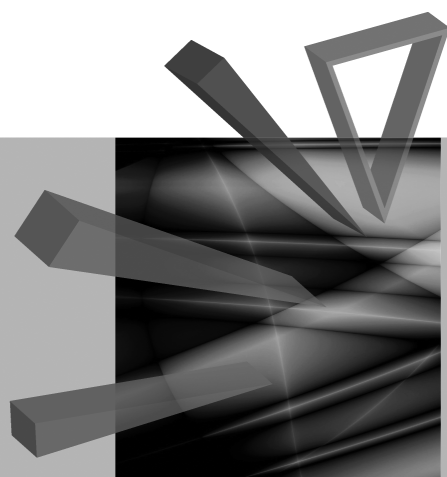


Prognozowanie w zarządzaniu firmą



Redaktorzy naukowi
Paweł Dittmann
Aleksandra Szpulak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2011

Senacka Komisja Wydawnicza

Zdzisław Pisz (przewodniczący),

*Andrzej Bąk, Krzysztof Jajuga, Andrzej Matysiak, Waldemar Podgórski,
Mieczysław Przybyła, Aniela Styś, Stanisław Urban*

Recenzenci

Włodzimierz Szkutnik, Jan Zawadzki

Redakcja wydawnicza

Barbara Majewska

Redakcja techniczna i korekta

Barbara Łopusiewicz

Skład i łamanie

Comp-rajt

Projekt okładki

Beata Dębska

Kopiowanie i powielanie w jakiegokolwiek formie
wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu
Wrocław 2011

ISSN 1899-3192

ISBN 978-83-7695-141-6

Druk: Drukarnia TOTEM

Spis treści

Wstęp	7
Agnieszka Przybylska-Mazur: Optymalne zasady polityki pieniężnej w prognozowaniu wskaźnika inflacji	9
Alicja Wolny-Dominiak: Zmodyfikowana regresja Poissona dla danych ubezpieczeniowych z dużą liczbą zer	21
Andrzej Gajda: Doświadczenia i metody pozyskiwania danych eksperckich na potrzeby badań z wykorzystaniem metod foresight	30
Anna Gondek: Prognozy rozwoju gospodarczego Polski z użyciem metody analogii przestrzenno-czasowych	41
Bartosz Lawędziak: Sekurytyzacja papierów wartościowych opartych na hipotece odwrotnej	50
Filip Chybalski: Prakseologiczne aspekty prognozowania	59
Ireneusz Kuroпка, Paweł Lenczewski: Możliwość zastosowania modeli ekonometrycznych do prognozowania w przedsiębiorstwie Brenntag Polska	69
Jacek Szandula: Wyszukiwanie formacji w kursach giełdowych przy użyciu metod klasyfikacji danych	82
Joanna Perzyńska: Zastosowanie sztucznych sieci neuronowych do wyznaczania nieliniowych prognoz kombinowanych	94
Konstancja Poradowska, Tomasz Szkutnik, Mirosław Wójciak: Scenariusze rozwoju wybranych technologii oszczędności energii w życiu codziennym	102
Maciej Oesterreich: Wykorzystanie pakietu statystycznego R w prognozowaniu na podstawie danych w postaci szeregów czasowych z wahaniami sezonowymi	113
Marcin Błażejowski, Paweł Kufel, Tadeusz Kufel: Algorytm zgodnego modelowania i prognozowania procesów ekonomicznych jako pakiet funkcji <i>Congruent Specification</i> programu Gretl	125
Marcin Błażejowski: Stacjonarność szeregów czasowych o wysokiej częstotliwości obserwowania – implementacja testu stacjonarności Dickeya w programie Gretl	137
Mirosław Wójciak: Wpływ czynników i zdarzeń kluczowych na rozwój nowych technologii – wybrane metody korygowania prognoz na przykładzie technologii energooszczędnych	149
Monika Dyduch: Grupowanie produktów strukturyzowanych	159
Piotr Bernat: Planowanie działalności przedsiębiorstwa wspomagane prognozowaniem	170

Roman Pawlukowicz: Informacje prognostyczne w rynkowych sposobach wyceny nieruchomości – identyfikacja i pozyskiwanie	182
Wojciech Zatoń: Uwarunkowania psychologiczne w prognozowaniu	189

