniowe prognozy wartości zagrożonych.

Na podstawie wyników badania empirycznego przeprowadzonego w podrozdziale piątym, zostanie udzielona odpowiedź na pytanie, czy w przypadku konstrukcji portfela o zadanej stopie zwrotu, zastosowanie zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych w części centralnej rozkładów pozwala na wybór nie gorszego modelu oceny ryzyka portfela papierów wartościowych niż model wybrany na podstawie testów zgodności dla całych rozkładów. W podrozdziale piątym badaniom będzie podlegała jakość portfeli dwuskładnikowych. Badanie zostanie przeprowadzone w oparciu o rozkłady dwuwymiarowe,

utworzone poprzez połączenie rozkładów brzegowych kopulami. W tej części badania rozkłady brzegowe będą wybierane na podstawie testów do badania rozkładów obustronnie uciętych, z progami ucięcia $z_{H_1} \neq 0$ i $z_{H_2} \neq 1$ oraz na podstawie testów do badania całych rozkładów.

We wszystkich podrozdziałach oprócz czwartego, badania empiryczne oparto na tych samych danych. Dane te zastosowano wcześniej w rozdziale trzecim do wyznaczania rozmiarów i mocy testów. Szczegółowy opis danych znajduje się w podrozdziale 3.4 zatytułowanym „Charakterystyka danych”. Badanie zawarte w podrozdziale czwartym wymagało użycia dłuższych szeregów czasowych, dane te zostaną opisane bezpośrednio w tym podrozdziale.

W eksperymentach przeprowadzonych w podrozdziałach drugim, trzecim i piątym, wykorzystano te same wyniki weryfikacji zgodności rozkładów empirycznych z wybranymi rozkładami teoretycznymi, przy zastosowaniu wybranych statystyk testowych. Weryfikacji dokonano na podstawie opisanego w podrozdziale 3.2 układu hipotez (3.2.2), w oparciu o procedurę I, również opisaną w wymienionym podrozdziale. Uzyskane podczas weryfikacji hipotez wartości p znajdują się w załączniku 4 umieszczonym w aneksie. W badaniach przyjęto poziom istotności $\alpha = 0,05$. Ze względu na dużą liczbę hipotez testowanych w podrozdziale czwartym, dane te nie zostaną zaprezentowane. W dalszej części rozdziału bezpośrednio po nazwach statystyk testowych zostaną podane przedziały, wskazujące na fragmenty dystrybuanty, których test zgodności dotyczył.

Rozkłady dwuwymiarowe w podrozdziałach trzecim i piątym wyznaczono łącząc rozkłady brzegowe kopulami. Tak powstałe rozkłady są w pełni parametryczne, a do estymacji ich parametrów zastosowano jednostopniową metodę największej wiarygodności. Do badań wybrano następujące rozkłady brzegowe: normalny (N), normalny odwrotny rozkład Gaussowski (NIG), hiperboliczny (Hyp), skośny uogólniony rozkład błędu ($SGED$) i skośny rozkład t -Studenta (ST). Wśród kopuli znalazły się natomiast kopule: Gaussa, t -Studenta, Claytona, Gumbela, Franka i Joe. Możliwe było utworzenie 150 rozkładów dwuwymiarowych dla każdego okresu (5 rozkładów brzegowych dla indeksu $WIG20$ x 5 rozkładów brzegowych dla indeksu $FTSE100$ x 6 kopul). Wybrane rozkłady prawdopodobieństwa oraz kopule zostały opisane w rozdziale drugim. Oszacowania parametrów rozkładów brzegowych oraz kopul znajdują się w załączniku 5 umieszczonym w aneksie.

4.2. Wpływ wyboru testu zgodności na oszacowania wartości zagrożonej na przykładzie rozkładów jednowymiarowych

W tym podrozdziale wartości zagrożone zostaną wyznaczone jako kwantyl rozkładu na podstawie rozkładów jednowymiarowych, a dokładność oszacowań będzie badana na podstawie tej samej próby, która była podstawą do szacowania wartości zagrożonej. Do badania wybrano następujące rozkłady bezwarunkowe: N , NIG , Hyp , $SGED$ i ST . Zbiór oszacowań wartości zagrożonej dla wybranego kwantyla będzie się więc składał z pięciu elementów, oznaczmy ten zbiór przez A_0 . Niech zbiór A_1 , gdzie $A_1 \subset A_0$, oznacza zbiór tych oszacowań wartości zagrożonych, które na podstawie wyników testowania wstecznego, opisanego w dalszej części podrozdziału, określono jako dokładne. Zbiór A_2 natomiast, gdzie $A_2 \subset A_0$, niech oznacza zbiór wartości zagrożonych oszacowanych na podstawie tych rozkładów teoretycznych, dla których nie było podstaw do odrzucenia hipotezy zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym. Jeśli stwierdzono zgodność rozkładu empirycznego z wybranym rozkładem teoretycznym, to szacunki wartości zagrożonych wyznaczone na podstawie rozkładu teoretycznego ze zbioru A_2 powinny zawierać się w zbiorze A_1 .

Wartość zagrożoną definiujemy dla pozycji długiej jako kwantyl rzędu q : $VaR_q = -F^{-1}(q)$, a dla pozycji krótkiej jako kwantyl rzędu $1 - q$: $VaR_{1-q} = F^{-1}(1 - q)$. Przedmiotem niniejszej pracy jest ocena poprawności pewnego wybranego modelu, dlatego wartość zagrożona zostanie oszacowana metodą Monte Carlo, która polega na generowaniu stóp zwrotu z założonego modelu (Just, 2012, s. 184).

Oceny jakości wartości zagrożonej dokonuje się stosując tzw. testowanie wsteczne (*backtesting*), poprzez wyznaczenie liczby przekroczeń oszacowanej wartości zagrożonej. Funkcję charakterystyczną zbioru określającą sekwencję przekroczeń wartości zagrożonej można określić dla pozycji długiej następująco (analogicznie postępuje się w przypadku pozycji krótkiej):

$$I_t = \begin{cases} 1, & r_t < -VaR_q, \\ 0, & r_t \geq -VaR_q. \end{cases} \quad (4.2.1.)$$

Udział liczby przekroczeń w całkowitej liczbie obserwacji powinien być zbliżony do założonego poziomu tolerancji q . Udział przekroczeń znacznie przekraczający założony poziom tolerancji wskazuje na niedoszacowanie ryzyka, a znacznie od niego mniejszy oznacza przeszacowanie

ryzyka. Najczęściej stosowanym testem statystycznym oceniającym jakość wartości zagrożonej jest test Kupca (Kupiec, 1995), zwany również bezwarunkowym testem pokrycia.

Układ weryfikowanych hipotez w przypadku testu Kupca ma postać:

$$\begin{aligned} H_0: \frac{k}{n} &= q, \\ H_1: \frac{k}{n} &\neq q, \end{aligned} \tag{4.2.2}$$

gdzie k oznacza liczbę przekroczeń, n oznacza liczbę obserwacji, a q oznacza przyjęty poziom tolerancji. Statystyka testowa dla testu Kupca ma postać:

$$LR = 2 \left(\ln \left(\left(\frac{k}{n} \right)^k \left(1 - \frac{k}{n} \right)^{n-k} \right) - \ln((q)^k (1 - q)^{n-k}) \right). \tag{4.2.3}$$

W przypadku prawdziwości hipotezy zerowej statystyka LR ma asymptotyczny rozkład χ^2 z jednym stopniem swobody. Test Kupca weryfikuje jedynie udział przekroczeń, nie sprawdzając niezależności pomiędzy przekroczeniami. W niniejszej pracy badaniom podlegają rozkłady bezwarunkowe, w tej sytuacji jakość wartości zagrożonej można ocenić testem Kupca. Jeśli jednak istnieje podejrzenie, że pomiędzy przekroczeniami występuje zależność, można wykonać np. test Christoffersena (Christoffersen, 1998). Więcej na temat testów oceniających wartość zagrożoną można znaleźć m. in. w pracach Echausta i Just (2020), Jeziorskiego (2014) oraz Just (2012).

Badanie empiryczne, którego wyniki zaprezentowano w tym podrozdziale, składa się z następujących etapów:

1. Określenie źródła danych.

Danymi są dzienne procentowe logarytmiczne stopy zwrotu, wyliczone na podstawie kursów zamknięcia indeksów *WIG20* oraz *FTSE100* z lat 2019-2020, szczegółowo opisane w podrozdziale 3.4.

2. Zdefiniowanie modeli do generowania symulacji Monte Carlo.

Modelami wartości zagrożonej, z których będą generowane dane do symulacji Monte Carlo, są jednowymiarowe rozkłady N , NIG , Hyp , $SGED$ i ST . Parametry rozkładów zostały ustalone na podstawie dziennych logarytmicznych procentowych stóp zwrotu wyliczonych dla kursów zamknięcia indeksów *WIG20* i *FTSE100* w latach 2019-2020. Rozkłady oraz ich parametry znajdują się w tabeli 3.4.2. Łącznie utworzono 20 wariantów (2 indeksy x 2 okresy x 5 rozkładów).

3. Oszacowanie wartości zagrożonych.

Z każdego rozkładu jednowymiarowego utworzonego w punkcie 2 losowano próbę o liczebności 10 000. Dla każdej wylosowanej próby wyznaczano kwantyle rzędu 0,01, 0,05, 0,1, 0,9, 0,95 i 0,99 (tabela 4.2.1).

4. Ocena jakości oszacowań wartości zagrożonych wyznaczonych w punkcie 3.

Oceny jakości dokonano w oparciu o test Kupca. Udziały liczby przekroczeń w całkowitej liczbie obserwacji zostały zaprezentowane w tabeli 4.2.2. Odrzucenie hipotezy H_0 (4.2.2) przez test Kupca na poziomie istotności 0,1, 0,05 i 0,01 oznaczono w tabelach 4.2.1. i 4.2.2. odpowiednio przez *, **, ***.

5. Ocena zgodności rozkładu rzeczywistego z wybranym rozkładem teoretycznym.

Oceny zgodności dokonano w oparciu o statystyki dla całych rozkładów (KS^{**} , V^{**} , AD^{**} , W^{2**} , AD^{2**}), statystyki dla dolnych ogonów rozkładów, z punktami ucięcia 0,05 i 0,1 (KS^{**} , V^{**} , AD^{**} , AD_{down}^{**} , W^{2**} , AD^{2**} , AD_{down}^{2**}) oraz statystyki dla górnych ogonów rozkładów, z punktami ucięcia 0,9 i 0,95 (KS^{**} , V^{**} , AD^{**} , AD_{up}^{**} , W^{2**} , AD^{2**} , AD_{up}^{2**}). Wyniki oceny znajdują się w załączniku 4 w aneksie.

Tabela 4.2.1. Oszacowania wartości zagrożonych dla rozkładów jednowymiarowych ⁽¹⁾

Rok	Indeks	Rozkład	Dolny ogon			Górny ogon		
			Rząd kwantyla (poziom tolerancji dla VaR)					
			0,1	0,05	0,01	0,9	0,95	0,99
2019	WIG20	N	-1,2295	-1,5729	-2,1853	1,1935	1,5266	2,1874
		NIG	-1,1772	-1,6050	-2,5680	1,0969	1,5175	2,5410
		Hyp	-1,1703	-1,5917	-2,5445	1,1064	1,5404	2,4595
		SGED	-1,1642	-1,5624	-2,3985	1,1467	1,5737	2,5399
		ST	-1,1232	-1,4985	-2,4966	1,0824	1,4708	2,3939
	FTSE100	N	-0,8978	-1,1799	-1,7160	0,9849	1,2739	1,7910
		NIG	-0,8751	-1,2149	-1,9590	0,9178	1,1783	1,8028
		Hyp	-0,8392	-1,1610	-1,8444	0,9151	1,1905	1,8142
		SGED	-0,8927	-1,2194	-1,8699	0,9202	1,2044	1,8423
		ST	-0,8243	-1,1556	-1,8966	0,8977	1,1864	1,8548
2020	WIG20	N	-2,6632	-3,4198	-4,8265	2,6039	3,3339	4,8412
		NIG	-2,2009	-3,2774	-5,9429	2,2061	3,1905	6,0791
		Hyp	-2,2469	-3,2071	-5,2223	2,1360	3,0143	5,1473
		SGED	-2,2754	-3,3670	-6,0825	2,1275	3,1217	5,6778
		ST	-2,0484	-2,9362	-5,9734	2,1027	3,1080	6,5647 **
	FTSE100	N	-2,4126	-3,0356	-4,2966	2,3474 ***	3,0249 **	4,2791
		NIG	-2,0990	-3,1403	-5,6669	1,8040	2,5262	4,4403
		Hyp	-2,1673	-3,0641	-5,2954	1,8712	2,5911	4,2661
		SGED	-2,1074	-3,0756	-5,2191	1,9048	2,7949 *	4,6659
		ST	-1,9960	-2,9544	-6,0202	1,8305	2,4774	4,5589

(1) odrzucenie hipotezy H_0 (4.2.2) na podstawie testu Kupca na poziomie istotności 0,1, 0,05 i 0,01 oznaczono odpowiednio przez *, **, ***

Źródło: obliczenia własne

Tabela 4.2.2. Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów jednowymiarowych ⁽¹⁾

Okres	Indeks	Rozkłady	Dolny ogon			Górny ogon		
			Rząd kwantyla (poziom tolerancji dla VaR)					
			0,1	0,05	0,01	0,9	0,95	0,99
2019	WIG20	N	0,0887	0,0565	0,0202	0,0806	0,0605	0,0161
		NIG	0,0968	0,0524	0,0040	0,0847	0,0605	0,0121
		Hyp	0,0968	0,0524	0,0081	0,0847	0,0565	0,0121
		SGED	0,0968	0,0565	0,0121	0,0847	0,0524	0,0121
		ST	0,1129	0,0685	0,0081	0,0887	0,0605	0,0161
	FTSE100	N	0,0711	0,0316	0,0198	0,0791	0,0316	0,0198
		NIG	0,0751	0,0316	0,0158	0,1028	0,0395	0,0198
		Hyp	0,0870	0,0316	0,0158	0,1067	0,0395	0,0158
		SGED	0,0711	0,0316	0,0158	0,1028	0,0356	0,0158
		ST	0,0909	0,0356	0,0158	0,1067	0,0395	0,0158
2020	WIG20	N	0,0754	0,0397	0,0159	0,0754	0,0476	0,0079
		NIG	0,0913	0,0516	0,0119	0,1032	0,0556	0,0040
		Hyp	0,0833	0,0556	0,0159	0,1111	0,0595	0,0040
		SGED	0,0833	0,0437	0,0079	0,1111	0,0556	0,0040
		ST	0,0952	0,0595	0,0119	0,1111	0,0556	0,0000 **
	FTSE100	N	0,0748	0,0630	0,0118	0,0512 ***	0,0197 **	0,0118
		NIG	0,0945	0,0630	0,0079	0,0984	0,0433	0,0079
		Hyp	0,0906	0,0630	0,0118	0,0945	0,0394	0,0118
		SGED	0,0945	0,0630	0,0118	0,0866	0,0276 *	0,0039
		ST	0,0945	0,0669	0,0079	0,0945	0,0433	0,0079

(1) odrzucenie hipotezy H_0 (4.2.2) na podstawie testu Kupca na poziomie istotności 0,1, 0,05 i 0,01 oznaczono odpowiednio przez *, **, ***

Źródło: obliczenia własne

Należy zaznaczyć, że punkty 4 i 5 opisanego powyżej badania zostały przeprowadzone niezależnie. Testy zgodności zostały przeprowadzone dla wszystkich rozkładów, a test Kupca dla wszystkich oszacowań wartości zagrożonej.

W każdym okresie, na dowolnym poziomie tolerancji, oszacowania wartości zagrożonej były wyznaczane na podstawie pięciu rozkładów. W wielu analizowanych przypadkach, wszystkie oszacowania wartości zagrożonej zostały uznane, na podstawie testu Kupca, jako dokładne. Oznacza to, że modele wartości zagrożonej, wybrane na podstawie dowolnego testu zgodności, również generują oszacowania uznane za dokładne. Nie ma tu zatem znaczenia, czy zostanie zastosowany test weryfikujący zgodność całego rozkładu, czy test weryfikujący zgodność jedynie w ogonach rozkładu. Z tej przyczyny interesujące są te przypadki, gdy nie wszystkie oszacowania wartości zagrożonej zostały uznane za dokładne. Spośród wszystkich wyznaczonych oszacowań wartości zagrożonej, na podstawie testu Kupca, na poziomie istotności 0,05, tylko trzy z nich uznano za nieprawidłowe. Były to oszacowania

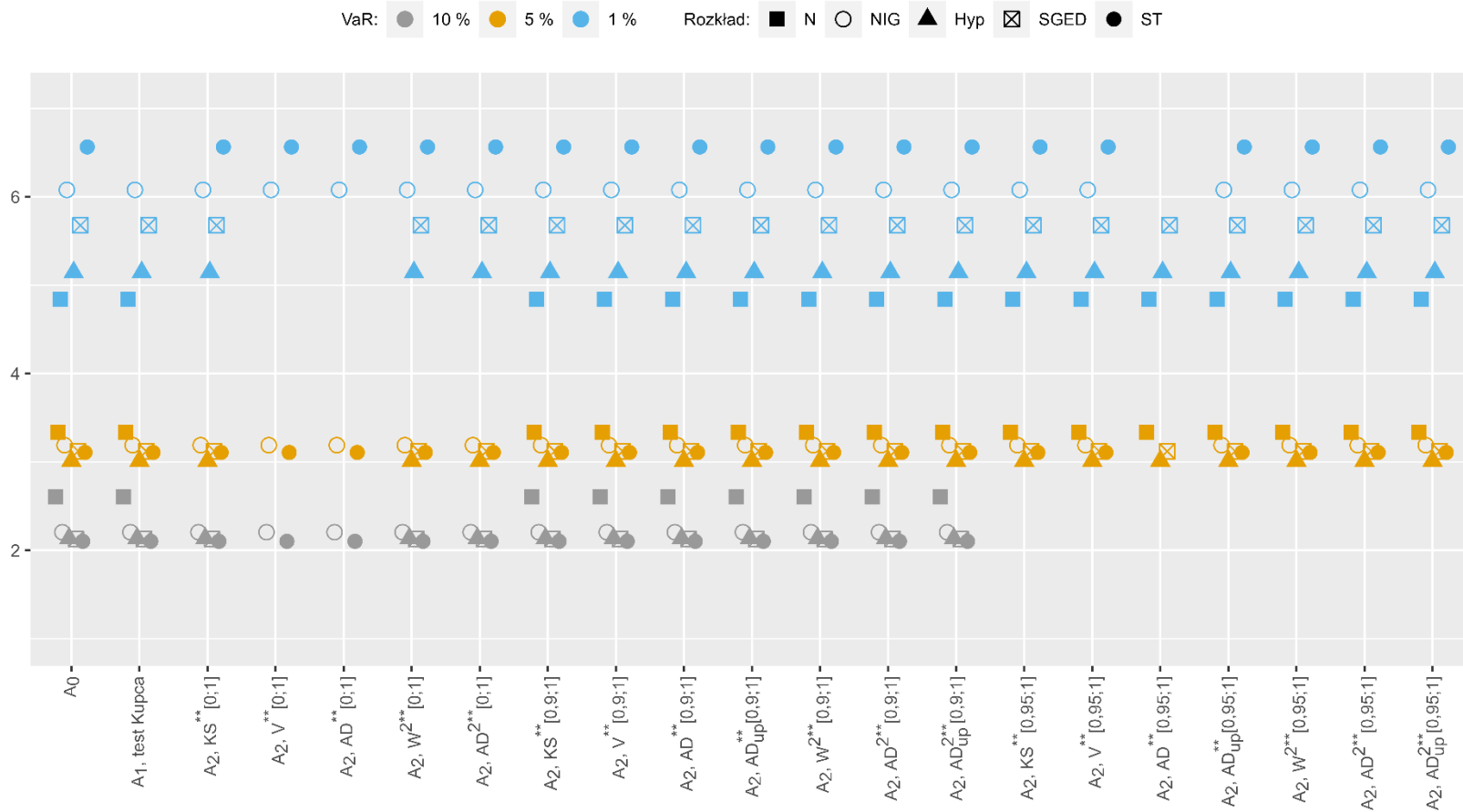
wartości zagrożonej dla kwantyli rzędu 0,9 i 0,95, wygenerowane na podstawie rozkładu normalnego, którego parametry wzorowano na stopach zwrotu z indeksu *FTSE100* w roku 2020 oraz oszacowania dla kwantyla rzędu 0,99, wygenerowane na podstawie rozkładu *ST*, którego parametry wzorowano na stopach zwrotu z indeksu *WIG20* również w roku 2020. W wymienionych powyżej trzech przypadkach ryzyko zostało przeszacowane. Przypadki, gdy nie wszystkie oszacowania wartości zagrożonej zostały uznane za dokładne, przedstawiono na rysunkach 4.2.1 i 4.2.2. Na obydwu rysunkach, dla każdego uwzględnionego w badaniu poziomu tolerancji, znajdują się zbiory oszacowań wartości zagrożonych oznaczone jako A_0 , A_1 i A_2 , opisane na początku tego podrozdziału. Zbiórów A_2 jest tyle, ile zastosowano statystyk w testach zgodności. W celu ich rozróżnienia, do nazwy zbioru A_2 dołączano również nazwę statystyki, na podstawie której przeprowadzono test zgodności oraz oznaczenie fragmentu dystrybuanty, której test dotyczył.

Na rysunku 4.2.1 przedstawiono zbiory oszacowań wartości zagrożonych wyznaczonych na podstawie modeli wartości zagrożonych dla indeksu *WIG20* w roku 2020, dla pozycji krótkiej. Na poziomach tolerancji 0,05 i 0,1 wszystkie oszacowania wartości zagrożonej zostały uznane, na podstawie testu Kupca, jako dokładne. Na poziomie tolerancji 0,01 oszacowania wyznaczone na podstawie rozkładu *ST* zostały uznane na podstawie testu Kupca za niepoprawne. Hipotezę zakładającą zgodność rozkładu empirycznego z rozkładem *ST* odrzucono stosując tylko jedną statystykę, przeznaczoną do badania górnych ogonów rozkładów, a mianowicie statystykę $AD^{**}[0,95; 1]$. Tylko modele wartości zagrożonej wyselekcjonowane na podstawie tej statystyki generowały zbiory oszacowań wartości zagrożonych zawierające poprawne wartości.

Na rysunku 4.2.2 przedstawiono zbiory oszacowań wartości zagrożonych wyznaczonych na podstawie modeli wartości zagrożonych dla indeksu *FTSE100* w roku 2020, dla pozycji krótkiej. Na poziomie tolerancji 0,01 wszystkie oszacowania wartości zagrożonej zostały uznane, na podstawie testu Kupca, jako dokładne. Stąd w tym przypadku wybór testu do weryfikacji zgodności rozkładu nie miał znaczenia. Na poziomach tolerancji 0,05 i 0,1, oszacowania wyznaczone na podstawie rozkładu normalnego zostały uznane na podstawie testu Kupca za niepoprawne. Hipotezy zakładającej zgodność rozkładu empirycznego z rozkładem normalnym nie można było natomiast odrzucić stosując jedną z pięciu statystyk przeznaczonych do badania górnych ogonów rozkładów: $KS^{**}[0,9; 1]$, $V^{**}[0,9; 1]$, $W^{2**}[0,9; 1]$, $AD^{2**}[0,9; 1]$ oraz $W^{2**}[0,95; 1]$. Stąd, modele wartości zagrożonej

wyselekcjonowane na podstawie tych statystyk, generowały zbiory oszacowań wartości zagrożonych zawierające niepoprawne wartości.

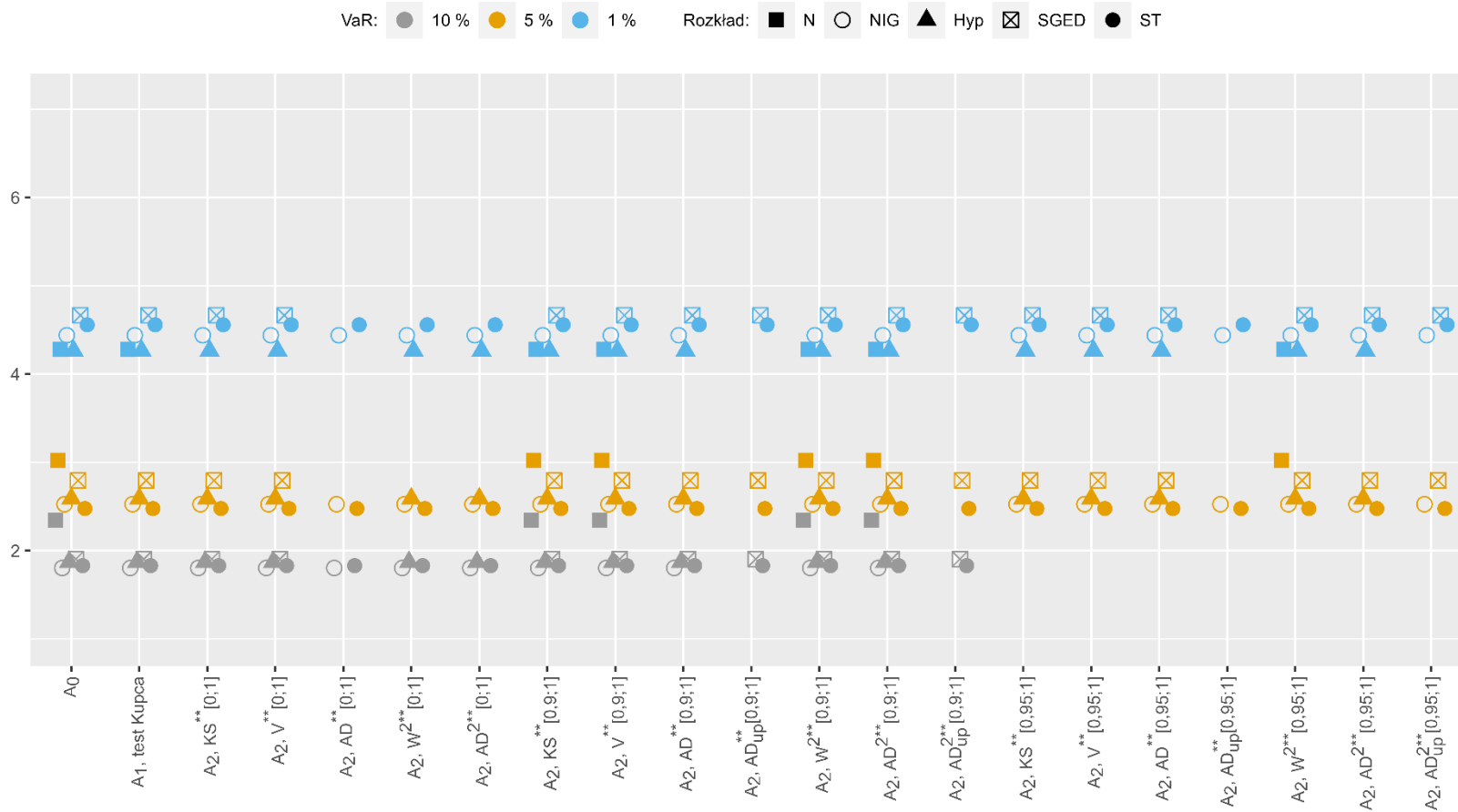
Można założyć, że zastosowanie testu zgodności powinno prowadzić do wyboru tych rozkładów, na podstawie których generowane są oszacowania wartości zagrożonej uznane w oparciu o test Kupca za prawidłowe. Kierując się tą zasadą można z kolei wskazać statystyki, które prowadzą do wyboru takich modeli. Zbiór modeli generujących prawidłowe oszacowania wartości zagrożonej dla ogona dolnego otrzymywano stosując dowolną statystykę. Zbiór modeli generujących prawidłowe oszacowania wartości zagrożonej dla ogona górnego, we wszystkich analizowanych przypadkach, otrzymywano jedynie w wyniku zastosowania statystyki $AD^{**}[0,95; 1]$. Stąd, zastosowanie testu zgodności dla uciętych rozkładów ciągłych w ogonach rozkładów pozwala na wybór nie gorszego modelu wartości zagrożonej niż model wartości zagrożonej wybrany na podstawie testów zgodności dla całych rozkładów. Podsumowując ten podrozdział należy podkreślić, że ostateczny zbiór oszacowań wartości zagrożonych jest częścią wspólną dwóch zbiorów, zbioru utworzonego na podstawie tych modeli wartości zagrożonych, które zostały uznane na podstawie testu Kupca za poprawne oraz zbioru utworzonego na podstawie tych rozkładów, dla których w wyniku przeprowadzenia wybranego testu zgodności nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.



Rysunek 4.2.1 Oszacowania wartości zagrożonej dla pozycji krótkiej (2020, WIG20) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_0 – zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie wszystkich modeli wartości zagrożonej; A_1 – zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie modeli, które w oparciu o test Kupca uznano za dokładne; A_2 – zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie rozkładów, dla których w wyniku przeprowadzenia wybranego testu zgodności nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z wybranym rozkładem teoretycznym.

Źródło: obliczenia własne



Rysunek 4.2.2 Oszacowania wartości zagrożonej dla pozycji krótkiej (2020, FTSE100) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_0 – zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie wszystkich modeli wartości zagrożonej; A_1 - zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie modeli, które w oparciu o test Kupca uznano za dokładne; A_2 - zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie rozkładów, dla których w wyniku przeprowadzenia wybranego testu zgodności nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z wybranym rozkładem teoretycznym.

Źródło: obliczenia własne

4.3. Wpływ wyboru testu zgodności na oszacowania wartości zagrożonej na przykładzie rozkładów dwuwymiarowych

W tym podrozdziale wartości zagrożone zostaną wyznaczone jako kwantyl rozkładu na podstawie rozkładów dwuwymiarowych, a jakość oszacowań będzie badana na podstawie tej samej próby, która była podstawą do szacowania wartości zagrożonej. Rozkłady dwuwymiarowe będą tworzone poprzez połączenie rozkładów brzegowych kopułami. Rozważane będą rozkłady w pełni parametryczne, a ich parametry będą estymowane jednostopniową metodą największej wiarygodności. Rozkłady utworzone za pomocą kopuli mogą być stosowane do szacowania wartości zagrożonej metodą Monte Carlo, a miary ryzyka są w ich przypadku często funkcjami zarówno kopuł, jak i rozkładów brzegowych (Patton, 2013, s. 918). Według Choi i Min (2013) przy stosowaniu takiej metody decydujący wpływ na poziom VaR ma jednak wybór rozkładów brzegowych.

W tym podrozdziale w każdym badanym okresie będzie rozważanych 150 modeli wartości zagrożonej (5 rozkładów brzegowych dla indeksu *WIG20* x 5 rozkładów brzegowych dla indeksu *FTSE100* x 6 kopuł). Zbiór oszacowań wartości zagrożonych dla każdego badanego okresu i wybranego poziomu tolerancji, oznaczony przez A_0 , będzie się więc składał ze 150 elementów. Tutaj, podobnie jak w przypadku rozkładów jednowymiarowych, również zostaną wyróżnione dwa dodatkowe zbiory. Przez A_1 zostanie oznaczony zbiór oszacowań wartości zagrożonych, które na podstawie testu Kupca określono jako dokładne. Przez zbiór A_2 natomiast zostanie oznaczony zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie tych rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym. Dla każdego poziomu tolerancji zbiorów A_2 będzie tyle, ile zastosowano testów statystycznych. Gdy będą analizowane przekroczenia wartości zagrożonych, do indeksów dolnych wyżej opisanych zbiorów zostanie dodana litera P, oznaczenia będą więc następujące: A_{0P} , A_{1P} , A_{2P} .

Badanie empiryczne składa się z następujących etapów:

1. Określenie źródła danych.

Danymi są dzienne procentowe logarytmiczne stopy zwrotu, wyliczone na podstawie kursów zamknięcia indeksów *WIG20* oraz *FTSE100* z lat 2019-2020, szczegółowo opisane w podrozdziale 3.4.

2. Zdefiniowanie modeli do generowania symulacji Monte Carlo.
 Na podstawie danych określonych w punkcie 1, oddzielnie dla roku 2019 i roku 2020, oszacowano parametry rozkładów dwuwymiarowych utworzonych poprzez połączenie rozkładów brzegowych (N , NIG , Hyp , $SGED$, ST) kopulami (Gaussa, t -Studenta, Clayтона, Gumbela, Franka i Joe). Oszacowania parametrów rozkładów brzegowych oraz kopul znajdują się w załączniku 5 umieszczonym w aneksie. W załączniku 5 zamieszczono również następujące miary zależności kopul: τ Kendalla, współczynnik λ^L zależności w dolnym ogonie oraz współczynnik λ^U zależności w górnym ogonie. Rozkłady z oszacowanymi parametrami są modelami do generowania symulacji. Łącznie utworzono 300 modeli (5 rozkładów brzegowych dla indeksu $WIG20$ x 5 rozkładów brzegowych dla indeksu $FTSE100$ x 6 kopul x 2 okresy).
3. Oszacowanie wartości zagrożonych.
 Z każdego rozkładu dwuwymiarowego utworzonego w punkcie 2 losowano próbę o liczebności 10 000. Następnie tworzone portfel dwuskładnikowy o wagach $w_1 = w_2 = 0,5$. Dla każdego portfela wyznaczano kwantyle rzędu 0,01, 0,05, 0,1, 0,9, 0,95 i 0,99. Otrzymane w ten sposób oszacowania wartości zagrożonych znajdują się w załączniku 6 umieszczonym w aneksie.
4. Ocena jakości oszacowań wartości zagrożonych wyznaczonych w punkcie 3.
 Oceny jakości dokonano w oparciu o test Kupca. W celu przeprowadzenia testu Kupca należało utworzyć portfel dwuskładnikowy o równych udziałach z danych rzeczywistych i wyznaczyć dla niego odpowiednie kwantyle. Odrzucenie hipotezy H_0 (4.2.2) przez test Kupca na poziomie istotności 0,1, 0,05 i 0,01 oznaczono w załączniku 6 odpowiednio przez *, **, ***.
5. Ocena zgodności rozkładów rzeczywistych z wybranymi rozkładami teoretycznymi.
 W tym punkcie wykorzystano wyniki testów zgodności z podrozdziału 4.2. Wyniki te zawierają oceny zgodności dokonane w oparciu o statystyki dla całych rozkładów (KS^{**} , V^{**} , AD^{**} , W^{2**} , AD^{2**}), statystyki dla dolnych ogonów rozkładów, z punktami ucięcia dystrybuanty 0,05 i 0,1 (KS^{**} , V^{**} , AD^{**} , AD_{down}^{**} , W^{2**} , AD^{2**} , AD_{down}^{2**}) oraz statystyki dla górnych ogonów rozkładów, z punktami ucięcia dystrybuanty 0,9 i 0,95 (KS^{**} , V^{**} , AD^{**} , AD_{up}^{**} , W^{2**} , AD^{2**} , AD_{up}^{2**}). Ocenie podlegały obydwa rozkłady brzegowe.

Należy zaznaczyć, że oszacowania parametrów rozkładów brzegowych z rozkładu dwuwymiarowego (etap 2 badania), nieznacznie różnią się od oszacowań parametrów rozkładów jednowymiarowych (przyjętych jako wzorce w rozdziale trzecim oraz podlegających badaniu w poprzednim podrozdziale). Różnice te wynikają z dwóch przyczyn. Po pierwsze, do estymacji parametrów rozkładu dwuwymiarowego zastosowano jednostopniową metodę największej wiarygodności. Po drugie, utworzenie rozkładów dwuwymiarowych wymagało połączenia szeregów stóp zwrotu indeksów *WIG20* i *FTSE100*. Wartości brakujące w obydwu szeregach uzupełniono metodą wstawiania ostatniej zaobserwowanej wartości, stąd otrzymane w ten sposób szeregi różnią się od szeregów oryginalnych. W roku 2019 uzupełniono w ten sposób 8 brakujących wartości dla indeksu *WIG20* i 3 dla indeksu *FTSE100*, a w 2020 roku 6 brakujących wartości dla indeksu *WIG20* i 4 dla indeksu *FTSE100*. Z powodu niewielkich różnic pomiędzy oszacowaniami parametrów oraz w związku z czasochłonnością obliczeń, w etapie 5 wyżej opisanego badania, posłużono się wynikami testów zgodności opracowanymi dla rozkładów jednowymiarowych z poprzedniego podrozdziału (umieszczonymi w załączniku 4 w aneksie).

Zbiory oszacowań wartości zagrożonych mogą liczyć do 150 elementów, dlatego zostaną omówione ich rozkłady. Nastąpi to w oparciu o wykresy pudełkowe. Dolny bok prostokąta na każdym wykresie odpowiada pierwszemu kwartyłowi, a górny bok trzeciemu kwartyłowi. Wysokość pudełka odpowiada wartości rozstępu ćwiartkowego. Odcinki odbiegające od pudełek wyznaczają najmniejszą i największą wartość w zbiorze, pod warunkiem, że dolny koniec jest nie mniejszy niż $Q_1 - 1,5 * (Q_3 - Q_1)$, a górny koniec nie większy niż $Q_3 + 1,5 * (Q_3 - Q_1)$. Obserwacje spoza tego przedziału są uznawane za odstające i umieszczane na wykresie w postaci punktów. Linia przecinająca pudełko odpowiada medianie. Rozkłady oszacowań wartości zagrożonych umieszczono na rysunkach: 4.3.1 (rok 2019, pozycja długa), 4.3.2 (rok 2019, pozycja krótka), 4.3.3 (rok 2020, pozycja długa) i 4.3.4 (rok 2020, pozycja krótka). Odpowiadające oszacowaniom wartości zagrożonych udziały przekroczeń w całkowitej liczbie obserwacji umieszczono na rysunkach: 4.3.5 (rok 2019, pozycja długa), 4.3.6 (rok 2019, pozycja krótka), 4.3.7 (rok 2020, pozycja długa) i 4.3.8 (rok 2020, pozycja krótka).

W pierwszej kolejności zostaną krótko omówione oszacowania wartości zagrożonych (rysunki 4.3.1-4.3.4). Wyniki na poziomie tolerancji 0,1 zostały zamieszczone na rysunkach jedynie w celach informacyjnych. Rok 2019 był rokiem małej zmienności stóp zwrotu, co

znalazło swoje odbicie w oszacowaniach wartości zagrożonych. Różnice pomiędzy maksymalnymi i minimalnymi oszacowaniami wartości zagrożonych, wygenerowanymi ze wszystkich modeli wartości zagrożonej, dla kwantyli rzędu 0,05 i 0,01 wynosiły odpowiednio 0,2555 i 0,6809. Analogiczne różnice dla kwantyli rzędu 0,95 i 0,99 wynosiły natomiast 0,3100 i 0,9362. Rok 2020 był z kolei rokiem pandemii COVID-19 i cechował się dużą zmiennością stóp zwrotu. Różnice pomiędzy maksymalnymi i minimalnymi oszacowaniami wartości zagrożonych, wygenerowanymi ze wszystkich modeli wartości zagrożonej, dla kwantyli rzędu 0,05 i 0,01 wynosiły odpowiednio 0,9349 i 3,6022. Analogiczne różnice dla kwantyli rzędu 0,95 i 0,99 wynosiły natomiast 1,3414 i 4,6329.

Przejdźmy teraz do analizy rozkładów udziałów przekroczeń w całkowitej liczbie obserwacji (rysunki 4.3.5-4.3.8). Udziały te powinny być zbliżone do założonego poziomu tolerancji, stąd można wnioskować, że mediany tych rozkładów powinny być również zbliżone do założonego poziomu tolerancji. Zestawienie wartości median dla zbiorów udziałów przekroczeń, wyznaczonych w oparciu o modele wartości zagrożonej wyselekcjonowane na podstawie różnych testów zgodności dla rozkładów brzegowych, zamieszczono w tabeli 4.3.1. W badaniu pominięto wyniki na poziomie tolerancji 0,1. Zaczniemy od roku 2019. W przypadku kwantyli rzędu 0,05, 0,95 oraz 0,99, wybór testu zgodności nie ma znaczenia. Mediany udziałów przekroczeń wyznaczone na podstawie modeli wyselekcjonowanych w oparciu o test zgodności są takie same, bez względu na wybór statystyki testowej (przeznaczonej do badania całego rozkładu lub ogona rozkładu). W przypadku kwantyla rzędu 0,01, mediany udziałów przekroczeń są najbliższe założonemu poziomowi tolerancji, gdy są wyznaczone na podstawie modeli wartości zagrożonej wyselekcjonowanych w oparciu o statystyki dla ogonów rozkładów, $AD^{**}[0; 0,1]$ lub $AD_{down}^{2**}[0; 0,05]$. Przejdźmy teraz do roku 2020. W przypadku kwantyla rzędu 0,99, wybór testu zgodności nie ma znaczenia. W przypadku kwantyla rzędu 0,05, mediany udziałów przekroczeń są najbliższe założonemu poziomowi tolerancji, gdy są wyznaczone na podstawie modeli wartości zagrożonej wyselekcjonowanych w oparciu o statystyki dla ogonów rozkładów: $KS^{**}[0; 0,1]$, $V^{**}[0; 0,1]$, $W^{2**}[0; 0,1]$, $AD^{2**}[0; 0,1]$, $KS^{**}[0; 0,05]$, $V^{**}[0; 0,05]$, $W^{2**}[0; 0,1]$ oraz $AD^{2**}[0; 0,1]$. W przypadku kwantyla rzędu 0,01, mediana udziałów przekroczeń jest najbliższa założonemu poziomowi tolerancji, gdy jest wyznaczona na podstawie modeli wartości zagrożonej wyselekcjonowanych w oparciu o statystykę dla ogona rozkładu $V^{**}[0; 0,05]$. W przypadku kwantyla rzędu 0,95, mediana udziałów przekroczeń jest najbliższa założonemu poziomowi tolerancji, gdy jest wyznaczona

na podstawie modeli wartości zagrożonej wyselekcjonowanych w oparciu o statystyki dla całych rozkładów.

