



DOI:10.18276/sip.2016.45/2-06

Jan Purczyński*
Kamila Bednarz-Okrzyńska**
Uniwersytet Szczeciński

WYBRANE METODY ESTYMACJI PARAMETRÓW FUNKCJI LOGISTYCZNEJ

Streszczenie

W opracowaniu rozpatrzono przydatność przybliżonych metod estymacji parametrów funkcji logistycznej: Hotellinga, Tintnera, Muchy, metody trzech sum oraz zmodyfikowanej metody Hotellinga. Jako przykłady uwzględniono wyposażenie gospodarstw domowych w Polsce w: telefony komórkowe, automaty pralnicze, chłodziarki oraz samochody osobowe – dane za lata 2000–2014. Na podstawie wartości mierników zgodności funkcji logistycznej z danymi empirycznymi (współczynnik zbieżności, współczynnik zmienności) dokonano oceny wyników poszczególnych metod.

Słowa kluczowe: estymacja parametrów funkcji logistycznej, metody przybliżone

Wstęp

Funkcję logistyczną wykorzystuje się do modelowania danych empirycznych odnoszących się do przypadków monotonicznego wzrostu obserwowanego zjawiska, przy czym tempo wzrostu maleje do zera. Jak stwierdza A. Zeliaś (1997, s. 123), funkcja logistyczna jest przydatna do opisu popytu na dobra trwałego użytku, na przykład samochody, telewizory, motocykle, w krajach wysoko rozwiniętych z gospodarką ustabilizowaną.

* Adres e-mail: jan.purczynski@wzieu.pl.

** Adres e-mail: kamila.bednarz@wzieu.pl.

T. Stanisław (1986, s. 122–168) szczegółowo omawia przykłady zastosowań funkcji logistycznej, a także wymienia związaną z tym literaturę. Spośród artykułów, które ukazały się w bieżącym wieku, można wymienić pracę Z. Muchy (2005, s. 17–26), w której zaproponowano metodę estymacji parametrów funkcji logistycznej bazującą na rozkładzie funkcji na składową parzystą i składową nieparzystą. Obszerny przegląd metod estymacji parametrów funkcji logistycznej zawiera praca J. Żurowskiej (2007, s. 143–158), gdzie na przykładzie prognozowania wskaźnika motoryzacji w Polsce przetestowano skuteczność 9 metod przybliżonych estymacji parametrów.

Funkcja logistyczna jest opisana wzorem:

$$y(t) = \frac{a}{1 + b \cdot e^{-c \cdot t}} \quad (1)$$

gdzie : $a > 0$, $b > 1$, $c > 0$.

Funkcja logistyczna narasta do wartości ustalonej a . Zachodzą następujące zależności:

$$t_p = \frac{\ln b}{c}; \quad y_p = \frac{a}{2} \quad (2)$$

gdzie: t_p – punkt przegięcia.

Wychodząc ze wzoru (1), uzyskuje się następujące równanie różniczkowe:

$$y'_t = c \cdot y_t \left(1 - \frac{y_t}{a} \right) \quad (3)$$

1. Wybrane metody estymacji parametrów funkcji logistycznej

1.1. Metoda Hotellinga (Stanisław, 1986)

W równaniu (3) pochodną y'_t zastępuje się ilorazem różnicowym:

$$y'_t = \frac{y_{t+1} - y_t}{h} \quad (4)$$

Wprowadzając nową zmienną:

$$u_t = \frac{y_{t+1} - y_t}{h \cdot y_t} \quad (5)$$

ze wzorów (3) i (4), otrzymuje się:

$$u_t = c + d \cdot y_t \quad (6)$$

gdzie: $d = -\frac{c}{a}$.

Dla trendu liniowego (6) wyznacza się metodą najmniejszych kwadratów (MNK) oszacowania parametrów \hat{c} i \hat{d} , a następnie oszacowanie $\hat{a} = -\frac{\hat{c}}{\hat{d}}$. Oszacowanie parametru \hat{b} wyznacza się ze wzoru:

$$\hat{b} = \frac{1}{N} \cdot \sum_{t=1}^N e^{\hat{c} \cdot t} \cdot \left(\frac{\hat{a}}{y_t} - 1 \right) \quad (7)$$

1.2. Metoda Tintnera (Stanisz, 1986)

W metodzie tej wprowadza się zmienną z_t będącą odwrotnością zmiennej y_t

$$z_t = \frac{1}{y_t} = \frac{1}{a} + \frac{b}{a} \cdot e^{-c \cdot t} \quad (8)$$

w wyniku czego uzyskuje się równanie trendu liniowego:

$$z_{t+1} = a_0 + a_1 z_t \quad (9)$$

Stosując MNK, otrzymuje się oszacowania parametrów \hat{a}_0 i \hat{a}_1 , a także oszacowania:

$$\hat{a} = \frac{1 - \hat{a}_1}{\hat{a}_0}; \quad \hat{c} = -\ln \hat{a}_1 \quad (10)$$

Oszacowanie parametru \hat{b} określa wzór (7).