Summaries

Agnieszka Przybylska-Mazur: Optimal monetary policy rules in forecasting of inflation rate	20
Alicja Wolny-Dominiak: Zero-inflated Poisson Model for insurance data with a large number of zeros	29
Andrzej Gajda: Experience and methods of data collection from experts for research using foresight methods	40
Anna Gondek: Economic growth forecasts for Poland using the time-space analogy method	49
Bartosz Lawędziak: Securitization of survivor bonds based on the reverse mortgage	58
Filip Chybalski: Praxiological aspects of forecasting	68
Ireneusz Kuropka, Paweł Lenczewski: Econometric models usage feasibility in Brenntag Poland forecasting	81
Jacek Szandula: Searching for technical analysis formations in stock prices with the use of cluster analysis methods	93
Joanna Perzyńska: Application of artificial neural networks to build the nonlinear combined forecasts	101
Konstancja Poradowska, Tomasz Szkutnik, Mirosław Wójciak: The scenarios of development of selected technologies related to energy saving in everyday life	112
Maciej Oesterreich: The R application in forecasting unsystematic lacks in seasonal time series	124
Marcin Błażejowski, Paweł Kufel, Tadeusz Kufel: Congruent modelling and forecasting algorithm as function package Congruent Specification in GRETL	136
Marcin Błażejowski: Stationarity of high-frequency time series – implementation of Dickey’s stationarity test in GRETL	148
Mirosław Wójciak: The influence of key and events factors on the development of new technologies – selected methods of forecast correction on the example of energy-saving technologies	158
Monika Dyduch: Ranking of structured products	169
Piotr Bernat: Forecasting assisted business management planning	181
Roman Pawlukowicz: Prognostic data in market ways of property valuation – identification and acquisition	188
Wojciech Zatoń: Psychological aspects of forecasting	199

Marcin Błażejowski

Wyższa Szkoła Bankowa w Toruniu

STACJONARNOŚĆ SZEREGÓW CZASOWYCH O WYSOKIEJ CZĘSTOTLIWOŚCI OBSERWOWANIA – IMPLEMENTACJA TESTU STACJONARNOŚCI DICKEYA W PROGRAMIE GRETL

Streszczenie: Celem artykułu jest ocena mocy testu stacjonarności szeregów czasowych o wysokiej częstotliwości obserwowania zaproponowanego przez D.A. Dickeya w 2009 r., weryfikującego hipotezę o sezonowej integracji procesu $SI_d(1)$. Test ten rozszerza zastosowanie testu sezonowego pierwiastka jednostkowego DHF o przypadki częstotliwości cyklu dla $d = 5, 6, 7, 21, 24, 26, 31, 48, 52, 168, 365...$ Ponadto zaprezentowano implementację testu stacjonarności Dickeya i testu DHF jako pakietu funkcji w programie Gretl.

Słowa kluczowe: procesy sezonowo zintegrowane, test stacjonarności Dickeya, modelowanie cykliczności dla danych o wysokiej częstotliwości obserwowania, Gretl.

1. Wstęp

„W analizie sezonowości szeregów czasowych o wysokiej częstotliwości obserwowania badacz często staje przed problemem określenia tego, czy wahania sezonowe mają charakter deterministyczny i są stałe w czasie, czy mają charakter stochastyczny i przez to zmieniają się w czasie. Prawidłowe rozróżnienie charakteru cykliczności ma zasadnicze znaczenie dla wyników analizy oraz jakości generowanych prognoz” [Dickey 2009; Dickey, Zhang 2010]. Dotychczas znane były dwa testy sezonowych pierwiastków jednostkowych:

1. Test DHF dla danych półrocznych, kwartalnych i miesięcznych [Dickey i in. 1984].
2. Test HEGY dla danych kwartalnych [Hylleberg i in. 1990].

W 2009 r. D.A. Dickey opublikował nowy test pozwalający na weryfikację występowania pierwiastka jednostkowego dla cykli o okresie innym niż 2, 4 i 12. Szczególnie interesująca jest możliwość testowania pierwiastków jednostkowych w danych o wysokiej częstotliwości obserwowania o cyklach 5, 6 i 7 (cykl o okresie tygodnia dla danych dziennych), 52 (sezonowy pierwiastek jednostkowy dla danych tygodniowych) czy 365 (sezonowy pierwiastek jednostkowy dla danych dziennych) oraz cykle o różnym okresie dla danych godzinowych.

W artykule przedstawione zostaną podstawowe informacje o konstrukcji tego testu, o jego implementacji w programie Gretl oraz wyniki weryfikacji hipotezy o sezonowym pierwiastku jednostkowym dla wybranych danych empirycznych.

2. Test stacjonarności Dickeya dla danych o wysokiej częstotliwości obserwowania

Zaproponowany przez Dickeya test stacjonarności dla danych o wysokiej częstotliwości obserwowania w swojej konstrukcji bardzo przypomina wcześniejsze testy pierwiastka jednostkowego, których współautorem był Dickey: ADF (*Augmented Dickey Fuller*) do weryfikacji hipotezy o niesezonowym pierwiastku jednostkowym oraz DHF (Dickey, Hasza, Fuller) do weryfikacji hipotezy o sezonowym pierwiastku jednostkowym dla danych półrocznych, kwartalnych oraz miesięcznych. W teście stawia się następujące hipotezy:

$$H_0: \delta_d = 0,$$

$$H_1: \delta_d < 0.$$

Statystyka testu τ , którą uzyskuje się w czteroetapowej procedurze testowej, ma niestandardowy rozkład (podobnie jak w testach ADF oraz DHF), jednak po pewnej korekcie sprowadza się ją do zmiennej o asymptotycznym rozkładzie $N(0,1)$. Procedura składa się z następujących czterech etapów.