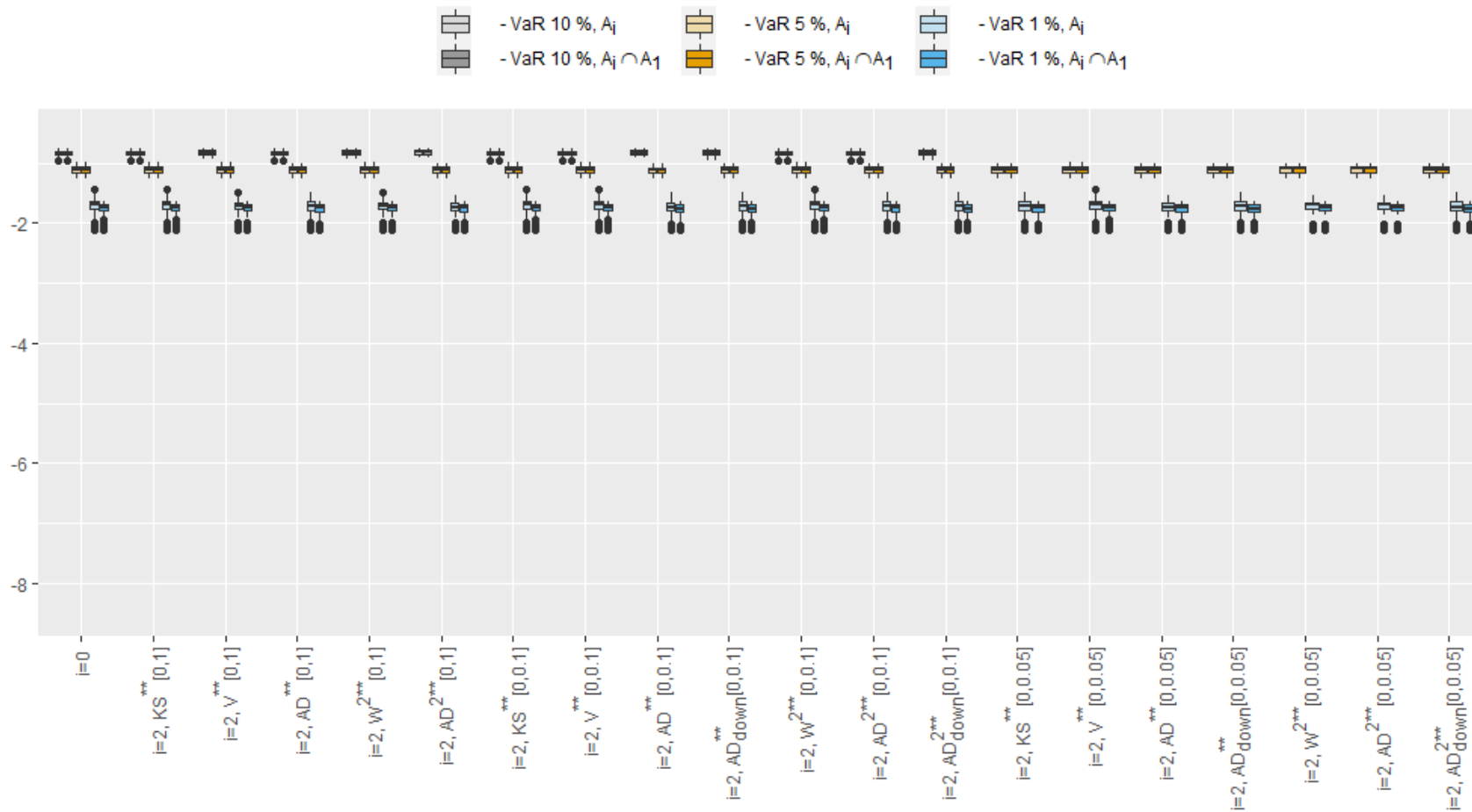
Tabela 4.3.1 Mediany rozkładów udziałów przekroczeń w całkowitej liczbie obserwacji (wyznaczone na podstawie modeli wartości zagrożonych wybranych w oparciu o różne testy zgodności dla rozkładów brzegowych)

Rok	2019				2020			
Rząd kwantyla	0,05	0,01	0,95	0,99	0,05	0,01	0,95	0,99
Wszystkie modele wartości zagrożonej:								
Brak testu zgodności	0,0508	0,0234	0,0430	0,0078	0,0659	0,0116	0,0426	0,0078
Modele wartości zagrożonej wybrane w oparciu o test Kupca:								
Brak testu zgodności	0,0508	0,0195	0,0430	0,0078	0,0659	0,0116	0,0426	0,0078
Modele wartości zagrożonej wybrane na podstawie testów zgodności w oparciu o statystyki testowe dla całych rozkładów:								
$KS^{**}[0; 1]$	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0465	0,0039
$V^{**}[0; 1]$	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0465	0,0039
$AD^{**}[0; 1]$	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0504	0,0078
$W^{2**}[0; 1]$	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0504	0,0078
$AD^{2**}[0; 1]$	0,0508	0,0234	0,0430	0,0059	0,0698	0,0078	0,0504	0,0078
Wartość najbardziej zbliżona do założonego poziomu tolerancji	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0504	0,0078
Modele wartości zagrożonej wybrane na podstawie testów zgodności w oparciu o statystyki testowe dla rozkładów uciętych:								
$KS^{**}[0; 0,1], [0,9; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0465	0,0078
$V^{**}[0; 0,1], [0,9; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0465	0,0078
$AD^{**}[0; 0,1], [0,9; 1]$	0,0508	0,0215	0,0430	0,0078	0,0698	0,0078	0,0426	0,0078
$AD_{down/up}^{**}[0; 0,1], [0,9; 1]$	0,0508	0,0234	0,0430	0,0078	0,0717	0,0078	0,0465	0,0078
$W^{2**}[0; 0,1], [0,9; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0426	0,0078
$AD^{2**}[0; 0,1], [0,9; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0465	0,0078
$AD_{down/up}^{2**}[0; 0,1], [0,9; 1]$	0,0508	0,0234	0,0430	0,0078	0,0717	0,0078	0,0465	0,0078
$KS^{**}[0; 0,05], [0,95; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0426	0,0078
$V^{**}[0; 0,05], [0,95; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0116	0,0426	0,0078
$AD^{**}[0; 0,05], [0,95; 1]$	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0465	0,0078
$AD_{down/up}^{**}[0; 0,05], [0,95; 1]$	0,0508	0,0234	0,0430	0,0078	0,0698	0,0078	0,0426	0,0078
$W^{2**}[0; 0,05], [0,95; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0426	0,0078
$AD^{2**}[0; 0,05], [0,95; 1]$	0,0508	0,0234	0,0430	0,0078	0,0659	0,0078	0,0426	0,0078
$AD_{down/up}^{2**}[0; 0,05], [0,95; 1]$	0,0508	0,0215	0,0430	0,0078	0,0698	0,0078	0,0426	0,0078
Wartość najbardziej zbliżona do założonego poziomu tolerancji	0,0508	0,0215	0,0430	0,0078	0,0659	0,0116	0,0465	0,0078

Źródło: obliczenia własne

W przypadku szacowania wartości zagrożonej metodą Monte Carlo na podstawie rozkładów dwuwymiarowych, tworzonych z kopul oraz rozkładów brzegowych metodą w pełni

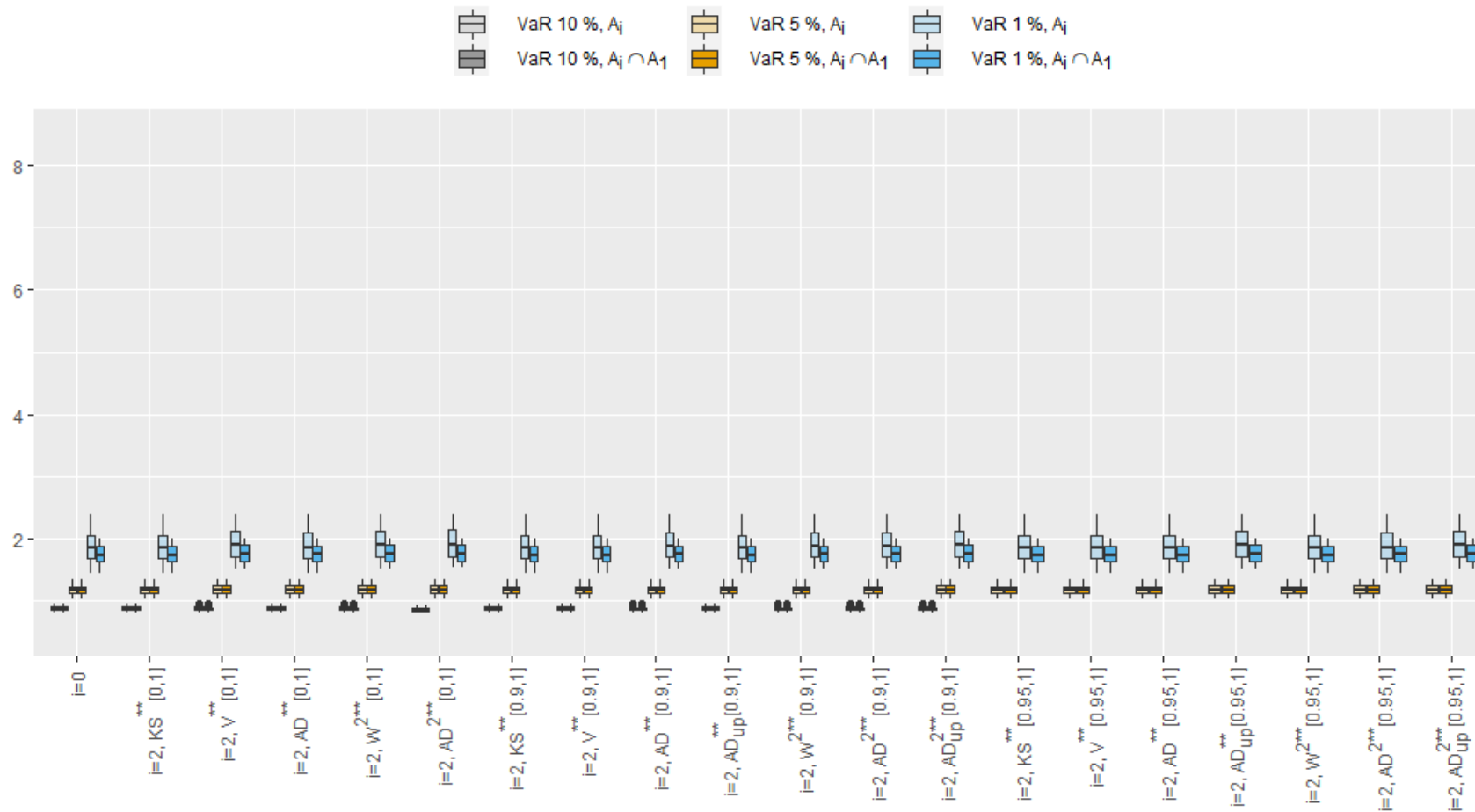
parametryczną, wskazane jest zastosowanie testów zgodności dla rozkładów brzegowych. Wybór statystyki testowej może istotnie wpłynąć na ostateczny zbiór dopuszczalnych oszacowań wartości zagrożonych oraz na rozkład udziałów ich przekroczeń. Podsumowując ten rozdział można stwierdzić, że w każdym okresie, dla kwantyli rzędu 0,01, 0,05 oraz 0,99 można wybrać statystykę testową dla rozkładu uciętego, która prowadzi do wyboru modeli wartości zagrożonej niemniej dokładnych, niż modele wybrane przy użyciu statystyk testowych dla całego rozkładu, przy czym dokładność jest tu mierzona medianą uzyskanych w oparciu o wyselekcjonowane modele udziałów przekroczeń. Stwierdzenie to nie jest prawdziwe w odniesieniu do górnego ogona rozkładu, do kwantyla rzędu 0,95.



Rysunek 4.3.1 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja długa, rok 2019) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_0 – zbiór oszacowań wartości zagrożonych wygenerowanych ze wszystkich rozkładów dwuwymiarowych; A_1 - zbiór oszacowań wartości zagrożonych wygenerowanych z modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_2 - zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

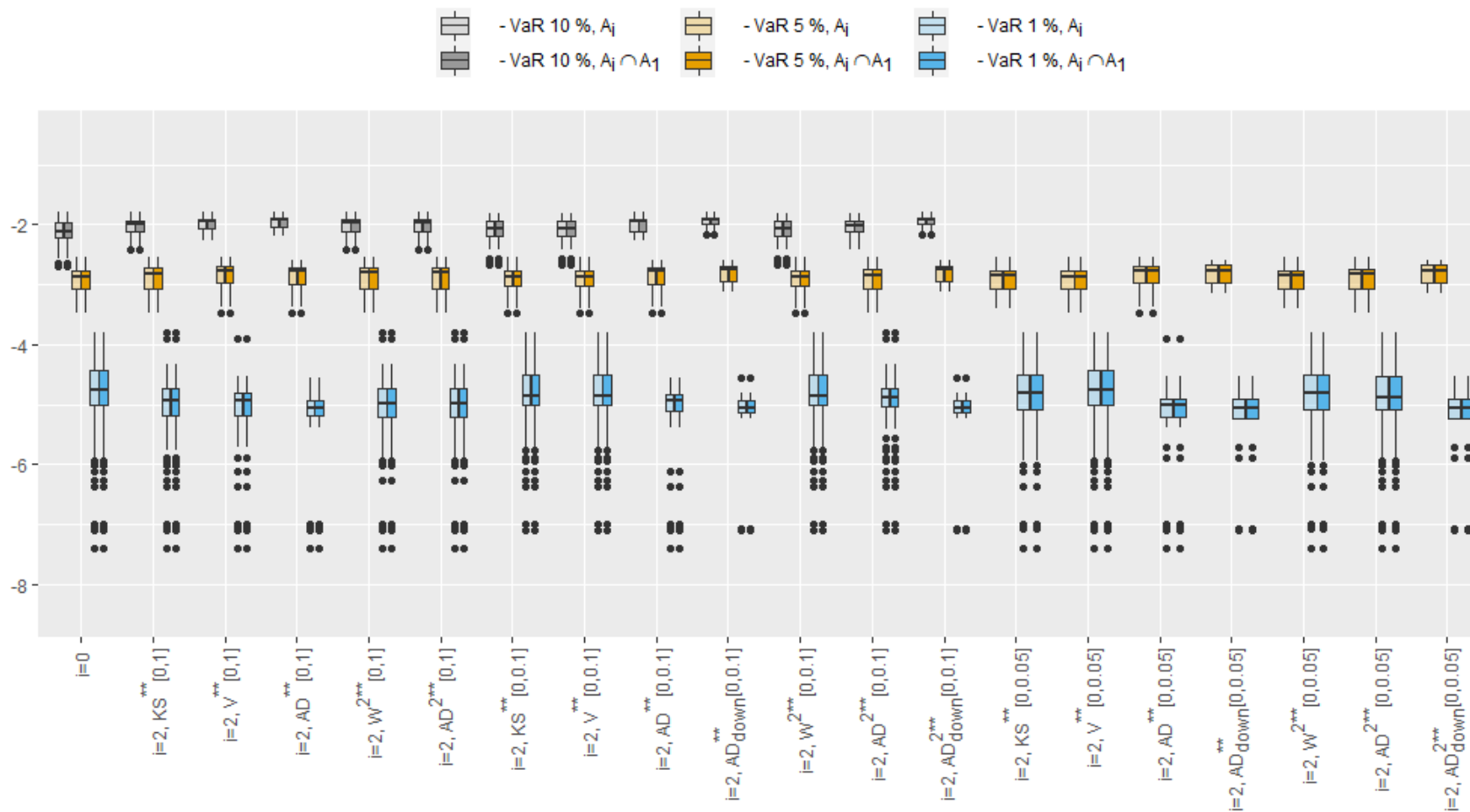
Źródło: obliczenia własne



Rysunek 4.3.2 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja krótka, rok 2019) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_0 – zbiór oszacowań wartości zagrożonych wygenerowanych ze wszystkich rozkładów dwuwymiarowych; A_1 - zbiór oszacowań wartości zagrożonych wygenerowanych z modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_2 - zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

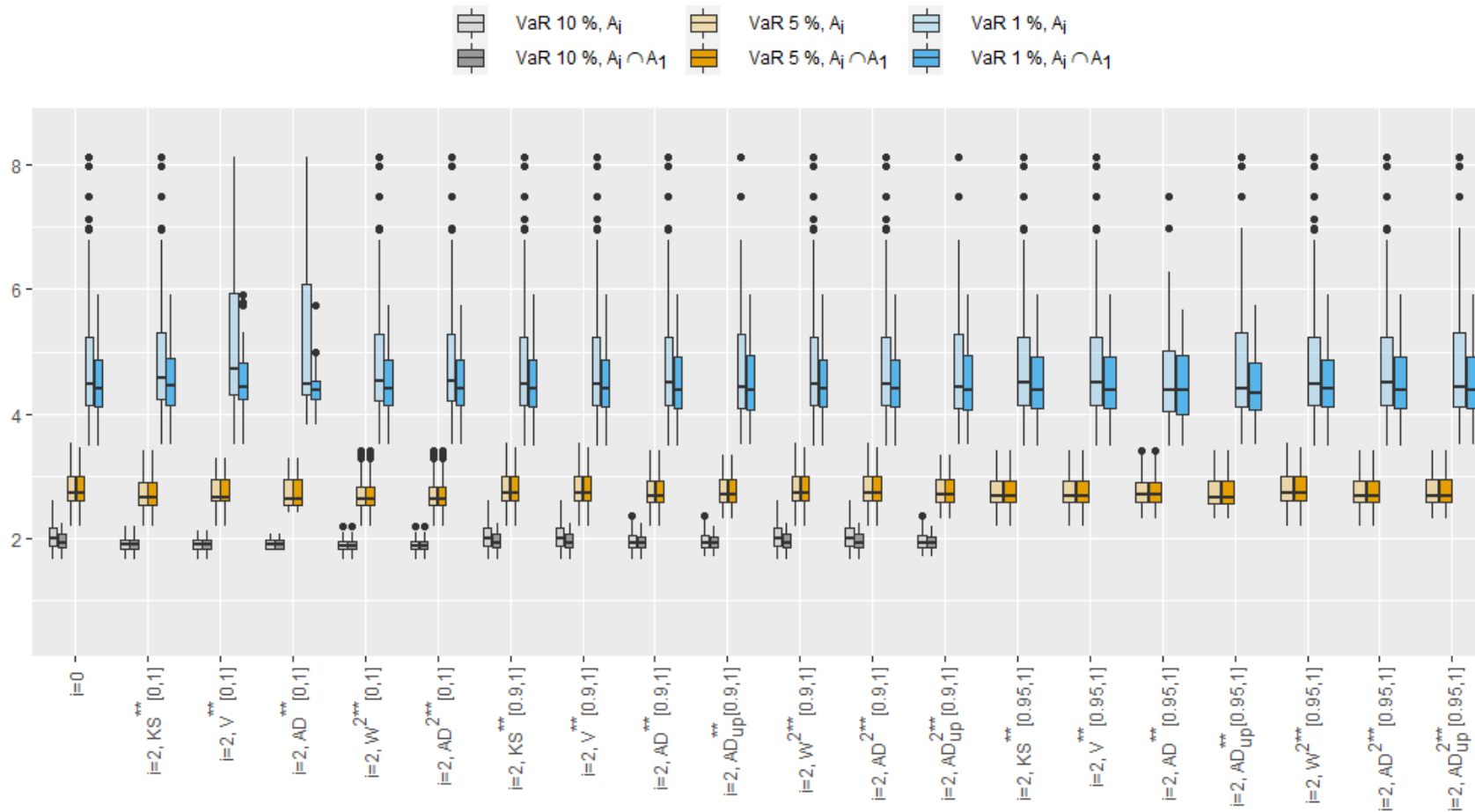
Źródło: obliczenia własne



Rysunek 4.3.3 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja długa, rok 2020) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_0 – zbiór oszacowań wartości zagrożonych wygenerowanych ze wszystkich rozkładów dwuwymiarowych; A_1 - zbiór oszacowań wartości zagrożonych wygenerowanych z modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_2 - zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

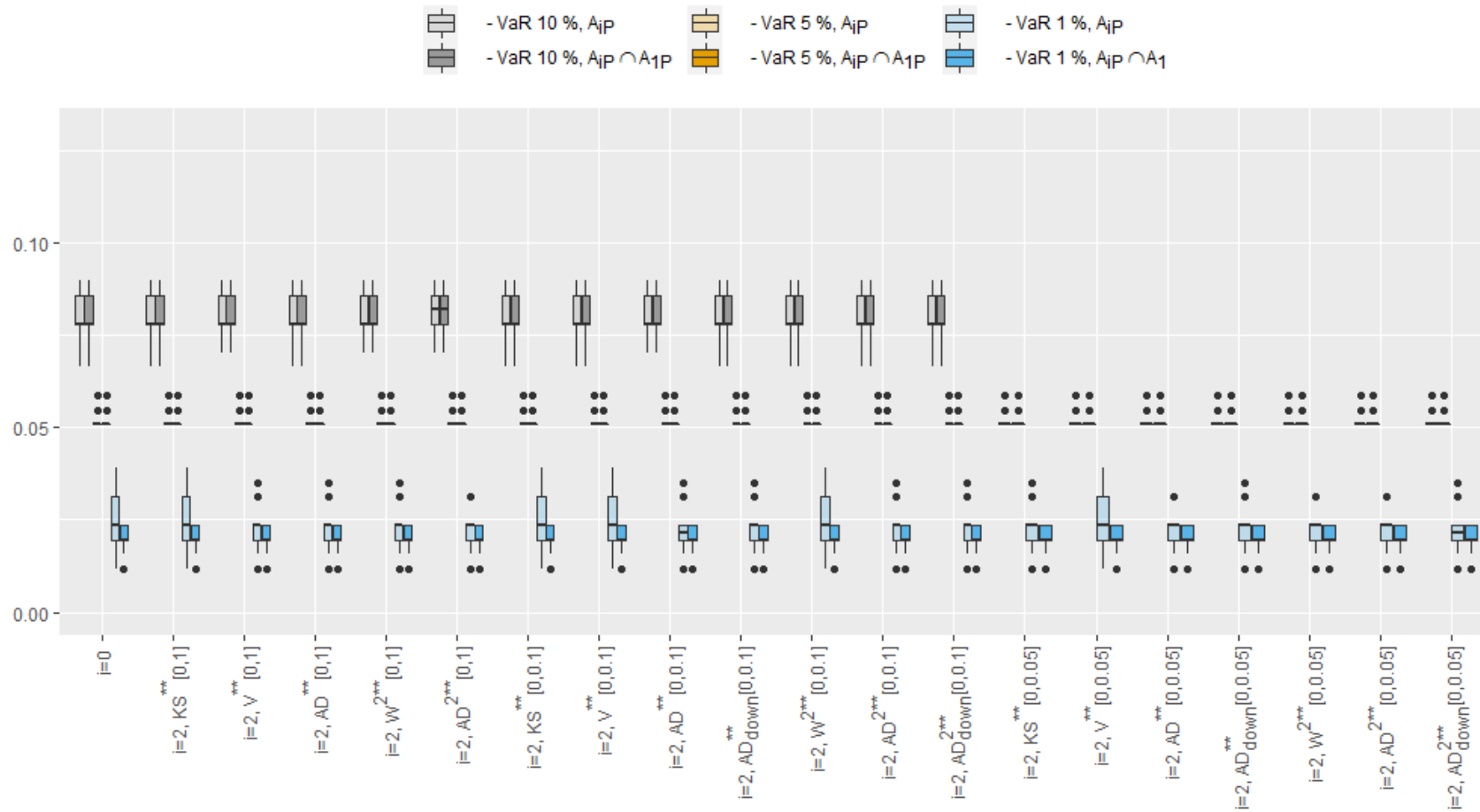
Źródło: obliczenia własne



Rysunek 4.3.4 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja krótka, rok 2020) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_0 – zbiór oszacowań wartości zagrożonych wygenerowanych ze wszystkich rozkładów dwuwymiarowych; A_1 - zbiór oszacowań wartości zagrożonych wygenerowanych z modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_2 - zbiór oszacowań wartości zagrożonych wyznaczonych na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

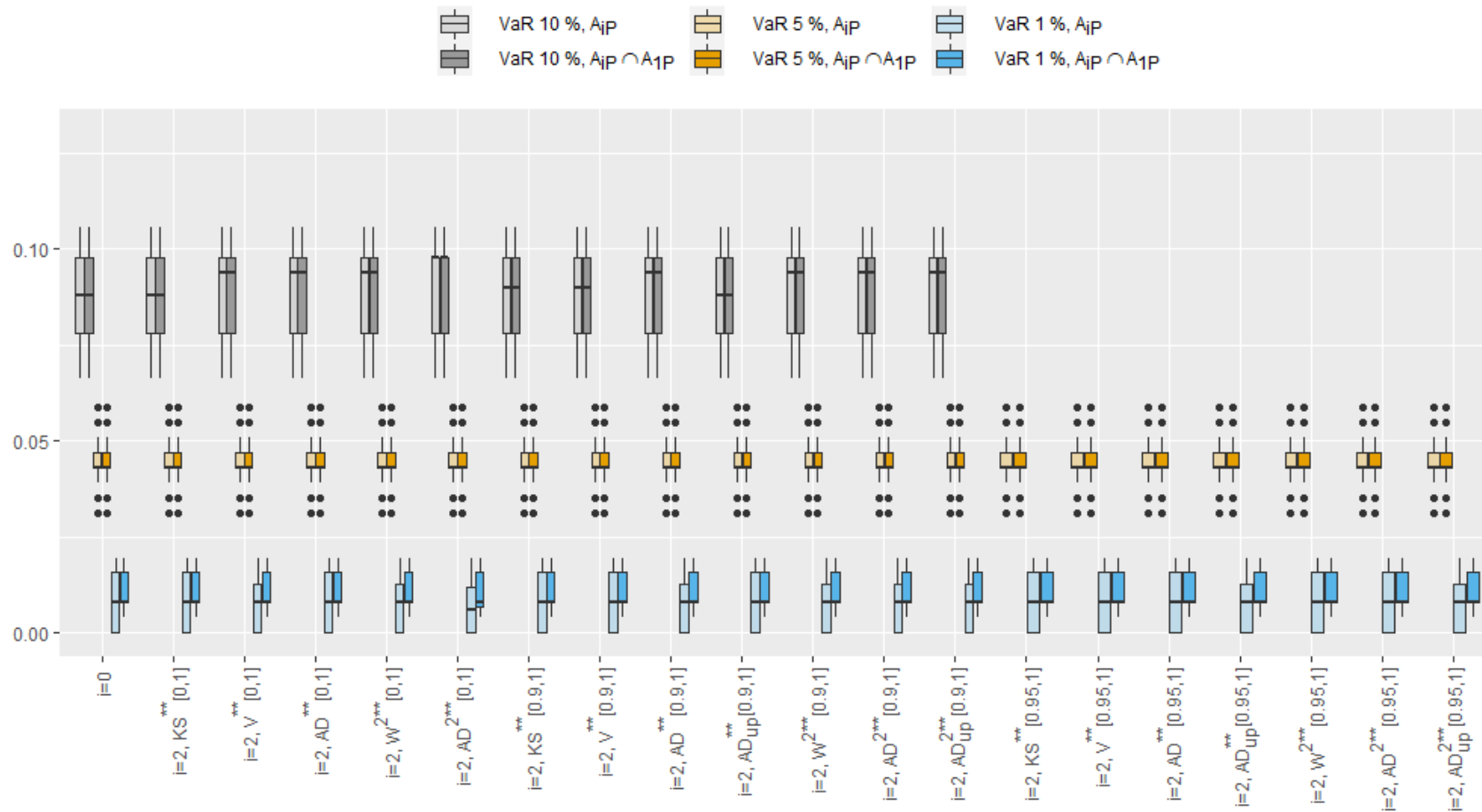
Źródło: obliczenia własne



Rysunek 4.3.5 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja długa, rok 2019) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_{0P} – udziały przekroczeń wyznaczone dla wszystkich rozkładów dwuwymiarowych; A_{1P} – udziały przekroczeń wyznaczone na podstawie tych modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_{2P} – udziały przekroczeń wyznaczone na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

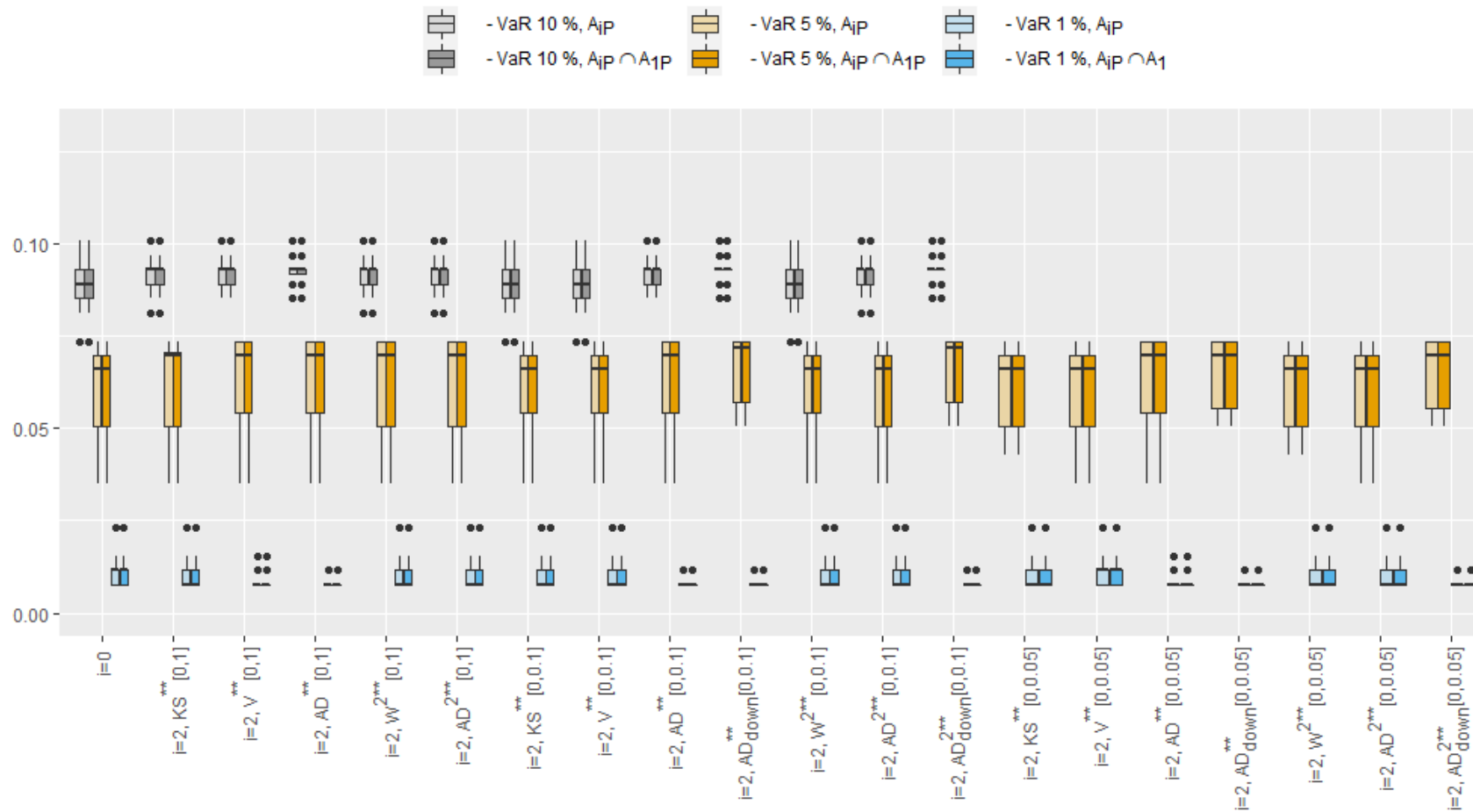
Źródło: obliczenia własne



Rysunek 4.3.6 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja krótka, rok 2019) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_{0P} – udziały przekroczeń wyznaczone dla wszystkich rozkładów dwuwymiarowych; A_{1P} – udziały przekroczeń wyznaczone na podstawie tych modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_{2P} – udziały przekroczeń wyznaczone na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

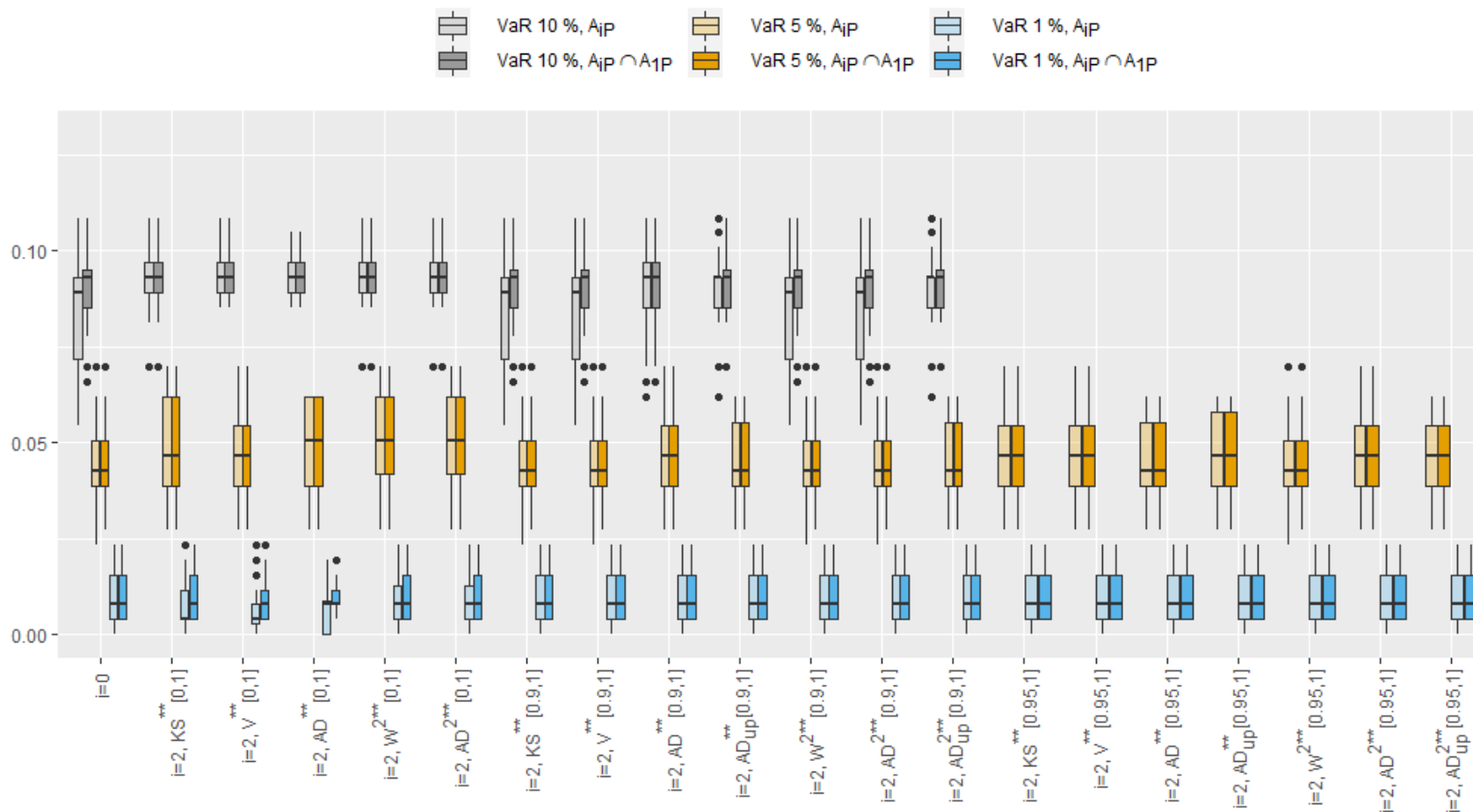
Źródło: obliczenia własne



Rysunek 4.3.7 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja długa, rok 2020) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_{0P} – udziały przekroczeń wyznaczone dla wszystkich rozkładów dwuwymiarowych; A_{1P} – udziały przekroczeń wyznaczone na podstawie tych modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_{2P} – udziały przekroczeń wyznaczone na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

Źródło: obliczenia własne



Rysunek 4.3.8 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja krótka, rok 2020) ⁽¹⁾

(1) Przyjęto następujące oznaczenia dla zbiorów: A_{0p} – udziały przekroczeń wyznaczone dla wszystkich rozkładów dwuwymiarowych; A_{1p} – udziały przekroczeń wyznaczone na podstawie tych modeli wartości zagrożonej, które na podstawie testu Kupca uznano za dokładne; A_{2p} – udziały przekroczeń wyznaczone na podstawie rozkładów dwuwymiarowych, dla których, w przypadku obydwu rozkładów brzegowych, nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym.

Źródło: obliczenia własne

4.4. Wpływ wyboru testu zgodności na oszacowania wartości zagrożonej na przykładzie rozkładów jednowymiarowych – wyniki spoza próby

W poprzednich podrozdziałach przekroczenia wartości zagrożonej były badane w próbie, która była podstawą do szacowania wartości zagrożonej. W tym podrozdziale wartości zagrożone będą weryfikowane poza próbą. Zweryfikowana zostanie zdolność omawianych dotąd rozkładów bezwarunkowych do generowania jednodniowych prognoz wartości zagrożonych. Klasyczne podejście polegające na wyznaczeniu kwantyla rozkładu zostanie zmodyfikowane poprzez wykorzystanie testów zgodności. Rozważone zostaną trzy, opisane poniżej, modele. W pierwszym przypadku nie zostaną zastosowane żadne testy zgodności, w kolejnych dwóch zostaną zastosowane testy zgodności oparte na wybranych statystykach przedstawionych w niniejszej pracy. W każdym przypadku o jakości zastosowanego podejścia decydować będzie test Kupca. Ze względu na fakt, iż badanie to koncentruje się na rozkładach bezwarunkowych, nie zostaną użyte inne metody testowania wstecznego koncentrujące się na niezależności przekroczeń VaR.

Zgodnie z wytycznymi Komitetu Bazylejskiego, wartość zagrożona może być wyznaczana poprzez zastosowanie podejścia standardowego lub poprzez wykorzystanie własnych modeli. Wartość zagrożona na poziomie tolerancji 0,01 powinna być wyznaczana na podstawie danych dziennych, a okres jej utrzymywania powinien wynosić 10 dni. Zgodnie z zaleceniami, jeśli wartość zagrożona została obliczona dla krótszego okresu, powinna zostać przeskalowana do okresu 10-dniowego za pomocą metody „*square root of time*”. Taką skalowalność zapewniają rozkłady eliptyczne, w tym normalny, jednak często nie są one odpowiednie do opisu rozkładów stóp zwrotu (Echaust i Just, 2020a). W tej sytuacji banki mogą chcieć zastosować własne modele.

W modelu I wartości zagrożone są wyznaczane dla każdego okresu na podstawie pięciu rozkładów bezwarunkowych: N , NIG , Hyp , $SGED$ lub ST . Wartości zagrożone są odpowiednimi kwantylami tych rozkładów bezwarunkowych. Kolejne dwa modele to autorskie modyfikacje standardowej metody obliczania wartości zagrożonej.

W modelu II założono, że w każdym okresie jest wybierany tylko jeden z rozkładów: N , NIG , Hyp , $SGED$ lub ST . Rozważono tutaj dwie wersje modelu. W pierwszej wersji modelu rozkład jest wybierany na podstawie testów zgodności opartych na statystykach testowych dla ogonów rozkładu, opisanych w rozdziale trzecim. W drugiej wersji modelu rozkład jest

wybijerany na podstawie testów zgodności dla całych rozkładów. Dla każdej wybranej statystyki sprawdzane są wartości p testów zgodności rozkładu empirycznego z rozkładami teoretycznymi (N , NIG , Hyp , $SGED$, ST) i dla danego okresu jest wybierany rozkład z maksymalną wartością p . W każdym oknie szacowania VaR może być zatem inny rozkład, ale zawsze taki, który jest dobrze dopasowany w danym oknie obserwacji.

W modelu III podstawowym rozkładem jest rozkład normalny. W każdym okresie wykonywane są testy zgodności rozkładu empirycznego z rozkładem normalnym, oparte na wybranych statystykach. Jeśli wartość p wybranego testu zgodności znajduje się poniżej założonego poziomu istotności (przyjęto 0,05), rozkład normalny jest zastępowany rozkładem NIG . Jeśli jednak wartość p wybranego testu zgodności znajduje się poniżej założonego poziomu istotności również wtedy, gdy sprawdzana jest zgodność rozkładu empirycznego z rozkładem NIG , wtedy zostawiany jest rozkład normalny. Proces ten będzie powtarzany dla każdego okresu i dla każdej wybranej statystyki osobno. Tutaj również rozważono dwie wersje modelu. W pierwszej wersji modelu testy zgodności są oparte na statystykach dla ogona górnego lub dolnego. W drugiej wersji modelu testy zgodności są oparte na statystykach dla całego rozkładu.

Badanie empiryczne, dla wszystkich modeli, składa się z następujących etapów:

1. Określenie źródła danych.

Danymi są dzienne procentowe logarytmiczne stopy zwrotu, wyliczone na podstawie kursów zamknięcia indeksów $WIG20$ oraz $FTSE100$ z okresu obejmującego lata 2017-2022, powiększonego o 250 notowań poprzedzających 01.01.2017. Łącznie liczba danych początkowych dla indeksu $WIG20$ wyniosła 1749, a dla indeksu $FTSE100$ 1765.

2. Oznaczenie przez T długości badanego okresu oraz przez t numeru kolejnego okresu. Niech $t = 1$.
3. Określenie parametrów wybranego modelu na podstawie okresu $[t, t + 249]$.
4. Wyznaczenie wartości zagrożonej na wybranym poziomie tolerancji dla ogona górnego i dolnego na podstawie modelu z punktu 3. Otrzymane liczby są wartościami zagrożonymi dla okresu $t + 250$ dla pozycji krótkiej oraz długiej.
5. Przesunięcie badanego okresu do przodu o jeden dzień, $t = t + 1$.
6. Powtarzanie etapów 3, 4 i 5, dopóki $t < T$.
7. Ocena testem Kupca oszacowań wartości zagrożonych wyznaczonych w punkcie 4.

Wyniki badania empirycznego dla całego badanego okresu umieszczono w tabelach 4.4.1-4.4.4. Wyniki obejmują: oczekiwaną liczbę przekroczeń (równą iloczynowi poziomu tolerancji oraz liczby obserwacji w badanym okresie), rzeczywistą liczbę przekroczeń, wartość statystyki testowej testu Kupca oraz wartość p testu Kupca. Gdy wartość p testu Kupca znajdowała się poniżej założonego poziomu istotności wynoszącego 0,05, trzy ostatnie wymienione wartości umieszczano na szarym tle. Na wykresach 4.4.1-4.4.2 przedstawiono przekroczenia VaR w badanym okresie.