1. W pierwszym etapie następuje oczyszczenie testowanego szeregu Y_t ze składników deterministycznych, takich jak trend oraz cykliczność. Ze względu na możliwe wysokie wartości d i tym samym potencjalnie znaczną liczbę zmiennych zero-jedynkowych Dickey sugeruje, aby deterministyczną cykliczność filtrować wybranymi harmonikami. Niemniej jednak wykorzystanie zmiennych zero-jedynkowych jest jak najbardziej możliwe. Do estymacji wykorzystuje się KMNK i w efekcie uzyskuje się nowy proces y_t .

2. W drugim etapie szacuje się model autoregresyjny $AR(p)$ dla sezonowych różnic reszt po wyeliminowaniu elementów deterministycznych z testowanego szeregu, czyli dla procesu $\Delta_d y_t$. Model ten powinien się charakteryzować resztami o własnościach białego szumu. Do estymacji wykorzystuje się algorytmy stosowane w estymacji modeli ARIMA (czyli *de facto* szacuje się odpowiedni model $ARMA(p,0)$).

3. W trzecim etapie wykorzystuje się oceny oszacowanego w etapie 2 modelu $AR(p)$ do przefiltrowania reszt po usunięciu składników deterministycznych, czyli procesu y_t . Jeżeli H_0 jest prawdziwa, to uzyskany na drodze tej filtracji nowy proces z_t powinien mieć charakter sezonowego błędzenia przypadkowego.

4. W czwartym etapie za pomocą KMNK szacuje się regresję sezonowych różnic procesu z_t uzyskanego w etapie 3 jako funkcji opóźnienia rzędu d poziomów tego procesu oraz opóźnionych sezonowych różnic procesu uzyskanego w etapie 1 (czyli procesu y_t) dla opóźnień $s = 1, \dots, p$ według następującej formuły:

$$\Delta_d z_t = \alpha z_{t-d} + \sum_{s=1}^p \beta_s \Delta_d y_{t-s} + v_t,$$

gdzie v_t jest składnikiem losowym. Statystykę τ uzyskuje się według formuły:

$$\tau = \frac{\hat{\beta}}{S(\hat{\beta})},$$

gdzie $\hat{\beta}$ jest oceną parametru β , natomiast $S(\hat{\beta})$ jest błędem standardowym tej oceny.

Uzyskana statystyka τ ma niestandardowy rozkład, jednak dokonując następującej korekty, uzyskuje się zmienną o rozkładzie $N(0,1)$:

$$u = \tau + \frac{1 + k\sqrt{2}}{2\sqrt{d}}, \quad (1)$$

gdzie k oznacza liczbę deterministycznych składników wykorzystanych w etapie 1 procedury testowej.

Dowód prawdziwości korekty (1) znajduje się z dodatku do artykułu [Dickey 2009]. Należy zauważyć, że każda uwzględniona zmienna deterministyczna zwiększa wartość korekty i tym samym przesuwą wartość statystyki testu w kierunku obszaru braku podstaw do odrzuceń H_0 . W celu sprawdzenia wpływu włączenia składników do opisu cyklu deterministycznego przeprowadzono symulację Monte Carlo dla danych o częstotliwości 7 (dziennych) oraz 52 (tygodniowych) według scenariuszy symulacyjnych zamieszczonych w tabelach 1 i 2.

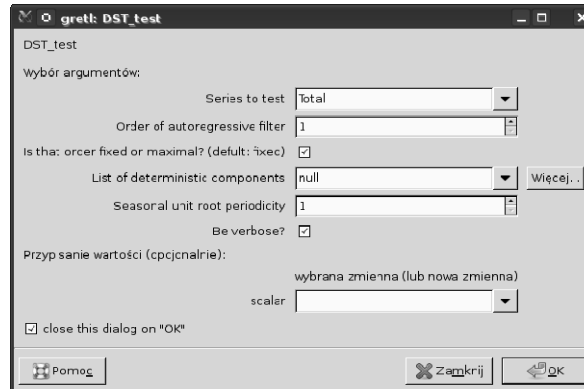
3. Implementacja testu Dickeya w programie Gretl

Implementacja testu Dickeya w programie Gretl jest bardzo nowym elementem, gdyż pierwsza stabilna wersja tego testu, oznaczona numerem 1.0, ukazała się 20 sierpnia 2010 r. Procedura ta, nazwana *DST_test*, została opracowana w formie pakietu funkcji Gretla, dostępnego na publicznym serwerze pakietów przez Marcina Błażejowskiego¹. Rysunek 1 prezentuje okno wyboru opcji w tym pakiecie.

W wierszu „Series to test” wskazuje się szereg czasowy, w którym testowany jest sezonowy pierwiastek jednostkowy. Natomiast w wierszu „Order of autoregressive fil-

¹ Kod tego pakietu został napisany na podstawie artykułu [Dickey 2009] oraz oryginalnego kodu dla programu SAS napisanego przez Dickeya dla szeregu „Natural Gas”. W tym miejscu chciałbym serdecznie podziękować prof. Dickeyowi za udostępnienie mi swojego kodu.

ter” – rząd p model autoregresyjnego $AR(p)$, który zostanie wykorzystany do prze-filtrowania oryginalnego szeregu (poziomów) w celu uzyskania procesu sezonowego błędzenia przypadkowego (przy założeniu prawdziwości H_0).



Rys. 1. Okno wyboru opcji testu stacjonarności Dickeya w programie Gretl

Źródło: opracowanie własne.

Opcja „Is that order fixed or maximal? (default: fixed)” służy do określenia, czy wskazany w poprzednim wierszu rząd p jest dokładną wartością, czy jest to maksymalny możliwy rząd modelu filtrującego. W pierwszym (domyślnym) przypadku do filtracji zostanie wykorzystany model $ARMA(p,0)$, natomiast w drugim zostaną oszacowane wszystkie możliwe modele $ARMA(p_i,0)$ dla $p_i = \{1,2,\dots,p\}$ i do filtracji będzie automatycznie wybrana specyfikacja charakteryzująca się najmniejszą wartością kryterium informacyjnego Schwarza BIC. Od strony praktycznej można tutaj zalecać wykorzystanie tego automatycznego mechanizmu przy jednoczesnym ustawieniu stosunkowo wysokiej wartości p dochodzącej nawet do $p = 2 \cdot d$. Tak wysoka wartość może co prawda skutkować długim czasem wykonywania się procedury lub błędów związanych z brakiem zbieżności algorytmu do estymacji modeli klasy ARIMA, jednak maksymalizuje prawdopodobieństwo uzyskania modelu, którego reszty będą miały własności procesu białego szumu, co w efekcie zapewnia, że uzyskany model faktycznie będzie filtrował testowany szereg do postaci procesu sezonowego błędzenia przypadkowego.

W wierszu „List of deterministic components” należy wskazać wszystkie składniki deterministyczne wchodzące w skład analizowanego szeregu (jest to argument opcjonalny, który domyślnie ma wartość *null*). Mogą to być takie składniki, jak deterministyczny trend oraz deterministyczna cykliczność (w postaci zmiennych zero-jedynkowych lub harmonik). Należy pamiętać, że wewnątrz pro-

cedury *DST_test* nie jest weryfikowana statystyczna istotność tych składników, w związku z czym weryfikację ich istotności należy przeprowadzić niezależnie. Badany szereg czasowy zostanie na wstępie procedury testowej oczyszczony ze wskazanych składników deterministycznych.

W kolejnym wierszu „Seasonal unit root periodicity” należy wskazać długość cyklu d , dla którego weryfikowane jest występowanie pierwiastka jednostkowego.

Ostatnim argumentem procedury jest opcja „Be verbose?”, które określa, czy powinny zostać pokazane wszystkie etapy procedury *DST_test* (ustawienie domyślne), czy jedynie powinna być zwrócona wartość p -value wyznaczonej statystyki testu. Wartość ta może być przypisana do określonej zmiennej, którą można wskazać w wierszu „scalar”. Jest to ustawienie opcjonalne.

Ze względu na to, że całość procedury *DST_test* została napisana w wewnętrznym języku skryptowym Gretla, każdy użytkownik ma możliwość wglądu do kodu tej procedury i jego ewentualnej modyfikacji. Całość pakietu składa się z 7 funkcji, z których tylko jedna – *DST_test* – jest funkcją publiczną dostępną dla użytkownika na zewnątrz pakietu. Pozostałych 6 funkcji to pomocnicze funkcje prywatne, które dostępne są tylko wewnątrz procedury *DST_test*. Instalacja pakietu wymaga od użytkownika pobrania go z serwera pakietów funkcji Gretla, co realizuje się poprzez wybranie z menu głównego okna programu: Plik/Pliki pakietów funkcji/na serwerze gretla i wskazanie pakietu *DST_test*.

4. Symulacyjna ocena mocy testu Dickeya dla danych o wybranych częstotliwościach obserwowania

Tabele 1-3 zawierają opisy scenariuszy symulacji Monte Carlo oraz ich wyników dla danych dziennych oraz tygodniowych. Każdy wariant powtarzany był 1000 razy.

W eksperymencie A wygenerowano procesy sezonowego błędzenia przypadkowego (wariant 1) oraz błędzenia przypadkowego z trendem liniowym i deterministyczną cyklicznością o okresie $d = 7$ (wariant 2) dla danych dziennych w układzie tygodnia 7-dniowego. Każdy wygenerowany szereg miał długość $n = 1820$. Weryfikowano hipotezę o występowaniu sezonowego pierwiastka jednostkowego dla $d = 7$, $(SI_7(1))$ procedurą *DST_test* w dwóch wariantach: w pierwszym rząd modelu autoregresyjnego był wybierany na podstawie kryterium informacyjnego BIC, w drugim wariantcie rząd modelu autoregresyjnego był ustalony na $p = 7$. Ponadto test przeprowadzany był w kolejnych dwóch wariantach: bez elementów deterministycznych oraz z uwzględnieniem trendu liniowego i zestawu 7 zmiennych zero-jedynkowych do opisu cyklu o okresie tygodnia. Łącznie dało to 8 scenariuszy symulacyjnych, z których każdy powtarzany był 1000 razy.