Tabela 4.4.1. Wyniki testu Kupca dla modelu I (poziom tolerancji 0,01)

Indeks	Rozkład	Oczekiwana liczba przekroczeń	Rząd kwantyla (poziom tolerancji dla VaR)					
			0,01			0,99		
			Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca	Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca
WIG20	<i>N</i>	15	27	7,85	0,0051	22	2,89	0,0889
	<i>NIG</i>	15	24	4,63	0,0315	16	0,07	0,7954
	<i>Hyp</i>	15	26	6,70	0,0096	17	0,26	0,6095
	<i>SGED</i>	15	25	5,62	0,0177	17	0,26	0,6095
	<i>ST</i>	15	24	4,63	0,0315	16	0,07	0,7954
FTSE100	<i>N</i>	15	38	24,54	0,0000	21	2,04	0,1535
	<i>NIG</i>	15	24	4,43	0,0352	21	2,04	0,1535
	<i>Hyp</i>	15	28	8,81	0,0030	22	2,75	0,0975
	<i>SGED</i>	15	27	7,60	0,0058	19	0,91	0,3389
	<i>ST</i>	15	26	6,46	0,0110	20	1,42	0,2326

Źródło: obliczenia własne

Tabela 4.4.2. Wyniki testu Kupca dla modelu I (poziom tolerancji 0,05)

Indeks	Rozkład	Oczekiwana liczba przekroczeń	Rząd kwantyla (poziom tolerancji dla VaR)					
			0,05			0,95		
			Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca	Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca
WIG20	<i>N</i>	75	71	0,22	0,6369	88	2,27	0,1318
	<i>NIG</i>	75	77	0,06	0,8089	92	3,82	0,0507
	<i>Hyp</i>	75	77	0,06	0,8089	93	4,26	0,0389
	<i>SGED</i>	75	76	0,02	0,9012	91	3,40	0,0653
	<i>ST</i>	75	82	0,68	0,4102	94	4,73	0,0296
FTSE100	<i>N</i>	76	94	4,31	0,0378	56	5,94	0,0148
	<i>NIG</i>	76	91	3,04	0,0810	84	0,91	0,3388
	<i>Hyp</i>	76	90	2,67	0,1024	81	0,37	0,5403
	<i>SGED</i>	76	91	3,04	0,0810	76	0,00	0,9765
	<i>ST</i>	76	97	5,79	0,0162	84	0,91	0,3388

Źródło: obliczenia własne

Wyniki testu Kupca dla modelu I, na poziomach tolerancji 0,01 i 0,05, znajdują się odpowiednio w tabelach 4.4.1 i 4.4.2. Na poziomie tolerancji 0,01, w przypadku ogona dolnego, ryzyko zostało niedoszacowane we wszystkich analizowanych wersjach modelu. Z kolei na poziomie tolerancji 0,01, w przypadku ogona górnego, ryzyko zostało prawidłowo oszacowane we wszystkich analizowanych wersjach modelu. Modelami generującymi poprawne, na podstawie testu Kupca, oszacowania wartości zagrożonej na poziomie tolerancji 0,05, są rozkłady N , NIG oraz $SGED$ w przypadku indeksu $WIG20$ oraz rozkłady NIG , Hyp oraz $SGED$ w przypadku indeksu $FTSE100$. W dalszej części podrozdziału uwaga zostanie skupiona na wyznaczeniu wartości zagrożonej na poziomie tolerancji 0,01.

Wyniki testu Kupca dla modelu II znajdują się w tabeli 4.4.3, a dla modelu III w tabeli 4.4.4. Dla obydwu modeli uzyskano podobne wyniki. Prawidłowe wyniki dla ogona dolnego uzyskano jedynie wtedy, gdy rozkład był wybierany na podstawie testów zgodności opartych na statystyce AD_{down}^{**} , ale tylko w przypadku indeksu $WIG20$. W przypadku ogona górnego prawidłowe wyniki otrzymano dla wszystkich analizowanych wersji modelu II i modelu III. Zastosowanie modelu II lub modelu III pozwala zatem na prawidłowe oszacowanie wartości zagrożonej dla ogona dolnego w przypadku indeksu $WIG20$, co nie było możliwe w wyniku zastosowania modelu I.

Na wykresach 4.4.1 i 4.4.2 przedstawiono oszacowania wartości zagrożonych na poziomie tolerancji 0,01 dla indeksów $WIG20$ oraz $FTSE100$, zestawione z ich stopami zwrotu dla wybranych przykładów. Przedstawienie wszystkich wyników znacznie zmniejszyłoby czytelność wykresów. W przypadku modelu I przedstawiono wyniki tylko dla dwóch rozkładów bezwarunkowych, N oraz NIG . Rozkłady te są elementami pozostałych dwóch modeli. W przypadku modelu II oraz modelu III przedstawiono tylko te wersje, w których zastosowano testy zgodności oparte o statystykę testową AD_{down}^{**} dla ogona dolnego oraz o statystykę AD_{up}^{**} dla ogona górnego, natomiast punkty ucięcia dystrybuanty wynosiły 0,05 dla ogona dolnego oraz 0,95 dla ogona górnego. Modele te określono jako poprawne, ale tylko w odniesieniu do indeksu $WIG20$.

Na obydwu wykresach można zauważyć kilka prawidłowości. W modelu I, wartości zagrożone prognozowane na podstawie rozkładu normalnego są zazwyczaj mniejsze w przypadku ogona górnego i większe w przypadku ogona dolnego niż te, które są prognozowane na podstawie rozkładu NIG . W modelu II występuje bardzo duża zmienność, tutaj bowiem w każdym okresie wartość zagrożona może być prognozowana na podstawie

innego rozkładu bezwarunkowego. W modelu III zmienność jest mniejsza niż w modelu II, ale większa niż w modelu I.

Tabela 4.4.3. Wyniki testu Kupca dla modelu II (poziom tolerancji 0,01)

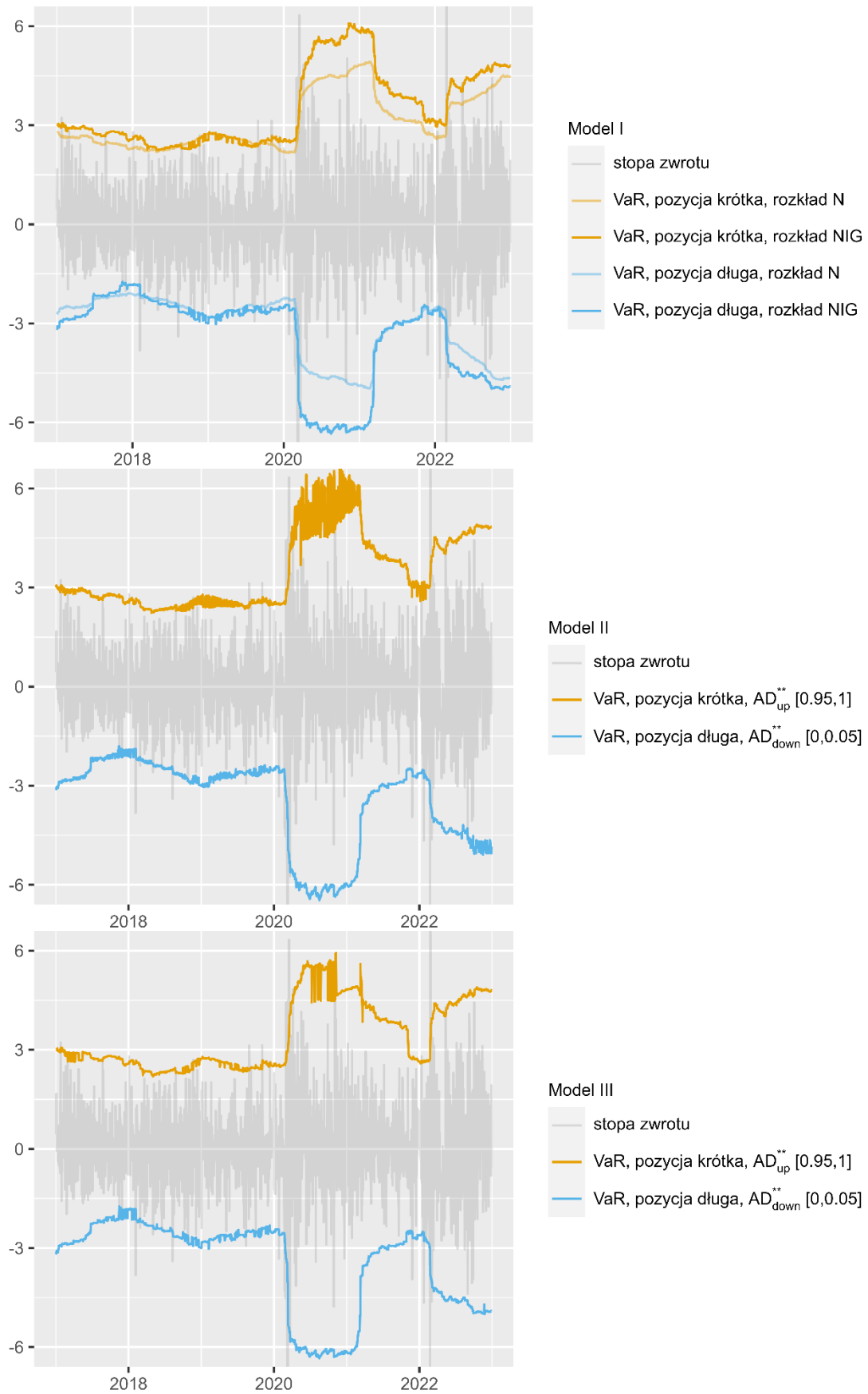
Indeks	Poziom ucięcia rozkładu	Statystyka testu zgodności	Oczekiwana liczba przekroczeń	Rząd kwantyla (poziom tolerancji dla VaR)						
				0,01			0,99			
				Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca	Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca	
<i>WIG20</i>	0,05 (ogon dolny) / 0,95 (ogon górny)	<i>KS**</i>	15	24	4,63	0,0315	21	2,16	0,1413	
		<i>V**</i>	15	26	6,70	0,0096	21	2,16	0,1413	
		<i>AD**</i>	15	24	4,63	0,0315	18	0,57	0,4488	
		<i>AD_{down/up}**</i>	15	22	2,89	0,0889	17	0,26	0,6095	
		<i>W^{2**}</i>	15	24	4,63	0,0315	21	2,16	0,1413	
		<i>AD^{2**}</i>	15	24	4,63	0,0315	21	2,16	0,1413	
		<i>AD_{down/up}^{2**}</i>	15	25	5,62	0,0177	18	0,57	0,4488	
	0,1 (ogon dolny) / 0,9 (ogon górny)	<i>KS**</i>	15	24	4,63	0,0315	20	1,53	0,2160	
		<i>V**</i>	15	24	4,63	0,0315	21	2,16	0,1413	
		<i>AD**</i>	15	24	4,63	0,0315	18	0,57	0,4488	
		<i>AD_{down/up}**</i>	15	23	3,72	0,0539	16	0,07	0,7954	
		<i>W^{2**}</i>	15	24	4,63	0,0315	20	1,53	0,2160	
		<i>AD^{2**}</i>	15	24	4,63	0,0315	19	1,00	0,3176	
		<i>AD_{down/up}^{2**}</i>	15	24	4,63	0,0315	18	0,57	0,4488	
	brak ucięcia	<i>KS**</i>	15	25	5,62	0,0177	16	0,07	0,7954	
		<i>V**</i>	15	25	5,62	0,0177	16	0,07	0,7954	
		<i>AD**</i>	15	24	4,63	0,0315	16	0,07	0,7954	
		<i>W^{2**}</i>	15	25	5,62	0,0177	16	0,07	0,7954	
		<i>AD^{2**}</i>	15	25	5,62	0,0177	16	0,07	0,7954	
	<i>FTSE100</i>	0,05 (ogon dolny) / 0,95 (ogon górny)	<i>KS**</i>	15	30	11,44	0,0007	19	0,91	0,3389
			<i>V**</i>	15	28	8,81	0,0030	21	2,04	0,1535
			<i>AD**</i>	15	26	6,46	0,0110	19	0,91	0,3389
			<i>AD_{down/up}**</i>	15	25	5,41	0,0200	20	1,42	0,2326
			<i>W^{2**}</i>	15	30	11,44	0,0007	20	1,42	0,2326
<i>AD^{2**}</i>			15	28	8,81	0,0030	20	1,42	0,2326	
<i>AD_{down/up}^{2**}</i>			15	26	6,46	0,0110	20	1,42	0,2326	
0,1 (ogon dolny) / 0,9 (ogon górny)		<i>KS**</i>	15	27	7,60	0,0058	20	1,42	0,2326	
		<i>V**</i>	15	31	12,86	0,0003	19	0,91	0,3389	
		<i>AD**</i>	15	26	6,46	0,0110	19	0,91	0,3389	
		<i>AD_{down/up}**</i>	15	26	6,46	0,0110	20	1,42	0,2326	
		<i>W^{2**}</i>	15	28	8,81	0,0030	20	1,42	0,2326	
		<i>AD^{2**}</i>	15	26	6,46	0,0110	19	0,91	0,3389	
		<i>AD_{down/up}^{2**}</i>	15	26	6,46	0,0110	18	0,51	0,4748	
brak ucięcia		<i>KS**</i>	15	26	6,46	0,0110	20	1,42	0,2326	
		<i>V**</i>	15	26	6,46	0,0110	21	2,04	0,1535	
		<i>AD**</i>	15	26	6,46	0,0110	20	1,42	0,2326	
		<i>W^{2**}</i>	15	26	6,46	0,0110	20	1,42	0,2326	
		<i>AD^{2**}</i>	15	26	6,46	0,0110	21	2,04	0,1535	

Źródło: obliczenia własne

Tabela 4.4.4. Wyniki testu Kupca dla modelu III (poziom tolerancji 0,01)

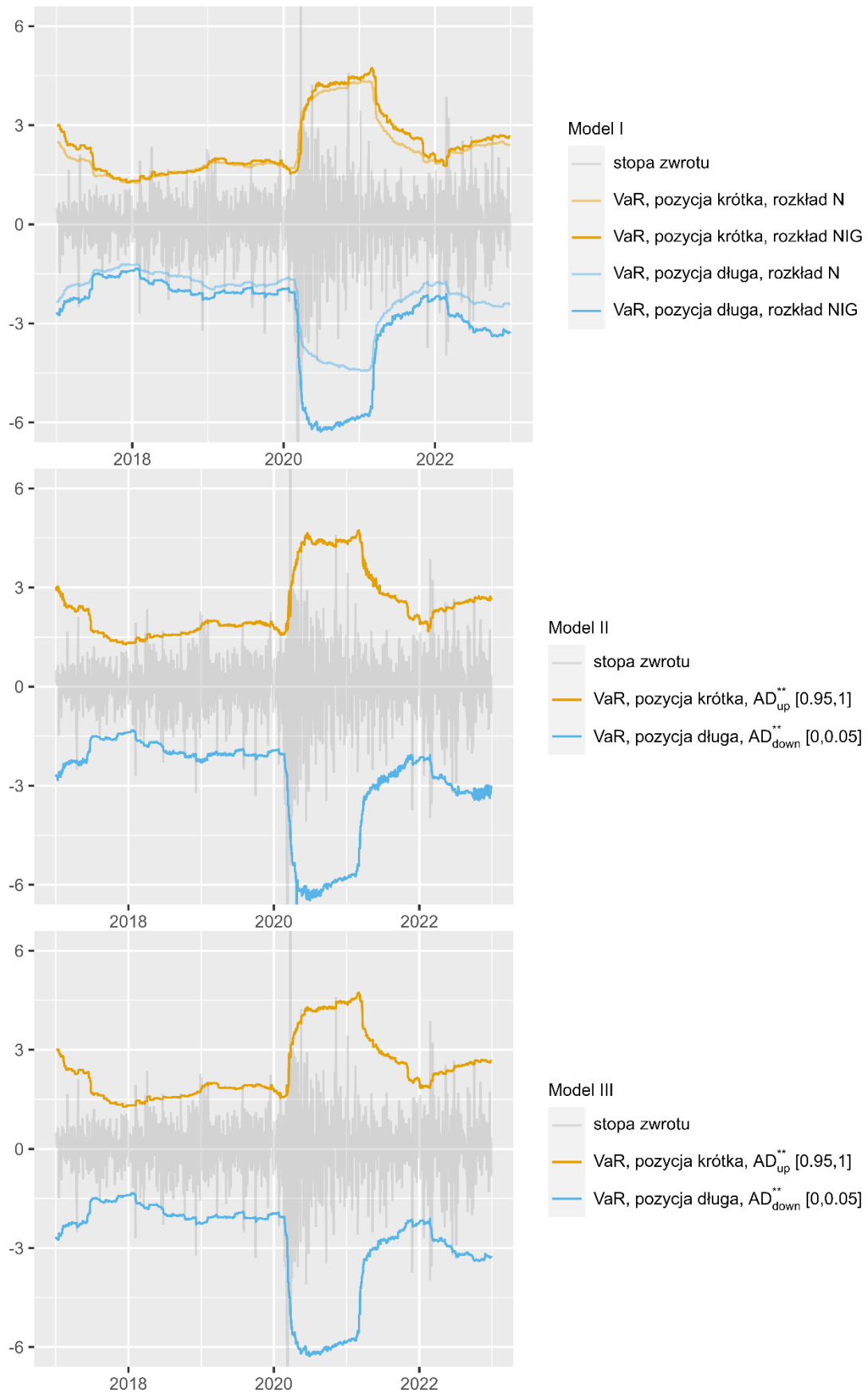
Indeks	Poziom ucięcia rozkładu	Statystyka testu zgodności	Oczekiwana liczba przekroczeń	Rząd kwantyla (poziom tolerancji dla VaR)						
				0,01			0,99			
				Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca	Rzeczywista liczba przekroczeń	Statystyka testowa Kupca	Wartość p testu Kupca	
WIG20	0,05 (ogon dolny) / 0,95 (ogon górny)	KS**	15	25	5,62	0,0177	21	2,16	0,1413	
		V**	15	26	6,70	0,0096	21	2,16	0,1413	
		AD**	15	24	4,63	0,0315	17	0,26	0,6095	
		AD** _{down/up}	15	22	2,89	0,0889	16	0,07	0,7954	
		W ^{2**}	15	24	4,63	0,0315	21	2,16	0,1413	
		AD ^{2**}	15	24	4,63	0,0315	21	2,16	0,1413	
		AD ^{2**} _{down/up}	15	24	4,63	0,0315	19	1,00	0,3176	
	0,1 (ogon dolny) / 0,9 (ogon górny)	KS**	15	25	5,62	0,0177	19	1,00	0,3176	
		V**	15	24	4,63	0,0315	20	1,53	0,2160	
		AD**	15	26	6,70	0,0096	17	0,26	0,6095	
		AD** _{down/up}	15	22	2,89	0,0889	16	0,07	0,7954	
		W ^{2**}	15	26	6,70	0,0096	19	1,00	0,3176	
		AD ^{2**}	15	26	6,70	0,0096	19	1,00	0,3176	
		AD ^{2**} _{down/up}	15	24	4,63	0,0315	17	0,26	0,6095	
	brak ucięcia	KS**	15	24	4,63	0,0315	16	0,07	0,7954	
		V**	15	24	4,63	0,0315	16	0,07	0,7954	
		AD**	15	24	4,63	0,0315	16	0,07	0,7954	
		W ^{2**}	15	26	6,70	0,0096	16	0,07	0,7954	
		AD ^{2**}	15	25	5,62	0,0177	16	0,07	0,7954	
	FTSE100	0,05 (ogon dolny) / 0,95 (ogon górny)	KS**	15	31	12,86	0,0003	19	0,91	0,3389
			V**	15	30	11,44	0,0007	21	2,04	0,1535
			AD**	15	24	4,43	0,0352	19	0,91	0,3389
			AD** _{down/up}	15	24	4,43	0,0352	20	1,42	0,2326
			W ^{2**}	15	30	11,44	0,0007	20	1,42	0,2326
AD ^{2**}			15	28	8,81	0,0030	20	1,42	0,2326	
AD ^{2**} _{down/up}			15	24	4,43	0,0352	20	1,42	0,2326	
0,1 (ogon dolny) / 0,9 (ogon górny)		KS**	15	26	6,46	0,0110	20	1,42	0,2326	
		V**	15	30	11,44	0,0007	19	0,91	0,3389	
		AD**	15	24	4,43	0,0352	19	0,91	0,3389	
		AD** _{down/up}	15	24	4,43	0,0352	20	1,42	0,2326	
		W ^{2**}	15	27	7,60	0,0058	20	1,42	0,2326	
		AD ^{2**}	15	26	6,46	0,0110	20	1,42	0,2326	
		AD ^{2**} _{down/up}	15	24	4,43	0,0352	19	0,91	0,3389	
brak ucięcia		KS**	15	24	4,43	0,0352	21	2,04	0,1535	
		V**	15	24	4,43	0,0352	21	2,04	0,1535	
		AD**	15	24	4,43	0,0352	21	2,04	0,1535	
		W ^{2**}	15	24	4,43	0,0352	21	2,04	0,1535	
		AD ^{2**}	15	24	4,43	0,0352	21	2,04	0,1535	

Źródło: obliczenia własne



Rysunek 4.4.1 Oszacowania wartości zagrożonej na poziomie tolerancji 0,01 (WIG20)

Źródło: obliczenia własne



Rysunek 4.4.2 Oszacowania wartości zagrożonej na poziomie tolerancji 0,01 (FTSE100)

Źródło: obliczenia własne

Dalszej analizie zostaną poddane modele II i III, ponieważ tylko w tych modelach stosowano testy zgodności. Modele będą analizowane w podziale na grupy, które wyznaczono w zależności od tego, czy model jest oparty na rozkładzie uciętym, czy na całym rozkładzie, co z kolei determinuje wybór odpowiednich statystyk w testach zgodności. W pierwszej grupie znajdują się modele, w których zastosowano testy zgodności oparte na statystykach przeznaczonych do badania rozkładów uciętych, a dokładniej w tym przypadku do badania ogonów rozkładów. W drugiej grupie znajdują się modele, w których zastosowano testy zgodności oparte na statystykach przeznaczonych do badania całych rozkładów. W oparciu o tabele 4.4.3-4.4.4, w ramach każdej grupy statystyk, wyznaczono te wersje modeli wartości zagrożonej, które uzyskały maksymalne wartości p w przebiegu testu Kupca (tabela 4.4.5). Analizując tabelę można stwierdzić, że wartości p wyznaczone na podstawie testu Kupca dla modeli opartych na rozkładach uciętych w każdym analizowanym przypadku są nie mniejsze niż wartości p wyznaczone dla modeli opartych na całych rozkładach. Przyjmując, że im większa wartość p , tym lepsza jest jakość modelu, można stwierdzić, że modele oparte o testy zgodności dla rozkładów uciętych są nie gorsze niż modele oparte o testy zgodności dla całych rozkładów.

Tabela 4.4.5. Maksymalne wartości p testu Kupca uzyskane dla modeli II i III w podziale na wybrane typy rozkładów

L.p.	Indeks	Model	Typ rozkładu	Rząd kwantyla (poziom tolerancji dla VaR)	
				0,01	0,99
1	WIG20	Model II	rozkład ucięty	0,0889	0,7954
2			cały rozkład	0,0315	0,7954
3	L.p. 1 - L.p. 2			0,0574	0,0000
4	WIG20	Model III	rozkład ucięty	0,0889	0,7954
5			cały rozkład	0,0315	0,7954
6	L.p. 4 - L.p. 5			0,0574	0,0000
7	FTSE100	Model II	rozkład ucięty	0,0200	0,4748
8			cały rozkład	0,0110	0,2326
9	L.p. 7 - L.p. 8			0,0090	0,2422
10	FTSE100	Model III	rozkład ucięty	0,0352	0,3389
11			cały rozkład	0,0352	0,1535
12	L.p. 10 - L.p. 11			0,0000	0,1854

Źródło: obliczenia własne

Podsumowując uzyskane wyniki można stwierdzić, że w przypadku ogona górnego, stosując dowolny model, możliwe było skonstruowanie wielu prawidłowych strategii wyznaczania

wartości zagrożonej. W przypadku ogona dolnego akceptowalne prognozy wartości zagrożonej, ale tylko dla indeksu *WIG20*, uzyskano korzystając z modelu II oraz z modelu III. Wyniki te były możliwe do uzyskania dzięki zastosowaniu testów zgodności ze statystyką dedykowaną do badania dolnych ogonów rozkładów, AD_{down}^{**} . Prognozy sporządzone na podstawie modelu II, w porównaniu do prognoz bazujących na modelu III, charakteryzowały się jednak znacznie większą zmiennością. Wyniki uzyskane na podstawie modelu II oraz modelu III stanowią racjonalną alternatywę dla modelu I, ale tylko w odniesieniu do indeksu *WIG20*. Przyczyną niezadowolających wyników większości modeli są trudności w prognozowaniu wartości zagrożonych w roku 2020, będącym rokiem pandemii COVID-19. Istotne jest również spostrzeżenie, że modele oparte o testy zgodności dla rozkładów uciętych są nie gorsze niż modele oparte o testy zgodności dla całych rozkładów. Modele, w których zastosowano zmodyfikowane testy zgodności, będą przedmiotem dalszych badań autorki.

4.5. Wpływ wyboru testu zgodności na jakość portfeli na przykładzie rozkładów dwuwymiarowych

Niech R_p oznacza stopę zwrotu z portfela papierów wartościowych, natomiast v_p pewną przyjętą miarę ryzyka. Problem optymalizacji portfela w sposób ogólny można przedstawić na dwa sposoby. Polega on na minimalizacji ryzyka przy zadanej stopie zwrotu lub na maksymalizacji stopy zwrotu przy zadanym ryzyku (Jurek, 2001, s.88). W klasycznym modelu Markowitza ryzyko portfela jest wyrażone jego wariancją, a portfele utworzone poprzez rozwiązanie wymienionych wyżej dwóch zadań są określane jako efektywne. Tego typu portfele, dla różnych wymaganych przez inwestora poziomów zwrotu, rozważano na przykład w pracy Lach i Sikory (2010). W tym podrozdziale zostaną przedstawione pewne modyfikacje tych zadań, opisane w następnym paragrafie. Selekcja modeli generujących dane do zadań nastąpi w oparciu o wybrane testy zgodności dla całych rozkładów oraz dla części centralnych rozkładów. Celem podrozdziału jest porównanie wybranych charakterystyk rozkładów stóp zwrotu z portfeli, utworzonych na podstawie modeli generujących dane, wybranych w oparciu o: (1) testy zgodności dla części centralnych rozkładów, (2) testy zgodności dla całych rozkładów. Wybranymi charakterystykami rozkładów stóp zwrotu z portfeli są: odchylenie standardowe, skośność oraz kurtoza. Inwestorzy preferują te portfele, które charakteryzują się wyższymi wartościami momentów rzędu nieparzystego (wartość oczekiwana, skośność)

i niższymi rzędu parzystego (wariancja, kurtoza) (Dudzińska-Baryła, Kopańska-Bródka i Michalska, 2017). Dodatkowo dla każdego portfela zostanie wyznaczony wskaźnik Sharpe'a.

Podstawowy model Markowitza minimalizujący ryzyko portfela jest następujący:

$$v_p = \sum_{i=1}^n \sum_{j=1}^n cov_{ij} w_i w_j \Rightarrow \min, \quad (4.5.1)$$

$$R_p = \sum_{j=1}^n R_j w_j = r_k, \quad (4.5.2)$$

$$\sum_{j=1}^n w_j = 1, \quad (4.5.3)$$

$$0 \leq w_j \leq u_j \quad (j = 1, \dots, n), \quad (4.5.4)$$

gdzie cov_{ij} oznacza kowariancję pomiędzy stopami zwrotu z akcji i oraz j , R_p oznacza stopę zwrotu z portfela, r_k to wymagany przez inwestora poziom zwrotu, natomiast w_j , R_j oraz u_j oznaczają odpowiednio udział, oczekiwaną stopę zwrotu oraz maksymalny udział j -tej akcji w portfelu. W dalszej części będzie rozważany przypadek dla portfela dwuskładnikowego, gdy $n = 2$. Pierwsza różnica pomiędzy podstawowym modelem Markowitza a rozważaną w tej pracy jego modyfikacją dotyczy rodzaju stóp zwrotu. W podstawowym modelu Markowitza zakłada się stosowanie zwykłych stóp zwrotu i są one w tym przypadku rekomendowane (Dorfleitner, 2002), w niniejszej pracy zostaną jednak zastosowane logarytmiczne stopy zwrotu (r). Logarytmiczne stopy zwrotu mogą być stosowane w optymalizacji portfela (zob. np.: Gurgul i Syrek, 2007; Lindberg, 2009), jednakże ich stosowanie powinno być uzasadnione i wyraźnie zaakcentowane (Dorfleitner, 2002). Wszystkie badania empiryczne zostały przeprowadzone w niniejszej pracy na podstawie logarytmicznych stóp zwrotu, co tłumaczy ich zastosowanie również w tym badaniu empirycznym. Druga różnica pomiędzy podstawowym modelem Markowitza a rozważaną w tej pracy jego modyfikacją dotyczy sposobu modelowania zależności pomiędzy walorami. W podstawowym modelu Markowitza zależność jest wyrażona poprzez kowariancję. W niniejszej pracy zależność będzie modelowana poprzez połączenie rozkładów brzegowych kopulami. Kopule w problemie optymalizacji portfela stosowano do modelowania zależności np. w pracach Gurgula i Syrka (2007), Lach (2015) oraz Stachury i Wodeckiej (2013). W rozważanym modelu stopy zwrotu mają być na wymaganym przez inwestora, ale nie maksymalnym, poziomie, stąd większe

znaczenie może mieć dopasowanie rozkładów w części ograniczonej punktami $z_{H1} > 0$ i $z_{H2} < 1$, niż w ich ogonach.

Uwzględniając powyższe modyfikacje, badanie empiryczne składa się z następujących etapów:

1. Określenie źródła danych.

Danymi są dzienne procentowe logarytmiczne stopy zwrotu, wyliczone na podstawie kursów zamknięcia indeksów *WIG20* oraz *FTSE100* z lat 2019-2020, szczegółowo opisane w podrozdziale 3.4.

2. Zdefiniowanie modeli do generowania symulacji Monte Carlo.

Modele generujące dane są takie same, jak w podrozdziale 4.3. Są to dwuwymiarowe rozkłady, utworzone poprzez połączenie rozkładów brzegowych (*N*, *NIG*, *Hyp*, *SGED*, *ST*) kopulami (Gaussa, *t*-Studenta, Claytona, Gumbela, Franka i Joe). Rozkłady wyznaczono oddzielnie dla roku 2019 i roku 2020, stąd dla każdego wyróżnionego okresu utworzono po 150 modeli generujących dane.

3. Losowanie danych.

Z każdego rozkładu dwuwymiarowego losowano po 10 prób o liczebności równej liczbie obserwacji w badanym okresie.

4. Utworzenie portfela.

Portfel tworzący rozwiązując następujące zadanie, dopuszczające krótką sprzedaż, w którym udziały aktywów w portfelu sumują się do jedności:

$$v_p = \sigma_p^2 \Rightarrow \min, \tag{4.5.5}$$

$$R_p = \sum_{j=1}^2 r_j w_j = r_k, \tag{4.5.6}$$

$$\sum_{j=1}^2 w_j = 1, \tag{4.5.7}$$

$$-2 \leq w_j \leq 2 \quad (j = 1, 2), \tag{4.5.8}$$

gdzie r_j oznacza oczekiwaną dzienną logarytmiczną stopę zwrotu z j -tej akcji w portfelu. Zadania rozwiązywano dla 5 przyjętych poziomów stopy zwrotu z inwestycji $r_k \in \{0,02; 0,025; 0,03; 0,035; 0,04\}$. W każdym okresie wygenerowano więc po 7500 portfeli (150 rozkładów x 5 przyjętych stóp zwrotu x 10 prób).

5. Wyznaczenie mierników jakości portfela.

Wyznaczono następujące charakterystyki rozkładu stopy zwrotu z każdego portfela: odchylenie standardowe, skośność oraz kurtozę. Dodatkowo w każdym przypadku wyznaczono wskaźnik Sharpe'a, będący ilorazem oczekiwanej stopy zwrotu oraz odchylenia standardowego stopy zwrotu z portfela.

6. Ocena zgodności rozkładu rzeczywistego z wybranym rozkładem teoretycznym.

Selekcji modeli generujących dane dokonano na podstawie testów zgodności przeprowadzonych dla obydwu rozkładów brzegowych. Model generujący dane uznawano za poprawny, jeśli dla obydwu rozkładów brzegowych nie było podstaw do odrzucenia hipotezy zerowej zakładającej zgodność rozkładu empirycznego z rozkładem teoretycznym w oparciu o wybrany test zgodności. W tym punkcie wykorzystano wyniki testów zgodności zawarte w załączniku 4. Wyniki te zawierają oceny zgodności dokonane w oparciu o statystyki dla całych rozkładów (KS^{**} , V^{**} , AD^{**} , W^{2**} , AD^{2**}) oraz statystyki dla rozkładów obustronnie uciętych, z punktami ucięcia dystrybuanty 0,1 i 0,9 oraz 0,05 i 0,95 (KS^{**} , V^{**} , AD^{**} , W^{2**} , AD^{2**}).

W wyniku przeprowadzenia testów zgodności powstało tyle zbiorów modeli generujących dane, ile rozważano statystyk testowych. Na podstawie każdego wyselekcjonowanego modelu generującego dane tworzono 50 portfeli. Rozkład stopy zwrotu z każdego portfela scharakteryzowano za pomocą odchylenia standardowego, skośności, kurtozy i wskaźnika Sharpe'a. Dla powstałych w ten sposób zbiorów oszacowań odchyłeń standardowych, skośności, kurtozy i wskaźnika Sharpe'a wyznaczono mediany. Mediany te znajdują się w tabeli 4.5.1.

W pierwszej kolejności zostaną omówione wyniki uzyskane na podstawie stóp zwrotu z roku 2019, w którym indeksy charakteryzowały się niską zmiennością. W odniesieniu do każdej statystyki dla całych rozkładów ($KS^{**}[0; 1]$, $V^{**}[0; 1]$, $AD^{**}[0; 1]$, $W^{2**}[0; 1]$, $AD^{2**}[0; 1]$), zastąpienie jej w testach zgodności analogiczną statystyką dla rozkładów uciętych (w punktach 0,05 i 0,95 lub 0,1 i 0,9), powoduje w konsekwencji uzyskanie nie gorszych median rozkładów wybranych charakterystyk portfela w co najmniej trzech przypadkach z czterech rozważanych (rozważane charakterystyki to: odchylenie standardowe, skośność, kurtoza i wskaźnik Sharpe'a). „Nie gorsze” oznacza tutaj, że mediana jest nie większa w przypadku rozkładów odchylenia standardowego i kurtozy oraz nie mniejsza w przypadku rozkładów skośności i

wskaźnika Sharpe'a. Natomiast zastosowanie statystyk $KS^{**}[0,05; 0,95]$, $KS^{**}[0,1; 0,9]$, $V^{**}[0,05; 0,95]$, $AD^{**}[0,05; 0,95]$ zamiast ich odpowiedników dla całych rozkładów powoduje w konsekwencji uzyskanie nie gorszych median rozkładów wybranych charakterystyk portfela we wszystkich rozważanych czterech przypadkach.

Przejdźmy teraz do omówienia wyników uzyskanych na podstawie stóp zwrotu z roku 2020, w którym indeksy charakteryzowały się wysoką zmiennością. Tutaj nie gorsze mediany rozkładów wybranych charakterystyk portfeli w dokładnie trzech przypadkach uzyskano w konsekwencji zastąpienia w testach zgodności statystyki $V^{**}[0; 1]$ przez statystykę $V^{**}[0,05; 0,95]$ lub $V^{**}[0,1; 0,9]$. Nie gorsze mediany rozkładów wybranych charakterystyk portfeli w dokładnie czterech przypadkach uzyskano w konsekwencji zastąpienia w testach zgodności: statystyki $AD^{**}[0; 1]$ przez statystykę $AD^{**}[0,05; 0,95]$ lub $AD^{**}[0,1; 0,9]$, statystyki $W^{2**}[0; 1]$ przez statystykę $W^{2**}[0,1; 0,9]$ oraz statystyki $AD^{2**}[0; 1]$ przez statystykę $AD^{2**}[0,1; 0,9]$.

Tabela 4.5.1. Mediany rozkładów wybranych charakterystyk portfela ⁽¹⁾

Statystyka testowa będąca podstawą wyboru rozkładów	Rok 2019				Rok 2020			
	Mediana rozkładu odchylenia standardowego	Mediana rozkładu skośności	Mediana rozkładu kurtozy	Mediana rozkładu wskaźnika Sharpe'a	Mediana rozkładu odchylenia standardowego	Mediana rozkładu skośności	Mediana rozkładu kurtozy	Mediana rozkładu wskaźnika Sharpe'a
$KS^{**}[0; 1]$	0,7729	0,0245	1,0761	3,8462	2,0526	-0,0616	3,5906	1,4563
$KS^{**}[0,05; 0,95]$	0,7729	0,0245	1,0761	3,8462	2,0912	-0,1044	4,4664	1,4218
$KS^{**}[0,1; 0,9]$	0,7712	0,0279	1,0568	3,8462	2,0572	-0,0213	4,0137	1,4493
$V^{**}[0; 1]$	0,7772	0,0426	1,2372	3,7975	2,0556	-0,0035	4,6394	1,4563
$V^{**}[0,05; 0,95]$	0,7748	0,0459	1,2194	3,8278	2,0489	-0,0302	3,7857	1,4563
$V^{**}[0,1; 0,9]$	0,7729	0,0245	1,0761	3,8462	2,0489	-0,0302	3,7857	1,4563
$AD^{**}[0; 1]$	0,7722	0,0154	1,2562	3,8462	2,1073	-0,1287	5,1316	1,4218
$AD^{**}[0,05; 0,95]$	0,7702	0,0225	1,0145	3,8462	2,0370	-0,0200	2,6010	1,4599
$AD^{**}[0,1; 0,9]$	0,7729	0,0245	1,0761	3,8462	2,0385	0,0001	2,4326	1,4599
$W^{2**}[0; 1]$	0,7772	0,0426	1,2372	3,7975	2,0653	-0,0681	3,7565	1,4493
$W^{2**}[0,05; 0,95]$	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.	b.d.
$W^{2**}[0,1; 0,9]$	0,7712	0,0279	1,0568	3,8462	2,0653	-0,0681	3,7565	1,4493
$AD^{2**}[0; 1]$	0,7762	0,0345	1,4048	3,7975	2,0653	-0,0681	3,7565	1,4493
$AD^{2**}[0,05; 0,95]$	0,7772	0,0426	1,2372	3,7975	2,1347	-0,1079	4,6155	1,4019
$AD^{2**}[0,1; 0,9]$	0,7729	0,0245	1,0761	3,8462	2,0653	-0,0681	3,7565	1,4493

(1) Gdy mediana określonej charakterystyki, wyznaczona na podstawie modeli wybranych w oparciu o statystykę dla rozkładów uciętych, była nie gorsza od mediany wyznaczonej na podstawie modeli wybranych w oparciu o analogiczną statystykę dla całych rozkładów, wartość mediany umieszczano na szarym tle. „Nie gorsza” oznacza nie większą w przypadku odchylenia standardowego i kurtozy oraz nie mniejszą w przypadku skośności i wskaźnika Sharpe'a.

Źródło: obliczenia własne

Inwestor, decydując się na w pełni parametryczną metodę konstrukcji rozkładu, powinien w przypadku obydwu rozkładów brzegowych ocenić zgodność rozkładu empirycznego z rozkładem teoretycznym w oparciu o wybrany test zgodności. Wybór statystyki testowej determinuje wybór modeli generujących dane, a wygenerowane dane są podstawą do wyznaczania portfeli optymalnych. W przypadku konstrukcji portfela o zadanej stopie zwrotu, zgodność rozkładów w części centralnej może mieć większe znaczenie, niż zgodność całych rozkładów. Na podstawie badań symulacyjnych, przeprowadzonych w oparciu o stopy zwrotu z indeksów *WIG20* i *FTSE100* z lat 2019-2020 stwierdzono, że stosowanie statystyk dla rozkładów uciętych zamiast statystyk dla całych rozkładów w testach zgodności dla rozkładów brzegowych, może w konsekwencji przynieść inwestorowi wymierne korzyści. W szczególności, zastąpienie statystyk $V^{**}[0; 1]$, $AD^{**}[0; 1]$, $W^{2**}[0; 1]$ oraz $AD^{2**}[0; 1]$ ich odpowiednikami dla rozkładów uciętych w punktach 0,1 i 0,9, prowadzi w rezultacie do poprawy mediany rozkładów przynajmniej trzech z czterech badanych charakterystyk portfela (odchylenia standardowego, skośności, kurtozy, wskaźnika Sharpe'a).

Zakończenie

Głównym celem rozprawy było „opracowanie modyfikacji wybranych testów zgodności dla uciętych rozkładów ciągłych i ocena ich przydatności w analizie rozkładów stóp zwrotu z akcji”. Cel główny miał zostać osiągnięty przez realizację pięciu celów pomocniczych.

Zgodnie z pierwszym celem pomocniczym, należało wyprowadzić formuły obliczeniowe statystyk dla prawostronnie i obustronnie uciętych rozkładów ciągłych, gdy liczba obserwacji uciętych nie jest znana. Wybrano następujące statystyki: Kołmogorowa-Smirnowa, Kuipera, Craméra-von Misesa, Andersona-Darlinga w wersjach supremum i kwadratowej oraz Andersona-Darlinga w wersjach supremum i kwadratowej z funkcją wagową dla ogona dolnego. Formuły te dla rozkładów obustronnie uciętych zostały wyprowadzone w załączniku 1 umieszczonym w aneksie i przedstawione w podrozdziale 3.1. Statystyki dla rozkładów prawostronnie lub lewostronnie uciętych są otrzymywane poprzez zastosowanie odpowiednich progów ucięcia. Gdy $z_{H1} > 0$ oraz $z_{H2} = 1$, otrzymuje się wzory dla rozkładów lewostronnie uciętych, gdy $z_{H1} = 0$ oraz $z_{H2} < 1$, wzory dla rozkładów prawostronnie uciętych, a gdy $z_{H1} > 0$ oraz $z_{H2} < 1$, wzory dla rozkładów obustronnie uciętych. Dodatkowo, gdy przyjmiemy, że progi ucięcia wynoszą $z_{H1} = 0$ i $z_{H2} = 1$, wyprowadzone przez autorkę formuły obliczeniowe redukują się do klasycznych statystyk przeznaczonych do badania całych rozkładów. Według najlepszej wiedzy nie istnieją w literaturze statystyki dla rozkładów obustronnie uciętych oraz prawostronnie uciętych, gdy liczba obserwacji poza progami ucięcia nie jest znana. Wyznaczone formuły obliczeniowe mogą być postrzegane jako uogólnienie statystyk dla rozkładów lewostronnie uciętych, które zostały wcześniej opisane w literaturze.

Zgodnie z drugim celem pomocniczym, należało opracować procedurę wnioskowania statystycznego w oparciu o statystyki dla obustronnie uciętych rozkładów ciągłych, która uwzględnia informację dotyczącą liczby uciętych danych. Procedura taka została przedstawiona w podrozdziale 3.2 (procedura I). Cel ten zawiera również opracowanie metody losowania danych z rozkładu, zapewniającej stałą liczbę losowanych danych z wybranego fragmentu dystrybuanty. Metoda taka, stanowiąca uszczegółowienie punktu 2 procedury I, została przedstawiona w podrozdziale 3.3. Należy zaznaczyć, że procedury wnioskowania statystycznego w oparciu o statystyki dla lewostronnie uciętych rozkładów ciągłych, gdy liczba uciętych danych nie jest znana, zostały już przedstawione w literaturze. Procedurę tę,

z niewielkimi zmianami, można z powodzeniem stosować również wraz ze statystykami dla rozkładów obustronnie uciętych, przedstawionych w niniejszej pracy.

Zgodnie z trzecim celem pomocniczym, należało ocenić jakość testów zgodności dla uciętych rozkładów ciągłych w wybranych fragmentach rozkładów bezwarunkowych, na podstawie wybranych przykładów. Ocena jakości miała objąć ocenę rozmiarów i mocy testów. Wybranymi fragmentami były natomiast ogony rozkładów oraz część dystrybuanty pozostająca po odrzuceniu ogonów. W przypadku rozkładów prawostronnie i obustronnie uciętych, przedstawionych w niniejszej pracy, takich badań dotychczas nie przeprowadzono. Empiryczne rozmiary i moc testów zgodności dla rozkładów lewostronnie uciętych były natomiast w literaturze przedstawiane, jednak dla innych rozkładów lub innych parametrów rozkładów. Metody zastosowane do oceny jakości zmodyfikowanych testów zgodności zostały opisane w podrozdziale 3.3, a wybrane przykłady w podrozdziale 3.4. W podrozdziale 3.5 przedstawiono wyniki oszacowań rozmiarów, a w podrozdziale 3.6 wyniki oszacowań mocy opracowanych testów zgodności. Testy w większości przypadków posiadały właściwe oszacowania rozmiarów, oszacowania mocy natomiast zależały od wielkości próby, weryfikowanej hipotezy, badanego fragmentu dystrybuanty oraz od zastosowanych poziomów ucięcia.