Eksperyment B był analogiczny do eksperymentu A, z tym, że został przeprowadzany dla danych tygodniowych. Każdy wygenerowany proces miał długość $n = 572$.

Weryfikowano hipotezę o występowaniu sezonowego pierwiastka jednostkowego dla $d = 52$, $(SI_{52}(1))$. Procedurę DST_test wykonywano w wariancie z automatycznym wyborem rzędu modelu autoregresyjnego oraz z rzędem ustalonym na $p = 4$. Ponadto część deterministyczna, którą uwzględniono w 4 z 8 scenariuszy, składała się z trendu liniowego oraz 4 harmonik dla częstości $\omega_1 = 2 \cdot \pi \cdot t / 52$ i $\omega_2 = 4 \cdot \pi \cdot t / 52$. Wyniki zrealizowanych symulacji zawarte są w tabeli 3.

Tabela 1. Eksperyment symulacyjny dla danych dziennych, proces błędzenia przypadkowego, $n = 1820$ liczba replikacji $m = 1000$

Model generujący	
$y_t = 1 \cdot y_{t-7} + \varepsilon_t$ (wariant 1)	$\varepsilon_t \sim IID(0,1)$
$y_t = 1 \cdot y_{t-7} + \alpha \cdot t + \sum \beta_i d_{it} + \varepsilon_t$ (wariant 2)	$\varepsilon_t \sim IID(0,1)$
	$t = 1, 2, \dots, 1820$
Eksperyment A	
$p = \{\text{ustalony na podstawie BIC}, 7\}$	
$d = 7$	
składniki deterministyczne {brak, 7 zmiennych zero-jedynkowych}	

Tabela 2. Eksperyment symulacyjny dla danych tygodniowych, $n = 572$ liczba replikacji $m = 1000$

Model generujący	
$y_t = 1 \cdot y_{t-52} + \varepsilon_t$ (wariant 1)	$\varepsilon_t \sim IID(0,1)$
$y_t = 1 \cdot y_{t-52} + \alpha \cdot t + \sum \beta_i d_{it} + \varepsilon_t$ (wariant 2)	$\varepsilon_t \sim IID(0,1)$
	$t = 1, 2, \dots, 572$
Eksperyment B	
$p = \{\text{ustalony na podstawie BIC}, 4\}$	
$d = 52$	
składniki deterministyczne {brak, 4 harmoniki dla częstości $\omega_1 = 2 \cdot \pi \cdot t / 52$ i $\omega_2 = 4 \cdot \pi \cdot t / 52$ }	

Wyniki zaprezentowane w tabeli 3 wskazują, że procedura DST_test ma większą moc dla wyższego d . Dla danych tygodniowych uzyskano tylko nieznaczne przekroczenie poziomu istotności $\alpha = 0,1$ w przypadku procesów błędzenia przypadkowego. W przypadku procesów z trendem i deterministyczną sezonowością nie odnotowano przekroczenia poziomu istotności. Dla danych tygodniowych uzyskano znacznie gorsze wyniki, szczególnie gdy generowane były procesy błędzenia przypadkowego, a w procedurze DST_test uwzględniono deterministyczną cykliczność. Tutaj błędy 1. rodzaju dochodziły do 55%, co jest wynikiem nieakceptowalnym. Potwierdza to konieczność ostrożnego włączania do procedury testowej składników deterministycznej cykliczności dla małych wartości d . Można także

stwierdzić, że sposób ustalania rzędu autoregresyjnego modelu filtrującego nie ma znaczącego wpływu na uzyskiwane wyniki.

Tabela 3. Procent odrzuceń H_0 w zrealizowanych eksperymentach symulacyjnych Monte Carlo według scenariuszy A i B, liczba replikacji każdego wariantu $m = 1000$

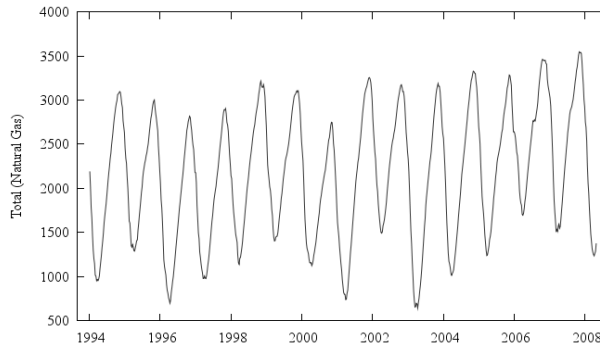
Uwzględnione w teście składniki deterministyczne	d	α	Proces sezonowego błędzenia przypadkowego (w %)		Proces sezonowego błędzenia przypadkowego z trendem i cyklem o okresie d (w %)	
			BIC	$p = 4$ lub 7	BIC	$p = 4$ lub 7
Brak	52	0,1	10,6	10,5	9,1	8,0
		0,05	4,4	3,9	3,5	3,2
		0,01	0,6	0,6	0,7	0,6
	7	0,1	17,2	17,2	7,1	7,1
		0,05	8,3	8,1	3,7	3,7
		0,01	1,4	1,3	0,5	0,5
Zmienne 0-1	52	0,1	10,2	10,3	4,9	4,3
		0,05	5,0	4,8	2,4	2,7
		0,01	1,0	1,2	0,2	0,1
	7	0,1	54,9	54,0	1,8	1,8
		0,05	37,4	37,2	0,5	0,5
		0,01	13,0	12,6	0,0	0,0

Źródło: opracowanie własne.