Czwarty cel pomocniczy zakładał wykorzystanie zmodyfikowanych testów zgodności dla rozkładów lewostronnie i prawostronnie uciętych do weryfikacji zgodności rozkładu empirycznego z rozkładem teoretycznym w ogonach rozkładów stóp zwrotu z akcji. Wyniki weryfikacji, wyznaczone na podstawie wybranych przykładów, znajdują się w załączniku 4 w aneksie. Cel ten obejmował również ocenę jakości oszacowań wartości zagrożonej w zależności od wyboru testu zgodności oraz opracowanie autorskiej procedury wyznaczania wartości zagrożonej w horyzoncie jednodniowym wykorzystującej testy zgodności. Oceny jakości oszacowań wartości zagrożonej w zależności od wyboru testu zgodności dokonano w podrozdziałach 4.2, 4.3 oraz 4.4. W podrozdziałach 4.2 i 4.3 przekroczenia wartości zagrożonej były badane na podstawie tej samej próby, która była podstawą do szacowania modelu. W podrozdziale 4.4 przekroczenia wartości zagrożonej były weryfikowane poza próbą. Dwie autorskie procedury wyznaczania wartości zagrożonej w horyzoncie jednodniowym, w których dzięki zastosowaniu wybranych testów zgodności uzyskano poprawę jakości prognoz, zostały przedstawione w podrozdziale 4.4.

Ostatni, piąty cel pomocniczy, zakładał zastosowanie zmodyfikowanych testów zgodności dla rozkładów obustronnie uciętych do weryfikacji zgodności rozkładu empirycznego z rozkładem teoretycznym w centralnej części rozkładów stóp zwrotu z akcji. Wyniki weryfikacji, wyznaczone na podstawie wybranych przykładów, znajdują się w załączniku 4 w aneksie. Cel ten obejmował również ocenę jakości portfela papierów wartościowych, mierzoną wybranymi charakterystykami rozkładu stopy zwrotu z portfela (odchyleniem standardowym, skośnością, kurtozą oraz wskaźnikiem Sharpe'a), w zależności od wyboru testu zgodności. Cel ten zrealizowano w podrozdziale 4.5. Uzyskane wyniki pozwalają stwierdzić, że poprzez wybór statystyki testowej można odpowiednio wpływać na charakterystyki portfela, to znaczy zwiększać wartości momentów rzędu nieparzystego (skośność) oraz zmniejszać wartości momentów rzędu parzystego (wariancja, kurtoza).

Wszystkie sformułowane cele pomocnicze zostały zrealizowane, można więc stwierdzić, że cel główny pracy również został osiągnięty. Przedstawione w pracy modyfikacje wybranych testów zgodności dla uciętych rozkładów ciągłych okazały się przydatne w analizie rozkładów stóp zwrotu z akcji, w szczególności w analizie ogonów rozkładów oraz części centralnej rozkładów.

W części empirycznej zweryfikowano hipotezę główną, zgodnie z którą zastosowanie opracowanych w pracy, zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych, pozwala na wybór nie gorszego modelu pomiaru ryzyka rynkowego niż model pomiaru ryzyka rynkowego wybrany na podstawie testów zgodności dla całych rozkładów. Weryfikację hipotezy głównej przeprowadzono poprzez weryfikację dwóch hipotez szczegółowych.

Najpierw zbadano pierwszą hipotezę szczegółową mówiącą, że zastosowanie opracowanych w pracy, zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych, w ogonach rozkładów, pozwala na wybór nie gorszego modelu wartości zagrożonej niż model wartości zagrożonej wybrany na podstawie testów zgodności dla całych rozkładów. W przypadku rozkładów jednowymiarowych (podrozdział 4.2), zbiór modeli generujących prawidłowe oszacowania wartości zagrożonej dla ogona dolnego otrzymywano stosując dowolną statystykę. Zbiór modeli generujących prawidłowe oszacowania wartości zagrożonej dla ogona górnego otrzymywano jedynie w wyniku zastosowania statystyki $AD^{**}[0,95; 1]$. W przypadku rozkładów dwuwymiarowych (podrozdział 4.3), w każdym badanym okresie, dla kwantyli rzędu 0,01, 0,05 oraz 0,99 można było wybrać statystykę testową dla rozkładu uciętego, która prowadziła do wyboru modeli wartości zagrożonej niemniej dokładnych, niż

modele wybrane przy użyciu statystyk testowych dla całego rozkładu, przy czym dokładność była mierzona medianą uzyskanych w oparciu o wyselekcjonowane modele udziałów przekroczeń. Stwierdzenie to nie było prawdziwe w odniesieniu do kwantyla rzędu 0,95 z górnego ogona rozkładu. Ponadto, w podrozdziale 4.4 pokazano, że jedyne akceptowalne strategie do wyznaczania jednodniowych prognoz wartości zagrożonej dla kwantyla rzędu 0,01, dla indeksu *WIG20*, można utworzyć stosując testy zgodności ze statystyką dedykowaną do badania dolnych ogonów rozkładów, AD_{down}^{**} . Pierwszą hipotezę pomocniczą można zatem w odniesieniu do kwantyli rzędu 0,01, 0,05 oraz 0,99 zweryfikować pozytywnie.

Następnie zbadano drugą hipotezę szczegółową, zgodnie z którą, w przypadku konstrukcji portfela o zadanej stopie zwrotu, zastosowanie opracowanych w pracy, zmodyfikowanych testów zgodności dla uciętych rozkładów ciągłych, w części centralnej rozkładów, pozwala na wybór nie gorszego modelu pomiaru ryzyka portfela papierów wartościowych niż model pomiaru ryzyka portfela papierów wartościowych wybrany na podstawie testów zgodności dla całych rozkładów. Na podstawie badań symulacyjnych, przeprowadzonych w oparciu o stopy zwrotu z indeksów *WIG20* i *FTSE100* z lat 2019-2020 stwierdzono, że stosowanie statystyk dla rozkładów uciętych zamiast statystyk dla całych rozkładów w testach zgodności dla rozkładów brzegowych, może w konsekwencji przynieść inwestorowi wymierne korzyści. W szczególności, zastąpienie statystyk $V^{**}[0; 1]$, $AD^{**}[0; 1]$, $W^{2**}[0; 1]$ oraz $AD^{2**}[0; 1]$ ich odpowiednikami dla rozkładów uciętych w punktach 0,1 i 0,9, prowadziło w rezultacie do poprawy mediany rozkładów przynajmniej trzech z czterech badanych charakterystyk portfela (odchylenia standardowego, skośności, kurtozy i wskaźnika Sharpe'a). Uzyskane wyniki pozwalają na pozytywną weryfikację drugiej hipotezy szczegółowej w odniesieniu do statystyk V^{**} , AD^{**} , W^{2**} oraz AD^{2**} . Pozytywna weryfikacja obydwu hipotez szczegółowych pozwala natomiast na pozytywną weryfikację hipotezy głównej.

Można stwierdzić, że wszystkie cele pomocnicze, a zatem również cel główny pracy, zostały zrealizowane. Ponadto pozytywnie zweryfikowano obydwie hipotezy szczegółowe, co pozwoliło na pozytywną weryfikację hipotezy głównej. Można zatem stwierdzić, że opracowane modyfikacje wybranych testów zgodności dla uciętych rozkładów ciągłych są przydatne w analizie rozkładów stóp zwrotu z akcji.

Bibliografia

- Aas, K. i Haff, D. H. (2008). *The generalized hyperbolic skew student's t-distribution* (SSRN Scholarly Paper Nr ID 922919). Pobrane z <https://papers.ssrn.com/abstract=922919>
- Ahmad, M. I., Sinclair, C. D. i Spurr, B. D. (1988). Assessment of flood frequency models using empirical distribution function statistics. *Water Resources Research*, 24(8), 1323–1328. <https://doi.org/10.1029/WR024i008p01323>
- Anderson, T. W. i Darling, D. A. (1952). Asymptotic theory of certain „goodness of fit” criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2), 193–212.
- Anderson, T. W. i Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, 49(268), 765–769. <https://doi.org/10.2307/2281537>
- Ayebo, A. i Kozubowski, T. J. (2003). An Asymmetric Generalization of Gaussian and Laplace Laws. *Journal of Probability and Statistical Science*, 187–210.
- Azzalini, A. i Arellano-Valle, R. B. (2013). Maximum penalized likelihood estimation for skew-normal and skew-t distributions. *Journal of Statistical Planning and Inference*, 143(2), 419–433. <https://doi.org/10.1016/j.jspi.2012.06.022>
- Azzalini, A. i Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2), 367–389. <https://doi.org/10.1111/1467-9868.00391>
- Babu, G. J. i Rao, C. R. (2004). Goodness-of-fit tests when parameters are estimated. *Sankhyā: The Indian Journal of Statistics (2003-2007)*, 66(1), 63–74. <https://doi.org/10.2307/25053332>
- Baringhaus, L., Ebner, B. i Henze, N. (2017). The limit distribution of weighted L^2 -goodness-of-fit statistics under fixed alternatives, with applications. *Annals of the Institute of Statistical Mathematics*, 69(5), 969–995. <https://doi.org/10.1007/s10463-016-0567-8>
- Barndorff-Nielsen, O. (1977). Exponentially Decreasing Distributions for the Logarithm of Particle Size. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 353(1674), 401–419.

- Barndorff-Nielsen, O. E. i Stelzer, R. (2005). Absolute Moments of Generalized Hyperbolic Distributions and Approximate Scaling of Normal Inverse Gaussian Lévy Processes. *Scandinavian Journal of Statistics*, 32(4), 617–637. <https://doi.org/10.1111/j.1467-9469.2005.00466.x>
- Barr, D. R. i Davidson, T. (1973). A Kolmogorov-Smirnov test for censored samples. *Technometrics*, 15(4), 739–757. <https://doi.org/10.1080/00401706.1973.10489108>
- Boos, D. D. (1984). Using extreme value theory to estimate large percentiles. *Technometrics*, 26(1), 33. <https://doi.org/10.1080/00401706.1984.10487919>
- Boucher, C. M., Daniélsson, J., Kouontchou, P. S. i Maillet, B. B. (2014). Risk models-at-risk. *Journal of Banking and Finance*, 44, 72–92. <https://doi.org/10.1016/j.jbankfin.2014.03.019>
- Broda, S. A., Krause, J. i Paoletta, M. S. (2018). Approximating expected shortfall for heavy-tailed distributions. *Econometrics and Statistics*, 8(C), 184–203.
- Brunlid, H. (2007). A Comparative Analysis of Hyperbolic Copulas Induced by a One Factor Lévy Model. <https://www.semanticscholar.org/paper/A-Comparative-Analysis-of-Hyperbolic-Copulas-by-a-Brunlid/5cdebfee05946d862276dabd4f421e839147757f>
- Caeiro, F. i Gomes, M. I. (2016). Threshold selection in extreme value analysis. W: Dey, D.K. i Yan, J. (Eds.), *Extreme value modeling and risk analysis: methods and applications* (s. 21–42). Boca Raton: Chapman and Hall/CRC. <https://doi.org/10.1201/b19721-5>
- Caers, J. i Maes, M. A. (1998). Identifying tails, bounds and end-points of random variables. *Structural Safety*, 20(1), 1–23. [https://doi.org/10.1016/S0167-4730\(97\)00036-2](https://doi.org/10.1016/S0167-4730(97)00036-2)
- Chapman, D. G. (1958). A comparative study of several one-sided goodness-of-fit tests. *The Annals of Mathematical Statistics*, 29(3), 655–674. <https://doi.org/10.1214/aoms/1177706529>
- Chernobai, A., Burnecki, K., Rachev, S.T., Trück, S. i Weron, R. (2006). Modelling catastrophe claims with left-truncated severity distributions. *Computational Statistics*, 21(3), 537–555. <https://doi.org/10.1007/s00180-006-0011-2>
- Chernobai, A., Menn, C., Rachev, S.T. i Trück, S. (2010). *Estimation of operational value-at-risk in the presence of minimum collection threshold: an empirical study* (Working Paper Series in Economics Nr 4). Pobrane z <https://econpapers.repec.org/paper/zbwkitwps/4.htm>

- Chernobai, A., Rachev, S. i Fabozzi, F. (2015). Composite goodness-of-fit tests for left truncated loss samples. W: C. Lee i J. Lee (Eds.), *Handbook of Financial Econometrics and Statistics* (s. 575-596). New York: Springer.
- Choi, J., Du, Y. i Song, Q. (2021). Inverse Gaussian quadrature and finite normal-mixture approximation of the generalized hyperbolic distribution. *Journal of Computational and Applied Mathematics*, 388, 113302. <https://doi.org/10.1016/j.cam.2020.113302>
- Choi, P. i Min, I. (2013). Marginal or copula: which one is critical? *Applied Economics Letters*, 20(16), 1462–1465. <https://doi.org/10.1080/13504851.2013.826870>
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, 39(4), 841–862. <https://doi.org/10.2307/2527341>
- Cizek, P., Härdle, W. K. i Weron, R. (2011). *Statistical tools for finance and insurance*. Berlin: Springer. Pobrane z <https://econpapers.repec.org/bookchap/wuuhsbook/hsbook0501.htm>
- Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1, 223–236.
- Cruz, M. G., Shevchenko, P. V. i Peters, G. W. (2015). *Fundamental aspects of operational risk and insurance analytics: a handbook of operational risk*. Hoboken, New Jersey: Wiley.
- Czapkiewicz, A. i Basiura, B. (2014). The position of the Wig index in comparison with selected market indices in boom and bust periods. *Statistics in Transition. New Series*, (3), 427–436.
- Duda, J., Gurgul, H. i Syrek, R. (2022). Multi-feature evaluation of financial contagion. *Central European Journal of Operations Research*, 30(4), 1167–1194. <https://doi.org/10.1007/s10100-021-00756-3>
- Danielsson, J. i de Vries, C. (1998). *Value-at-risk and extreme returns* (Tinbergen Institute Discussion Paper Nr 98-017/2). Pobrane z <https://papers.tinbergen.nl/98017.pdf>
- Danielsson, J., Ergun, L. M., de Haan, L. i de Vries, C. G. (2016). *Tail index estimation: quantile driven threshold selection* (SSRN Scholarly Paper Nr ID 2717478). <https://doi.org/10.2139/ssrn.2717478>
- Danielsson, J., James, K. R., Valenzuela, M. i Zer, I. (2016). Model risk of risk models. *Journal of Financial Stability*, 23, 79–91. <https://doi.org/10.1016/j.jfs.2016.02.002>

- Davison, A. C. i Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge Series in Statistical and Probabilistic Mathematics, 1.
<https://doi.org/10.1017/CBO9780511802843>
- Deluca, A. i Corral, A. (2013). Fitting and goodness-of-fit test of non-truncated and truncated power-law distributions. *Acta Geophysica*, 61(6), 1351–1394.
<https://doi.org/10.2478/s11600-013-0154-9>
- Doman, M. i Doman, R. (2009). *Modelowanie zmienności i ryzyka: Metody ekonometrii finansowej*. Oficyna a Wolters Kluwer business.
- Doman, R. (2011). *Zastosowania kopuli w modelowaniu dynamiki zależności na rynkach finansowych*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego.
- Domański, C. (1979). *Statystyczne testy nieparametryczne*. Warszawa: Państwowe Wydawnictwo Ekonomiczne.
- Domański, C. (2014). *Testy statystyczne w procesie podejmowania decyzji*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Domański, C. i Pruska, K. (2000). *Nieklasyczne metody statystyczne*. Warszawa: Polskie Wydawnictwo Ekonomiczne.
- Dorfleitner, G. (2002). *Why the Return Notion Matters* (SSRN Scholarly Paper Nr ID 302811).
Pobrane z <https://doi.org/10.2139/ssrn.302811>
- Dudzińska-Baryła, R., Kopańska-Bródka, D. i Michalska, E. (2017). Struktura portfeli efektywnych w modelach średnia-wariancja-skośność. *Finanse, Rynki Finansowe, Ubezpieczenia*, 2(86), 185–196. <https://doi.org/10.18276/frfu.2017.86-15>
- Dufour, R. i Maag, U. R. (1978). Distribution results for modified Kolmogorov-Smirnov statistics for truncated or censored samples. *Technometrics*, 20(1), 29–32.
<https://doi.org/10.1080/00401706.1978.10489613>
- Eberlein, E. i Hammerstein, E. A. (2004). *Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes* (T. 58, s. 264). Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-7943-9_15
- Eberlein, E. i Keller, U. (1995). Hyperbolic Distributions in Finance. *Bernoulli*, 1(3), 281–299.
<https://doi.org/10.2307/3318481>
- Eberlein, E. i Prause, K. (2002). The generalized hyperbolic model: financial derivatives and risk measures. W: Geman, H., Madan, D., Pliska, S. R. i Vorst, T. (Eds.), *Mathematical Finance—Bachelier Congress 2000: selected papers from the first world congress of the*

- bachelier finance society, Paris, June 29–July 1, 2000* (s. 245–267). Berlin: Springer.
https://doi.org/10.1007/978-3-662-12429-1_12
- Echaust, K. (2014). *Ryzyko zdarzeń ekstremalnych na rynku kontraktów futures w Polsce*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego.
- Echaust, K. i Just, M. (2020a). A Comparison of Conditional and Unconditional VaR Models. University of Hradec Kralove. <https://doi.org/10.36689/uhk/hed/2020-01-014>
- Echaust, K. i Just, M. (2020b). Value at risk estimation using the GARCH-EVT approach with optimal tail selection. *Mathematics*, 8(1), 114. <https://doi.org/10.3390/math8010114>
- Echaust, K. i Lach, A. (2017). Goodness-of-fit tests for truncated distributions. *35th International Conference Mathematical Methods in Economics: Conference Proceedings*.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1), 1–26. <https://doi.org/10.1214/aos/1176344552>
- Eicker, F. (1979). The asymptotic distribution of the suprema of the standardized empirical processes. *The Annals of Statistics*, 7(1), 116–138.
<https://doi.org/10.1214/aos/1176344559>
- Fang, H.-B., Fang, K.-T. i Kotz, S. (2002). The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis*, 82(1), 1–16.
- Fernández, C., Osiewalski, J. i Steel, M. F. J. (1995). Modeling and Inference with u-Spherical Distributions. *Journal of the American Statistical Association*, 90(432), 1331–1340.
<https://doi.org/10.1080/01621459.1995.10476637>
- Fernandez, C. i Steel, M. F. J. (1998). On bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93(441), 359–371. <https://doi.org/10.2307/2669632>
- Fischer, M. (2011). Generalized hyperbolic distributions. W: *International Encyclopedia of Statistical Science*. Berlin: Springer.
- Fischer, Matthias i Jakob, K. (2016). PTAS distributions with application to risk management. *Journal of Statistical Distributions and Applications*, 3(1), 1–18.
<https://doi.org/10.1186/s40488-016-0049-9>
- Fisz, M. (1969). *Rachunek prawdopodobieństwa i statystyka matematyczna* (wyd. 4). Warszawa: Państwowe Wydawnictwo Naukowe.
- Gaunt, R. E. (2014). Variance-Gamma approximation via Stein's method. *Electronic Journal of Probability*, 19. <https://doi.org/10.1214/EJP.v19-3020>

- Gerstenkorn, T. i Śródka, T. (1976). *Kombinatoryka i rachunek prawdopodobieństwa*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Górecki, T. (2011). *Podstawy statystyki z przykładami w R*. Legionowo: Wydawnictwo BTC.
- Guegan, D. i Hassani, B. K. (2018). More accurate measurement for enhanced controls: VaR vs ES? *Journal of International Financial Markets, Institutions and Money*, 54, 152–165. <https://doi.org/10.1016/j.intfin.2017.06.002>
- Gurgul, H. i Syrek, R. (2007). Wykorzystanie kopul do konstrukcji portfeli inwestycyjnych. *Ekonomia Menedżerska*, 2, 31–44.
- Haas, M. i Pigorsch, C. (2009). Financial economics, fat-tailed distributions. W: *Encyclopedia of complexity and systems science* (T. 4, s. 3404–3435).
- Halphen, E. (1941). Sur un nouveau type de courbe de frequence. *Comptes Rendus de l'Academie des Sciences*, 213, 633–635 (opublikowano pod pseudonimem Dugue).
- Halphen, E. (1953). Un exemple d'application des méthodes statistiques: Le problème du plan pour l'équipement électrique français. *Revue de Statistique Appliquée*, 1(1), 39–46.
- Heilpern, S. (2007). *Funkcje łączące*. Wrocław: Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu.
- Heilpern, S. (2014). Zależność: fakty i mity. *Studia Ekonomiczne / Uniwersytet Ekonomiczny w Katowicach*, 207, 85–98.
- Janssen, A. (2000). Global power functions of goodness of fit tests. *Annals of Statistics*, 28(1), 239–253. <https://doi.org/10.1214/aos/1016120371>
- Jeziorski, P. (2014). Wybrane testy nieobciążoności miar ryzyka na przykładzie value at risk. *Studia Ekonomiczne / Uniwersytet Ekonomiczny w Katowicach*, 207, 110–118.
- Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman and Hall/CRC.
- Jurek, W. (2001). *Konstrukcja i analiza portfela papierów wartościowych o zmiennym dochodzie*. Poznań: Wydawnictwo Akademii Ekonomicznej w Poznaniu.
- Just, M. (2012). Porównanie metod szacowania wartości zagrożonej na rynku metali szlachetnych. *Zeszyty Naukowe SGGW - Ekonomia i Organizacja Gospodarki Żywnościowej*, (96), 181–193.
- Karl, A. T., Eubank, R., Milovanovic, J., Reiser, M. i Young, D. (2014). Using RngStreams for parallel random number generation in C++ and R. *Computational Statistics*, 29(5), 1301. <https://doi.org/10.1007/s00180-014-0492-3>

- Kębłowski, P. (2013). Właściwości wybranych metod małopróbkowego wnioskowania o rzędnice kointegracji. *Przegląd Statystyczny*, 60(2), 163–185.
- Kim, G., Silvapulle, M. J. i Silvapulle, P. (2007). Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, 51(6), 2836–2850. <https://doi.org/10.1016/j.csda.2006.10.009>
- Komunjer, I. (2007). Asymmetric power distribution: theory and applications to risk measurement. *Journal of Applied Econometrics*, 22(5), 891–921. <https://doi.org/10.1002/jae.961>
- Krysicki, W. (Red.). (2004). *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Cz. 1: Rachunek prawdopodobieństwa* (Wyd. 9). Warszawa: Wydawnictwo Naukowe PWN.
- Krysicki, W. (Red.). (2012). *Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Cz. 2: Statystyka matematyczna* (Wyd. 8). Warszawa: Wydawnictwo Naukowe PWN.
- Krzyśko, M. (2004). *Statystyka matematyczna* (wyd. 2 popr.). Poznań: Wydawnictwo Naukowe UAM.
- Küchler, U., Neumann, K., Sørensen, M. i Streller, A. (1999). Stock returns and hyperbolic distributions. *Mathematical and Computer Modelling*, 29(10), 1–15. [https://doi.org/10.1016/S0895-7177\(99\)00088-6](https://doi.org/10.1016/S0895-7177(99)00088-6)
- Kupiec, P. (1995). *Techniques for verifying the accuracy of risk measurement models* (SSRN Scholarly Paper Nr ID 6697). Pobrane z <https://papers.ssrn.com/abstract=6697>
- Lach, A. (2015). Zastosowanie kopuli w zadaniu minimalizacji wartości zagrożonej portfela. *Optimum. Studia Ekonomiczne*, nr 2(74).
- Lach, A. (2017). *Testy zgodności z rozkładami uciętymi* (Niepublikowana praca licencjacka). Poznań: Uniwersytet im. Adama Mickiewicza.
- Lach, A. i Sikora, W. (2010). Analiza efektywności sekwencyjnego inwestowania w akcje. *Zeszyty Naukowe / Uniwersytet Ekonomiczny w Poznaniu*, 138, 46–59.
- Lach, A. i Smaga, Ł. (2018). Comparison of the goodness-of-fit tests for truncated distributions. *Przegląd Statystyczny*, 65(3), 296–313.
- L'Ecuyer, P. (1999). Good parameters and implementations for combined multiple recursive random number generators. *Operations Research*, 47(1), 159.

- Lee, C.-F. i Lee, J. C. (2015). *Handbook of financial econometrics and statistics*. Pobrane z <https://doi.org/10.1007/978-1-4614-7750-1>
- Lehmann, E. i Romano, J. (2005). *Testing statistical hypotheses*. New York: Springer New York.
- Lindberg, C. (2009). Portfolio optimization when expected stock returns are determined by exposure to risk. *Bernoulli*, 15(2), 464–474. <https://doi.org/10.3150/08-BEJ163>
- Łuczak, A. i Just, M. (2020a). A complex MCDM procedure for the assessment of economic development of units at different government levels. *Mathematics*, 8(7), 1067. <https://doi.org/10.3390/math8071067>
- Łuczak, A. i Just, M. (2020b). The positional MEF-TOPSIS method for the assessment of complex economic phenomena in territorial units. *Statistics in Transition New Series*, 21(2). <https://doi.org/10.21307/stattrans-2020-018>
- Magiera, R. (2002). *Modele i metody statystyki matematycznej*. Wrocław: Oficyna Wydawnicza GiS.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, 36. Pobrane z https://econpapers.repec.org/article/ucpjnbus/v_3a36_3ay_3a1963_3ap_3a394.htm
- Marín, J. M. i Sucarrat, G. (2012). Modelling the skewed exponential power distribution in finance. W: Perna, C. i Sibillo, M. (Eds.), *Mathematical and statistical methods for actuarial sciences and finance* (s. 279–286). Milano: Springer Milan. https://doi.org/10.1007/978-88-470-2342-0_33
- Martinez, W. L. i Martinez, A. R. (2008). *Computational statistics handbook with MATLAB*. New York: Chapman and Hall/CRC.
- Nelsen, R. B. (1999). *An introduction to copulas*. New York: Springer.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2), 347–370. <https://doi.org/10.2307/2938260>
- Orzeszko, W. (2014). Symulacyjna ocena rozmiaru testu BDS. *Przegląd Statystyczny*, 61(4), 335–361.
- Pasieczna, A. H. i Szydłowska, M. J. (2021). Estimating Model Risk of VaR under Different Approaches: Study on European Banks. *Współczesne Problemy Zarządzania*, 9(2), 65–76. <https://doi.org/10.52934/wpz.151>

- Patton, A. (2013). Chapter 16: copula methods for forecasting multivariate time series. W: Elliott, G. i Timmermann, A. (Eds.), *Handbook of economic forecasting*, 2, 899–960. <https://doi.org/10.1016/B978-0-444-62731-5.00016-6>
- Pavia, J. M. (2015). Testing Goodness-of-Fit with the Kernel Density Estimator: GoFKernel. *Journal of Statistical Software*, 66(1), 1–27. <https://doi.org/10.18637/jss.v066.c01>
- Pearson, E. (1963). Comparison of tests for randomness of points on a line. *Biometrika*, 50(3/4), 315.
- Pettitt, A. N. i Stephens, M. A. (1976). Modified Cramér-von Mises statistics for censored data. *Biometrika*, 63(2), 291–298. <https://doi.org/10.2307/2335622>
- Pewsey, A. (2018). Parametric bootstrap edf-based goodness-of-fit testing for sinh–arcsinh distributions. *Test*, 27(1), 147–172. <https://doi.org/10.1007/s11749-017-0538-2>
- Prause, K. (1999). *The generalized hyperbolic model: estimation, financial derivatives, and risk measures* (Unpublished doctoral dissertation). Freiburg: Universität Freiburg. Pobrane z <https://d-nb.info/961152192/34>
- Qin, J. (2019). Probabilistic modeling of random variables with inconsistent data. *Applied Mathematical Modelling*, 73, 401–411. <https://doi.org/10.1016/j.apm.2019.04.017>
- Rodríguez, J. C. i Viollaz, A. J. (1995). A Cramér-von Mises type goodness of fit test with asymmetric weight function. *Communications in Statistics - Theory and Methods*, 24(4), 1095–1120. <https://doi.org/10.1080/03610929508831542>
- Rodríguez, J. C. i Viollaz, A. J. (1999). A weighted Cramér-von Mises statistic derived from a decomposition of the Anderson-Darling statistic. *Communications in Statistics - Theory and Methods*, 28(10), 2333–2346. <https://doi.org/10.1080/03610929908832424>
- Shapiro, S. S., Wilk, M. B. i Chen, H. J. (1968). A comparative study of various tests for normality. *Journal of the American Statistical Association*, 63(324), 1343. <https://doi.org/10.2307/2285889>
- Sinclair, C. D., Spurr, B. D. i Ahmad, M. I. (1987). *Modified Anderson-Darling test*. Scotland. Sobczyk, M. (2010). *Statystyka matematyczna*. Warszawa: Wydawnictwo C.H. Beck.
- Stephens, M. A. (1965). The goodness-of-fit statistic V_N : distribution and significance points. *Biometrika*, 52(3/4), 309–321. <https://doi.org/10.2307/2333685>
- Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69(347), 730–737. <https://doi.org/10.1080/01621459.1974.10480196>

- Stoyanov, S. V., Rachev, S., Racheva-Iotova, B. i Fabozzi, F. J. (2010). *Fat-tailed models for risk estimation* (SSRN Scholarly Paper Nr ID 1729040). Pobrane z <https://papers.ssrn.com/abstract=1729040>
- Stute, W., Manteiga, W. G. i Quindimil, M. P. (1993). Bootstrap based goodness-of-fit-tests. *Metrika*, 40(1), 243–256. <https://doi.org/10.1007/BF02613687>
- Szkutnik, T. (2016). Zastosowanie rozkładów uciętych i cenzurowanych w kwantyfikacji ryzyka operacyjnego. Badania symulacyjne. *Studia Ekonomiczne / Uniwersytet Ekonomiczny w Katowicach*, 264, 109–134.
- Thas, O. (2010). *Comparing distributions*. New York: Springer.
- Theodossiou, P. (2008). *Financial data and the skewed generalized t distribution* (SSRN Scholarly Paper Nr ID 65037). <https://doi.org/10.2139/ssrn.65037>
- Tollenaar, N. i Mooijaart, A. (2003). Type I errors and power of the parametric bootstrap goodness-of-fit test: Full and limited information. *British Journal of Mathematical & Statistical Psychology*, 56(2), 271–288. <https://doi.org/10.1348/000711003770480048>
- Trybuła, S. (2004). *Statystyka matematyczna z elementami teorii decyzji* (Wyd. 2 popr.). Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.
- Vilijandas, B., Kruopis, J. i Nikulin, M. S. (2011). *Non-parametric tests for complete data*. London: Wiley.
- Zhu, D., & Zinde-Walsh, V. (2009). Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics*, 148(1), 86–99. <https://doi.org/10.1016/j.jeconom.2008.09.038>
- Zieliński, R. i Zieliński, W. (1990). *Tablice statystyczne* (wyd. 2 zm.). Warszawa: Państwowe Wydawnictwo Naukowe.
- Zielonka, P. (2006). *Behawioralne aspekty inwestowania na rynku papierów wartościowych*. Warszawa: CeDeWu.

Spis tabel

Tabela 1.4.1. Statystyki dla rozkładów lewostronnie uciętych	31
Tabela 2.1.1. Przykłady zastosowania rozkładów <i>GH</i> , <i>Hyp</i> , <i>N</i> , <i>NIG</i> , <i>SGED</i> oraz <i>ST</i> w badaniach empirycznych	38
Tabela 3.1.1. Wzory oraz formuły obliczeniowe statystyk dla obustronnie uciętych rozkładów ciągłych	65
Tabela 3.3.1. Porównanie wariancji estymatora wartości p dla podstawowej i zmodyfikowanej metody bootstrapu	74
Tabela 3.4.1. Statystyki opisowe badanych szeregów	74
Tabela 3.4.2. Wzorce rozkładów rzeczywistych biorących udział w badaniu rozmiarów i mocy testów	76
Tabela 3.5.1. Udziały odrzuconych hipotez zerowych testu dla wskaźnika struktury, dla $\alpha =$ 0,05 (w %)	78
Tabela 3.5.2. Udziały odrzuconych hipotez zerowych testu dla wskaźnika struktury, w podziale na przekroczenia dolnej lub górnej granicy 95% przedziału ufności, statystyki testowe oraz fragment dystrybuanty, dla $\alpha = 0,05$ (w %)	79
Tabela 3.5.3. Udziały odrzuconych hipotez zerowych testu dla wskaźnika struktury, w podziale na przekroczenia dolnej lub górnej granicy 95% przedziału ufności, statystyki testowe oraz rozkłady prawdopodobieństwa, dla $\alpha = 0,05$ (w %)	79
Tabela 4.2.1. Oszacowania wartości zagrożonych dla rozkładów jednowymiarowych	106
Tabela 4.2.2. Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów jednowymiarowych	107
Tabela 4.3.1. Mediany rozkładów udziałów przekroczeń w całkowitej liczbie obserwacji (wyznaczone na podstawie modeli wartości zagrożonych wybranych w oparciu o różne testy zgodności dla rozkładów brzegowych)	116
Tabela 4.4.1. Wyniki testu Kupca dla modelu I (poziom tolerancji 0,01)	128
Tabela 4.4.2. Wyniki testu Kupca dla modelu I (poziom tolerancji 0,05)	128
Tabela 4.4.3. Wyniki testu Kupca dla modelu II (poziom tolerancji 0,01)	130
Tabela 4.4.4. Wyniki testu Kupca dla modelu III (poziom tolerancji 0,01)	131
Tabela 4.4.5. Maksymalne wartości p testu Kupca uzyskane dla modeli II i III w podziale na wybrane typy rozkładów	134
Tabela 4.5.1. Mediany rozkładów wybranych charakterystyk portfela	139

Spis rysunków

Rysunek 2.1.1. Funkcja gęstości rozkładu normalnego $N(\mu; \sigma^2)$	40
Rysunek 2.1.2. Funkcja gęstości rozkładu $GH(\mu = 0; \delta = 1; \alpha; \beta; \lambda)$ dla różnych wartości parametrów α, β, λ	43
Rysunek 2.1.3. Funkcja gęstości rozkładu $Hyp(\mu = 0; \delta = 1; \alpha; \beta)$ dla różnych wartości parametrów α i β	45
Rysunek 2.1.4. Funkcja gęstości rozkładu $NIG(\mu = 0; \delta = 1; \alpha; \beta)$ dla różnych wartości parametrów α i β	47
Rysunek 2.1.5. Funkcja gęstości rozkładu $SGED(\mu = 0; \delta = 1; \alpha; \kappa)$ dla różnych wartości parametrów α i κ	49
Rysunek 2.1.6. Funkcja gęstości rozkładu $ST(\xi = 0; \omega = 1; \alpha; \nu)$ dla różnych wartości parametrów α i ν	50
Rysunek 2.2.1. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Gaussa ($r = 0,5$)	56
Rysunek 2.2.2. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą t-Studenta ($r = 0,5, \nu = 3$)	57
Rysunek 2.2.3. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Franka ($\alpha = -5$)	59
Rysunek 2.2.4. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Claytona ($\alpha = 2$)	60
Rysunek 2.2.5. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Gumbela ($\alpha = 3$)	61
Rysunek 2.2.6. Wykres perspektywiczny i konturowy dla funkcji gęstości rozkładu dwuwymiarowego utworzonego z połączenia rozkładów brzegowych $N(0; 1)$ i $N(1; 1)$ kopulą Joe ($\alpha = 3$)	62
Rysunek 3.4.1. Notowania i stopy zwrotu dla indeksów WIG20 oraz FTSE100 w latach 2019-2020	75

Rysunek 3.6.1. Oszacowania mocy testów zgodności dla $H_0: F = N$ (ogon górny)	87
Rysunek 3.6.2. Oszacowania mocy testów zgodności dla $H_0: F = N$ (ogon dolny)	88
Rysunek 3.6.3. Oszacowania mocy testów zgodności dla $H_0: F = N$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład).....	89
Rysunek 3.6.4. Oszacowania mocy testów zgodności dla $H_0: F = NIG$ (ogon górny)	90
Rysunek 3.6.5. Oszacowania mocy testów zgodności dla $H_0: F = NIG$ (ogon dolny)	91
Rysunek 3.6.6. Oszacowania mocy testów zgodności dla $H_0: F = NIG$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład).....	92
Rysunek 3.6.7. Oszacowania mocy testów zgodności dla $H_0: F = Hyp$ (ogon górny).....	93
Rysunek 3.6.8. Oszacowania mocy testów zgodności dla $H_0: F = Hyp$ (ogon dolny)	94
Rysunek 3.6.9. Oszacowania mocy testów zgodności dla $H_0: F = Hyp$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład).....	95
Rysunek 3.6.10. Oszacowania mocy testów zgodności dla $H_0: F = SGED$ (ogon górny)	96
Rysunek 3.6.11. Oszacowania mocy testów zgodności dla $H_0: F = SGED$ (ogon dolny)	97
Rysunek 3.6.12. Oszacowania mocy testów zgodności dla $H_0: F = SGED$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład).....	98
Rysunek 3.6.13. Oszacowania mocy testów zgodności dla $H_0: F = ST$ (ogon górny).....	99
Rysunek 3.6.14. Oszacowania mocy testów zgodności dla $H_0: F = ST$ (ogon dolny).....	100
Rysunek 3.6.15. Oszacowania mocy testów zgodności dla $H_0: F = ST$ (część rozkładu po odrzuceniu ogonów oraz cały rozkład).....	101
Rysunek 4.2.1 Oszacowania wartości zagrożonej dla pozycji krótkiej (2020, WIG20).....	110
Rysunek 4.2.2 Oszacowania wartości zagrożonej dla pozycji krótkiej (2020, FTSE100)	111
Rysunek 4.3.1 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja długa, rok 2019).....	118
Rysunek 4.3.2 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja krótka, rok 2019)	119
Rysunek 4.3.3 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja długa, rok 2020).....	120
Rysunek 4.3.4 Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych (pozycja krótka, rok 2020)	121
Rysunek 4.3.5 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja długa, rok 2019).....	122

Rysunek 4.3.6 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja krótka, rok 2019)	123
Rysunek 4.3.7 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja długa, rok 2020).....	124
Rysunek 4.3.8 Udziały liczby przekroczeń w całkowitej liczbie obserwacji dla rozkładów dwuwymiarowych (pozycja krótka, rok 2020)	125
Rysunek 4.4.1 Oszacowania wartości zagrożonej na poziomie tolerancji 0,01 (<i>WIG20</i>).....	132
Rysunek 4.4.2 Oszacowania wartości zagrożonej na poziomie tolerancji 0,01 (<i>FTSE100</i>)	133

Aneks

Załącznik 1. Wyprowadzenie wzorów wybranych statystyk dla rozkładów obustronnie uciętych.....	160
Załącznik 2. Empiryczne rozmiary wybranych testów zgodności dla poziomu istotności $\alpha = 0,05$	180
Załącznik 3. Empiryczne moce wybranych testów zgodności dla poziomu istotności $\alpha = 0,05$	184
Załącznik 4. Oszacowania wartości p testów zgodności dla indeksów <i>WIG20</i> i <i>FTSE100</i> (poziom istotności $\alpha = 0,05$)	204
Załącznik 5. Oszacowania parametrów kopuli oraz rozkładów brzegowych	206
Załącznik 6. Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych	212

Załącznik 1. Wyprowadzenie wzorów wybranych statystyk dla rozkładów obustronnie uciętych

1. Oznaczenia i przekształcenia pomocnicze

W tej sekcji zostaną objaśnione użyte oznaczenia oraz zostaną przedstawione wybrane przekształcenia zastosowanych w obliczeniach całek i sum. Umieszczenie ich w tym podrozdziale ma na celu zwiększenie przejrzystości obliczeń w kolejnych podrozdziałach.

Tabela 1.1. Przyjęte oznaczenia

Oznaczenie	Opis oznaczenia
$\hat{F}_0(x)$	dystrybuanta teoretyczna całego rozkładu (z parametrami oszacowanymi na podstawie próby)
$H1, H2$	punkty ucięcia dystrybuanty teoretycznej \hat{F}_0
$\hat{F}_0^{**}(x) = \begin{cases} 0 & \text{dla } x \leq H1, \\ \frac{\hat{F}_0(x) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \lim_{x \rightarrow H1^-} \hat{F}_0(H1)} & \text{dla } H1 < x \leq H2, \\ 1 & \text{dla } x > H2. \end{cases}$	dystrybuanta teoretyczna rozkładu obustronnie uciętego
$\mathbf{X} = (X_1, X_2, \dots, X_n)'$	n -elementowa próba prosta
$\mathbf{x} = (x_1, x_2, \dots, x_n)'$	realizacja próby \mathbf{X}
$X_{(j)}$	j -ta statystyka pozycyjna
$x_{(j)}$	realizacja j -tej statystyki pozycyjnej
$F_n(x) = \frac{\#\{1 \leq j \leq n: X_j \leq x\}}{n}$	dystrybuanta empiryczna
$z = \hat{F}_0(x)$	wartość dystrybuanty teoretycznej \hat{F}_0 w punkcie x
$z_j = \hat{F}_0(x_{(j)})$	wartość dystrybuanty teoretycznej \hat{F}_0 w punkcie $x_{(j)}$
$z_{H1} = \hat{F}_0(H1)$	wartość dystrybuanty teoretycznej \hat{F}_0 w punkcie $H1$
$z_{H2} = \hat{F}_0(H2)$	wartość dystrybuanty teoretycznej \hat{F}_0 w punkcie $H2$
n	liczba elementów w próbie z rozkładu obustronnie uciętego
j	numer elementu w próbie, $j = 1, \dots, n$

Źródło: opracowanie własne

Podczas wyprowadzania formuł obliczeniowych wielokrotnie będą użyte następujące przekształcenia:

$$\frac{j}{n} - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} = \frac{1}{z_{H2} - z_{H1}} \left(\frac{j}{n} (z_{H2} - z_{H1}) - (z_j - z_{H1}) \right) = \frac{1}{z_{H2} - z_{H1}} \left(z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right) \quad (1)$$

$$\begin{aligned} \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} - \frac{j-1}{n} &= \frac{1}{z_{H2} - z_{H1}} \left(z_j - z_{H1} - \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \\ &= \frac{1}{z_{H2} - z_{H1}} \left(z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right) \end{aligned} \quad (2)$$

$$\frac{z - z_{H1}}{z_{H2} - z_{H1}} \left(1 - \frac{z - z_{H1}}{z_{H2} - z_{H1}} \right) = \frac{z - z_{H1}}{z_{H2} - z_{H1}} \left(\frac{z_{H2} - z_{H1} - z + z_{H1}}{z_{H2} - z_{H1}} \right) = \frac{(z - z_{H1})(z_{H2} - z)}{(z_{H2} - z_{H1})^2} \quad (3)$$

W trakcie wyprowadzania formuł w wersji kwadratowej zostaną wykorzystane rozwiązania poniższych całek nieoznaczonych (ich szczegółowe wyprowadzenie znajduje się pod listą):

$$I_1 = \int \frac{1}{(z - z_{H1})(z_{H2} - z)} dz = \frac{1}{z_{H2} - z_{H1}} (\ln(z - z_{H1}) - \ln(z_{H2} - z)) + c_1 \quad (4)$$

$$I_2 = \int \frac{1}{(z_{H2} - z)} dz = -\ln(z_{H2} - z) + c_2 \quad (5)$$

$$I_3 = \int \frac{(z - z_{H1})}{(z_{H2} - z)} dz = -z - (z_{H2} - z_{H1}) \ln(z_{H2} - z) + c_3 \quad (6)$$

$$I_4 = \int \frac{(z_{H2} - z)}{(z - z_{H1})} dz = -z + (z_{H2} - z_{H1}) \ln(z - z_{H1}) + c_4 \quad (7)$$

$$\begin{aligned} I_5 &= \int \frac{(z_{H1} - z)^2}{(z_{H2} - z)^2} dz = \frac{(z_{H1}^2 - z_{H2}^2)}{z_{H2} - z} + 2(z_{H2} - z_{H1}) \left(\ln(z_{H2} - z) + \frac{z_{H2}}{z_{H2} - z} \right) + z + c_5 \\ &= \frac{(z_{H1} - z_{H2})^2}{z_{H2} - z} + 2(z_{H2} - z_{H1}) \ln(z_{H2} - z) + z + c_5 \end{aligned} \quad (8)$$

$$I_6 = \int \frac{1}{(z_{H2} - z)^2} dz = \frac{1}{z_{H2} - z} + c_6 \quad (9)$$

$$I_7 = \int \frac{z}{(z_{H2} - z)^2} dz = \ln(z_{H2} - z) + \frac{z_{H2}}{z_{H2} - z} + c_7 \quad (10)$$

$$I_8 = \int \frac{(z_{H1} - z)}{(z_{H2} - z)^2} dz = -\ln(z_{H2} - z) + \frac{z_{H1} - z_{H2}}{z_{H2} - z} + c_8 \quad (11)$$

Całkę I_1 rozwiązuje się poprzez jej rozkład na ułamki proste:

$$\frac{M_{11}}{(z - z_{H1})} + \frac{M_{12}}{(z_{H2} - z)} = \frac{1}{(z - z_{H1})(z_{H2} - z)}$$

Po przyrównaniu współczynników stojących przy kolejnych wyrazach wielomianów otrzymujemy:

$$M_{11} = M_{12} = \frac{1}{z_{H2} - z_{H1}}$$

Korzystając z tego, że $\int \frac{1}{x-a} dx = \ln|x-a|$ oraz zauważając, że $(z - z_{H1}) > 0$ i $(z_{H2} - z) > 0$, całkę można obliczyć następująco:

$$I_1 = \int \frac{1}{\frac{z_{H2} - z_{H1}}{z - z_{H1}}} dz + \int \frac{1}{\frac{z_{H2} - z_{H1}}{z_{H2} - z}} dz = \frac{1}{z_{H2} - z_{H1}} (\ln(z - z_{H1}) - \ln(z_{H2} - z)) + c_1.$$

W rozwiązaniu całki I_2 skorzystano z tego, że $\int \frac{1}{a-x} dx = -\ln|a-x|$ oraz $(z_{H2} - z) > 0$.

Całkę I_3 można rozwiązać następująco:

$$I_3 = \int \frac{(z - z_{H1})}{(z_{H2} - z)} dz = \int \frac{-(z_{H2} - z) + z_{H2} - z_{H1}}{z_{H2} - z} dz = -\int dz + (z_{H2} - z_{H1}) \int \frac{dz}{z_{H2} - z} \\ = -z - (z_{H2} - z_{H1}) \ln(z_{H2} - z) + c_3.$$

Obliczanie całki I_4 przebiega podobnie do obliczania całki I_3 :

$$I_4 = \int \frac{(z_{H2} - z)}{(z - z_{H1})} dz = \int \frac{-(z - z_{H1}) - z_{H1} + z_{H2}}{(z - z_{H1})} dz = -\int dz + (z_{H2} - z_{H1}) \int \frac{dz}{z - z_{H1}} \\ = -z + (z_{H2} - z_{H1}) \ln(z - z_{H1}) + c_4.$$

Całkę I_5 można zapisać następująco:

$$I_5 = \int \frac{(z_{H1} - z)^2}{(z_{H2} - z)^2} dz = \int \frac{z_{H1}^2 - 2z_{H1}z + (z_{H2} - z)^2 - z_{H2}^2 + 2z_{H2}z}{(z_{H2} - z)^2} dz \\ = (z_{H1}^2 - z_{H2}^2) \int \frac{1}{(z_{H2} - z)^2} dz + 2(z_{H2} - z_{H1}) \int \frac{z}{(z_{H2} - z)^2} dz + \int dz.$$

Do wyliczenia całki I_5 potrzebne jest rozwiązanie całek $I_6 = \int \frac{1}{(z_{H2}-z)^2} dz$ oraz $I_7 = \int \frac{z}{(z_{H2}-z)^2} dz$. Do całki $I_6 = \int \frac{1}{(z_{H2}-z)^2} dz$ zastosujemy podstawienie $u = z_{H2} - z$, stąd $-du = dz$:

$$I_6 = -\int u^{-2} du = \frac{1}{u} + c_6 = \frac{1}{z_{H2} - z} + c_6.$$

Całkę $I_7 = \int \frac{z}{(z_{H2}-z)^2} dz$ rozwiążmy następująco:

$$I_7 = \int \frac{z}{(z_{H2} - z)^2} dz = \int \frac{-(z_{H2} - z) + z_{H2}}{(z_{H2} - z)^2} dz = -\int \frac{dz}{z_{H2} - z} + z_{H2} \int \frac{dz}{(z_{H2} - z)^2} dz.$$

Korzystając z rozwiązania całki I_2 oraz I_6 otrzymujemy:

$$I_7 = \ln(z_{H2} - z) + \frac{z_{H2}}{z_{H2} - z} + c_7.$$

Całka I_5 jest zatem równa:

$$\begin{aligned} I_5 &= \frac{(z_{H1}^2 - z_{H2}^2)}{z_{H2} - z} + 2(z_{H2} - z_{H1}) \left(\ln(z_{H2} - z) + \frac{z_{H2}}{z_{H2} - z} \right) + z + c_5 = \\ &= \frac{(z_{H1} - z_{H2})^2}{z_{H2} - z} + 2(z_{H2} - z_{H1}) \ln(z_{H2} - z) + z + c_5. \end{aligned}$$

Całkę I_8 można przedstawić następująco:

$$I_8 = \int \frac{(z_{H1} - z)}{(z_{H2} - z)^2} dz = \int \frac{(z_{H2} - z) + (z_{H1} - z_{H2})}{(z_{H2} - z)^2} dz = \int \frac{1}{(z_{H2} - z)} dz + (z_{H1} - z_{H2}) \int \frac{1}{(z_{H2} - z)^2} dz.$$

Korzystając z I_2 oraz I_6 otrzymujemy:

$$I_8 = -\ln(z_{H2} - z) + \frac{z_{H1} - z_{H2}}{z_{H2} - z} + c_8.$$

Podczas wyprowadzania formuł obliczeniowych wykorzystane zostaną następujące sumy (również wylistowane zgodnie z kolejnością ich użycia, pod listą przedstawiono ich wyprowadzenie):

$$\sum_{j=1}^{n-1} j^2 (\ln(z_{j+1} - z_{H1}) - \ln(z_j - z_{H1})) = \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) + n^2 * \ln(z_n - z_{H1}) \quad (12)$$

$$\sum_{j=1}^{n-1} j^2 (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) = \sum_{j=1}^n (1 - 2j) \ln(z_{H2} - z_j) + n^2 * \ln(z_{H2} - z_n) \quad (13)$$

$$\sum_{j=1}^{n-1} 2nj (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) = \sum_{j=1}^n (-2n) \ln(z_{H2} - z_j) + 2n^2 * \ln(z_{H2} - z_n) \quad (14)$$

$$\sum_{j=1}^{n-1} (z_{j+1} - z_j) = z_n - z_1 \quad (15)$$

$$\sum_{j=1}^{n-1} (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) = \ln(z_{H2} - z_n) - \ln(z_{H2} - z_1) \quad (16)$$

$$\sum_{j=1}^{n-1} j^2 (z_{j+1} - z_j) = \sum_{j=1}^n (1 - 2j) z_j + n^2 * z_n \quad (17)$$

$$\sum_{j=1}^{n-1} j (z_{j+1} - z_j) = - \sum_{j=1}^n z_j + n z_n \quad (18)$$

$$\sum_{j=1}^{n-1} j (z_{j+1}^2 - z_j^2) = - \sum_{j=1}^n z_j^2 + n z_n^2 \quad (19)$$

$$\sum_{j=1}^{n-1} (z_{j+1}^2 - z_j^2) = z_n^2 - z_1^2 \quad (20)$$

$$\sum_{j=1}^{n-1} (z_{j+1}^3 - z_j^3) = z_n^3 - z_1^3 \quad (21)$$

$$\sum_{j=1}^{n-1} (n-j)^2 \left(\frac{1}{z_{H2} - z_{j+1}} - \frac{1}{z_{H2} - z_j} \right) = \sum_{j=1}^n \frac{1 + 2(n-j)}{z_{H2} - z_j} - \frac{n^2}{z_{H2} - z_1} \quad (22)$$

$$\sum_{j=1}^{n-1} (n-j) (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) = \sum_{j=1}^n \ln(z_{H2} - z_j) - n * \ln(z_{H2} - z_1) \quad (23)$$

$$\sum_{j=1}^{n-1} j^2 \left(\frac{1}{z_{H1} - z_{j+1}} - \frac{1}{z_{H1} - z_j} \right) = \sum_{j=1}^n (1-2j) \frac{1}{z_{H1} - z_j} + n^2 * \frac{1}{z_{H1} - z_n} \quad (24)$$

$$\sum_{j=1}^{n-1} j (\ln(z_j - z_{H1}) - \ln(z_{j+1} - z_{H1})) = \sum_{j=1}^n \ln(z_j - z_{H1}) - n \ln(z_n - z_{H1}) \quad (25)$$

Szczegółowe wyprowadzenie (12):

$$\begin{aligned} \sum_{j=1}^{n-1} j^2 (\ln(z_{j+1} - z_{H1}) - \ln(z_j - z_{H1})) &= 1 * \ln(z_2 - z_{H1}) - 1 * \ln(z_1 - z_{H1}) + 4 * \ln(z_3 - z_{H1}) - 4 * \\ &\ln(z_2 - z_{H1}) + \dots + (n-2)^2 * \ln(z_{n-1} - z_{H1}) - (n-2)^2 * \ln(z_{n-2} - z_{H1}) + (n-1)^2 * \ln(z_n - z_{H1}) - \\ &(n-1)^2 * \ln(z_{n-1} - z_{H1}) = -1 * \ln(z_1 - z_{H1}) - 3 * \ln(z_2 - z_{H1}) + \dots + (1-2(n-2)) * \ln(z_{n-2} - z_{H1}) + \\ &(1-2(n-1)) * \ln(z_{n-1} - z_{H1}) + (n-1)^2 * \ln(z_n - z_{H1}) = \sum_{j=1}^{n-1} (1-2j) \ln(z_j - z_{H1}) + \\ &(1-2n) * \ln(z_n - z_{H1}) + n^2 * \ln(z_n - z_{H1}) = \sum_{j=1}^n (1-2j) \ln(z_j - z_{H1}) + n^2 * \ln(z_n - z_{H1}). \end{aligned}$$

Szczegółowe wyprowadzenie (13):

$$\begin{aligned} \sum_{j=1}^{n-1} j^2 (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) &= 1 * \ln(z_{H2} - z_2) - 1 * \ln(z_{H2} - z_1) + 4 * \ln(z_{H2} - z_3) - 4 * \\ &\ln(z_{H2} - z_2) + \dots + (n-2)^2 * \ln(z_{H2} - z_{n-1}) - (n-2)^2 * \ln(z_{H2} - z_{n-2}) + (n-1)^2 * \ln(z_{H2} - z_n) - \\ &(n-1)^2 * \ln(z_{H2} - z_{n-1}) = -1 * \ln(z_{H2} - z_1) - 3 * \ln(z_{H2} - z_2) + \dots + (1-2(n-2)) * \ln(z_{H2} - z_{n-2}) + \\ &(1-2(n-1)) * \ln(z_{H2} - z_{n-1}) + (n-1)^2 * \ln(z_{H2} - z_n) = \sum_{j=1}^{n-1} (1-2j) \ln(z_{H2} - z_j) + \\ &(1-2n) * \ln(z_{H2} - z_n) + n^2 * \ln(z_{H2} - z_n) = \sum_{j=1}^n (1-2j) \ln(z_{H2} - z_j) + n^2 * \ln(z_{H2} - z_n) \end{aligned}$$

Szczegółowe wyprowadzenie (14):

$$\begin{aligned} \sum_{j=1}^{n-1} 2nj (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) &= 2n * \ln(z_{H2} - z_2) - 2n * \ln(z_{H2} - z_1) + 4n * \ln(z_{H2} - z_3) - \\ &4n * \ln(z_{H2} - z_2) + \dots + 2n(n-2) * \ln(z_{H2} - z_{n-1}) - 2n(n-2) * \ln(z_{H2} - z_{n-2}) + 2n(n-1) * \\ &\ln(z_{H2} - z_n) - 2n(n-1) * \ln(z_{H2} - z_{n-1}) = -2n * \ln(z_{H2} - z_1) - 2n * \ln(z_{H2} - z_2) + \dots - \\ &2n * \ln(z_{H2} - z_{n-2}) - 2n * \ln(z_{H2} - z_{n-1}) + 2n(n-1) * \ln(z_{H2} - z_n) = \sum_{j=1}^{n-1} (-2n) \ln(z_{H2} - z_j) - \\ &2n * \ln(z_{H2} - z_n) + 2n^2 * \ln(z_{H2} - z_n) = \sum_{j=1}^n (-2n) \ln(z_{H2} - z_j) + 2n^2 * \ln(z_{H2} - z_n) \end{aligned}$$

Szczegółowe wyprowadzenie (15):

$$\sum_{j=1}^{n-1} (z_{j+1} - z_j) = z_2 - z_1 + z_3 - z_2 + \dots + z_{n-1} - z_{n-2} + z_n - z_{n-1} = z_n - z_1$$

Szczegółowe wyprowadzenie (16):

$$\begin{aligned} \sum_{j=1}^{n-1} (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) &= \ln(z_{H2} - z_2) - \ln(z_{H2} - z_1) + \ln(z_{H2} - z_3) - \ln(z_{H2} - z_2) + \dots + \\ &\ln(z_{H2} - z_{n-1}) - \ln(z_{H2} - z_{n-2}) + \ln(z_{H2} - z_n) - \ln(z_{H2} - z_{n-1}) = \ln(z_{H2} - z_n) - \ln(z_{H2} - z_1) \end{aligned}$$

Szczegółowe wyprowadzenie (17):

$$\begin{aligned} \sum_{j=1}^{n-1} j^2 (z_{j+1} - z_j) &= 1 * z_2 - 1 * z_1 + 4 * z_3 - 4 * z_2 + \dots + (n-2)^2 * z_{n-1} - (n-2)^2 * z_{n-2} + (n-1)^2 * \\ &z_n - (n-1)^2 * z_{n-1} = -1 * z_1 - 3 * z_2 - \dots + (1-2(n-2)) * z_{n-2} + (1-2(n-1)) * z_{n-1} + (n-1)^2 * \\ &z_n = \sum_{j=1}^{n-1} (1-2j) z_j + (1-2n) * z_n + n^2 * z_n = \sum_{j=1}^n (1-2j) z_j + n^2 * z_n \end{aligned}$$

Szczegółowe wyprowadzenie (18):

$$\sum_{j=1}^{n-1} j(z_{j+1} - z_j) = 1 * z_2 - 1 * z_1 + 2 * z_3 - 2 * z_2 + \dots + (n-2) * z_{n-1} - (n-2) * z_{n-2} + (n-1) * z_n - (n-1) * z_{n-1} = -z_1 - z_2 - \dots - z_{n-1} - z_n + nz_n = -\sum_{j=1}^n z_j + nz_n$$

Szczegółowe wyprowadzenie (19):

$$\sum_{j=1}^{n-1} j(z_{j+1}^2 - z_j^2) = 1 * z_2^2 - 1 * z_1^2 + 2 * z_3^2 - 2 * z_2^2 + \dots + (n-2) * z_{n-1}^2 - (n-2) * z_{n-2}^2 + (n-1) * z_n^2 - (n-1) * z_{n-1}^2 = -z_1^2 - z_2^2 - \dots - z_{n-1}^2 - z_n^2 + nz_n^2 = -\sum_{j=1}^n z_j^2 + nz_n^2$$

Szczegółowe wyprowadzenie (20):

$$\sum_{j=1}^{n-1} (z_{j+1}^2 - z_j^2) = z_2^2 - z_1^2 + z_3^2 - z_2^2 + \dots + z_{n-1}^2 - z_{n-2}^2 + z_n^2 - z_{n-1}^2 = z_n^2 - z_1^2$$

Szczegółowe wyprowadzenie (21):

$$\sum_{j=1}^{n-1} (z_{j+1}^3 - z_j^3) = z_2^3 - z_1^3 + z_3^3 - z_2^3 + \dots + z_{n-1}^3 - z_{n-2}^3 + z_n^3 - z_{n-1}^3 = z_n^3 - z_1^3$$

Szczegółowe wyprowadzenie (22):

$$\begin{aligned} \sum_{j=1}^{n-1} (n-j)^2 \left(\frac{1}{z_{H2-z_{j+1}}} - \frac{1}{z_{H2-z_j}} \right) &= \frac{(n-1)^2}{z_{H2-z_2}} - \frac{(n-1)^2}{z_{H2-z_1}} + \frac{(n-2)^2}{z_{H2-z_3}} - \frac{(n-2)^2}{z_{H2-z_2}} + \dots + \frac{(n-(n-2))^2}{z_{H2-z_{n-1}}} - \frac{(n-(n-2))^2}{z_{H2-z_{n-2}}} + \\ \frac{(n-(n-1))^2}{z_{H2-z_n}} - \frac{(n-(n-1))^2}{z_{H2-z_{n-1}}} &= -\frac{n^2}{z_{H2-z_1}} + \frac{1+2(n-1)}{z_{H2-z_1}} + \frac{1+2(n-2)}{z_{H2-z_2}} + \dots + \frac{1+2(n-(n-2))}{z_{H2-z_{n-2}}} + \frac{1+2(n-(n-1))}{z_{H2-z_{n-1}}} + \frac{1+2(n-n)}{z_{H2-z_n}} = \\ \sum_{j=1}^n \frac{1+2(n-j)}{z_{H2-z_j}} - \frac{n^2}{z_{H2-z_1}} \end{aligned}$$

Szczegółowe wyprowadzenie (23):

$$\begin{aligned} \sum_{j=1}^{n-1} (n-j) (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) &= (n-1) * \ln(z_{H2} - z_2) - (n-1) * \ln(z_{H2} - z_1) + (n-2) * \\ \ln(z_{H2} - z_3) - (n-2) * \ln(z_{H2} - z_2) + \dots + (n-(n-2)) * \ln(z_{H2} - z_{n-1}) - (n-(n-2)) * \\ \ln(z_{H2} - z_{n-2}) + (n-(n-1)) * \ln(z_{H2} - z_n) - (n-(n-1)) * \ln(z_{H2} - z_{n-1}) &= n * \ln(z_{H2} - z_2) - 1 * \\ \ln(z_{H2} - z_2) - n * \ln(z_{H2} - z_1) + 1 * \ln(z_{H2} - z_1) + n * \ln(z_{H2} - z_3) - 2 * \ln(z_{H2} - z_3) - n * \\ \ln(z_{H2} - z_2) + 2 * \ln(z_{H2} - z_2) + \dots + 2 * \ln(z_{H2} - z_{n-1}) - 2 * \ln(z_{H2} - z_{n-2}) + 1 * \ln(z_{H2} - z_n) - 1 * \\ \ln(z_{H2} - z_{n-1}) &= -n * \ln(z_{H2} - z_1) + 1 * \ln(z_{H2} - z_1) + 1 * \ln(z_{H2} - z_2) + \dots + 1 * \ln(z_{H2} - z_{n-2}) + 1 * \\ \ln(z_{H2} - z_{n-1}) + 1 * \ln(z_{H2} - z_n) &= \sum_{j=1}^n \ln(z_{H2} - z_j) - n * \ln(z_{H2} - z_1) \end{aligned}$$

Szczegółowe wyprowadzenie (24):

$$\begin{aligned} \sum_{j=1}^{n-1} j^2 \left(\frac{1}{z_{H1-z_{j+1}}} - \frac{1}{z_{H1-z_j}} \right) &= 1 * \frac{1}{z_{H1-z_2}} - 1 * \frac{1}{z_{H1-z_1}} + 4 * \frac{1}{z_{H1-z_3}} - 4 * \frac{1}{z_{H1-z_2}} + \dots + (n-2)^2 * \frac{1}{z_{H1-z_{n-1}}} - \\ (n-2)^2 * \frac{1}{z_{H1-z_{n-2}}} + (n-1)^2 * \frac{1}{z_{H1-z_n}} - (n-1)^2 * \frac{1}{z_{H1-z_{n-1}}} &= -1 * \frac{1}{z_{H1-z_1}} - 3 * \frac{1}{z_{H1-z_2}} - \dots + \\ (1-2(n-2)) * \frac{1}{z_{H1-z_{n-2}}} + (1-2(n-1)) * \frac{1}{z_{H1-z_{n-1}}} + (n-1)^2 * \frac{1}{z_{H1-z_n}} &= \sum_{j=1}^{n-1} (1-2j) \frac{1}{z_{H1-z_j}} + \\ (1-2n) * \frac{1}{z_{H1-z_n}} + n^2 * \frac{1}{z_{H1-z_n}} &= \sum_{j=1}^n (1-2j) \frac{1}{z_{H1-z_j}} + n^2 * \frac{1}{z_{H1-z_n}} \end{aligned}$$

Szczegółowe wyprowadzenie (25):

$$\begin{aligned} \sum_{j=1}^{n-1} j (\ln(z_j - z_{H1}) - \ln(z_{j+1} - z_{H1})) &= 1 * \ln(z_1 - z_{H1}) - 1 * \ln(z_2 - z_{H1}) + 2 * \ln(z_2 - z_{H1}) - 2 * \\ \ln(z_3 - z_{H1}) + \dots + (n-2) * \ln(z_{n-2} - z_{H1}) - (n-2) * \ln(z_{n-1} - z_{H1}) &+ (n-1) * \ln(z_{n-1} - z_{H1}) - \\ (n-1) * \ln(z_n - z_{H1}) &= \ln(z_1 - z_{H1}) + \ln(z_2 - z_{H1}) + \dots + \ln(z_{n-1} - z_{H1}) + \ln(z_n - z_{H1}) - \\ n \ln(z_n - z_{H1}) &= \sum_{j=1}^n \ln(z_j - z_{H1}) - n \ln(z_n - z_{H1}) \end{aligned}$$

2. Wyprowadzenie formuły obliczeniowej dla statystyki KS^{**}

Statystyka KS^{**} ma postać:

$$KS^{**} = \sqrt{n} \sup_x |F_n(x) - \hat{F}_0^{**}(x)|.$$

Formułę obliczeniową dla statystyki KS^{**} można wyprowadzić następująco (wykorzystano przekształcenia pomocnicze 1 i 2):

$$\begin{aligned} KS^{**} &= \sqrt{n} \max \left\{ \sup_j \{F_n(x_{(j)}) - \hat{F}_0^{**}(x_{(j)})\}, \sup_j \{\hat{F}_0^{**}(x_{(j)}) - F_n(x_{(j)})\} \right\} \\ &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{j}{n} - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \right\}, \sup_j \left\{ \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} - \frac{j-1}{n} \right\} \right\} \\ &= \frac{\sqrt{n}}{z_{H2} - z_{H1}} \max \left\{ \sup_j \left\{ z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right\}, \right. \\ &\quad \left. \sup_j \left\{ z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right\} \right\}. \end{aligned}$$

3. Wyprowadzenie formuły obliczeniowej dla statystyki V^{**}

Statystyka V^{**} ma postać:

$$V^{**} = \sqrt{n} \left(\sup_x \{F_n(x) - \hat{F}_0^{**}(x)\} + \sup_x \{\hat{F}_0^{**}(x) - F_n(x)\} \right).$$

Formułę obliczeniową dla statystyki V^{**} można wyprowadzić następująco (wykorzystano przekształcenia pomocnicze 1 i 2):

$$\begin{aligned} V^{**} &= \sqrt{n} \left(\sup_j \{F_n(x_{(j)}) - \hat{F}_0^{**}(x_{(j)})\} + \sup_j \{\hat{F}_0^{**}(x_{(j)}) - F_n(x_{(j)})\} \right) \\ &= \sqrt{n} \left(\sup_j \left\{ \frac{j}{n} - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \right\} + \sup_j \left\{ \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} - \frac{j-1}{n} \right\} \right) \\ &= \frac{\sqrt{n}}{z_{H2} - z_{H1}} \left(\sup_j \left\{ z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right\} + \sup_j \left\{ z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right\} \right). \end{aligned}$$

4. Wyprowadzenie formuły obliczeniowej dla statystyki AD^{**}

Statystyka AD^{**} ma postać:

$$AD^{**} = \sqrt{n} \sup_x \left| \frac{F_n(x) - \hat{F}_0^{**}(x)}{\sqrt{\hat{F}_0^{**}(x) (1 - \hat{F}_0^{**}(x))}} \right|.$$

Formułę obliczeniową dla statystyki AD^{**} można wyprowadzić następująco (wykorzystano przekształcenia pomocnicze 1 i 2):

$$\begin{aligned}
AD^{**} &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{F_n(x_{(j)}) - \hat{F}_0^{**}(x_{(j)})}{\sqrt{\hat{F}_0^{**}(x_{(j)})(1 - \hat{F}_0^{**}(x_{(j))}}} \right\}, \sup_j \left\{ \frac{\hat{F}_0^{**}(x_{(j)}) - F_n(x_{(j)})}{\sqrt{\hat{F}_0^{**}(x_{(j)})(1 - \hat{F}_0^{**}(x_{(j))}}} \right\} \right\} \\
&= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\frac{j}{n} - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}}}{\sqrt{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \left(1 - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}}\right)}} \right\}, \sup_j \left\{ \frac{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} - \frac{j-1}{n}}{\sqrt{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \left(1 - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}}\right)}} \right\} \right\} \\
&= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\frac{1}{z_{H2} - z_{H1}} \left(z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right)}{\sqrt{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \left(\frac{z_{H2} - z_j}{z_{H2} - z_{H1}} \right)}} \right\}, \right. \\
&\quad \left. \sup_j \left\{ \frac{\frac{1}{z_{H2} - z_{H1}} \left(z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right)}{\sqrt{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \left(\frac{z_{H2} - z_j}{z_{H2} - z_{H1}} \right)}} \right\} \right\} \\
&= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j}{\sqrt{(z_j - z_{H1})(z_{H2} - z_j)}} \right\}, \sup_j \left\{ \frac{z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right)}{\sqrt{(z_j - z_{H1})(z_{H2} - z_j)}} \right\} \right\}.
\end{aligned}$$

5. Wyprowadzenie formuły obliczeniowej dla statystyki AD_{up}^{**}

Statystyka AD_{up}^{**} ma postać:

$$AD_{up}^{**} = \sqrt{n} \sup_x \left| \frac{F_n(x) - \hat{F}_0^{**}(x)}{1 - \hat{F}_0^{**}(x)} \right|.$$

Formułę obliczeniową dla statystyki AD_{up}^{**} można wyprowadzić następująco (wykorzystano przekształcenia pomocnicze 1 i 2):

$$\begin{aligned}
AD_{up}^{**} &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{F_n(x_{(j)}) - \hat{F}_0^{**}(x_{(j)})}{1 - \hat{F}_0^{**}(x_{(j)})} \right\}, \sup_j \left\{ \frac{\hat{F}_0^{**}(x_{(j)}) - F_n(x_{(j)})}{1 - \hat{F}_0^{**}(x_{(j)})} \right\} \right\} \\
&= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\frac{j}{n} - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}}}{1 - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}}} \right\}, \sup_j \left\{ \frac{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} - \frac{j-1}{n}}{1 - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}}} \right\} \right\} \\
&= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\frac{1}{z_{H2} - z_{H1}} \left(z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right)}{\frac{z_{H2} - z_j}{z_{H2} - z_{H1}}} \right\}, \right. \\
&\quad \left. \sup_j \left\{ \frac{\frac{1}{z_{H2} - z_{H1}} \left(z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right)}{\frac{z_{H2} - z_j}{z_{H2} - z_{H1}}} \right\} \right\} = \\
&= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j}{z_{H2} - z_j} \right\}, \sup_j \left\{ \frac{z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right)}{z_{H2} - z_j} \right\} \right\}.
\end{aligned}$$

6. Wyprowadzenie formuły obliczeniowej dla statystyki AD_{down}^{**}

Statystyka AD_{down}^{**} ma postać:

$$AD_{down}^{**} = \sqrt{n} \sup_x \left| \frac{F_n(x) - \hat{F}_0^{**}(x)}{\hat{F}_0^{**}(x)} \right|.$$

Formułę obliczeniową dla statystyki AD_{down}^{**} można wyprowadzić następująco (wykorzystano przekształcenia pomocnicze 1 i 2):

$$\begin{aligned} AD_{down}^{**} &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{F_n(x_{(j)}) - \hat{F}_0^{**}(x_{(j)})}{\hat{F}_0^{**}(x_{(j)})} \right\}, \sup_j \left\{ \frac{\hat{F}_0^{**}(x_{(j)}) - F_n(x_{(j)})}{\hat{F}_0^{**}(x_{(j)})} \right\} \right\} \\ &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\left(\frac{j}{n} - \frac{z_j - z_{H1}}{z_{H2} - z_{H1}} \right)}{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}}} \right\}, \sup_j \left\{ \frac{\left(\frac{z_j - z_{H1}}{z_{H2} - z_{H1}} - \frac{j-1}{n} \right)}{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}}} \right\} \right\} = \\ &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\left(\frac{1}{z_{H2} - z_{H1}} \left(z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right) \right)}{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}}} \right\}, \right. \\ &\quad \left. \sup_j \left\{ \frac{\left(\frac{1}{z_{H2} - z_{H1}} \left(z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right) \right)}{\frac{z_j - z_{H1}}{z_{H2} - z_{H1}}} \right\} \right\} = \\ &= \sqrt{n} \max \left\{ \sup_j \left\{ \frac{\left(z_{H1} + \frac{j}{n} (z_{H2} - z_{H1}) - z_j \right)}{z_j - z_{H1}} \right\}, \sup_j \left\{ \frac{\left(z_j - \left(z_{H1} + \frac{j-1}{n} (z_{H2} - z_{H1}) \right) \right)}{z_j - z_{H1}} \right\} \right\}. \end{aligned}$$

7. Wyprowadzenie formuły obliczeniowej dla statystyki AD^{2**}

Statystyka AD^{2**} ma postać:

$$AD^{2**} = n \int_{H1}^{H2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{\hat{F}_0^{**}(x)(1 - \hat{F}_0^{**}(x))} d\hat{F}_0^{**}(x).$$

Podczas wyznaczania formuły obliczeniowej statystyki AD^{2**} wzorowano się metodą zastosowaną w Chernobai, Rachev i Fabozzi (2005, s. 16). Zastosujemy podstawienie $z^{**} = \hat{F}_0^{**}(x)$ i zauważmy, że $F_n(x) = F_n(z^{**})$. Jednocześnie zmienmy granice całkowania

$$\hat{F}_0^{**}(H1) = \frac{\hat{F}_0(H1) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 0 \text{ oraz } \hat{F}_0^{**}(H2) = \frac{\hat{F}_0(H2) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 1:$$

$$AD^{2**} = n \int_{H1}^{H2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{\hat{F}_0^{**}(x)(1 - \hat{F}_0^{**}(x))} d\hat{F}_0^{**}(x) = n \int_0^1 \frac{(F_n(z^{**}) - z^{**})^2}{z^{**}(1 - z^{**})} dz^{**}.$$

Następnym krokiem jest zamiana zmiennej z^{**} na z . Zgodnie z definicją rozkładu obustronnie uciętego $z^{**} = \frac{z-z_{H1}}{z_{H2}-z_{H1}}$, stąd $z = z^{**}(z_{H2} - z_{H1}) + z_{H1}$. Granice całkowania zmienią się na: $0 * (z_{H2} - z_{H1}) + z_{H1} = z_{H1}$ oraz $1 * (z_{H2} - z_{H1}) + z_{H1} = z_{H2}$:

$$AD^{2**} = n \int_{z_{H1}}^{z_{H2}} \frac{\left(F_n(z^{**}) - \frac{z - z_{H1}}{z_{H2} - z_{H1}}\right)^2}{\frac{z - z_{H1}}{z_{H2} - z_{H1}} \left(1 - \frac{z - z_{H1}}{z_{H2} - z_{H1}}\right)} d \frac{z - z_{H1}}{z_{H2} - z_{H1}}.$$

Korzystając z przekształcenia 3:

$$AD^{2**} = n \int_{z_{H1}}^{z_{H2}} \frac{\frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z_{H1})^2}}{\frac{(z - z_{H1})(z_{H2} - z)}{(z_{H2} - z_{H1})^2}} d \frac{z - z_{H1}}{z_{H2} - z_{H1}}.$$

Ostatecznie korzystając z tego, że $\left(\frac{z - z_{H1}}{z_{H2} - z_{H1}}\right)' = \frac{1}{z_{H2} - z_{H1}}$:

$$AD^{2**} = \frac{n}{z_{H2} - z_{H1}} \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz.$$

Stąd:

$$\frac{z_{H2} - z_{H1}}{n} AD^{2**} = \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz.$$

Prawą stronę ostatniej równości można przybliżyć poprzez wyliczenie sumy całek w przedziałach wyznaczonych punktami skoku dystrybuanty. Podział wynika z definicji dystrybuanty empirycznej:

$$\begin{aligned} \frac{z_{H2} - z_{H1}}{n} AD^{2**} &= \int_{z_{H1}}^{z_1} \frac{(0 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n} * (z_{H2} - z_{H1}) + z_{H1} - z\right)^2}{(z - z_{H1})(z_{H2} - z)} dz \\ &+ \int_{z_n}^{z_{H2}} \frac{(1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz. \end{aligned}$$

Wyliczmy teraz powyższe trzy całki składowe. Do wyznaczenia całki A użyto (6).

$$\begin{aligned} A &= \int_{z_{H1}}^{z_1} \frac{(0 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz \\ &= \int_{z_{H1}}^{z_1} \frac{(z - z_{H1})^2}{(z - z_{H1})(z_{H2} - z)} dz = \int_{z_{H1}}^{z_1} \frac{z - z_{H1}}{z_{H2} - z} dz = [-z - (z_{H2} - z_{H1}) \ln(z_{H2} - z)]_{z_{H1}}^{z_1} \\ &= z_{H1} - z_1 + (z_{H2} - z_{H1})(\ln(z_{H2} - z_{H1}) - \ln(z_{H2} - z_1)). \end{aligned}$$

$$\begin{aligned}
B &= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n}(z_{H2} - z_{H1}) + z_{H1} - z\right)^2}{(z - z_{H1})(z_{H2} - z)} dz = \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n}(z_{H2} - z_{H1}) - (z - z_{H1})\right)^2}{(z - z_{H1})(z_{H2} - z)} dz \\
&= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\frac{j^2}{n^2}(z_{H2} - z_{H1})^2 - 2\frac{j}{n}(z_{H2} - z_{H1})(z - z_{H1}) + (z - z_{H1})^2}{(z - z_{H1})(z_{H2} - z)} dz \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \int_{z_j}^{z_{j+1}} \frac{1}{(z - z_{H1})(z_{H2} - z)} dz - 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j \int_{z_j}^{z_{j+1}} \frac{1}{(z_{H2} - z)} dz \\
&\quad + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{(z - z_{H1})}{(z_{H2} - z)} dz.
\end{aligned}$$

Na podstawie (4), (5) i (6):

$$\begin{aligned}
B &= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \left[\frac{1}{z_{H2} - z_{H1}} (\ln(z - z_{H1}) - \ln(z_{H2} - z)) \right]_{z_j}^{z_{j+1}} - 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j [-\ln(z_{H2} - z)]_{z_j}^{z_{j+1}} \\
&\quad + \sum_{j=1}^{n-1} [-z - (z_{H2} - z_{H1}) \ln(z_{H2} - z)]_{z_j}^{z_{j+1}} \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 (\ln(z_{j+1} - z_{H1}) - \ln(z_{H2} - z_{j+1}) - \ln(z_j - z_{H1}) + \ln(z_{H2} - z_j)) \\
&\quad + \frac{2(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) \\
&\quad - \sum_{j=1}^{n-1} (z_{j+1} + (z_{H2} - z_{H1}) \ln(z_{H2} - z_{j+1}) - z_j - (z_{H2} - z_{H1}) \ln(z_{H2} - z_j)) \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 (\ln(z_{j+1} - z_{H1}) - \ln(z_j - z_{H1})) \\
&\quad - \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) \\
&\quad + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} 2nj (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) - \sum_{j=1}^{n-1} (z_{j+1} - z_j) \\
&\quad - (z_{H2} - z_{H1}) \sum_{j=1}^{n-1} (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)).
\end{aligned}$$

Następnie korzystając z sum (12)-(16):

$$\begin{aligned}
B &= \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) + (z_{H2} - z_{H1}) \ln(z_n - z_{H1}) \\
&\quad - \frac{(z_{H2} - z_{H1})}{n^2} \left(\sum_{j=1}^n (1 - 2j) \ln(z_{H2} - z_j) + n^2 * \ln(z_{H2} - z_n) + \sum_{j=1}^n (2n) \ln(z_{H2} - z_j) \right. \\
&\quad \left. - 2n^2 * \ln(z_{H2} - z_n) \right) + z_1 - z_n + (z_{H2} - z_{H1})(\ln(z_{H2} - z_1) - \ln(z_{H2} - z_n)) \\
&= \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) + (z_{H2} - z_{H1}) \ln(z_n - z_{H1}) \\
&\quad - \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 + 2(n - j)) \ln(z_{H2} - z_j) + z_1 - z_n + (z_{H2} - z_{H1}) \ln(z_{H2} - z_1).
\end{aligned}$$

Całka C zostanie wyznaczona korzystając z wyniku (7):

$$\begin{aligned}
C &= \int_{z_n}^{z_{H2}} \frac{(1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz = \int_{z_n}^{z_{H2}} \frac{(z_{H2} - z)^2}{(z - z_{H1})(z_{H2} - z)} dz = \int_{z_n}^{z_{H2}} \frac{(z_{H2} - z)}{(z - z_{H1})} dz = \\
&= [-z]_{z_n}^{z_{H2}} + (z_{H2} - z_{H1}) [\ln(z - z_{H1})]_{z_n}^{z_{H2}} \\
&= z_n - z_{H2} + (z_{H2} - z_{H1})(\ln(z_{H2} - z_{H1}) - \ln(z_n - z_{H1})).
\end{aligned}$$

Ostatecznie statystyka AD^{2**} jest sumą całek A, B i C przemnożonych przez stałą:

$$\begin{aligned}
AD^{2**} &= \frac{n}{z_{H2} - z_{H1}} \left(z_{H1} - z_1 + (z_{H2} - z_{H1})(\ln(z_{H2} - z_{H1}) - \ln(z_{H2} - z_1)) \right. \\
&\quad + \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) + (z_{H2} - z_{H1}) \ln(z_n - z_{H1}) \\
&\quad - \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 + 2(n - j)) \ln(z_{H2} - z_j) + z_1 - z_n + (z_{H2} - z_{H1}) \ln(z_{H2} - z_1) + z_n \\
&\quad \left. - z_{H2} + (z_{H2} - z_{H1})(\ln(z_{H2} - z_{H1}) - \ln(z_n - z_{H1})) \right) \\
&= \frac{n}{z_{H2} - z_{H1}} \left((z_{H1} - z_{H2}) + 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_{H1}) \right. \\
&\quad \left. + \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) - \frac{(z_{H2} - z_{H1})}{n^2} \sum_{j=1}^n (1 + 2(n - j)) \ln(z_{H2} - z_j) \right) \\
&= -n + 2n \ln(z_{H2} - z_{H1}) + \frac{1}{n} \sum_{j=1}^n (1 - 2j) \ln(z_j - z_{H1}) - \frac{1}{n} \sum_{j=1}^n (1 + 2(n - j)) \ln(z_{H2} - z_j).
\end{aligned}$$

8. Wyprowadzenie formuły obliczeniowej dla statystyki W^{2**}

Statystyka W^{2**} ma postać:

$$W^{2**} = n \int_{H1}^{H2} (F_n(x) - \hat{F}_0^{**}(x))^2 d\hat{F}_0^{**}(x).$$

Podczas wyznaczania formuły obliczeniowej wzorowano się metodą zastosowaną w Chernobai i in. (2005, s. 16).