5. Przykłady weryfikacji sezonowych pierwiastków jednostkowych w empirycznych szeregach czasowych

W części empirycznej zaprezentowane zostaną wyniki weryfikacji hipotezy o sezonowym pierwiastku jednostkowym w trzech empirycznych szeregach czasowych. Dla danych miesięcznych oraz kwartalnych weryfikację przeprowadzono w oparciu o omawianą procedurę DST_test oraz procedurę DHF_test , będącą implementacją testu DHF. Autorem tej procedury jest Marcin Błazejowski.

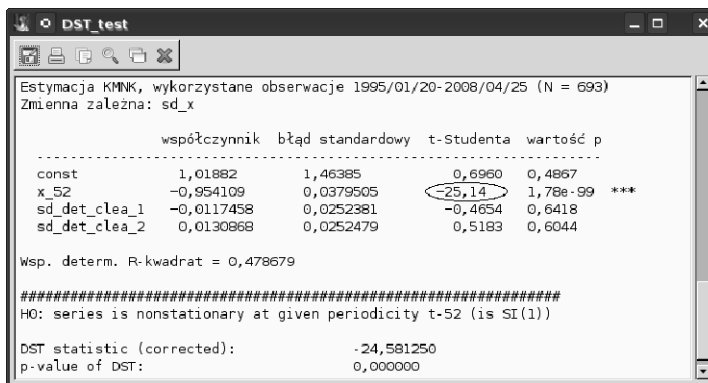
Przykład 1. Szereg tygodniowych stanów zasobów gazu ziemnego w USA. Pierwszy przykład empiryczny dotyczy szeregu danych tygodniowych dotyczących stanów zasobów gazu ziemnego w USA. Jest to ten sam szereg, którego Dickey użył w swoim przykładzie zaprezentowanym w [Dickey 2009]. Plik z danymi dla Gretla można pobrać ze strony internetowej: http://www.gretl.pl/DST/natural_gas.gdt. Na rysunku 2 przedstawiony jest przebieg tego szeregu.



Rys. 2. Przebieg procesu tygodniowych stanów zasobów gazu ziemnego w USA w okresie od 7.01.1994 do 25.04.2008, $n = 747$

Źródło: opracowanie własne.

Za pomocą procedury *DST_test* została zweryfikowana hipoteza o występowaniu sezonowego pierwiastka jednostkowego dla cyklu o okresie $d = 52$. Jako autoregresyjny model filtrujący przyjęto model $AR(2)$, natomiast jako składniki deterministyczne – trend liniowy i 4 harmoniki o częstościach $\omega_1 = 2 \cdot \pi \cdot t \cdot 7 / 365.25$ oraz $\omega_2 = 4 \cdot \pi \cdot t \cdot 7 / 365.25$. Rysunek 3 przedstawia uzyskany wynik procedury *DST_test* dla procesu „Natural Gas”.



Rys. 3. Wyniki testu stacjonarności Dickeya dla procesu tygodniowych zasobów gazu ziemnego w USA

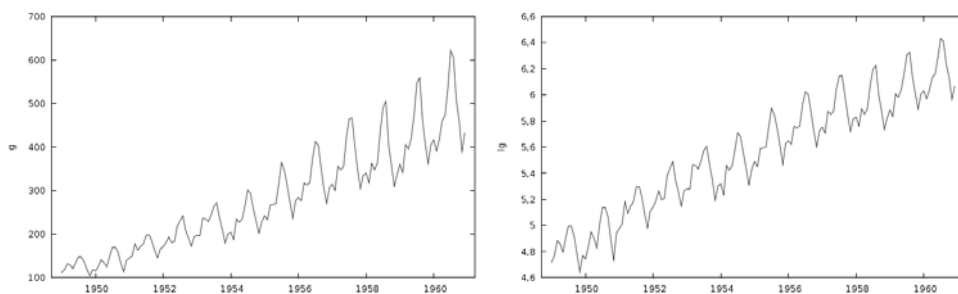
Źródło: opracowanie własne.

Zaznaczona elipsą na rysunku 3 wartość $\tau = -25,14$ to wartość statystyki *t*-Studenta i jest ona minimalnie wyższa niż wartość $-26,25$ uzyskana przez Dickeya

[2009]. Wartość p -value tej statystyki skorygowanej i przekształconej do rozkładu $N(0,1)$ wynosi p -value = 0,00, co wskazuje na konieczność odrzucenia H_0 – badany proces nie posiada sezonowego pierwiastka jednostkowego dla cyklu o okresie $d = 52$.

Przykład 2. Szereg Box-Jenkins G (airline passengers).

Box i Jenkins w 1976 r. (polskie wydanie 1983) opracowali dwuparametryczny model $SARIMA(0,1,1)(0,1,1)$ dla logarytmów miesięcznej liczby przelotów pasażerskich w USA [Box, Jenkins 1983]. Rysunek 4 przedstawia oryginalny proces tych przelotów – g (rysunek po lewej stronie) oraz proces logarytmów tych przelotów – lg (rysunek po prawej stronie).



Rys. 4. Proces miesięcznej liczby przelotów pasażerskich oraz jego logarytmów naturalnych w okresie od stycznia 1949 r. do grudnia 1960 r. ($n = 144$)

Źródło: opracowanie własne na podstawie [Box, Jenkins 1983].

Przebieg procesu g wskazuje na rosnącą amplitudę wahań sezonowych, co sugeruje występowanie sezonowego pierwiastka jednostkowego. Weryfikację hipotezy o sezonowej integracji obu procesów przeprowadzono za pomocą procedur DST_test oraz DHF_test w dwóch wersjach: bez sezonowych zmiennych zero-jedynkowych oraz z ich uwzględnieniem. Wyniki prezentuje tabela 4.

Jak można zauważyć test DHF (procedura DHF_test) w żadnym wariancie nie odrzuca hipotezy o sezonowym pierwiastku jednostkowym. Procedura DST_test w przypadku weryfikacji dla procesu oryginalnego także ani razu nie odrzuca H_0 , jednak dla procesu logarytmów aż w 6 przypadkach na 9 odrzuca hipotezę o pierwiastku jednostkowym. Uzyskane wyniki potwierdzają sugestie Dickeya [2009], aby dla danych półrocznych, kwartalnych oraz miesięcznych stosować klasyczny test DHF ze względu na jego większą moc w przypadku tych częstotliwości.

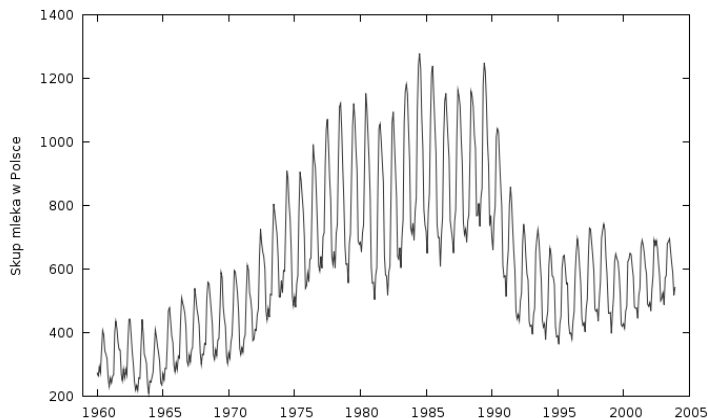
Tabela 4. Wyniki weryfikacji hipotezy o sezonowym pierwiastku jednostkowym w procesie miesięcznej liczby przelotów pasażerskich oraz jego logarytmów za pomocą procedur DST_test oraz DHF_test

Rząd modelu filtrującego $AR(p)$	Sezonowe zmienne 0-1	Proces g			Proces lg		
		α			α		
		0,1	0,05	0,01	0,1	0,05	0,01
Procedura DST_test							
$p = 12$	nie	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$
$p = 12$	tak	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(0)$	$SI_d(0)$	$SI_d(1)$
p ustalone na podstawie kryterium BIC	nie	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(0)$	$SI_d(1)$	$SI_d(1)$
p ustalone na podstawie kryterium BIC	tak	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(0)$	$SI_d(0)$	$SI_d(0)$
Procedura DHF_test							
$p = 12$	nie	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$
$p = 12$	tak	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$

Źródło: opracowanie własne.

Przykład 3. Szereg miesięcznego skupu mleka w Polsce.

Kolejny przykład dotyczy procesu miesięcznego skupu mleka w Polsce, znany z podręcznika [Kufel 2007]. Rysunek 5 przedstawia przebieg tego procesu w okresie od stycznia 1960 do grudnia 2003 r.



Rys. 5. Przebieg procesu miesięcznego skupu mleka w Polsce w okresie od stycznia 1960 do grudnia 2003 ($n = 528$)

Źródło: opracowanie własne na podstawie [Kufel 2007].

Tabela 5 zawiera wyniki weryfikacji hipotezy o sezonowym pierwiastku jednostkowym w analizowanym procesie. Oznaczenia są analogiczne jak w tabeli 4.