Zastosujmy podstawienie $z^{**} = \hat{F}_0^{**}(x)$ i zauważmy, że $F_n(x) = F_n(z^{**})$. Jednocześnie zmieńmy granice całkowania: $\hat{F}_0^{**}(H1) = \frac{\hat{F}_0(H1) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 0$ oraz $\hat{F}_0^{**}(H2) = \frac{\hat{F}_0(H2) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 1$:

$$W^{2**} = n \int_{H1}^{H2} (F_n(x) - \hat{F}_0^{**}(x))^2 d\hat{F}_0^{**}(x) = n \int_0^1 (F_n(z^{**}) - z^{**})^2 dz^{**}.$$

Następnym krokiem jest zamiana zmiennej z^{**} na z . Zgodnie z definicją rozkładu obustronnie uciętego $z^{**} = \frac{z - z_{H1}}{z_{H2} - z_{H1}}$, stąd $z = z^{**}(z_{H2} - z_{H1}) + z_{H1}$. Granice całkowania zmienią się na: $0 * (z_{H2} - z_{H1}) + z_{H1} = z_{H1}$ oraz $1 * (z_{H2} - z_{H1}) + z_{H1} = z_{H2}$.

$$W^{2**} = n \int_{z_{H1}}^{z_{H2}} \left(F_n(z^{**}) - \frac{z - z_{H1}}{z_{H2} - z_{H1}} \right)^2 d \frac{z - z_{H1}}{z_{H2} - z_{H1}} = n \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z_{H1})^2} d \frac{z - z_{H1}}{z_{H2} - z_{H1}}.$$

Ostatecznie korzystając z tego, że $\left(\frac{z - z_{H1}}{z_{H2} - z_{H1}} \right)' = \frac{1}{z_{H2} - z_{H1}}$:

$$W^{2**} = \frac{n}{z_{H2} - z_{H1}} \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z_{H1})^2} dz.$$

Stąd:

$$\frac{(z_{H2} - z_{H1})^3}{n} W^{2**} = \int_{z_{H1}}^{z_{H2}} (F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2 dz.$$

Prawą stroną ostatniej równości można przybliżyć poprzez wyliczenie sumy całek w przedziałach wyznaczonych punktami skoku dystrybuanty. Podział wynika z definicji dystrybuanty empirycznej.

$$\begin{aligned} \frac{(z_{H2} - z_{H1})^3}{n} W^{2**} &= \int_{z_{H1}}^{z_1} (0 * (z_{H2} - z_{H1}) + z_{H1} - z)^2 dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \left(\frac{j}{n} * (z_{H2} - z_{H1}) + z_{H1} - z \right)^2 dz \\ &+ \int_{z_n}^{z_{H2}} (1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2 dz. \end{aligned}$$

Wyliczmy teraz powyższe trzy całki składowe.

$$\begin{aligned}
A &= \int_{z_{H1}}^{z_1} (0 * (z_{H2} - z_{H1}) + z_{H1} - z)^2 dz \\
&= \int_{z_{H1}}^{z_1} (z_{H1} - z)^2 dz = \int_{z_{H1}}^{z_1} (z_{H1}^2 - 2z_{H1}z + z^2) dz = z_{H1}^2 \int_{z_{H1}}^{z_1} dz - 2z_{H1} \int_{z_{H1}}^{z_1} z dz + \int_{z_{H1}}^{z_1} z^2 dz \\
&= z_{H1}^2 [z]_{z_{H1}}^{z_1} - 2z_{H1} \left[\frac{1}{2} z^2 \right]_{z_{H1}}^{z_1} + \left[\frac{1}{3} z^3 \right]_{z_{H1}}^{z_1} = -\frac{z_{H1}^3}{3} + z_{H1}^2 z_1 - z_{H1} z_1^2 + \frac{z_1^3}{3}.
\end{aligned}$$

$$\begin{aligned}
B &= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \left(\frac{j}{n} * (z_{H2} - z_{H1}) + (z_{H1} - z) \right)^2 dz \\
&= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \left(\frac{j^2}{n^2} (z_{H2} - z_{H1})^2 \right) dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \left(2 \frac{j}{n} (z_{H2} - z_{H1})(z_{H1} - z) \right) dz \\
&\quad + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} (z_{H1} - z)^2 dz \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \int_{z_j}^{z_{j+1}} dz + \frac{2(z_{H2} - z_{H1})z_{H1}}{n} \sum_{j=1}^{n-1} j \int_{z_j}^{z_{j+1}} dz - \frac{2(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j \int_{z_j}^{z_{j+1}} z dz \\
&\quad + z_{H1}^2 \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} dz - 2z_{H1} \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} z dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} z^2 dz \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 [z]_{z_j}^{z_{j+1}} + \frac{2(z_{H2} - z_{H1})z_{H1}}{n} \sum_{j=1}^{n-1} j [z]_{z_j}^{z_{j+1}} - \frac{2(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j \left[\frac{1}{2} z^2 \right]_{z_j}^{z_{j+1}} \\
&\quad + z_{H1}^2 \sum_{j=1}^{n-1} [z]_{z_j}^{z_{j+1}} - 2z_{H1} \sum_{j=1}^{n-1} \left[\frac{1}{2} z^2 \right]_{z_j}^{z_{j+1}} + \sum_{j=1}^{n-1} \left[\frac{1}{3} z^3 \right]_{z_j}^{z_{j+1}} \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 (z_{j+1} - z_j) + \frac{2(z_{H2} - z_{H1})z_{H1}}{n} \sum_{j=1}^{n-1} j (z_{j+1} - z_j) \\
&\quad - \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j (z_{j+1}^2 - z_j^2) + z_{H1}^2 \sum_{j=1}^{n-1} (z_{j+1} - z_j) - z_{H1} \sum_{j=1}^{n-1} (z_{j+1}^2 - z_j^2) \\
&\quad + \frac{1}{3} \sum_{j=1}^{n-1} (z_{j+1}^3 - z_j^3).
\end{aligned}$$

Następnie wykorzystując sumy (15), (17)-(21) otrzymujemy:

$$\begin{aligned}
B &= \frac{(z_{H2} - z_{H1})^2}{n^2} \left(\sum_{j=1}^n (1 - 2j)z_j + n^2 z_n \right) - \frac{2(z_{H2} - z_{H1})z_{H1}}{n} \left(\sum_{j=1}^n z_j - n z_n \right) + \frac{(z_{H2} - z_{H1})}{n} \left(\sum_{j=1}^n z_j^2 - n z_n^2 \right) \\
&\quad + z_{H1}^2 (z_n - z_1) - z_{H1} (z_n^2 - z_1^2) + \frac{1}{3} (z_n^3 - z_1^3) = \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^n (1 - 2j)z_j - \frac{2(z_{H2} - z_{H1})z_{H1}}{n} \sum_{j=1}^n z_j + \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n z_j^2 + z_{H1}^2 z_n - z_{H2} z_n^2 - z_{H1}^2 z_1 \\
&\quad + z_{H1} z_1^2 + \frac{z_n^3}{3} - \frac{z_1^3}{3}.
\end{aligned}$$

$$\begin{aligned}
C &= \int_{z_n}^{z_{H2}} (1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2 dz = \int_{z_n}^{z_{H2}} (z_{H2} - z)^2 dz \\
&= \int_{z_n}^{z_{H2}} (z_{H2}^2 - 2z_{H2}z + z^2) dz = z_{H2}^2 [z]_{z_n}^{z_{H2}} - 2z_{H2} \left[\frac{1}{2} z^2 \right]_{z_n}^{z_{H2}} + \left[\frac{1}{3} z^3 \right]_{z_n}^{z_{H2}} \\
&= \frac{z_{H2}^3}{3} - z_{H2}^2 z_n + z_{H2} z_n^2 - \frac{z_n^3}{3}.
\end{aligned}$$

Ostatecznie statystyka jest sumą całek A , B i C przemnożonych przez stałą:

$$\begin{aligned}
W^{2**} &= \frac{n}{(z_{H2} - z_{H1})^3} \left(-\frac{z_{H1}^3}{3} + z_{H1}^2 z_1 - z_{H1} z_1^2 + \frac{z_1^3}{3} + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^n (1 - 2j) z_j - \frac{2(z_{H2} - z_{H1})z_{H1}}{n} \sum_{j=1}^n z_j \right. \\
&\quad \left. + \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n z_j^2 + z_{H2}^2 z_n - z_{H2} z_n^2 - z_{H1}^2 z_1 + z_{H1} z_1^2 + \frac{z_n^3}{3} - \frac{z_1^3}{3} + \frac{z_{H2}^3}{3} - z_{H2}^2 z_n + z_{H2} z_n^2 \right. \\
&\quad \left. - \frac{z_n^3}{3} \right) \\
&= \frac{n(z_{H2}^3 - z_{H1}^3)}{3(z_{H2} - z_{H1})^3} + \frac{1}{n(z_{H2} - z_{H1})} \sum_{j=1}^n (1 - 2j) z_j - \frac{2z_{H1}}{(z_{H2} - z_{H1})^2} \sum_{j=1}^n z_j + \frac{1}{(z_{H2} - z_{H1})^2} \sum_{j=1}^n z_j^2 \\
&= \frac{n(z_{H2}^2 + z_{H1}z_{H2} + z_{H1}^2)}{3(z_{H2} - z_{H1})^2} + \frac{1}{n(z_{H2} - z_{H1})} \sum_{j=1}^n (1 - 2j) z_j \\
&\quad + \frac{1}{(z_{H2} - z_{H1})^2} \left(\sum_{j=1}^n z_j^2 - \sum_{j=1}^n 2z_{H1}z_j \right) \\
&= \frac{n}{3} + \frac{nz_{H1}z_{H2}}{(z_{H2} - z_{H1})^2} + \frac{1}{n(z_{H2} - z_{H1})} \sum_{j=1}^n (1 - 2j) z_j + \frac{1}{(z_{H2} - z_{H1})^2} \left(\sum_{j=1}^n (z_j - z_{H1})^2 \right) \\
&\quad - \frac{nz_{H1}^2}{(z_{H2} - z_{H1})^2} \\
&= \frac{n}{3} + \frac{nz_{H1}}{(z_{H2} - z_{H1})} + \frac{1}{n(z_{H2} - z_{H1})} \sum_{j=1}^n (1 - 2j) z_j + \frac{1}{(z_{H2} - z_{H1})^2} \left(\sum_{j=1}^n (z_j - z_{H1})^2 \right).
\end{aligned}$$

9. Wyprowadzenie formuły obliczeniowej dla statystyki AD_{up}^{2**}

Statystyka AD_{up}^{2**} ma postać:

$$AD_{up}^{2**} = n \int_{H1}^{H2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{(1 - \hat{F}_0^{**}(x))^2} d\hat{F}_0^{**}(x).$$

Podczas wyznaczania formuły obliczeniowej wzorowano się metodą zastosowaną w Chernobai i in. (2005, s. 16).

Zastosujmy podstawienie $z^{**} = \hat{F}_0^{**}(x)$ i zauważmy, że $F_n(x) = F_n(z^{**})$. Jednocześnie zmieńmy granice całkowania: $\hat{F}_0^{**}(H1) = \frac{\hat{F}_0(H1) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 0$ oraz $\hat{F}_0^{**}(H2) = \frac{\hat{F}_0(H2) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 1$:

$$AD_{up}^{2**} = n \int_{H_1}^{H_2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{(1 - \hat{F}_0^{**}(x))^2} d\hat{F}_0^{**}(x) = n \int_0^1 \frac{(F_n(z^{**}) - z^{**})^2}{(1 - z^{**})^2} dz^{**}.$$

Następnym krokiem jest zamiana zmiennej z^{**} na z . Zgodnie z definicją rozkładu obustronnie uciętego $z^{**} = \frac{z - z_{H1}}{z_{H2} - z_{H1}}$, stąd $z = z^{**}(z_{H2} - z_{H1}) + z_{H1}$. Granice całkowania zmienią się na: $0 * (z_{H2} - z_{H1}) + z_{H1} = z_{H1}$ oraz $1 * (z_{H2} - z_{H1}) + z_{H1} = z_{H2}$.

$$AD_{up}^{2**} = n \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**}) - \frac{z - z_{H1}}{z_{H2} - z_{H1}})^2}{(1 - \frac{z - z_{H1}}{z_{H2} - z_{H1}})^2} d \frac{z - z_{H1}}{z_{H2} - z_{H1}} = n \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z_{H1})^2} \frac{dz - z_{H1}}{(z_{H2} - z_{H1})^2}.$$

Ostatecznie korzystając z tego, że $(\frac{z - z_{H1}}{z_{H2} - z_{H1}})' = \frac{1}{z_{H2} - z_{H1}}$:

$$AD_{up}^{2**} = \frac{n}{z_{H2} - z_{H1}} \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz.$$

Stąd:

$$\frac{z_{H2} - z_{H1}}{n} AD_{up}^{2**} = \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz.$$

Prawą stroną ostatniej równości można przybliżyć poprzez wyliczenie sumy całek w przedziałach wyznaczonych punktami skoku dystrybuanty. Podział wynika z definicji dystrybuanty empirycznej.

$$\begin{aligned} \frac{z_{H2} - z_{H1}}{n} AD_{up}^{2**} &= \int_{z_{H1}}^{z_1} \frac{(0 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{(\frac{j}{n} * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz \\ &\quad + \int_{z_n}^{z_{H2}} \frac{(1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz. \end{aligned}$$

Wyliczmy teraz powyższe trzy całki składowe. Całka A zostanie obliczona wykorzystując 8:

$$\begin{aligned} A &= \int_{z_{H1}}^{z_1} \frac{(0 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz \\ &= \int_{z_{H1}}^{z_1} \frac{(z_{H1} - z)^2}{(z_{H2} - z)^2} dz = (z_{H1}^2 - z_{H2}^2) \left[\frac{1}{z_{H2} - z} \right]_{z_{H1}}^{z_1} + 2(z_{H2} - z_{H1}) \left[\frac{z_{H2}}{z_{H2} - z} + \ln(z_{H2} - z) \right]_{z_{H1}}^{z_1} \\ &\quad + [z]_{z_{H1}}^{z_1} \\ &= \frac{(z_{H1}^2 - z_{H2}^2)}{z_{H2} - z_1} - \frac{(z_{H1}^2 - z_{H2}^2)}{z_{H2} - z_{H1}} + \frac{2(z_{H2} - z_{H1})z_{H2}}{z_{H2} - z_1} - \frac{2(z_{H2} - z_{H1})z_{H2}}{z_{H2} - z_{H1}} \\ &\quad + 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_1) - 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_{H1}) + z_1 - z_{H1} \\ &= z_1 - z_{H1} - (z_{H2} - z_{H1}) + \frac{(z_{H2} - z_{H1})^2}{(z_{H2} - z_1)} - 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_{H1}) \\ &\quad + 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_1). \end{aligned}$$

$$\begin{aligned}
B &= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n} * (z_{H2} - z_{H1}) + (z_{H1} - z)\right)^2}{(z_{H2} - z)^2} dz \\
&= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j^2}{n^2} (z_{H2} - z_{H1})^2\right)}{(z_{H2} - z)^2} dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(2 \frac{j}{n} (z_{H2} - z_{H1})(z_{H1} - z)\right)}{(z_{H2} - z)^2} dz \\
&\quad + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{(z_{H1} - z)^2}{(z_{H2} - z)^2} dz \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \int_{z_j}^{z_{j+1}} \frac{1}{(z_{H2} - z)^2} dz + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j \int_{z_j}^{z_{j+1}} \frac{(z_{H1} - z)}{(z_{H2} - z)^2} dz \\
&\quad + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{(z_{H1} - z)^2}{(z_{H2} - z)^2} dz.
\end{aligned}$$

Korzystając kolejno z (9), (11) oraz (8):

$$\begin{aligned}
B &= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \left[\frac{1}{z_{H2} - z} \right]_{z_j}^{z_{j+1}} + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j \left[-\ln(z_{H2} - z) + \frac{(z_{H1} - z_{H2})}{(z_{H2} - z)} \right]_{z_j}^{z_{j+1}} + \sum_{j=1}^{n-1} [z]_{z_j}^{z_{j+1}} \\
&\quad + 2(z_{H2} - z_{H1}) \sum_{j=1}^{n-1} [\ln(z_{H2} - z)]_{z_j}^{z_{j+1}} + (z_{H1} - z_{H2})^2 \sum_{j=1}^{n-1} \left[\frac{1}{z_{H2} - z} \right]_{z_j}^{z_{j+1}} \\
&= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \left(\frac{1}{z_{H2} - z_{j+1}} - \frac{1}{z_{H2} - z_j} \right) \\
&\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j (-\ln(z_{H2} - z_{j+1}) + \ln(z_{H2} - z_j)) \\
&\quad - \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} 2nj \left(\frac{1}{(z_{H2} - z_{j+1})} - \frac{1}{(z_{H2} - z_j)} \right) + \sum_{j=1}^{n-1} (z_{j+1} - z_j) \\
&\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} n (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)) \\
&\quad + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} n^2 \left(\frac{1}{z_{H2} - z_{j+1}} - \frac{1}{z_{H2} - z_j} \right) \\
&= \sum_{j=1}^{n-1} (z_{j+1} - z_j) + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} (n-j)^2 \left(\frac{1}{z_{H2} - z_{j+1}} - \frac{1}{z_{H2} - z_j} \right) \\
&\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} (n-j) (\ln(z_{H2} - z_{j+1}) - \ln(z_{H2} - z_j)).
\end{aligned}$$

Następnie korzystamy z sum (15), (22) i (23):

$$\begin{aligned}
 B &= z_n - z_1 + \frac{(z_{H2} - z_{H1})^2}{n^2} \left(-\frac{n^2}{z_{H2} - z_1} + \sum_{j=1}^n (1 + 2(n-j)) \frac{1}{z_{H2} - z_j} \right) \\
 &\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \left(-n \ln(z_{H2} - z_1) + \sum_{j=1}^n \ln(z_{H2} - z_j) \right) \\
 &= z_n - z_1 - \frac{(z_{H2} - z_{H1})^2}{z_{H2} - z_1} + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^n (1 + 2(n-j)) \frac{1}{z_{H2} - z_j} \\
 &\quad - 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_1) + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n \ln(z_{H2} - z_j).
 \end{aligned}$$

$$C = \int_{z_n}^{z_{H2}} \frac{(1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z_{H2} - z)^2} dz = \int_{z_n}^{z_{H2}} dz = z_{H2} - z_n.$$

Ostatecznie statystyka jest sumą całek A , B i C przemnożonych przez stałą:

$$\begin{aligned}
 AD_{up}^{2**} &= \frac{n}{z_{H2} - z_{H1}} \left(z_1 - z_{H1} - (z_{H2} - z_{H1}) + \frac{(z_{H2} - z_{H1})^2}{(z_{H2} - z_1)} - 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_{H1}) \right. \\
 &\quad + 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_1) + z_n - z_1 - \frac{(z_{H2} - z_{H1})^2}{z_{H2} - z_1} \\
 &\quad + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^n (1 + 2(n-j)) \frac{1}{z_{H2} - z_j} - 2(z_{H2} - z_{H1}) \ln(z_{H2} - z_1) \\
 &\quad \left. + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n \ln(z_{H2} - z_j) + z_{H2} - z_n \right) \\
 &= -2n \ln(z_{H2} - z_{H1}) + 2 \sum_{j=1}^n \ln(z_{H2} - z_j) + \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n (1 + 2(n-j)) \frac{1}{z_{H2} - z_j}.
 \end{aligned}$$

10. Wyprowadzenie formuły obliczeniowej dla statystyki AD_{down}^{2**}

Statystyka AD_{down}^{2**} ma postać:

$$AD_{down}^{2**} = n \int_{H1}^{H2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{(\hat{F}_0^{**}(x))^2} d\hat{F}_0^{**}(x).$$

Podczas wyznaczania formuły obliczeniowej wzorowano się metodą zastosowaną w (Chernobai i in., 2005, s. 16).

Zastosujmy podstawienie $z^{**} = \hat{F}_0^{**}(x)$ i zauważmy, że $F_n(x) = F_n(z^{**})$. Jednocześnie zmieńmy granice całkowania: $\hat{F}_0^{**}(H1) = \frac{\hat{F}_0(H1) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 0$ oraz $\hat{F}_0^{**}(H2) = \frac{\hat{F}_0(H2) - \hat{F}_0(H1)}{\hat{F}_0(H2) - \hat{F}_0(H1)} = 1$:

$$AD_{down}^{2**} = n \int_{H1}^{H2} \frac{(F_n(x) - \hat{F}_0^{**}(x))^2}{(\hat{F}_0^{**}(x))^2} d\hat{F}_0^{**}(x) = n \int_0^1 \frac{(F_n(z^{**}) - z^{**})^2}{(z^{**})^2} dz^{**}.$$

Następnym krokiem jest zamiana zmiennej z^{**} na z . Zgodnie z definicją rozkładu obustronnie uciętego $z^{**} = \frac{z-z_{H1}}{z_{H2}-z_{H1}}$, stąd $z = z^{**}(z_{H2} - z_{H1}) + z_{H1}$. Granice całkowania zmieniają się na: $0 * (z_{H2} - z_{H1}) + z_{H1} = z_{H1}$ oraz $1 * (z_{H2} - z_{H1}) + z_{H1} = z_{H2}$.

$$AD_{down}^{2**} = n \int_{z_{H1}}^{z_{H2}} \frac{\left(F_n(z^{**}) - \frac{z-z_{H1}}{z_{H2}-z_{H1}}\right)^2}{\left(\frac{z-z_{H1}}{z_{H2}-z_{H1}}\right)^2} d \frac{z-z_{H1}}{z_{H2}-z_{H1}} = n \int_{z_{H1}}^{z_{H2}} \frac{\frac{(F_n(z^{**})(z_{H2}-z_{H1}) + z_{H1} - z)^2}{(z_{H2}-z_{H1})^2}}{\frac{(z-z_{H1})^2}{(z_{H2}-z_{H1})^2}} d \frac{z-z_{H1}}{z_{H2}-z_{H1}}.$$

Ostatecznie korzystając z tego, że $\left(\frac{z-z_{H1}}{z_{H2}-z_{H1}}\right)' = \frac{1}{z_{H2}-z_{H1}}$:

$$AD_{down}^{2**} = \frac{n}{z_{H2}-z_{H1}} \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2}-z_{H1}) + z_{H1} - z)^2}{(z-z_{H1})^2} dz.$$

Stąd:

$$\frac{z_{H2}-z_{H1}}{n} AD_{down}^{2**} = \int_{z_{H1}}^{z_{H2}} \frac{(F_n(z^{**})(z_{H2}-z_{H1}) + z_{H1} - z)^2}{(z-z_{H1})^2} dz.$$

Prawą stronę ostatniej równości można przybliżyć poprzez wyliczenie sumy całek w przedziałach wyznaczonych punktami skoku dystrybuanty. Podział wynika z definicji dystrybuanty empirycznej.

$$\begin{aligned} & \frac{z_{H2}-z_{H1}}{n} AD_{down}^{2**} \\ &= \int_{z_{H1}}^{z_1} \frac{(0 * (z_{H2}-z_{H1}) + z_{H1} - z)^2}{(z-z_{H1})^2} dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n} * (z_{H2}-z_{H1}) + z_{H1} - z\right)^2}{(z-z_{H1})^2} dz \\ &+ \int_{z_n}^{z_{H2}} \frac{(1 * (z_{H2}-z_{H1}) + z_{H1} - z)^2}{(z-z_{H1})^2} dz. \end{aligned}$$

Wyliczmy teraz powyższe trzy całki składowe. Całka A:

$$\begin{aligned} A &= \int_{z_{H1}}^{z_1} \frac{(0 * (z_{H2}-z_{H1}) + z_{H1} - z)^2}{(z-z_{H1})^2} dz = \int_{z_{H1}}^{z_1} \frac{(z_{H1} - z)^2}{(z-z_{H1})^2} dz = \int_{z_{H1}}^{z_1} dz = [z]_{z_{H1}}^{z_1} = z_1 - z_{H1}. \\ B &= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n} * (z_{H2}-z_{H1}) + z_{H1} - z\right)^2}{(z-z_{H1})^2} dz \\ &= \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j^2}{n^2} (z_{H2}-z_{H1})^2\right)}{(z-z_{H1})^2} dz + \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(2 \frac{j}{n} (z_{H2}-z_{H1})(z_{H1}-z)\right)}{(z-z_{H1})^2} dz \\ &+ \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{(z_{H1}-z)^2}{(z-z_{H1})^2} dz \\ &= \frac{(z_{H2}-z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \int_{z_j}^{z_{j+1}} \frac{1}{(z-z_{H1})^2} dz + 2 \frac{(z_{H2}-z_{H1})}{n} \sum_{j=1}^{n-1} j \int_{z_j}^{z_{j+1}} \frac{(z_{H1}-z)}{(z-z_{H1})^2} dz \\ &+ \sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{(z_{H1}-z)^2}{(z-z_{H1})^2} dz. \end{aligned}$$

Korzystając kolejno z (9) oraz (5):

$$\begin{aligned}
 B &= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \left[\frac{1}{z_{H1} - z_j} \right]_{z_j}^{z_{j+1}} + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j [-\ln(z - z_{H1})]_{z_j}^{z_{j+1}} + \sum_{j=1}^{n-1} [z]_{z_j}^{z_{j+1}} \\
 &= \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^{n-1} j^2 \left(\frac{1}{z_{H1} - z_{j+1}} - \frac{1}{z_{H1} - z_j} \right) \\
 &\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^{n-1} j (\ln(z_j - z_{H1}) - \ln(z_{j+1} - z_{H1})) + \sum_{j=1}^{n-1} (z_{j+1} - z_j).
 \end{aligned}$$

Następnie korzystamy z sum (15), (24) i (25):

$$\begin{aligned}
 B &= \frac{(z_{H2} - z_{H1})^2}{n^2} \left(\sum_{j=1}^n (1 - 2j) \frac{1}{z_{H1} - z_j} + n^2 * \frac{1}{z_{H1} - z_n} \right) \\
 &\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \left(\sum_{j=1}^n \ln(z_j - z_{H1}) - n \ln(z_n - z_{H1}) \right) + z_n - z_1 \\
 &= z_n - z_1 + \frac{(z_{H2} - z_{H1})^2}{z_{H1} - z_n} + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^n (1 - 2j) \frac{1}{z_{H1} - z_j} - 2(z_{H2} - z_{H1}) \ln(z_n - z_{H1}) \\
 &\quad + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n \ln(z_j - z_{H1}).
 \end{aligned}$$

Całka C zostanie obliczona wykorzystując 8:

$$\begin{aligned}
 C &= \int_{z_n}^{z_{H2}} \frac{(1 * (z_{H2} - z_{H1}) + z_{H1} - z)^2}{(z - z_{H1})^2} dz = \\
 &= \int_{z_n}^{z_{H2}} \frac{(z_{H2} - z)^2}{(z - z_{H1})^2} dz = (z_{H2}^2 - z_{H1}^2) \left[\frac{1}{z_{H1} - z} \right]_{z_n}^{z_{H2}} + 2(z_{H1} - z_{H2}) \left[\frac{z_{H1}}{z_{H1} - z} + \ln(z - z_{H1}) \right]_{z_n}^{z_{H2}} \\
 &\quad + [z]_{z_n}^{z_{H2}} \\
 &= \frac{(z_{H2}^2 - z_{H1}^2)}{z_{H1} - z_{H2}} - \frac{(z_{H2}^2 - z_{H1}^2)}{z_{H1} - z_n} + \frac{2(z_{H1} - z_{H2})z_{H1}}{z_{H1} - z_{H2}} - \frac{2(z_{H1} - z_{H2})z_{H1}}{z_{H1} - z_n} \\
 &\quad + 2(z_{H1} - z_{H2}) \ln(z_{H2} - z_{H1}) - 2(z_{H1} - z_{H2}) \ln(z_n - z_{H1}) + z_{H2} - z_n \\
 &= z_{H2} - z_n + (z_{H1} - z_{H2}) - \frac{(z_{H1} - z_{H2})^2}{(z_{H1} - z_n)} - 2(z_{H1} - z_{H2}) \ln(z_n - z_{H1}) \\
 &\quad + 2(z_{H1} - z_{H2}) \ln(z_{H2} - z_{H1}).
 \end{aligned}$$

Ostatecznie statystyka jest sumą całek A, B i C przemnożonych przez stałą:

$$\begin{aligned}
 AD_{down}^{2**} &= \frac{n}{z_{H2} - z_{H1}} \left(z_1 - z_{H1} + (z_{H1} - z_{H2}) + \frac{(z_{H2} - z_{H1})^2}{(z_{H1} - z_n)} - 2(z_{H2} - z_{H1}) \ln(z_n - z_{H1}) \right. \\
 &\quad + 2(z_{H1} - z_{H2}) \ln(z_{H2} - z_{H1}) + z_n - z_1 - \frac{(z_{H1} - z_{H2})^2}{z_{H1} - z_n} + \frac{(z_{H2} - z_{H1})^2}{n^2} \sum_{j=1}^n (1 - 2j) \frac{1}{z_{H1} - z_j} \\
 &\quad \left. - 2(z_{H1} - z_{H2}) \ln(z_n - z_{H1}) + 2 \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n \ln(z_j - z_{H1}) + z_{H2} - z_n \right) = \\
 &= -2n \ln(z_{H2} - z_{H1}) + 2 \sum_{j=1}^n \ln(z_j - z_{H1}) + \frac{(z_{H2} - z_{H1})}{n} \sum_{j=1}^n (1 - 2j) \frac{1}{z_{H1} - z_j}.
 \end{aligned}$$

Załącznik 2. Empiryczne rozmiary wybranych testów zgodności dla poziomu istotności $\alpha=0,05^*$

Rozkład teoretyczny **	Przedział dystrybucyjny	n	KS**	V**	AD**	AD ^{down} ** AD ^{up} **	W ^{2**}	AD ^{2**}	AD ^{2**down} AD ^{2**up}	
N	[0; 0,1]	10	6,0	5,3	5,0	4,1	5,1	5,4	4,5	
		20	4,9	4,2	3,8 *	5,4	4,2	3,8 *	4,9	
		30	5,3	6,3 *	4,4	4,5	5,6	5,6	4,8	
		40	3,5 **	4,7	4,5	4,9	4,0	4,0	5,1	
		50	5,7	5,2	4,6	5,0	4,5	4,1	5,1	
		100	4,3	5,0	4,3	5,2	4,5	4,3	5,7	
	[0,9; 1]	10	4,6	4,9	5,0	4,2	4,9	4,5	4,3	
		20	4,5	4,0	5,0	4,9	4,9	5,0	4,5	
		30	4,7	4,9	5,1	3,7 *	4,6	4,7	3,8 *	
		40	3,7 *	4,5	5,3	3,7 *	4,4	3,6 *	3,9	
		50	4,7	4,8	5,8	5,4	4,6	4,2	4,8	
		100	3,6 *	4,1	3,7 *	4,2	3,9	3,9	3,7 *	
	[0,1; 0,9]	80	3,0 ***	4,3	4,8		2,8 ***	2,6 ***		
		160	4,0	6,2 *	4,5		2,6 ***	2,6 ***		
		240	3,9	5,1	6,2 *		3,5 **	3,7 *		
		320	2,1 ***	4,5	5,5		1,6 ***	1,9 ***		
		400	3,7 *	5,6	4,4		3,0 ***	2,6 ***		
		800	2,5 ***	3,6 *	4,8		2,0 ***	2,3 ***		
	[0; 1]	100	4,1	5,0	5,0		4,0	3,9		
		200	6,1	5,6	4,2		5,7	5,3		
		300	5,4	4,2	3,9		4,6	4,6		
		400	3,2 **	3,6 *	4,1		3,6 *	3,3 **		
		500	3,6 *	4,7	4,7		4,5	5,0		
		1000	3,9	4,9	4,3		3,6 *	3,8 *		
	N	[0; 0,05]	5	5,5	4,7	6,3 *	4,6	5,3	5,6	4,1
			10	4,3	4,4	4,3	4,4	4,4	4,5	4,5
			15	4,3	4,9	5,3	5,0	4,9	5,4	4,7
			20	5,2	5,4	4,3	4,2	4,9	4,8	4,1
			25	4,2	4,8	5,0	5,6	4,8	5,0	5,2
			50	4,1	4,7	4,6	5,5	5,1	4,6	4,7
[0,95; 1]		5	4,9	5,0	5,2	4,7	4,8	5,7	4,8	
		10	4,1	4,7	4,6	4,9	4,1	4,5	4,6	
		15	4,8	4,6	6,3 *	5,8	5,1	4,8	5,6	
		20	4,7	5,2	3,9	3,9	5,3	5,2	4,1	
		25	5,0	5,1	4,5	4,2	4,1	4,1	4,0	
		50	4,2	4,0	4,9	5,2	4,5	4,4	5,8	
[0,05; 0,95]		90	3,6 *	5,5	5,8		3,3 **	3,0 ***		
		180	2,7 ***	4,0	3,9		1,6 ***	2,7 ***		
		270	3,1 ***	5,3	4,3		2,2 ***	3,4 **		
		360	3,3 **	4,3	5,0		2,9 ***	2,6 ***		
		450	2,7 ***	3,9	5,9		2,6 ***	2,8 ***		
		900	2,7 ***	3,3 **	6,4 **		1,4 ***	1,7 ***		
NIG		[0; 0,1]	10	5,6	5,5	5,4	4,4	6,2 *	5,2	3,9
			20	5,1	6,0	4,6	3,3 **	5,0	4,6	3,5 **
			30	3,8 *	4,1	6,1	5,1	4,5	4,3	4,8
			40	4,3	4,7	5,5	5,4	4,2	4,2	5,0
			50	4,3	4,2	4,1	4,1	4,8	4,9	3,9
			100	4,8	4,0	4,8	5,2	4,3	4,0	5,2
	[0,9; 1]	10	5,0	5,9	6,2 *	4,2	5,5	4,9	3,6 *	
		20	4,5	5,4	6,0	5,0	3,8 *	3,7 *	4,8	
		30	4,0	4,0	4,6	4,8	3,6 *	3,7 *	5,0	
		40	4,4	4,8	5,5	5,1	4,0	4,1	6,0	
		50	3,9	3,7 *	4,9	6,0	4,0	3,7 *	5,9	
		100	5,4	5,6	4,5	4,3	5,1	5,3	4,5	
	[0,1; 0,9]	80	2,7 ***	3,5 **	5,6		2,3 ***	3,1 ***		
		160	4,3	5,2	4,4		2,9 ***	3,6 *		
		240	3,6 *	4,2	4,9		3,0 ***	3,3 **		
		320	3,5 **	4,6	4,0		3,8 *	4,2		
		400	4,5	4,8	4,5		3,9	3,8 *		
		800	4,2	5,0	3,8 *		2,8 ***	3,3 **		
	[0; 1]	100	3,6 *	3,8 *	5,2		3,4 **	3,4 **		
		200	4,7	6,1	5,4		4,4	3,7 *		
		300	4,4	3,8 *	6,2 *		3,9	3,6 *		
		400	5,1	5,5	6,2 *		5,5	4,8		
		500	5,5	5,5	5,5		5,4	4,7		
		1000	5,9	5,4	5,1		6,0	4,6		

Rozkład teoretyczny **	Przedział dystrybucyjny	n	KS**	V**	AD**	AD ^{down} ** AD ^{up} **	W ^{2**}	AD ^{2**}	AD ^{down} 2** AD ^{up} 2**	
NIG	[0; 0,05]	5	4,0	5,4	5,4	3,6 *	4,7	5,0	3,0 ***	
		10	4,9	4,2	5,1	3,2 **	5,0	4,3	3,4 **	
		15	4,6	5,4	4,4	2,9 ***	4,7	4,7	2,6 ***	
		20	5,1	5,4	4,3	3,3 **	4,5	4,7	2,9 ***	
		25	4,5	5,2	4,0	4,5	4,4	4,0	4,3	
	50	4,7	4,7	4,1	4,9	4,9	4,8	3,8 *		
	[0,95; 1]	5	4,3	4,4	4,4	4,3	4,6	4,5	3,6 *	
		10	5,6	4,4	5,7	4,3	4,8	4,2	4,1	
		15	4,7	5,1	5,0	4,4	4,0	4,1	4,1	
		20	4,0	5,7	4,8	4,3	4,3	4,8	4,1	
		25	4,9	4,7	5,2	3,9	4,7	4,5	4,0	
	[0,05; 0,95]	50	4,5	5,6	5,3	6,0	3,8 *	4,6	5,4	
		90	4,8	5,4	5,4		3,4 **	3,2 **		
		180	4,5	4,5	4,9		3,5 **	3,4 **		
		270	3,6 *	3,7 *	5,9		4,3	5,1		
		360	4,8	4,9	6,8 **		4,2	3,7 *		
		450	4,1	5,2	4,9		4,5	4,0		
		900	3,5 **	3,6 *	5,2		3,7 *	3,1 ***		
	Hyp	[0; 0,1]	10	4,9	4,6	4,8	5,1	4,8	4,6	4,2
			20	5,1	5,3	5,4	5,1	4,7	4,1	4,4
30			5,0	3,7 *	5,3	4,5	3,8 *	4,2	4,4	
40			3,7 *	4,6	4,5	5,0	3,8 *	3,9	3,6 *	
50			4,6	5,2	5,2	4,3	4,9	4,5	3,9	
100		4,8	4,6	5,2	3,8 *	4,5	4,4	4,0		
[0,9; 1]		10	4,9	5,7	3,5 **	3,7 *	4,2	4,8	3,4 **	
		20	4,5	5,2	4,4	3,9	4,6	4,0	3,6 *	
		30	3,9	5,2	4,3	4,3	4,6	4,7	4,3	
		40	4,4	4,4	5,1	3,9	4,2	3,8 *	4,4	
		50	4,5	4,4	5,0	5,7	5,4	4,8	5,2	
100		4,2	4,9	6,1	4,3	3,9	4,2	3,9		
[0,1; 0,9]		80	2,6 ***	3,0 ***	4,4		2,8 ***	3,2 **		
		160	2,8 ***	2,6 ***	5,6		2,0 ***	2,1 ***		
		240	3,2 **	2,5 ***	5,1		2,8 ***	3,5 **		
		320	4,0	1,9 ***	4,5		3,5 **	4,9		
		400	3,3 **	2,1 ***	4,2		3,7 *	3,7 *		
800		3,8 *	4,0	4,2		3,8 *	4,4			
[0; 1]		100	2,7 ***	3,1 ***	4,1		2,4 ***	2,4 ***		
		200	2,8 ***	2,8 ***	5,1		2,7 ***	3,1 ***		
	300	2,6 ***	2,7 ***	4,8		2,9 ***	2,8 ***			
	400	3,7 *	2,9 ***	3,8 *		3,7 *	3,7 *			
	500	3,8 *	2,9 ***	4,3		2,4 ***	2,8 ***			
1000	3,2 **	3,2 **	4,7		3,8 *	4,5				
Hyp	[0; 0,05]	5	3,5 **	3,6 *	4,5	4,8	3,7 *	4,1	3,9	
		10	4,0	4,5	4,9	4,3	3,7 *	3,8 *	3,9	
		15	5,3	4,5	4,6	4,6	5,0	5,1	4,5	
		20	5,6	5,0	4,3	4,3	5,8	5,1	3,8 *	
		25	5,4	4,5	5,2	4,8	5,9	6,1	4,8	
	50	5,0	3,6 *	5,9	5,5	5,2	5,2	5,8		
	[0,95; 1]	5	4,6	4,6	4,4	5,3	3,7 *	4,2	5,3	
		10	3,4 **	4,3	4,9	5,4	4,2	4,6	5,2	
		15	4,1	4,5	5,1	4,9	3,3 **	3,2 **	4,8	
		20	5,0	4,5	4,3	4,5	4,8	4,4	3,9	
		25	5,2	6,5 **	4,1	4,4	5,1	4,9	4,2	
	50	4,3	4,8	6,3 *	5,9	3,6 *	3,8 *	5,2		
	[0,05; 0,95]	90	4,5	5,1	4,1		3,3 **	3,5 **		
		180	3,1 ***	2,5 ***	4,6		3,3 **	3,5 **		
		270	3,5 **	3,2 **	5,1		3,7 *	3,4 **		
		360	3,2 **	3,5 **	4,8		2,9 ***	2,6 ***		
		450	3,5 **	3,5 **	5,4		3,1 ***	3,2 **		
	900	4,4	2,9 ***	3,6 *		2,8 ***	3,1 ***			

Rozkład teoretyczny **	Przedział dystrybucyjny	n	KS**	V**	AD**	AD** _{down} AD** _{up}	W ^{2**}	AD ^{2**}	AD ^{2**} _{down} AD ^{2**} _{up}	
SGED	[0; 0,1]	10	4,9	5,4	5,5	3,9	4,8	3,9	3,0 ***	
		20	5,1	5,8	5,0	4,4	5,2	4,5	4,8	
		30	3,9	3,9	3,4 **	3,8 *	4,2	4,2	3,7 *	
		40	4,3	5,5	5,2	4,0	3,9	4,1	4,1	
		50	5,0	5,7	4,8	4,4	5,6	5,1	4,5	
		100	5,1	5,7	6,4 **	4,5	4,6	4,6	5,0	
	[0,9; 1]	10	3,6 *	5,3	5,4	4,6	3,0 ***	3,8 *	4,3	
		20	5,5	5,4	5,2	4,3	5,2	4,9	3,9	
		30	5,2	5,3	4,1	5,3	4,4	4,8	4,5	
		40	4,3	5,0	4,8	5,1	4,0	3,9	5,2	
		50	3,8 *	5,6	4,6	5,0	4,4	4,8	4,9	
		100	3,6 *	4,4	4,8	4,9	3,7 *	3,5 **	5,6	
	[0,1; 0,9]	80	4,0	4,1	4,0		3,2 **	3,0 ***		
		160	4,0	4,2	5,0		3,9	4,4		
		240	3,8 *	3,9	4,7		3,9	3,9		
		320	4,4	3,8 *	4,9		4,3	4,0		
		400	3,4 **	3,6 *	4,9		3,9	3,7 *		
		800	3,1 ***	3,4 **	5,6		3,7 *	3,3 **		
	[0, 1]	100	4,2	3,7 *	4,1		2,6 ***	3,0 ***		
		200	4,3	3,8 *	4,2		4,3	5,0		
		300	4,6	4,4	4,2		3,9	3,2 **		
		400	5,4	4,5	4,9		4,9	4,2		
		500	4,8	3,7 *	5,6		4,7	4,5		
		1000	3,9	3,9	5,1		3,8 *	3,9		
	SGED	[0; 0,05]	5	4,7	4,6	4,8	5,0	4,7	4,6	5,3
			10	4,4	5,0	5,9	5,1	4,6	4,7	4,5
			15	4,5	4,1	5,0	4,0	3,2 **	3,8 *	3,8 *
			20	4,9	6,1	5,3	3,9	5,2	5,7	4,2
			25	4,5	3,8 *	4,7	4,3	3,9	4,2	3,4 **
			50	6,2 *	4,6	5,2	4,0	5,4	4,9	4,4
[0,95; 1]		5	4,6	5,1	5,5	5,3	5,0	5,3	4,5	
		10	6,0	4,6	5,2	5,2	5,4	4,3	4,5	
		15	5,9	5,9	4,8	3,9	5,3	6,4 **	3,4 **	
		20	4,9	4,6	4,8	3,2 **	5,4	5,0	3,6 *	
		25	6,2 *	7,5 ***	5,7	5,1	5,9	5,5	4,9	
		50	5,0	5,1	5,8	5,8	4,9	5,4	5,4	
[0,05; 0,95]		90	3,5 **	3,8 *	4,2		2,8 ***	2,6 ***		
		180	3,8 *	4,0	4,6		2,9 ***	2,5 ***		
		270	4,6	5,9	5,1		3,4 **	2,9 ***		
		360	3,5 **	3,8 *	5,0		2,8 ***	3,0 ***		
		450	3,8 *	4,4	5,0		3,9	4,0		
		900	4,1	3,5 **	5,7		2,8 ***	2,1 ***		
ST		[0; 0,1]	10	4,5	5,1	4,9	5,6	4,4	4,2	5,4
			20	4,7	5,2	4,3	4,5	4,6	4,5	4,2
			30	3,3 **	4,0	4,9	4,7	3,7 *	3,8 *	5,0
			40	3,7 *	4,9	6,8 **	6,3 *	3,8 *	4,0	5,4
			50	4,4	4,2	4,2	4,6	4,0	3,4 **	4,5
			100	4,3	5,0	4,2	3,8 *	4,2	4,2	3,9
	[0,9; 1]	10	5,2	5,1	4,0	4,9	4,8	4,7	5,2	
		20	4,7	5,5	4,8	5,0	4,9	4,7	5,0	
		30	4,0	5,1	4,9	4,3	3,2 **	3,7 *	4,2	
		40	5,9	6,8 **	4,7	4,9	5,1	5,8	4,9	
		50	4,0	4,8	4,5	5,3	4,8	4,4	4,8	
		100	4,6	4,6	5,2	4,4	4,4	4,9	4,5	
	[0,1; 0,9]	80	3,4 **	4,2	5,2		3,8 *	3,5 **		
		160	4,3	4,4	4,8		3,2 **	3,6 *		
		240	3,9	3,8 *	5,0		3,9	4,5		
		320	3,7 *	4,7	4,5		3,7 *	3,5 **		
		400	2,8 ***	2,7 ***	4,7		2,0 ***	2,3 ***		
		800	4,6	5,0	4,7		3,6 *	3,7 *		
	[0, 1]	100	4,1	4,3	5,7		3,8 *	4,0		
		200	5,4	4,0	4,9		3,9	4,4		
		300	4,4	4,8	4,4		5,2	5,5		
		400	4,3	4,3	6,4 **		4,4	4,6		
		500	3,9	2,8 ***	4,4		3,3 **	3,4 **		
		1000	4,8	4,9	3,8 *		5,2	4,8		

Rozkład teoretyczny **	Przedział dystrybucyjny	n	KS**	V**	AD**	AD ^{down} ** AD ^{up} **	W ^{2**}	AD ^{2**}	AD ^{2** down} AD ^{2** up}
ST	[0; 0,05]	5	4,1	4,4	4,7	4,9	4,3	4,3	5,0
		10	3,8 *	3,9	5,6	5,0	4,1	4,2	5,2
		15	4,7	5,2	4,8	5,6	5,2	5,3	5,3
		20	3,9	3,9	5,0	5,3	4,2	3,9	4,7
		25	4,4	4,3	4,4	5,1	4,4	4,7	4,4
		50	5,0	5,9	3,3 **	4,3	4,5	4,4	3,8 *
	[0,95; 1]	5	4,9	4,5	5,3	3,9	3,9	5,1	3,1 ***
		10	3,9	5,0	4,8	3,9	4,3	4,4	3,4 **
		15	4,3	3,6 *	4,5	4,7	3,6 *	4,0	4,4
		20	4,3	4,2	5,8	4,5	5,5	5,5	4,0
		25	5,4	5,6	5,6	5,2	5,3	5,4	4,9
		50	4,0	3,5 **	5,5	5,2	3,7 *	3,7 *	5,2
	[0,05; 0,95]	90	4,1	5,1	4,1		4,5	3,8 *	
		180	4,0	3,5 **	4,6		3,6 *	3,6 *	
		270	5,5	5,6	5,1		3,4 **	3,6 *	
		360	4,8	5,3	5,2		5,3	4,0	
		450	4,5	4,7	5,0		4,0	3,7 *	
		900	4,3	4,0	6,0		3,8 *	3,6 *	

* Wartości znajdujące się poza 90%, 95% i 99% przedziałem ufności zostały oznaczone odpowiednio przez *, **, ***.

** Wzorcami dla rozkładów rzeczywistych F są następujące rozkłady (z parametrami oszacowanymi na podstawie stóp zwrotu z indeksu *FTSE100* w roku 2020): $N(-0,06; 1,85)$, $NIG(0,24; 1,36; 0,45; -0,10)$, $Hyp(0,21; 0,30; 0,84; -0,09)$, $SGED(-0,06; 1,79; 0,94; 0,95)$, $ST(0,54; 1,24; -0,50; 3,24)$.

Źródło: obliczenia własne

Rozkład	Przedział teoretyczny	Przedział dystrybucyjny	ε	FTSE100, 2019				FTSE100, 2020				FTSE100, 2020				WIG20, 2019				WIG20, 2020												
				KS**	V**	AD** AD** AD** AD**	W**	AD** AD** AD** AD**	KS**	V**	AD** AD** AD** AD**	W**	AD** AD** AD** AD**	KS**	V**	AD** AD** AD** AD**	W**	AD** AD** AD** AD**	KS**	V**	AD** AD** AD** AD**	W**	AD** AD** AD** AD**									
rozkład rzeczywisty	[0;1]	NIG	10	4,2	5,6	4,9	5,1	3,4	3,4	4,4	4,4	1,9	4,4	3,6	1,6	5,1	4,9	5,7	3,3	3,3	5,5	4,7	3,3	4,4	5,3	4,4	2,2	5,3	4,6	2,8		
			20	4,1	4,8	4,6	3,5	4,2	4,4	3,7	4,4	4,4	1,4	5,1	4,9	2,0	5,1	5,4	3,9	3,2	3,2	4,7	5,3	2,9	4,6	7,1	3,6	1,5	4,0	3,6	1,5	
			30	5,2	5,7	3,3	3,4	4,8	4,7	2,7	3,3	5,1	3,9	1,5	3,2	3,0	1,5	4,7	4,0	3,4	3,3	4,0	3,6	3,0	5,7	7,0	3,3	1,0	5,6	5,2	1,4	
			40	4,1	4,2	4,9	2,7	4,1	3,9	2,3	4,9	7,5	2,8	0,7	4,3	4,4	0,7	4,0	4,6	3,8	2,9	4,7	4,6	3,2	3,3	6,9	3,8	1,4	4,2	4,3	1,6	
			100	4,6	5,6	4,6	3,0	5,1	4,1	2,6	4,6	7,9	3,0	0,6	4,1	4,6	1,7	5,1	6,2	3,6	3,9	4,5	4,1	3,3	5,5	7,8	3,2	0,6	5,1	5,4	1,2	
			200	4,1	4,3	3,9	4,4	4,9	4,4	4,5	4,8	5,1	4,5	1,9	4,8	4,6	2,0	4,0	5,1	6,4	3,7	4,1	4,5	3,8	5,3	6,0	3,9	1,7	3,8	3,6	2,4	
			300	4,1	4,0	5,5	4,7	4,5	4,3	4,5	5,5	4,8	4,0	1,7	5,0	4,3	1,6	4,7	4,1	4,6	2,2	4,3	3,8	2,6	3,6	5,0	5,0	1,8	4,2	3,2	1,8	
			400	5,7	5,6	4,5	3,4	4,7	4,4	3,4	4,9	5,3	3,5	2,2	4,9	4,8	2,4	3,3	3,7	4,6	2,7	4,4	4,2	2,6	4,6	5,3	3,9	1,4	4,0	4,4	1,1	
			500	4,3	4,9	3,5	3,5	4,7	4,7	3,3	4,3	5,6	3,9	3,5	5,2	5,1	2,7	4,1	6,0	3,6	1,4	4,2	3,8	2,1	3,9	5,5	3,3	1,9	3,6	3,8	1,5	
			1000	4,2	4,4	3,3	2,5	4,2	4,5	2,5	5,2	5,2	3,4	2,0	6,6	5,5	1,6	4,4	5,4	3,9	2,7	3,9	3,8	2,3	4,0	5,2	3,3	1,1	3,6	3,3	1,4	
			2000	3,2	3,2	4,5	2,8	2,2	4,3	5,5	5,1	5,1	3,3	3,3	3,8	4,9	4,3	3,4	3,0	4,4	4,4	1,1	2,7	2,7	4,3	5,7	5,3	4,0	3,2	4,0	3,2	
			3000	3,5	4,6	3,6	4,7	3,0	2,7	3,0	2,7	5,4	6,6	5,5	4,9	4,5	2,8	4,5	5,5	5,5	2,9	3,1	2,9	3,1	4,8	8,9	4,0	4,0	5,9	5,2	4,0	
			4000	3,2	3,3	4,5	2,1	2,1	3,0	8,3	12,6	4,9	10,3	6,5	11,2	7,9	3,9	5,0	5,5	5,0	4,7	2,4	3,0	3,0	8,6	13,0	6,2	10,1	8,6	10,1	8,6	
			5000	2,8	4,3	4,7	3,2	3,8	4,7	17,5	27,8	4,4	24,8	15,4	4,4	15,4	5,7	6,7	6,9	6,9	3,7	3,7	5,1	2,0	2,0	30,9	6,1	29,6	18,3	2,0	1,1	
			10000	4,8	4,6	2,5	3,8	3,8	2,7	2,7	2,7	5,6	5,8	2,0	4,9	4,3	3,4	3,0	4,4	4,4	4,4	2,7	2,7	2,7	4,4	5,6	2,4	4,4	5,6	4,4	4,4	
			5	6,0	7,9	3,6	1,9	6,3	4,6	2,1	5,4	5,4	5,2	0,7	5,4	4,6	1,0	4,9	5,2	3,5	3,5	5,2	4,2	2,6	3,9	3,9	5,3	3,0	1,1	4,1	3,4	1,4
			10	4,3	4,4	3,9	3,2	4,5	4,7	2,5	5,0	7,1	3,3	1,0	4,6	5,0	1,7	4,2	4,4	3,5	2,2	3,9	4,2	2,1	4,9	5,2	4,3	0,9	5,1	4,3	1,1	
			15	4,6	4,9	4,6	4,2	4,2	3,9	3,4	4,4	5,1	3,9	1,3	4,6	4,8	1,5	5,0	6,2	4,0	2,6	4,3	4,4	2,6	5,5	5,0	3,1	1,1	5,2	5,0	1,1	
			20	4,9	5,8	3,1	2,7	4,5	5,0	2,1	4,4	5,3	3,8	0,4	5,2	6,3	1,5	5,3	7,0	4,6	2,7	5,2	4,7	2,2	6,1	6,3	3,6	1,1	6,3	4,9	1,3	
			25	4,1	4,0	3,2	2,6	4,0	4,2	2,2	7,0	5,1	4,0	0,9	5,9	6,0	2,1	4,4	4,1	3,4	2,7	4,4	4,0	2,5	5,3	5,9	2,9	0,7	4,9	5,2	0,9	
50	3,7	5,3	4,1	2,8	4,1	3,6	2,2	7,5	7,2	4,6	0,6	9,4	9,5	2,1	4,3	3,7	3,7	2,4	4,7	4,6	2,3	5,5	6,5	3,6	0,5	7,7	8,0	1,1				
100	5,0	5,5	4,9	5,1	5,0	5,5	4,3	3,5	4,0	4,5	3,1	3,7	3,8	3,0	5,1	5,2	4,5	3,9	4,2	4,7	3,5	4,9	5,0	4,3	2,0	4,7	4,3	1,5				
150	5,5	6,0	4,6	2,9	5,4	4,9	2,5	4,9	4,6	4,3	2,0	3,9	3,9	2,2	5,8	5,7	3,8	2,5	5,4	4,9	2,1	4,2	6,2	3,2	1,8	4,2	3,9	1,5				
200	5,4	5,4	3,1	3,1	5,0	4,6	2,5	4,2	4,1	5,6	2,5	3,6	3,7	2,8	4,3	5,2	3,5	2,5	4,0	4,2	2,6	4,7	6,2	4,4	0,7	4,6	3,8	0,9				
250	4,2	2,9	5,0	3,9	3,8	3,9	3,8	4,1	4,6	4,0	2,0	5,1	4,9	1,8	5,4	4,2	4,2	1,6	4,6	4,7	1,5	3,3	5,0	3,9	1,5	4,1	4,3	1,2				
500	4,3	5,7	3,7	3,1	4,0	3,9	3,2	4,7	4,9	3,1	1,6	4,8	4,9	1,1	4,8	4,6	4,0	1,9	5,1	4,2	1,7	5,3	5,3	4,0	0,7	4,8	4,7	0,9				
900	3,8	4,0	4,8	3,6	4,0	3,6	4,0	5,5	6,3	7,1	5,4	5,0	5,4	5,0	3,7	5,4	4,3	4,3	4,2	4,5	4,4	6,1	4,7	4,7	4,2	4,2	4,1	4,1				
1800	5,1	6,3	5,4	4,5	4,6	4,6	6,2	7,9	6,9	7,3	6,0	7,3	6,0	4,3	5,8	5,4	5,4	4,9	4,9	4,1	5,5	6,9	6,5	5,4	6,1	6,3	6,3	6,3				
2700	3,4	4,1	5,1	3,3	3,9	3,9	9,6	11,0	5,7	9,9	8,2	9,9	8,2	5,2	5,1	3,6	3,6	3,7	3,9	3,7	3,9	6,1	9,6	5,4	8,4	6,5	8,4	6,5				
3600	4,2	4,9	4,8	4,0	4,1	4,1	8,9	11,5	5,4	9,9	7,9	9,9	7,9	4,7	4,6	3,8	3,8	4,6	5,3	4,6	5,3	9,8	13,5	4,3	10,7	8,5	10,7	8,5				
4500	4,4	4,0	4,2	3,7	3,8	3,8	11,8	15,4	5,3	14,1	11,0	14,1	11,0	3,8	4,9	5,5	5,5	4,3	4,3	4,3	3,8	10,9	15,9	5,8	14,7	10,5	14,7	10,5				
9000	5,6	5,4	2,8	4,6	4,6	3,5	18,6	26,4	5,3	26,5	20,0	26,5	20,0	5,3	6,3	6,3	4,3	5,1	4,8	5,1	4,8	21,3	30,5	5,5	28,6	18,4	28,6	18,4				

Rozkład teoretyczny	Przedział dystrybuanty	FTSE100, 2019										FTSE100, 2020										FTSE100, 2021									
		KS**	V**	AD** AD** AD**	AD** AD** AD**	W**	AD** AD** AD**	AD** AD** AD**	KS**	V**	AD** AD** AD**	AD** AD** AD**	W**	AD** AD** AD**	AD** AD** AD**	KS**	V**	AD** AD** AD**	AD** AD** AD**	W**	AD** AD** AD**	AD** AD** AD**	KS**	V**	AD** AD** AD**	AD** AD** AD**	W**	AD** AD** AD**	AD** AD** AD**		
rozkład rzeczywisty	SGED	[0,1]										[0,1]										[0,1]									
		10	20	30	40	50	60	70	80	90	100	10	20	30	40	50	60	70	80	90	100	10	20	30	40	50	60	70	80	90	100
rozkład rzeczywisty	SGED	[0,05;0,95]										[0,05;0,95]										[0,05;0,95]									
		10	20	30	40	50	60	70	80	90	100	10	20	30	40	50	60	70	80	90	100	10	20	30	40	50	60	70	80	90	100

Rozkład	Przedział teoretyczny	Przedział dystrybuanty	FTSE100, 2019		FTSE100, 2020		FTSE100, 2020		FTSE100, 2020		FTSE100, 2020		FTSE100, 2020		FTSE100, 2020		FTSE100, 2020		FTSE100, 2020														
			KS**	V**	AD** AD** _{up}	AD** AD** _{down}	W**	AD** AD** _{up}	AD** AD** _{down}	W**	AD** AD** _{up}	AD** AD** _{down}	W**	AD** AD** _{up}	AD** AD** _{down}	W**	AD** AD** _{up}	AD** AD** _{down}	W**	AD** AD** _{up}	AD** AD** _{down}												
ST	rozkład rzeczywisty	[0;0.1]	10	4.3	5.9	3.4	3.2	4.1	3.5	3.1	5.6	5.0	5.1	1.7	4.9	4.6	1.5	4.1	4.5	4.7	3.3	4.2	4.6	2.5	4.4	4.9	3.0	0.8	3.8	3.3	0.8		
			20	3.6	4.1	4.0	2.4	3.3	3.2	1.7	5.8	5.9	3.1	0.8	6.0	5.6	0.6	4.6	3.8	4.0	2.3	3.3	3.3	3.3	1.9	5.9	5.4	3.0	0.7	5.4	4.1	0.5	
			30	4.1	4.6	3.7	1.8	4.8	4.6	1.6	4.9	3.0	0.5	5.9	5.1	0.5	4.9	1.5	4.9	3.6	1.5	4.9	3.6	3.6	1.5	4.5	5.6	2.2	0.2	4.7	4.5	0.5	
			40	5.3	5.9	3.5	2.3	5.8	5.4	1.9	5.1	5.9	3.8	0.2	5.1	4.3	0.0	5.7	5.6	3.0	1.8	5.5	4.3	1.2	5.0	5.2	2.6	0.2	5.7	5.2	0.1		
			50	6.4	6.2	3.5	2.1	6.4	5.5	1.8	7.0	6.7	3.0	0.0	6.2	6.2	0.1	5.6	5.2	3.4	0.5	6.1	5.3	1.1	6.3	6.8	2.5	0.0	6.4	6.0	0.1		
			100	5.1	4.3	2.2	0.9	4.7	4.8	1.2	6.9	6.6	2.4	0.0	7.4	7.2	0.0	6.2	4.6	2.8	0.5	6.3	5.5	0.4	6.7	7.6	2.5	0.0	7.8	7.2	0.1		
			10	4.9	5.3	5.9	2.8	5.1	4.7	2.7	3.8	4.0	2.8	1.0	4.3	3.3	0.8	4.2	4.7	4.1	3.9	4.0	3.5	2.9	4.3	3.9	0.4	4.5	4.2	0.5			
			20	4.2	4.1	3.6	2.3	4.1	4.2	1.9	4.2	4.8	2.9	0.3	4.3	3.3	0.3	4.9	6.1	2.7	2.2	5.0	4.0	1.7	4.4	5.1	2.9	0.4	5.1	5.2	0.4		
			30	5.6	4.9	3.7	1.7	4.7	4.5	1.6	5.1	6.0	4.6	0.4	6.0	5.5	0.4	4.8	4.3	4.2	1.6	4.7	4.8	1.5	5.2	6.5	4.2	0.1	5.5	5.6	0.0		
			40	3.9	4.8	2.8	1.5	4.8	4.5	1.8	5.5	6.0	2.6	0.1	5.1	4.8	0.0	5.6	6.3	2.8	1.4	6.1	5.9	1.2	6.1	6.4	3.7	0.1	6.0	5.4	0.0		
			50	4.7	5.2	3.0	1.4	5.1	5.0	2.0	6.6	6.1	2.7	0.2	6.0	6.2	0.2	4.0	3.7	2.5	0.7	4.3	3.7	0.9	5.8	6.3	2.6	0.1	6.2	4.9	0.2		
			100	5.8	5.7	4.8	1.8	5.8	5.9	1.4	5.7	5.9	3.5	0.1	6.7	6.6	0.1	5.9	6.1	2.3	0.4	6.3	6.2	0.5	7.3	7.3	3.0	0.0	7.8	7.7	0.0		
			80	4.1	5.8	5.1																											
			160	3.9	4.3	4.0																											
			240	4.2	6.0	4.8																											
			320	3.8	4.4	4.3																											
			400	4.2	5.2	4.3																											
			800	4.6	5.7	6.0																											
			100	4.8	5.1	3.5																											
			200	4.6	4.4	1.1																											
300	5.3	4.9	1.5																														
400	4.1	4.8	1.6																														
500	5.1	5.0	1.3																														
1000	4.5	4.9	0.8																														
5	3.9	4.4	5.1																														
10	5.3	6.1	2.8																														
15	4.5	6.0	4.6																														
20	3.9	5.6	3.9																														
25	4.2	3.9	3.3																														
50	5.9	5.0	2.8																														
5	4.2	3.8	5.1																														
10	4.6	4.9	3.7																														
15	3.7	4.8	3.4																														
20	5.3	5.5	4.3																														
25	3.7	4.7	4.2																														
50	5.8	5.4	3.9																														
90	3.7	4.4	5.8																														
180	3.6	4.1	3.9																														
270	3.8	4.9	6.5																														
360	4.5	4.8	6.0																														
450	3.8	4.5	5.3																														
900	5.3	7.1	4.9																														

Załącznik 4. Oszacowania wartości p testów zgodności dla indeksów WIG20 i FTSE100 ($\alpha = 0,05$)*

Okres	Fragment dystrybuanty	Indeks							WIG20							FTSE100						
		Rozkład	KS**	V**	AD**	AD ^{down} AD ^{up}	W ^{2**}	AD ^{2**} AD ^{2**} AD ^{2**} up	KS**	V**	AD**	AD ^{down} AD ^{up}	W ^{2**}	AD ^{2**} AD ^{2**} AD ^{2**} up	KS**	V**	AD**	AD ^{down} AD ^{up}	W ^{2**}	AD ^{2**} AD ^{2**} AD ^{2**} up		
2019	[0, 1]	N	0,13	0,04	0,08	0,01	0,01	0,01	0,07	0,11	0,01	0,01	0,06	0,02	0,06	0,02	0,02	0,06	0,02	0,02	0,02	
		N/G	0,46	0,39	0,98	0,40	0,41	0,40	0,42	0,40	0,44	0,42	0,42	0,27	0,42	0,42	0,27	0,42	0,42	0,27	0,27	
		Hyp	0,41	0,46	0,98	0,45	0,46	0,45	0,46	0,46	0,43	0,23	0,37	0,18	0,37	0,18	0,37	0,18	0,37	0,18	0,37	0,18
		SGED	0,55	0,50	0,86	0,58	0,62	0,58	0,62	0,39	0,15	0,06	0,38	0,07	0,38	0,07	0,38	0,07	0,38	0,07	0,38	0,07
	[0,05, 0,95]	ST	0,69	0,62	0,98	0,57	0,56	0,57	0,56	0,38	0,39	0,51	0,43	0,37	0,43	0,37	0,43	0,37	0,43	0,37	0,43	0,37
		N	0,12	0,01	0,38	0,03	0,01	0,03	0,01	0,20	0,37	0,91	0,52	0,58	0,20	0,37	0,91	0,52	0,58	0,20	0,37	0,91
		N/G	0,66	0,64	0,49	0,45	0,25	0,45	0,25	0,11	0,10	0,64	0,22	0,24	0,11	0,10	0,64	0,22	0,24	0,11	0,10	0,64
		Hyp	0,81	0,55	0,73	0,54	0,42	0,54	0,42	0,16	0,15	0,61	0,21	0,27	0,16	0,15	0,61	0,21	0,27	0,16	0,15	0,61
	[0, 0,05]	SGED	0,60	0,43	0,88	0,52	0,48	0,52	0,48	0,13	0,02	0,49	0,08	0,09	0,13	0,02	0,49	0,08	0,09	0,13	0,02	0,49
		ST	0,64	0,59	0,02	0,37	0,27	0,37	0,27	0,15	0,15	0,38	0,21	0,21	0,15	0,15	0,38	0,21	0,21	0,15	0,15	0,38
		N	0,30	0,44	0,03	0,32	0,37	0,32	0,37	0,31	0,10	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
		N/G	0,97	0,91	0,97	0,59	0,97	0,59	0,97	0,64	0,48	0,23	0,09	0,02	0,64	0,48	0,23	0,09	0,02	0,64	0,48	0,23
	[0,95, 1]	Hyp	0,96	0,87	0,97	0,53	0,93	0,95	0,71	0,06	0,44	0,13	0,06	0,02	0,06	0,44	0,13	0,06	0,02	0,06	0,44	0,13
		SGED	0,89	0,87	0,98	0,60	0,97	0,97	0,91	0,07	0,49	0,08	0,04	0,01	0,07	0,49	0,08	0,04	0,01	0,07	0,49	0,08
		ST	0,98	0,91	0,97	0,50	0,95	0,95	0,68	0,30	0,43	0,29	0,11	0,12	0,30	0,43	0,29	0,11	0,12	0,30	0,43	0,29
		N	0,27	0,48	0,06	0,04	0,38	0,23	0,02	0,06	0,21	0,20	0,24	0,08	0,03	0,21	0,20	0,24	0,08	0,03	0,21	0,20
[0,1, 0,9]	N/G	0,97	0,96	0,17	0,55	0,92	0,83	0,85	0,27	0,21	0,50	0,35	0,30	0,27	0,21	0,50	0,35	0,30	0,27	0,21	0,50	
	Hyp	1,00	0,94	0,15	0,54	0,95	0,92	0,90	0,22	0,12	0,41	0,38	0,22	0,22	0,12	0,41	0,38	0,22	0,22	0,12	0,41	
	SGED	0,84	0,87	0,74	0,32	0,88	0,81	0,67	0,06	0,13	0,30	0,32	0,06	0,06	0,13	0,30	0,32	0,06	0,06	0,13	0,30	
	ST	0,98	0,95	1,00	0,57	0,97	0,98	0,78	0,30	0,20	0,36	0,37	0,34	0,29	0,20	0,36	0,37	0,34	0,29	0,20	0,36	
	N	0,39	0,16	0,66	0,23	0,14	0,23	0,14	0,36	0,50	0,70	0,77	0,89	0,36	0,50	0,70	0,77	0,89	0,36	0,50	0,70	
	N/G	0,15	0,46	0,10	0,45	0,41	0,45	0,41	0,11	0,33	0,33	0,18	0,16	0,11	0,33	0,33	0,18	0,16	0,11	0,33	0,33	
	Hyp	0,14	0,36	0,12	0,34	0,29	0,34	0,29	0,12	0,35	0,36	0,13	0,18	0,12	0,35	0,36	0,13	0,18	0,12	0,35	0,36	
	SGED	0,14	0,45	0,29	0,29	0,37	0,29	0,37	0,01	0,08	0,40	0,04	0,11	0,01	0,08	0,40	0,04	0,11	0,01	0,08	0,40	
	ST	0,37	0,48	0,08	0,53	0,45	0,53	0,45	0,20	0,36	0,52	0,38	0,35	0,20	0,36	0,52	0,38	0,35	0,20	0,36	0,52	
	N	0,33	0,71	0,19	0,25	0,13	0,12	0,17	0,30	0,17	0,01	0,01	0,27	0,02	0,01	0,01	0,27	0,02	0,01	0,01	0,01	
	N/G	0,41	0,58	0,05	0,49	0,50	0,50	0,65	0,58	0,45	0,23	0,14	0,24	0,41	0,45	0,23	0,14	0,24	0,41	0,45	0,23	
	Hyp	0,22	0,43	0,69	0,52	0,21	0,31	0,59	0,51	0,33	0,02	0,07	0,34	0,11	0,33	0,02	0,07	0,34	0,11	0,33	0,02	
	SGED	0,39	0,69	0,78	0,49	0,40	0,52	0,75	0,80	0,49	0,05	0,03	0,52	0,32	0,49	0,05	0,03	0,52	0,32	0,49	0,05	
	ST	0,16	0,30	0,55	0,54	0,14	0,21	0,48	0,45	0,33	0,40	0,20	0,37	0,29	0,33	0,40	0,20	0,37	0,29	0,33	0,40	
	N	0,08	0,36	0,05	0,06	0,03	0,02	0,03	0,03	0,11	0,19	0,22	0,32	0,16	0,11	0,19	0,22	0,32	0,16	0,11	0,19	
	N/G	0,10	0,38	0,45	0,36	0,07	0,08	0,37	0,08	0,12	0,44	0,56	0,07	0,11	0,12	0,44	0,56	0,07	0,11	0,12	0,44	
Hyp	0,28	0,54	0,65	0,33	0,25	0,28	0,39	0,09	0,11	0,37	0,51	0,07	0,09	0,11	0,37	0,51	0,07	0,09	0,11	0,37		
SGED	0,29	0,71	0,62	0,24	0,25	0,32	0,37	0,04	0,05	0,01	0,42	0,03	0,01	0,05	0,01	0,42	0,03	0,01	0,05	0,01		
ST	0,09	0,31	0,36	0,46	0,07	0,07	0,07	0,08	0,12	0,41	0,50	0,08	0,09	0,12	0,41	0,50	0,08	0,09	0,12	0,41		

Okres	Fragment dystrybuanty	Indeks		WIG20						FTSE100								
		Rozkład	KS**	V**	AD**	AD ^{down} AD ^{up} **	W ^{2**}	AD ^{2**}	AD ^{2**down} AD ^{2**up}	KS**	V**	AD**	AD ^{down} AD ^{up} **	W ^{2**}	AD ^{2**}	AD ^{2**down} AD ^{2**up}		
2020	[0, 1]	N	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	
		NIG	0,48	0,34	0,07			0,24	0,27		0,53	0,46	0,14		0,31	0,36		
		Hyp	0,10	0,04	0,01			0,11	0,11		0,20	0,35	0,04		0,15	0,11		
		SGED	0,07	0,04	0,01			0,09	0,13		0,11	0,12	0,05		0,05	0,03		
		ST	0,55	0,51	0,71			0,30	0,28		0,61	0,43	0,65		0,37	0,31		
	[0,05, 0,95]	N	0,01	0,01	0,12	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
		NIG	0,32	0,51	0,13	0,03	0,03	0,50	0,37	0,05	0,02	0,08	0,31	0,18	0,05	0,14	0,23	
		Hyp	0,61	0,29	0,03	0,03	0,02	0,61	0,33	0,02	0,14	0,11	0,07	0,06	0,27	0,29	0,06	
		SGED	0,61	0,47	0,01	0,01	0,01	0,66	0,38	0,01	0,13	0,17	0,05	0,02	0,34	0,39	0,02	
		ST	0,69	0,74	0,71	0,28	0,28	0,70	0,71	0,87	0,06	0,06	0,35	0,54	0,13	0,21	0,62	
	[0,95, 1]	N	0,78	0,92	0,92	0,18	0,18	0,72	0,82	0,76	0,03	0,03	0,01	0,01	0,06	0,02	0,01	
		NIG	0,22	0,66	0,02	0,51	0,51	0,30	0,19	0,22	0,24	0,19	0,21	0,07	0,22	0,26	0,08	
		Hyp	0,79	0,68	0,85	0,45	0,45	0,90	0,92	0,60	0,59	0,20	0,09	0,04	0,49	0,43	0,04	
		SGED	0,55	0,71	0,51	0,56	0,56	0,76	0,67	0,50	0,38	0,17	0,21	0,08	0,36	0,24	0,08	
		ST	0,12	0,47	0,01	0,51	0,51	0,14	0,09	0,13	0,61	0,49	0,78	0,18	0,62	0,69	0,60	
[0,1, 0,9]	N	0,01	0,01	0,10	0,10	0,10	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	
	NIG	0,15	0,33	0,43			0,39	0,57		0,50	0,35	0,02		0,40	0,27			
	Hyp	0,12	0,07	0,86			0,18	0,39		0,46	0,40	0,48		0,34	0,23			
	SGED	0,02	0,04	0,17			0,07	0,19		0,05	0,07	0,27		0,04	0,03			
	ST	0,34	0,49	0,98			0,27	0,51		0,65	0,39	0,91		0,40	0,53			
	[0, 0,1]	N	0,15	0,49	0,01	0,01	0,01	0,01	0,03	0,01	0,01	0,05	0,01	0,01	0,01	0,01	0,01	
		NIG	0,75	0,89	0,09	0,09	0,09	0,57	0,62	0,02	0,19	0,07	0,39	0,11	0,13	0,21	0,12	
		Hyp	0,48	0,80	0,02	0,02	0,02	0,14	0,12	0,02	0,30	0,20	0,03	0,02	0,31	0,29	0,03	
		SGED	0,45	0,85	0,02	0,02	0,02	0,19	0,12	0,02	0,13	0,10	0,08	0,04	0,15	0,19	0,04	
		ST	0,19	0,61	0,63	0,16	0,16	0,10	0,16	0,35	0,05	0,03	0,47	0,33	0,02	0,02	0,50	
	[0,9, 1]	N	0,27	0,63	0,61	0,23	0,23	0,13	0,12	0,23	0,55	0,21	0,01	0,01	0,58	0,25	0,01	
		NIG	0,73	0,77	0,71	0,47	0,47	0,85	0,70	0,29	0,50	0,47	0,17	0,04	0,49	0,60	0,05	
		Hyp	0,84	0,85	0,74	0,46	0,46	0,93	0,89	0,76	0,30	0,53	0,11	0,03	0,26	0,32	0,05	
		SGED	0,85	0,92	0,85	0,59	0,59	0,92	0,88	0,59	0,14	0,28	0,26	0,09	0,18	0,22	0,10	
		ST	0,59	0,83	0,66	0,54	0,54	0,80	0,68	0,25	0,61	0,50	0,89	0,19	0,69	0,82	0,65	

* Wartości p znajdujące się poniżej przyjętego poziomu istotności ($\alpha = 0,05$) zaznaczono pogrubioną czcionką.

Źródło: obliczenia własne

Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Rok	Parametry kopuli		Miary zależności kopuli			Parametry rozkładu indeksu WIG20				Parametry rozkładu indeksu FTSE100				
				1	2	τ Kendall	λ^L	λ^U	1	2	3	4	1	2	3	4	
t-Studenta (p, v)	N(μ, σ)	N(μ, σ)	2019	0,4466	14,5099	0,2947	0,0273	0,0273	-0,0097	0,9265			0,0532	0,7362			
			2020	0,7719	7,8050	0,5614	0,3153	0,3153	-0,0211	2,1265			-0,0442	1,8005			
		NIG($\mu, \delta, \alpha, \beta$)	2019	0,4409	9,1194	0,2907	0,0753	0,0753	-0,0079	0,9304			0,0713	0,9992	1,8972	-0,0161	
			2020	0,7260	120,5176	0,5173	0,0000	0,0000	-0,0186	2,0646			0,1369	1,8450	0,6408	-0,0535	
		Hyp($\mu, \delta, \alpha, \beta$)	2019	0,4411	9,3852	0,2908	0,0715	0,0715	-0,0079	0,9297			0,0790	0,6838	2,5386	-0,0327	
			2020	0,7403	7,2064	0,5307	0,2999	0,2999	-0,0302	2,1226			0,2069	1,2226	0,9813	-0,0581	
		SGED($\mu, \sigma, \alpha, \kappa$)	2019	0,4429	9,6710	0,2921	0,0680	0,0680	-0,0082	0,9300			0,0608	0,7284	1,3258	0,9805	
			2020	0,7328	35,3113	0,5235	0,0234	0,0234	-0,0283	2,0808			-0,0254	1,7209	1,1828	0,9624	
		ST(ξ, ω, α, ν)	2019	0,4401	8,7501	0,2901	0,0809	0,0809	-0,0071	0,9308			0,0429	0,6002	0,0418	6,2183	
			2020	0,7226	120,3338	0,5141	0,0000	0,0000	-0,0402	2,0655			0,4849	1,3950	-0,4417	5,0824	
		NIG($\mu, \delta, \alpha, \beta$)	N(μ, σ)	2019	0,4451	15,2937	0,2937	0,0234	0,0234	-0,1717	1,1396	1,3551	0,1981	0,0501	0,7349		
				2020	0,7249	49,2580	0,5162	0,0067	0,0067	-0,1791	1,8924	0,5172	0,0572	-0,0584	1,8411		
	NIG($\mu, \delta, \alpha, \beta$)		2019	0,4350	7,0420	0,2865	0,1129	0,1129	-0,1449	1,0707	1,2479	0,1680	0,0765	0,9440	1,7830	-0,0287	
			2020	0,7192	6,3519	0,5110	0,3078	0,3078	-0,0005	1,3704	0,3347	0,0014	0,1814	1,4028	0,4819	-0,0747	
	Hyp($\mu, \delta, \alpha, \beta$)		2019	0,4356	7,3880	0,2869	0,1052	0,1052	-0,1466	1,0761	1,2582	0,1702	0,0841	0,6166	2,4562	-0,0457	
			2020	0,7025	4,0538	0,4958	0,3900	0,3900	-0,0890	1,3686	0,3548	0,0248	0,0761	0,0020	0,8371	-0,0472	
	SGED($\mu, \sigma, \alpha, \kappa$)		2019	0,4379	7,5477	0,2885	0,1026	0,1026	-0,1450	1,0734	1,2523	0,1674	0,0590	0,7297	1,2863	0,9784	
			2020	0,7185	6,1925	0,5103	0,3127	0,3127	-0,0496	1,3994	0,3495	0,0129	-0,0521	1,7449	0,9507	0,9491	
	ST(ξ, ω, α, ν)		2019	0,4337	6,6403	0,2856	0,1223	0,1223	-0,1446	1,0659	1,2404	0,1688	0,0549	0,5900	0,0182	5,6927	
			2020	0,7198	6,7821	0,5116	0,2937	0,2937	0,0161	1,3508	0,3269	-0,0086	0,4950	1,2500	-0,4728	3,5740	
	Hyp($\mu, \delta, \alpha, \beta$)		N(μ, σ)	2019	0,4447	15,7890	0,2934	0,0213	0,0213	-0,1560	0,5272	1,7617	0,1780	0,0499	0,7352		
				2020	0,7428	6,2011	0,5330	0,3359	0,3359	-0,2950	1,2556	0,8625	0,0997	-0,0344	1,8787		
		NIG($\mu, \delta, \alpha, \beta$)	2019	0,4351	6,9610	0,2866	0,1148	0,1148	-0,1369	0,4290	1,6857	0,1574	0,0786	0,9383	1,7668	-0,0331	
			2020	0,7203	5,6367	0,5120	0,3353	0,3353	0,0506	0,1698	0,7162	-0,0071	0,1377	1,4843	0,5125	-0,0500	
		Hyp($\mu, \delta, \alpha, \beta$)	2019	0,4356	7,3220	0,2870	0,1067	0,1067	-0,1380	0,4376	1,6936	0,1590	0,0863	0,6089	2,4436	-0,0502	
			2020	0,7207	4,1466	0,5124	0,4016	0,4016	0,0501	0,0033	0,7078	-0,0165	0,0775	0,0012	0,8092	-0,0397	
		SGED($\mu, \sigma, \alpha, \kappa$)	2019	0,4379	7,4539	0,2886	0,1046	0,1046	-0,1369	0,4319	1,6882	0,1568	0,0586	0,7307	1,2813	0,9780	
			2020	0,7048	3,8844	0,4979	0,4009	0,4009	0,0554	0,0137	0,7200	-0,0167	-0,0348	1,7129	1,0132	0,9524	
		ST(ξ, ω, α, ν)	2019	0,4337	6,5541	0,2856	0,1246	0,1246	-0,1368	0,4226	1,6810	0,1587	0,0600	0,5894	0,0084	5,6291	
			2020	0,7028	3,9132	0,4961	0,3976	0,3976	0,0795	0,0002	0,7029	-0,0318	0,4007	1,2148	-0,3838	3,2515	
		SGED($\mu, \sigma, \alpha, \kappa$)	N(μ, σ)	2019	0,4444	15,6596	0,2932	0,0217	0,0217	-0,0020	0,9249	1,2885	1,0864	0,0513	0,7349		
				2020	0,7379	19,3705	0,5284	0,0947	0,0947	0,0376	1,9263	1,2396	1,0498	-0,0486	1,8472		
	NIG($\mu, \delta, \alpha, \beta$)		2019	0,4341	7,1870	0,2859	0,1091	0,1091	0,0041	0,9298	1,2372	1,0594	0,0826	0,9419	1,7774	-0,0355	
			2020	0,7232	7,0253	0,5147	0,2890	0,2890	0,0034	2,0030	0,9058	0,9727	0,1944	1,4462	0,5028	-0,0751	
	Hyp($\mu, \delta, \alpha, \beta$)		2019	0,4347	7,5473	0,2863	0,1014	0,1014	0,0039	0,9286	1,2410	1,0598	0,0902	0,6129	2,4508	-0,0526	
			2020	0,7043	3,9187	0,4975	0,3986	0,3986	-0,0067	2,0084	0,9066	0,9736	0,0995	0,0196	0,8093	-0,0397	
	SGED($\mu, \sigma, \alpha, \kappa$)		2019	0,4370	7,6881	0,2879	0,0993	0,0993	0,0031	0,9292	1,2392	1,0591	0,0610	0,7302	1,2836	0,9795	
			2020	0,7185	5,1724	0,5104	0,3525	0,3525	0,0046	2,0001	0,9204	0,9778	-0,0326	1,7424	0,9653	0,9558	
	ST(ξ, ω, α, ν)		2019	0,4329	6,7862	0,2850	0,1183	0,1183	0,0054	0,9305	1,2342	1,0601	0,0662	0,5898	0,0018	5,6708	
			2020	0,7216	7,2481	0,5132	0,2805	0,2805	-0,0217	2,0112	0,8937	0,9638	0,5154	1,2663	-0,4830	3,6934	
	ST(ξ, ω, α, ν)		N(μ, σ)	2019	0,4462	15,7842	0,2945	0,0215	0,0215	-0,3413	0,8136	0,5007	6,1115	0,0490	0,7347		
				2020	0,7202	84,8267	0,5119	0,0003	0,0003	-0,3635	1,5025	0,2762	4,7954	-0,0639	1,8437		
		NIG($\mu, \delta, \alpha, \beta$)	2019	0,4358	7,1406	0,2871	0,1108	0,1108	-0,2865	0,7825	0,4196	5,3586	0,0758	0,9480	1,7908	-0,0292	
			2020	0,7201	6,9575	0,5118	0,2883	0,2883	-0,0659	1,2585	0,0696	2,9450	0,1958	1,3655	0,4593	-0,0774	
		Hyp($\mu, \delta, \alpha, \beta$)	2019	0,4364	7,4801	0,2875	0,1035	0,1035	-0,2901	0,7843	0,4252	5,4139	0,0834	0,6216	2,4619	-0,0461	
			2020	0,7074	4,0384	0,5002	0,3951	0,3951	-0,1090	1,2634	0,0698	2,7795	0,0834	0,0000	0,7932	-0,0459	
		SGED($\mu, \sigma, \alpha, \kappa$)	2019	0,4385	7,6560	0,2890	0,1005	0,1005	-0,2862	0,7827	0,4180	5,3832	0,0580	0,7297	1,2893	0,9778	
			2020	0,7203	7,5387	0,5120	0,2705	0,2705	-0,1355	1,2742	0,1050	3,0482	-0,0579	1,7693	0,9271	0,9464	
ST(ξ, ω, α, ν)		2019	0,4345	6,7404	0,2862	0,1201	0,1201	-0,2866	0,7818	0,4216	5,3116	0,0545	0,5907	0,0169	5,7223		
		2020	0,7187	6,5458	0,5105	0,3006	0,3006	-0,0390	1,2424	0,0443	2,8656	0,4815	1,2295	-0,4457	3,4370		

Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Rok	Parametry kopuli		Miary zależności kopuli			Parametry rozkładu indeksu WIG20				Parametry rozkładu indeksu FTSE100				
				1	2	τ Kendall	χ^L	χ^U	1	2	3	4	1	2	3	4	
Claytona (θ)	N(μ, σ)	N(μ, σ)	2019	0,5850		0,2263	0,3058	0,0000	-0,0148	0,9099			0,0504	0,7464			
			2020	1,4627		0,4224	0,6226	0,0000	0,0316	2,1531			-0,0286	1,7843			
		NIG(μ, δ, α, β)	2019	0,7313		0,2677	0,3876	0,0000	-0,0188	0,9362			0,1630	0,9830	1,8854	-0,2266	
			2020	1,1506		0,3652	0,5475	0,0000	0,1101	2,2271			0,2508	2,2063	0,8738	-0,1005	
		Hyp(μ, δ, α, β)	2019	0,7265		0,2665	0,3852	0,0000	-0,0183	0,9346			0,1755	0,6517	2,5165	-0,2518	
			2020	1,0071		0,3349	0,5024	0,0000	0,1536	2,1715			0,2742	0,3011	0,9475	-0,1154	
		SGED(μ, σ, α, κ)	2019	0,7229		0,2655	0,3833	0,0000	-0,0198	0,9336			0,0407	0,7345	1,2917	0,9361	
			2020	1,2535		0,3853	0,5752	0,0000	0,0958	2,2223			-0,0223	1,6304	1,3727	0,9441	
		ST(ξ, ω, α, ν)	2019	0,7315		0,2678	0,3877	0,0000	-0,0181	0,9364			0,2432	0,6222	-0,3685	6,1572	
			2020	1,1905		0,3731	0,5586	0,0000	0,0859	2,2435			0,8701	1,6390	-0,8265	9,3827	
		NIG(μ, δ, α, β)	N(μ, σ)	2019	0,5623		0,2195	0,2915	0,0000	-0,1336	1,1936	1,4846	0,1508	0,0542	0,7454		
				2020	1,4503		0,4203	0,6201	0,0000	0,3725	6,6956	1,8745	-0,1069	0,0276	1,9505		
	NIG(μ, δ, α, β)		2019	0,7408		0,2703	0,3923	0,0000	-0,0489	1,1000	1,2533	0,0319	0,1814	0,9211	1,7429	-0,2548	
			2020	1,7620		0,4684	0,6748	0,0000	0,2244	1,3093	0,2754	-0,0562	0,4144	1,3384	0,4311	-0,1467	
	Hyp(μ, δ, α, β)		2019	0,7309		0,2676	0,3874	0,0000	-0,0541	1,1073	1,2697	0,0394	0,1919	0,5760	2,4139	-0,2762	
			2020	1,5096		0,4301	0,6318	0,0000	0,0899	1,5195	0,3985	-0,0158	0,0872	0,0059	0,8143	-0,0368	
	SGED(μ, σ, α, κ)		2019	0,7238		0,2657	0,3838	0,0000	-0,0588	1,1094	1,2771	0,0428	0,0420	0,7405	1,2478	0,9349	
			2020	1,6949		0,4587	0,6643	0,0000	0,1976	1,4134	0,3285	-0,0524	-0,0550	1,8861	0,9003	0,8674	
	ST(ξ, ω, α, ν)		2019	0,7418		0,2706	0,3928	0,0000	-0,0481	1,0976	1,2511	0,0323	0,2790	0,6214	-0,4351	5,5123	
			2020	1,7754		0,4703	0,6768	0,0000	0,2209	1,2707	0,2552	-0,0562	0,9238	1,3916	-0,8981	3,0884	
	Hyp(μ, δ, α, β)		N(μ, σ)	2019	0,5640		0,2200	0,2926	0,0000	-0,1226	0,6185	1,8735	0,1364	0,0541	0,7463		
				2020	1,1496		0,3650	0,5472	0,0000	-0,0690	0,5771	0,8212	0,0275	0,0639	1,9768		
		NIG(μ, δ, α, β)	2019	0,7441		0,2712	0,3940	0,0000	-0,0569	0,4727	1,6908	0,0404	0,1830	0,9176	1,7330	-0,2584	
			2020	1,5734		0,4403	0,6437	0,0000	0,2173	0,0264	0,6943	-0,0642	0,1043	1,1087	0,3278	-0,0468	
		Hyp(μ, δ, α, β)	2019	0,7342		0,2685	0,3890	0,0000	-0,0608	0,4837	1,7032	0,0462	0,1936	0,5707	2,4052	-0,2798	
			2020	1,5994		0,4443	0,6483	0,0000	0,2225	0,1651	0,7211	-0,0656	0,1872	0,0301	0,8046	-0,0708	
		SGED(μ, σ, α, κ)	2019	0,7268		0,2665	0,3853	0,0000	-0,0648	0,4864	1,7085	0,0488	0,0414	0,7416	1,2444	0,9343	
			2020	1,6918		0,4583	0,6638	0,0000	0,2441	0,0000	0,7015	-0,0688	-0,0479	1,7828	1,0451	0,8662	
		ST(ξ, ω, α, ν)	2019	0,7449		0,2714	0,3943	0,0000	-0,0565	0,4708	1,6905	0,0412	0,2826	0,6222	-0,4431	5,4755	
			2020	1,5686		0,4395	0,6428	0,0000	0,1938	0,0005	0,6923	-0,0512	0,4146	1,1574	-0,3419	2,7646	
		SGED(μ, σ, α, κ)	N(μ, σ)	2019	0,5679		0,2211	0,2951	0,0000	-0,0082	0,9000	1,3095	1,0468	0,0556	0,7467		
				2020	1,3754		0,4075	0,6041	0,0000	-0,0022	1,8326	1,4876	0,9906	0,0519	1,9654		
	NIG(μ, δ, α, β)		2019	0,7514		0,2731	0,3975	0,0000	-0,0192	0,9338	1,2292	1,0035	0,1860	0,9171	1,7299	-0,2613	
			2020	1,7483		0,4664	0,6727	0,0000	-0,0331	2,2105	0,8016	0,9094	0,4284	1,4315	0,4836	-0,1545	
	Hyp(μ, δ, α, β)		2019	0,7417		0,2705	0,3928	0,0000	-0,0176	0,9310	1,2353	1,0052	0,1968	0,5688	2,4011	-0,2828	
			2020	1,6102		0,4460	0,6502	0,0000	0,0111	2,0688	0,8748	0,9534	0,2834	0,1479	0,8263	-0,1217	
	SGED(μ, σ, α, κ)		2019	0,7339		0,2685	0,3889	0,0000	-0,0190	0,9290	1,2383	1,0050	0,0428	0,7424	1,2427	0,9347	
			2020	1,7141		0,4615	0,6674	0,0000	-0,0326	2,1180	0,8635	0,9241	-0,0517	1,8422	0,9588	0,8668	
	ST(ξ, ω, α, ν)		2019	0,7520		0,2733	0,3978	0,0000	-0,0180	0,9334	1,2277	1,0042	0,2867	0,6232	-0,4487	5,4616	
			2020	1,7552		0,4674	0,6737	0,0000	-0,0281	2,2568	0,7791	0,9119	0,9643	1,4448	-0,9400	3,4592	
	ST(ξ, ω, α, ν)		N(μ, σ)	2019	0,5647		0,2202	0,2931	0,0000	-0,2648	0,7889	0,3748	6,7868	0,0532	0,7454		
				2020	1,5814		0,4416	0,6451	0,0000	1,4012	2,3797	-1,1172	91,1922	0,0208	1,9625		
		NIG(μ, δ, α, β)	2019	0,7419		0,2706	0,3929	0,0000	-0,0653	0,7543	0,0597	5,4827	0,1802	0,9243	1,7484	-0,2543	
			2020	1,7345		0,4645	0,6706	0,0000	0,4417	1,2952	-0,3078	2,5422	0,3992	1,2766	0,3969	-0,1358	
		Hyp(μ, δ, α, β)	2019	0,7321		0,2680	0,3880	0,0000	-0,0790	0,7557	0,0808	5,5749	0,1908	0,5804	2,4181	-0,2758	
			2020	1,5694		0,4397	0,6430	0,0000	0,3822	1,2707	-0,3279	2,8350	0,3778	0,0089	0,8532	-0,1720	
		SGED(μ, σ, α, κ)	2019	0,7253		0,2661	0,3846	0,0000	-0,0869	0,7559	0,0895	5,6126	0,0411	0,7407	1,2503	0,9345	
			2020	1,7067		0,4604	0,6662	0,0000	0,4744	1,3599	-0,3394	2,9246	-0,0582	1,9279	0,8727	0,8419	
		ST(ξ, ω, α, ν)	2019	0,7430		0,2709	0,3934	0,0000	-0,0651	0,7535	0,0611	5,4628	0,2773	0,6219	-0,4335	5,5357	
			2020	1,7231		0,4628	0,6688	0,0000	0,3994	1,2622	-0,2685	2,4247	0,8621	1,3260	-0,8023	2,9413	

Załącznik 6. Oszacowania wartości zagrożonych dla rozkładów dwuwymiarowych

Rok	Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Dolny ogon			Górny ogon			
				Rząd kwantyla						
				0,1	0,05	0,01	0,9	0,95	0,99	
2019	Gaussa(ρ)	N(μ, σ)	N(μ, σ)	-0,8844 *	-1,1538	-1,6491 **	0,9128	1,1774	1,6738	
			NIG($\mu, \delta, \alpha, \beta$)	-0,8665	-1,1375	-1,6507 *	0,9117	1,1702	1,7294	
			Hyp($\mu, \delta, \alpha, \beta$)	-0,8670	-1,1599	-1,6606 *	0,8927	1,1428	1,6951	
			SGED($\mu, \sigma, \alpha, \kappa$)	-0,8623	-1,1275	-1,6705 *	0,9091	1,1434	1,6387	
			ST(ξ, ω, α, ν)	-0,8928 *	-1,1367	-1,6781 *	0,9090	1,1736	1,6758	
		NIG($\mu, \delta, \alpha, \beta$)	N(μ, σ)	-0,8576	-1,1019	-1,6777 *	0,8900	1,1877	1,9308	
			NIG($\mu, \delta, \alpha, \beta$)	-0,8356	-1,1373	-1,7675	0,8839	1,1711	1,8864	
			Hyp($\mu, \delta, \alpha, \beta$)	-0,8440	-1,1377	-1,7827	0,8854	1,1863	1,8468	
			SGED($\mu, \sigma, \alpha, \kappa$)	-0,8256	-1,1380	-1,7369	0,9068	1,1943	1,9236	
			ST(ξ, ω, α, ν)	-0,8016	-1,0690	-1,6864 *	0,8827	1,2068	1,8962	
		Hyp($\mu, \delta, \alpha, \beta$)	N(μ, σ)	-0,8536	-1,1279	-1,6647 *	0,8892	1,1807	1,7884	
			NIG($\mu, \delta, \alpha, \beta$)	-0,8523	-1,1450	-1,7744	0,8715	1,1781	1,8384	
			Hyp($\mu, \delta, \alpha, \beta$)	-0,8308	-1,0922	-1,7229 *	0,8836	1,2039	1,9245	
			SGED($\mu, \sigma, \alpha, \kappa$)	-0,8618	-1,1436	-1,8233	0,8918	1,2054	1,8409	
			ST(ξ, ω, α, ν)	-0,8327	-1,1086	-1,7502	0,8986	1,2009	1,9189	
		SGED($\mu, \sigma, \alpha, \kappa$)	N(μ, σ)	-0,8502	-1,1137	-1,7094 *	0,8968	1,1916	1,7485	
			NIG($\mu, \delta, \alpha, \beta$)	-0,8383	-1,1444	-1,8560	0,9016	1,1843	1,8629	
			Hyp($\mu, \delta, \alpha, \beta$)	-0,8345	-1,1506	-1,7536	0,8751	1,1549	1,7502	
			SGED($\mu, \sigma, \alpha, \kappa$)	-0,8597	-1,1473	-1,7049 *	0,9059	1,2496	1,8741	
			ST(ξ, ω, α, ν)	-0,7907	-1,0788	-1,6735 *	0,8887	1,1833	1,8753	
		ST(ξ, ω, α, ν)	N(μ, σ)	-0,8653	-1,1271	-1,6835 *	0,8947	1,1878	1,8835	
			NIG($\mu, \delta, \alpha, \beta$)	-0,8253	-1,1008	-1,7877	0,8501	1,1421	1,9102	
			Hyp($\mu, \delta, \alpha, \beta$)	-0,8389	-1,1089	-1,7285 *	0,8767	1,2001	1,8607	
			SGED($\mu, \sigma, \alpha, \kappa$)	-0,8423	-1,1104	-1,7214 *	0,8552	1,1793	1,9136	
			ST(ξ, ω, α, ν)	-0,8429	-1,1289	-1,7602	0,8646	1,1825	1,9166	
		t-Studenta (ρ, ν)	N(μ, σ)	N(μ, σ)	-0,8634	-1,1166	-1,6176 ***	0,9161	1,1882	1,7013
				NIG($\mu, \delta, \alpha, \beta$)	-0,8562	-1,1397	-1,7282 *	0,8915	1,1823	1,7749
				Hyp($\mu, \delta, \alpha, \beta$)	-0,8711	-1,1644	-1,6191 ***	0,9200	1,1776	1,7698
				SGED($\mu, \sigma, \alpha, \kappa$)	-0,8576	-1,1280	-1,7043 *	0,9221	1,1909	1,7472
				ST(ξ, ω, α, ν)	-0,8486	-1,0987	-1,6807 *	0,9005	1,2004	1,7315
			NIG($\mu, \delta, \alpha, \beta$)	N(μ, σ)	-0,8269	-1,0935	-1,6321 ***	0,8776	1,1630	1,8978
				NIG($\mu, \delta, \alpha, \beta$)	-0,7778	-1,0472	-1,7744	0,8892	1,2044	1,9784
				Hyp($\mu, \delta, \alpha, \beta$)	-0,7898	-1,0774	-1,7423	0,8715	1,1815	1,9097
				SGED($\mu, \sigma, \alpha, \kappa$)	-0,8187	-1,0956	-1,7183 *	0,8993	1,1967	1,9647
				ST(ξ, ω, α, ν)	-0,8001	-1,0846	-1,7647	0,8755	1,1725	1,9676
			Hyp($\mu, \delta, \alpha, \beta$)	N(μ, σ)	-0,8445	-1,1074	-1,6899 *	0,8857	1,2118	1,8810
				NIG($\mu, \delta, \alpha, \beta$)	-0,8034	-1,1082	-1,7522	0,8939	1,1847	1,8778
				Hyp($\mu, \delta, \alpha, \beta$)	-0,7883	-1,0752	-1,7862	0,8964	1,1964	1,9922
				SGED($\mu, \sigma, \alpha, \kappa$)	-0,8086	-1,1102	-1,8211	0,8653	1,1821	1,9629
				ST(ξ, ω, α, ν)	-0,7888	-1,0801	-1,7398	0,8686	1,1815	1,9315
			SGED($\mu, \sigma, \alpha, \kappa$)	N(μ, σ)	-0,8474	-1,1009	-1,6860 *	0,8996	1,1979	1,8288
				NIG($\mu, \delta, \alpha, \beta$)	-0,8122	-1,1125	-1,8011	0,8879	1,2219	1,9189
				Hyp($\mu, \delta, \alpha, \beta$)	-0,8053	-1,1000	-1,7833	0,8726	1,1760	1,8531
				SGED($\mu, \sigma, \alpha, \kappa$)	-0,7929	-1,0864	-1,7735	0,8838	1,1837	1,8758
				ST(ξ, ω, α, ν)	-0,7919	-1,0914	-1,6681 *	0,8721	1,1873	1,8348
			ST(ξ, ω, α, ν)	N(μ, σ)	-0,8343	-1,1023	-1,6873 *	0,9191	1,2142	1,8939
				NIG($\mu, \delta, \alpha, \beta$)	-0,7916	-1,0972	-1,7875	0,8598	1,1791	1,8410
				Hyp($\mu, \delta, \alpha, \beta$)	-0,7835	-1,0552	-1,6587 *	0,8545	1,1684	1,8975
				SGED($\mu, \sigma, \alpha, \kappa$)	-0,8164	-1,0746	-1,7182 *	0,8934	1,2113	1,9879
				ST(ξ, ω, α, ν)	-0,7832	-1,0479	-1,6820 *	0,8779	1,1888	1,9486

Rok	Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Dolny ogon			Górny ogon		
				Rząd kwantyla					
				0,1	0,05	0,01	0,9	0,95	0,99
2019	Claytona (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-0,8685	-1,1819	-1,8138	0,8389	1,0648	1,4724
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8996 *	-1,2360	-1,9029	0,8579	1,0726	1,4905
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8856 *	-1,2183	-1,9443	0,8521	1,0711	1,4631
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,9082 *	-1,2371	-1,9381	0,8695	1,0729	1,4940
			$ST(\xi, \omega, \alpha, \nu)$	-0,8957 *	-1,1904	-1,8671	0,8492	1,0705	1,4543
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,8238	-1,1142	-1,7625	0,8316	1,0772	1,5287
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8518	-1,2183	-2,0562	0,8270	1,0819	1,6397
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8563	-1,1971	-2,0493	0,8346	1,0764	1,6309
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8502	-1,2021	-1,9823	0,8265	1,0787	1,6407
			$ST(\xi, \omega, \alpha, \nu)$	-0,8286	-1,1720	-2,0274	0,8310	1,0876	1,6500
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,8352	-1,1494	-1,7479	0,8570	1,1017	1,5994
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8494	-1,2093	-2,0329	0,8281	1,0647	1,5953
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8707	-1,2350	-2,1232	0,8320	1,0971	1,6332
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8343	-1,1809	-2,0587	0,8303	1,0824	1,6094
			$ST(\xi, \omega, \alpha, \nu)$	-0,8408	-1,1893	-2,0486	0,8354	1,0860	1,6407
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-0,8231	-1,1296	-1,7823	0,8574	1,0842	1,6007
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8563	-1,1784	-1,9898	0,8497	1,1112	1,6599
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8203	-1,2088	-2,0991	0,8278	1,0634	1,5826
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8423	-1,1932	-1,9985	0,8382	1,0759	1,5790
			$ST(\xi, \omega, \alpha, \nu)$	-0,8257	-1,1753	-2,0145	0,8365	1,0866	1,5830
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-0,8440	-1,1177	-1,6946 *	0,8345	1,0636	1,5711
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8473	-1,1946	-2,0481	0,8214	1,0715	1,5773
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8535	-1,2105	-1,9199	0,8319	1,0511	1,5341
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8351	-1,2001	-2,0653	0,8230	1,0632	1,6052
	$ST(\xi, \omega, \alpha, \nu)$		-0,8729	-1,2537	-2,1156	0,8205	1,0456	1,6093	
	Gumbela (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-0,8769	-1,1309	-1,5891 ***	0,9693 *	1,2688	1,8634
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8401	-1,0921	-1,5872 ***	0,9391 *	1,2425	1,9498
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8552	-1,0948	-1,5928 ***	0,9433 *	1,2516	1,9209
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8336	-1,0781	-1,5592 ***	0,9194	1,2361	1,9974
			$ST(\xi, \omega, \alpha, \nu)$	-0,8387	-1,0822	-1,5675 ***	0,9551 *	1,2701	2,0032
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,8302	-1,1016	-1,6108 ***	0,9462 *	1,3102	2,1757 **
			$NIG(\mu, \delta, \alpha, \beta)$	-0,7998	-1,0684	-1,6429 ***	0,9310 *	1,2957	2,2538 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8242	-1,0972	-1,6615 *	0,9162	1,2755	2,2512 **
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8103	-1,0765	-1,6715 *	0,8905	1,2621	2,1817 **
			$ST(\xi, \omega, \alpha, \nu)$	-0,8095	-1,0637	-1,6511 *	0,9464 *	1,3470	2,2194 **
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,8191	-1,0720	-1,6820 *	0,9660 *	1,3164	2,1497 **
			$NIG(\mu, \delta, \alpha, \beta)$	-0,7928	-1,0871	-1,7504	0,9213	1,2877	2,1015 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8029	-1,0711	-1,6744 *	0,9471 *	1,3556	2,3065 **
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,7813	-1,0422	-1,6922 *	0,9346 *	1,2948	2,2081 **
			$ST(\xi, \omega, \alpha, \nu)$	-0,7985	-1,0679	-1,6903 *	0,9354 *	1,3051	2,2363 **
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-0,8315	-1,0914	-1,6243 ***	0,9364 *	1,2783	2,0503 **
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8084	-1,0828	-1,6540 *	0,9284 *	1,2565	2,1274 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8229	-1,0899	-1,6364 ***	0,9015	1,2583	2,0469 **
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,7973	-1,0787	-1,6925 *	0,9406 *	1,2805	2,0934 **
			$ST(\xi, \omega, \alpha, \nu)$	-0,8121	-1,1065	-1,6427 ***	0,9048	1,2670	2,1105 **
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-0,8067	-1,0410	-1,5946 ***	0,9543 *	1,2785	2,0667 **
			$NIG(\mu, \delta, \alpha, \beta)$	-0,7986	-1,0897	-1,6604 *	0,9101	1,2773	2,1666 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8036	-1,0634	-1,6881 *	0,9110	1,2876	2,3058 **
$SGED(\mu, \sigma, \alpha, \kappa)$			-0,8068	-1,0876	-1,6533 *	0,9063	1,2608	2,1646 **	
$ST(\xi, \omega, \alpha, \nu)$			-0,7790	-1,0429	-1,6831 *	0,9242	1,2697	2,3905 **	

Rok	Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Dolny ogon			Górny ogon				
				Rząd kwantyla							
				0,1	0,05	0,01	0,9	0,95	0,99		
2019	Franka (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-0,9650 *	-1,1885	-1,6021 ***	0,9637 *	1,2278	1,6493		
			$NIG(\mu, \delta, \alpha, \beta)$	-0,9092 *	-1,1739	-1,6453 ***	0,9053	1,1571	1,5721		
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,9570 *	-1,1923	-1,6296 ***	0,8910	1,1295	1,5683		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,9603 *	-1,2208	-1,7246 *	0,9177	1,1488	1,5697		
			$ST(\xi, \omega, \alpha, \nu)$	-0,9119 *	-1,1715	-1,6211 ***	0,9107	1,1600	1,5885		
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,8979 *	-1,1572	-1,7019 *	0,9024	1,1626	1,7463		
			$NIG(\mu, \delta, \alpha, \beta)$	-0,9023 *	-1,1889	-1,7076 *	0,8460	1,1112	1,6875		
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8811	-1,1319	-1,7293 *	0,8774	1,1415	1,7270		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8630	-1,1533	-1,7434	0,8758	1,1483	1,7614		
			$ST(\xi, \omega, \alpha, \nu)$	-0,8688	-1,1436	-1,7542	0,8446	1,1005	1,7352		
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,9269 *	-1,1903	-1,7099 *	0,9097	1,1810	1,7086		
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8821	-1,1684	-1,7568	0,8847	1,1667	1,7491		
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8686	-1,1562	-1,6962 *	0,8753	1,1559	1,7241		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,9039 *	-1,1769	-1,7042 *	0,8404	1,1172	1,6408		
			$ST(\xi, \omega, \alpha, \nu)$	-0,8764	-1,1393	-1,7635	0,8587	1,1552	1,7129		
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-0,9248 *	-1,1588	-1,6914 *	0,8936	1,1647	1,6437		
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8783	-1,1615	-1,7859	0,8466	1,1334	1,7048		
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,8832	-1,1843	-1,7295 *	0,8599	1,1285	1,6465		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8701	-1,1278	-1,5961 ***	0,8582	1,1375	1,6917		
			$ST(\xi, \omega, \alpha, \nu)$	-0,8707	-1,1581	-1,7542	0,8669	1,1536	1,7186		
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-0,9130 *	-1,1823	-1,6930 *	0,8733	1,1211	1,6447		
			$NIG(\mu, \delta, \alpha, \beta)$	-0,8842	-1,1612	-1,8237	0,8635	1,1204	1,7117		
			$Hyp(\mu, \delta, \alpha, \beta)$	-0,9071 *	-1,1544	-1,7420	0,8877	1,1730	1,7765		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8965 *	-1,1812	-1,7292 *	0,8691	1,1336	1,7301		
			$ST(\xi, \omega, \alpha, \nu)$	-0,8790	-1,1756	-1,7949	0,8511	1,1143	1,7526		
		Joe (θ)		$N(\mu, \sigma)$	$N(\mu, \sigma)$	-0,8084	-1,0317	-1,4423 ***	0,8813	1,1642	1,7280
					$NIG(\mu, \delta, \alpha, \beta)$	-0,8029	-1,0252	-1,4921 ***	0,8747	1,1712	1,8472
					$Hyp(\mu, \delta, \alpha, \beta)$	-0,8036	-1,0280	-1,4923 ***	0,8358	1,1225	1,7787
					$SGED(\mu, \sigma, \alpha, \kappa)$	-0,8175	-1,0416	-1,5414 ***	0,8778	1,1903	1,8566
					$ST(\xi, \omega, \alpha, \nu)$	-0,8039	-1,0373	-1,5470 ***	0,8454	1,1477	1,7938
				$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,7598	-0,9982	-1,5487 ***	0,9165	1,2822	2,1718 **
					$NIG(\mu, \delta, \alpha, \beta)$	-0,7914	-1,0669	-1,6426 ***	0,8581	1,2397	2,2144 **
					$Hyp(\mu, \delta, \alpha, \beta)$	-0,7853	-1,0303	-1,5354 ***	0,8819	1,2840	2,2601 **
					$SGED(\mu, \sigma, \alpha, \kappa)$	-0,7916	-1,0371	-1,5362 ***	0,9013	1,2914	2,2137 **
					$ST(\xi, \omega, \alpha, \nu)$	-0,7851	-1,0313	-1,6023 ***	0,8592	1,2403	2,1308 **
				$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-0,8001	-1,0408	-1,5680 ***	0,8768	1,2267	2,0528 **
					$NIG(\mu, \delta, \alpha, \beta)$	-0,7884	-1,0553	-1,6065 ***	0,8734	1,2700	2,2377 **
					$Hyp(\mu, \delta, \alpha, \beta)$	-0,7933	-1,0455	-1,6019 ***	0,8789	1,2694	2,2007 **
					$SGED(\mu, \sigma, \alpha, \kappa)$	-0,7717	-1,0196	-1,5977 ***	0,8770	1,2470	2,1657 **
					$ST(\xi, \omega, \alpha, \nu)$	-0,7717	-1,0176	-1,6279 ***	0,8862	1,2535	2,2850 **
				$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-0,7690	-1,0149	-1,4850 ***	0,8498	1,2087	1,9323
					$NIG(\mu, \delta, \alpha, \beta)$	-0,7922	-1,0363	-1,6217 ***	0,8707	1,2321	2,1040 **
					$Hyp(\mu, \delta, \alpha, \beta)$	-0,7908	-1,0543	-1,6028 ***	0,8578	1,2378	2,1274 **
					$SGED(\mu, \sigma, \alpha, \kappa)$	-0,7946	-1,0640	-1,5950 ***	0,8965	1,2979	2,1587 **
					$ST(\xi, \omega, \alpha, \nu)$	-0,7752	-1,0311	-1,5929 ***	0,8634	1,2679	2,2708 **
				$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-0,7866	-1,0170	-1,5770 ***	0,8769	1,2299	2,0999 **
					$NIG(\mu, \delta, \alpha, \beta)$	-0,7913	-1,0525	-1,6607 *	0,8470	1,1981	2,2573 **
					$Hyp(\mu, \delta, \alpha, \beta)$	-0,7912	-1,0544	-1,6758 *	0,8314	1,1801	2,2107 **
$SGED(\mu, \sigma, \alpha, \kappa)$	-0,7782				-1,0284	-1,6311 ***	0,8658	1,2626	2,2347 **		
$ST(\xi, \omega, \alpha, \nu)$	-0,7681				-1,0264	-1,6704 *	0,8440	1,2227	2,2608 **		

Rok	Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Dolny ogon			Górny ogon				
				Rząd kwantyla							
				0,1	0,05	0,01	0,9	0,95	0,99		
2020	Gaussa(ρ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-2,3541	-3,0386	-4,3490	2,2601 **	2,9011	4,2046		
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1975	-2,8597	-4,3780	2,1220	2,7247	4,1168		
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,1578	-2,8938	-4,4361	2,1929 *	2,9033	4,3911		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,2895	-3,0403	-4,5353	2,1218	2,7995	4,2399		
			$ST(\xi, \omega, \alpha, \nu)$	-2,2625	-2,9317	-4,5482	2,0894	2,7658	4,1631		
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,1010	-2,8176	-4,2710	2,0968	2,7950	4,3502		
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9758	-2,8706	-4,9599	1,8635	2,6186	4,3227		
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,9305	-2,7389	-4,6243	1,8109	2,6154	4,5111		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-1,9875	-2,7823	-4,8977	1,9194	2,7356	4,7566		
			$ST(\xi, \omega, \alpha, \nu)$	-1,9143	-2,7941	-5,0735	1,8352	2,5984	4,4130		
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,0962	-2,8430	-4,2312	2,1460	2,9286	4,5941		
			$NIG(\mu, \delta, \alpha, \beta)$	-1,8759	-2,6857	-4,7910	1,9119	2,6727	4,5140		
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,2594	-3,2377	-5,4241	1,7598	2,4749	4,1303		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0814	-2,8711	-4,9489	1,7674	2,5351	4,3896		
			$ST(\xi, \omega, \alpha, \nu)$	-1,9754	-2,8230	-5,0978	1,7926	2,4852	4,0920		
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-2,1535	-2,9606	-4,5427	2,1372	2,8195	4,3666		
			$NIG(\mu, \delta, \alpha, \beta)$	-2,0628	-3,0072	-5,1137	1,9039	2,6143	4,3196		
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,0396	-2,8614	-4,7468	1,7962	2,6079	4,4714		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0710	-2,9486	-4,9035	1,8949	2,6875	4,4691		
			$ST(\xi, \omega, \alpha, \nu)$	-2,0327	-2,9016	-5,4416	1,8196	2,5302	4,3371		
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-2,1331	-2,8077	-4,2386	2,0952	2,7355	4,2241		
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9415	-2,7727	-5,0470	1,8840	2,6588	4,6845		
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,9480	-2,7904	-4,9186	1,8562	2,6520	5,2332		
			$SGED(\mu, \sigma, \alpha, \kappa)$	-1,9378	-2,7240	-4,8189	1,9155	2,7202	4,6037		
			$ST(\xi, \omega, \alpha, \nu)$	-1,9064	-2,6940	-5,2440	1,8313	2,5430	4,3878		
		t-Studenta (ρ, ν)		$N(\mu, \sigma)$	$N(\mu, \sigma)$	-2,3771	-3,1371	-4,4011	2,3166 **	2,9868	4,2766
					$NIG(\mu, \delta, \alpha, \beta)$	-2,1433	-2,8430	-4,1201	2,1881 *	2,8397	4,2811
					$Hyp(\mu, \delta, \alpha, \beta)$	-2,2206	-2,9475	-4,4325	2,2071 *	2,9093	4,4581
					$SGED(\mu, \sigma, \alpha, \kappa)$	-2,1898	-2,9040	-4,4268	2,1952 *	2,8617	4,3583
					$ST(\xi, \omega, \alpha, \nu)$	-2,2149	-2,9093	-4,4645	2,0298	2,6483	3,9401
				$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,1209	-2,7654	-4,2199	2,1374	2,8769	4,5017
					$NIG(\mu, \delta, \alpha, \beta)$	-1,9020	-2,7707	-5,0078	1,8061	2,5435	4,4051
					$Hyp(\mu, \delta, \alpha, \beta)$	-1,8347	-2,6927	-4,9471	1,7999	2,6432	4,8926
					$SGED(\mu, \sigma, \alpha, \kappa)$	-1,8216	-2,6269	-4,6723	1,8909	2,7620	4,7562
					$ST(\xi, \omega, \alpha, \nu)$	-1,8933	-2,7786	-5,3862	1,8001	2,5086	4,5375
				$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,1235	-2,7964	-4,3086	2,2763 **	3,0270	4,7902
					$NIG(\mu, \delta, \alpha, \beta)$	-1,9366	-2,7576	-4,7328	1,9103	2,6817	4,6103
					$Hyp(\mu, \delta, \alpha, \beta)$	-2,0792	-2,9380	-4,9051	1,9046	2,6803	4,5855
					$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0181	-2,9123	-4,8609	1,9234	2,7258	4,5882
					$ST(\xi, \omega, \alpha, \nu)$	-1,9805	-2,8008	-4,7357	1,8423	2,5556	4,5399
				$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-2,1445	-2,8395	-4,2387	2,2165 *	2,9468	4,4586
					$NIG(\mu, \delta, \alpha, \beta)$	-1,9316	-2,7962	-4,8529	1,8820	2,5730	4,5910
					$Hyp(\mu, \delta, \alpha, \beta)$	-2,0045	-2,8596	-5,0421	1,8889	2,7382	4,8645
					$SGED(\mu, \sigma, \alpha, \kappa)$	-1,9457	-2,8402	-4,7876	1,8898	2,6570	4,8345
					$ST(\xi, \omega, \alpha, \nu)$	-1,9618	-2,8205	-4,9488	1,8631	2,6492	4,3778
				$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-2,1446	-2,8312	-4,2488	2,1341	2,8302	4,4805
					$NIG(\mu, \delta, \alpha, \beta)$	-1,9293	-2,7042	-5,0081	1,8356	2,6786	4,9954
					$Hyp(\mu, \delta, \alpha, \beta)$	-1,9300	-2,8071	-5,2019	1,8643	2,6922	4,8094
$SGED(\mu, \sigma, \alpha, \kappa)$	-1,9276				-2,7889	-4,8883	1,8847	2,6542	4,9432		
$ST(\xi, \omega, \alpha, \nu)$	-1,8853				-2,7143	-5,0903	1,8527	2,6222	4,9742		

Rok	Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Dolny ogon			Górny ogon		
				Rząd kwantyla					
				0,1	0,05	0,01	0,9	0,95	0,99
2020	Claytona (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-2,3564	-3,1653	-4,5670	2,1814 *	2,6810	3,5776 *
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1146	-2,9425	-4,5067	2,0779	2,5687	3,5839 *
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,0404	-2,8038	-4,4473	1,9958	2,4971	3,4817 *
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,1906	-2,9628	-4,4158	2,0844	2,5931	3,6237 *
			$ST(\xi, \omega, \alpha, \nu)$	-2,2150	-2,9871	-4,5465	2,0566	2,5844	3,5182 *
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,2671	-2,9983	-4,4391	2,0992	2,5946	3,5502 *
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1744	-3,4722	-7,0060	1,8121	2,4217	3,8270
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,9481	-2,8752	-5,3942	1,8195	2,4002	3,6481 *
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,1819	-3,2989	-6,1152	1,8286	2,4306	3,7183
			$ST(\xi, \omega, \alpha, \nu)$	-2,2016	-3,3823	-7,4112	1,8351	2,4336	4,0297
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,0905	-2,8412	-4,5753	2,0391	2,5674	3,6581 *
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1164	-3,1802	-5,9430	1,8484	2,4977	3,8217
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,1262	-3,1071	-5,7362	1,8224	2,3849	3,5796 *
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,2120	-3,2618	-5,5535	1,8291	2,4050	3,6311 *
			$ST(\xi, \omega, \alpha, \nu)$	-2,0228	-3,1500	-6,0171	1,8702	2,5170	3,8973
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-2,1563	-2,9054	-4,3858	2,0827	2,6454	3,6061 *
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1829	-3,3815	-6,2825	1,9096	2,5063	3,8506
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,0775	-3,0956	-5,7562	1,7912	2,3884	3,8322
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,1914	-3,2791	-5,9414	1,8675	2,5095	3,8697
			$ST(\xi, \omega, \alpha, \nu)$	-2,2708	-3,4048	-7,0076	1,8709	2,4883	3,9780
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-2,3237	-3,0835	-4,4844	2,1360	2,6338	3,5538 *
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9822	-3,1264	-7,1116	1,8587	2,4852	4,1360
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,0336	-2,9932	-5,8851	1,6749	2,1915	3,5172 *
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,1069	-3,2085	-6,3730	1,8491	2,4582	3,8362
	$ST(\xi, \omega, \alpha, \nu)$		-2,0122	-3,0879	-7,0803	1,8106	2,4329	4,0566	
	Gumbela (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-2,4027	-3,0839	-4,3674	2,6003 ***	3,3998 *	4,8332
			$NIG(\mu, \delta, \alpha, \beta)$	-2,2301	-2,9151	-4,4748	2,2652 **	3,1137 *	5,0405
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,1966	-2,9349	-4,3567	2,3770 **	3,2072 *	5,0601
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,2655	-2,9565	-4,4231	2,3604 **	3,1411 *	4,8731
			$ST(\xi, \omega, \alpha, \nu)$	-2,1832	-2,9017	-4,5638	2,2684 **	3,0238	5,0026
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,1394	-2,7930	-4,2476	2,4124 **	3,3942 *	5,8555
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9222	-2,7369	-4,7236	2,0252	3,0007	6,0616 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,8719	-2,6015	-4,5763	1,7380	2,6208	5,2930
			$SGED(\mu, \sigma, \alpha, \kappa)$	-1,9423	-2,7490	-4,8480	2,1224	3,1109 *	5,9195
			$ST(\xi, \omega, \alpha, \nu)$	-1,9157	-2,7197	-4,9272	1,9076	2,9362	6,1732 **
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,1930	-2,8865	-4,3817	2,3983 **	3,3237 *	5,2851
			$NIG(\mu, \delta, \alpha, \beta)$	-1,8460	-2,6233	-4,3393	1,8282	2,6805	4,9746
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,0162	-2,7702	-4,4604	1,9282	2,7879	4,9273
			$SGED(\mu, \sigma, \alpha, \kappa)$	-1,9776	-2,7418	-4,4613	2,1125	3,0925 *	5,2355
			$ST(\xi, \omega, \alpha, \nu)$	-1,9435	-2,6848	-4,6602	1,9522	2,7967	5,2970
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-2,2146	-2,8986	-4,5059	2,3350 **	3,1977 *	5,2377
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9744	-2,7860	-4,6349	1,9936	2,9133	5,2959
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,1004	-2,8895	-4,8956	1,7556	2,6268	4,6434
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0271	-2,8570	-4,7440	2,0818	3,0878 *	5,6256
			$ST(\xi, \omega, \alpha, \nu)$	-1,9257	-2,8333	-5,1064	1,9237	2,8564	5,2811
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-2,1105	-2,6750	-4,3160	2,3945 **	3,3576 *	6,4364 **
			$NIG(\mu, \delta, \alpha, \beta)$	-1,8539	-2,6302	-4,8090	1,9858	2,9455	6,1227 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,8307	-2,6056	-4,6306	2,0685	3,0264	6,0244 **
$SGED(\mu, \sigma, \alpha, \kappa)$			-1,9554	-2,7005	-4,7530	1,9800	3,0156	5,7807	
$ST(\xi, \omega, \alpha, \nu)$	-1,8000		-2,5890	-4,5667	1,9657	2,9432	5,7528		

Rok	Kopula	Rozkłady brzegowe (WIG20)	Rozkłady brzegowe (FTSE100)	Dolny ogon			Górny ogon		
				Rząd kwantyla					
				0,1	0,05	0,01	0,9	0,95	0,99
2020	Franka (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-2,6953	-3,3284	-4,4318	2,5363 **	3,1644 *	4,3085
			$NIG(\mu, \delta, \alpha, \beta)$	-2,5791	-3,3523	-4,8455	2,2745 **	2,8930	3,9822
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,6793	-3,4706	-4,9757	2,1960 *	2,8277	4,0684
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,6536	-3,3990	-4,8630	2,2521 **	2,8735	3,9378
			$ST(\xi, \omega, \alpha, \nu)$	-2,5487	-3,2719	-4,9226	2,2513 **	2,8313	3,8363
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,4691	-3,3001	-5,0110	2,2562 **	2,9179	4,4263
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1977	-3,0188	-5,0536	1,9569	2,6506	4,2342
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,1995	-3,0001	-5,0291	2,0385	2,7719	4,5538
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,2593	-3,0805	-4,7250	1,9185	2,6608	4,2957
			$ST(\xi, \omega, \alpha, \nu)$	-2,2048	-3,0702	-5,1822	1,9352	2,6386	4,2437
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,4304	-3,0910	-4,4068	2,2953 **	2,9620	4,1359
			$NIG(\mu, \delta, \alpha, \beta)$	-2,2339	-3,0931	-4,9514	1,9689	2,7021	4,2232
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,3595	-3,1057	-4,7724	2,0617	2,8153	4,2336
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,3410	-3,2035	-4,8504	2,0306	2,7936	4,2393
			$ST(\xi, \omega, \alpha, \nu)$	-2,2424	-3,1587	-5,6923	1,7165	2,3242	3,6624 *
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-2,5332	-3,2937	-4,6943	2,2731 **	2,9383	4,3486
			$NIG(\mu, \delta, \alpha, \beta)$	-2,3717	-3,2816	-5,1748	2,0085	2,6758	4,1208
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,4117	-3,2725	-5,0301	1,8754	2,5271	3,8512
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,3183	-3,1591	-4,8996	2,0158	2,6728	4,0893
			$ST(\xi, \omega, \alpha, \nu)$	-2,2349	-3,0958	-5,2016	1,9390	2,6153	4,1603
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-2,4322	-3,0856	-4,6090	2,2974 **	3,0089	4,5372
			$NIG(\mu, \delta, \alpha, \beta)$	-2,1675	-2,9771	-4,9309	1,9626	2,6698	4,3877
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,2273	-3,1530	-5,7111	1,9101	2,6156	4,4432
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,1292	-2,8799	-4,8202	1,9806	2,7113	4,3082
	$ST(\xi, \omega, \alpha, \nu)$		-2,1135	-2,9512	-5,0876	1,9289	2,6018	4,3831	
	Joe (θ)	$N(\mu, \sigma)$	$N(\mu, \sigma)$	-2,4696	-3,0319	-4,0594	2,5487 **	3,3234 *	4,8038
			$NIG(\mu, \delta, \alpha, \beta)$	-2,2540	-2,8738	-4,3795	2,2760 **	3,2281 *	5,3030
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,2718	-2,8483	-4,1062	2,2421 *	3,1381 *	5,0585
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,2432	-2,8848	-4,2152	2,2839 **	3,2412 *	5,2292
			$ST(\xi, \omega, \alpha, \nu)$	-2,2366	-2,8288	-4,2032	2,2031 *	3,0709 *	5,4096
		$NIG(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,1679	-2,7795	-4,2615	2,3053 **	3,5329 **	6,5206 **
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9962	-2,7222	-4,9003	2,0678	3,2067 *	6,9643 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,8771	-2,5373	-3,9092	2,0348	3,1145 *	6,2590 **
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0637	-2,8101	-4,7690	2,0482	3,2203 **	6,5345 **
			$ST(\xi, \omega, \alpha, \nu)$	-1,8897	-2,6544	-4,9137	1,9936	3,0750 *	6,7732 **
		$Hyp(\mu, \delta, \alpha, \beta)$	$N(\mu, \sigma)$	-2,2497	-2,8272	-3,9709	2,3387 **	3,3146 *	5,3737
			$NIG(\mu, \delta, \alpha, \beta)$	-2,0317	-2,7369	-4,3870	2,0162	2,9770	5,6746
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,9215	-2,5627	-3,8090 *	2,1081	2,9888	5,2279
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0831	-2,8699	-4,4831	2,0025	3,0207	5,6605
			$ST(\xi, \omega, \alpha, \nu)$	-1,9551	-2,6320	-4,5048	1,9520	2,9108	5,4734
		$SGED(\mu, \sigma, \alpha, \kappa)$	$N(\mu, \sigma)$	-2,1698	-2,7964	-4,2670	2,3116 **	3,2850 *	5,5629
			$NIG(\mu, \delta, \alpha, \beta)$	-2,0190	-2,7791	-4,8905	2,1941 *	3,4057 *	6,9885 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-2,0510	-2,7783	-4,5032	1,7199	2,6038	4,9422
			$SGED(\mu, \sigma, \alpha, \kappa)$	-2,0372	-2,8000	-4,4583	2,1364	3,3472 *	6,2742 **
			$ST(\xi, \omega, \alpha, \nu)$	-1,9568	-2,6802	-4,7384	2,1930 *	3,3475 *	7,4966 **
		$ST(\xi, \omega, \alpha, \nu)$	$N(\mu, \sigma)$	-2,1657	-2,7603	-4,4433	2,3517 **	3,4517 *	7,1349 **
			$NIG(\mu, \delta, \alpha, \beta)$	-1,9234	-2,7686	-4,9296	1,9755	3,2171 *	7,9822 **
			$Hyp(\mu, \delta, \alpha, \beta)$	-1,9141	-2,5933	-4,5463	1,9875	3,0074	6,6502 **
$SGED(\mu, \sigma, \alpha, \kappa)$			-1,9789	-2,7000	-4,9313	2,0020	3,0980 *	6,7915 **	
$ST(\xi, \omega, \alpha, \nu)$			-1,8809	-2,6585	-5,0576	2,0574	3,2888 *	8,1146 **	

Źródło: obliczenia własne