Tabela 5. Wyniki weryfikacji hipotezy o sezonowym pierwiastku jednostkowym w procesie miesięcznego skupu mleka w Polsce za pomocą procedur *DST_test* oraz *DHF_test*

Rząd modelu filtrującego $AR(p)$	Sezonowe zmienne 0-1	Proces skupu mleka		
		α		
		0,1	0,05	0,01
		Procedura <i>DST_test</i>		
$p = 12$	nie	$SI_d(0)$	$SI_d(0)$	$SI_d(0)$
$p = 12$	tak	$SI_d(0)$	$SI_d(0)$	$SI_d(0)$
p ustalone na podstawie kryterium BIC	nie	$SI_d(0)$	$SI_d(0)$	$SI_d(0)$
p ustalone na podstawie kryterium BIC	tak	$SI_d(0)$	$SI_d(0)$	$SI_d(0)$
		Procedura <i>DHF_test</i>		
$p = 12$	nie	$SI_d(0)$	$SI_d(0)$	$SI_d(1)$
$p = 12$	tak	$SI_d(1)$	$SI_d(1)$	$SI_d(1)$

Źródło: opracowanie własne.

Podobnie jak w przykładzie dla przelotów pasażerskich, wyniki dla procedur *DST_test* oraz *DHF_test* nie są jednoznaczne. Procedura *DST_test* we wszystkich przypadkach wskazuje na brak sezonowej integracji testowanego procesu, podczas gdy test DHF odrzuca hipotezę o pierwiastku jednostkowym tylko w 2 przypadkach na 6.

Więcej przykładów na weryfikację hipotez o występowaniu sezonowych pierwiastków jednostkowych w danych o wysokiej częstotliwości obserwowania można znaleźć w [Kufel 2010].

6. Podsumowanie

Zaprezentowany w 2009 r. przez Dickeya nowy test występowania sezonowych pierwiastków jednostkowych dla danych o wysokiej częstotliwości obserwowania wypełnia lukę, jaka istniała dotąd w tym zakresie. Pozwala on bowiem na uwzględnienie w modelowaniu i prognozowaniu stochastycznego charakteru obserwowanych cykliczności. Asymptotyczna zbieżność statystyki tego testu do rozkładu $N(0,1)$ stwarza duże możliwości aplikacyjne, gdyż nie wymaga posiadania tablic niestandardowego rozkładu dla wielu możliwych okresowości d . Niemniej jednak moc oraz rozmiar tego testu wymagają dalszych badań w tym zakresie. Wyniki analiz symulacyjnych zaprezentowane w artykule wskazują, że moc testu rośnie wraz z d (co wynika z szybkości zbiegania statystyki testu do rozkładu $N(0,1)$).

Zaprezentowane wyniki badań dla szeregów empirycznych także wskazują na mniejszą moc procedury *DST_test* w stosunku do procedury *DHF_test*. Należy jednak pamiętać, że wartości krytyczne w teście DHF wyznaczone są dokładnie dla

$d = 2, 4, 12$, podczas gdy w procedurze *DST_test* są przybliżone, przy czym zbieżność do rozkładu $N(0,1)$ rośnie wraz z wartością d .

Implementacja wieloetapowej procedury testu Dickeya w programie Gretl pozwala na weryfikację hipotezy o sezonowym pierwiastku jednostkowym bez konieczności pisania skomplikowanych skryptów. Okienkowy interfejs pakietu znakomicie ułatwia jego wykorzystanie w empirycznych analizach. Jest to także prawdopodobnie pierwsza (i jedyna) na świecie implementacja testu Dickeya wśród programów statystyczno-ekonometrycznych.

Literatura

- Box G.E.P., Jenkins G., *Analiza szeregów czasowych. Prognozowanie i sterowanie*, PWN, Warszawa 1983.
- Dickey D.A., *Stationarity Testing in High-Frequency Seasonal Time Series*, SAS Global Forum 2009: Statistics and Data Analysis, 2009.
- Dickey D.A., Hasza D., Fuller W., *Testing for unit roots in seasonal time series*, „Journal of the American Statistical Association” 1984, no. 79.
- Dickey D.A., Zhang Y., *Seasonal unit root tests in long periodicity cases*, „Journal of the Korean Statistical Society” 2010, no. 39.
- Hylleberg S., Engle R.F., Granger C.W., Yoo B., *Seasonal integration and cointegration*, „Journal of Econometrics” 1990, no. 44.
- Kufel T., *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, Wydawnictwo Naukowe PWN, Warszawa 2007.
- Kufel T., *Ekonometryczna analiza cykliczności procesów gospodarczych o wysokiej częstotliwości obserwowania*, Wydawnictwo Naukowe UMK, Toruń 2010.

STATIONARITY OF HIGH-FREQUENCY TIME SERIES – IMPLEMENTATION OF DICKEY'S STATIONARITY TEST IN GRETL

Summary: The aim of the paper is to assess power of the stationarity test for high frequency time series introduced by D.A. Dickey in 2009 verifying null hypothesis that given series is seasonally integrated (is $SI_d(1)$). This test extends the usage of standard DHF test for seasonal unit root [Dickey, Hasza, Fuller 1984] for cycles of frequency $d = 5, 6, 7, 21, 24, 26, 31, 48, 52, 168, 365...$ Furthermore, the implementation of these two tests (Dickey's stationarity of high-frequency time series and DHF) in GRETL is presented.