1.3. Zmodyfikowana metoda Hotellinga (Purczyński, 2003, s. 31–32)

Metoda Hotellinga bazowała na zastąpieniu wartości pochodnej ilorazem różnicowym [wzór (4)].

Modyfikacja metody Hotellinga polega na zastąpieniu pochodnej następującym wyrażeniem:

$$y'_t = \frac{y_{t+1} - y_{t-1}}{2 \cdot h} \quad (11)$$

Ze wzorów (3) i (11) otrzymuje się liniowe równanie rekurencyjne:

$$y_{t+2} = 2 \cdot A \cdot y_{t+1} + 2 \cdot B \cdot (y_{t+1})^2 + y_t \quad (12)$$

gdzie: $A = h \cdot c$; $B = -\frac{c}{a} \cdot h$.

Stosując MNK do równania (12), uzyskuje się oszacowania parametrów \hat{A} i \hat{B} , a następnie oszacowania parametrów:

$$\hat{c} = \frac{\hat{A}}{h}; \hat{a} = -\frac{\hat{c}}{\hat{B}}h \quad (13)$$

Oszacowanie parametru \hat{b} wyznacza się ze wzoru (7), gdzie \hat{a} i \hat{c} są określone zależnością (13).

1.4. Pozostałe metody

W opracowaniu będą uwzględnione dwie dalsze metody. Pierwszą z nich jest metoda trzech sum opisana między innymi w cytowanych pracach (Stanisz, 1986; Żurowska, 2007).

Kolejną metodą podlegającą weryfikacji była metoda Muchy (2005). Pierwszy krok w tej metodzie pokrywa się z metodą Tintnera, a mianowicie wprowadza się funkcję z_t będącą odwrotnością zmiennej y_t [wzór (8)]. Następnie tworzy się funkcję symetryczną i antysymetryczną dla funkcji z_t . Korzystając z właściwości tych funkcji, wyznaczono wzory określające oszacowania parametrów funkcji logistycznej. Dla osób, które będą chciały wdrożyć proponowaną metodę, istotna może być informacja, że we wzorze (20) pracy (Mucha, 2005) w miejsce zależności $\hat{a} = \frac{1}{\hat{A}}$ powinno być $\hat{a} = \frac{2}{\hat{A}}$.

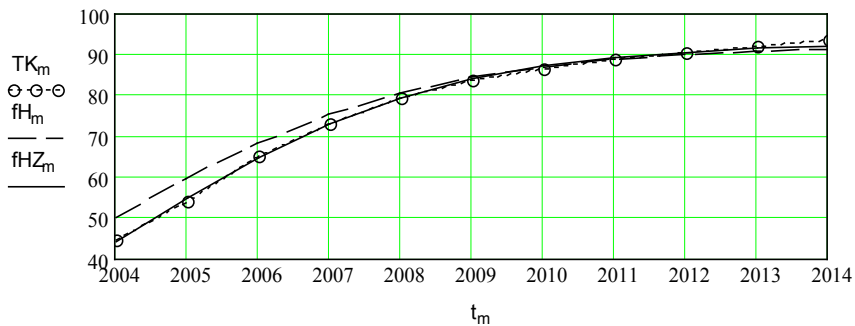
2. Przykłady estymacji funkcji logistycznej

Celem artykułu jest ocena przydatności przybliżonych metod estymacji parametrów funkcji logistycznej (Hotellinga, Tintnera, trzech sum, Muchy oraz zmodyfikowanej metody Hotellinga). Badania wykonano na podstawie empirycznych szeregów czasowych zaczerpniętych z tabeli pod tytułem „Gospodarstwa domowe wyposażone w niektóre przedmioty trwałego użytkowania” zamieszczonej w kolejnych rocznikach statystycznych (Rocznik statystyczny Rzeczypospolitej Polskiej; lata 2006–2014). Ocena jakości wyników poszczególnych metod zostanie wykonana na podstawie wartości mierników zgodności funkcji logistycznej z danymi empirycznymi: współczynnika zbieżności oraz współczynnika zmienności.

Jako pierwsze zostanie rozpatrzone wyposażenie gospodarstw domowych w telefony komórkowe, wyrażone w procentach – na rysunku 1 dane wyjściowe oznaczono jako TK_m – linia kropkowana z kółkami. Estymację parametrów funkcji

logistycznej wykonano następującymi metodami: Hotellinga (H), zmodyfikowaną metodą Hotellinga (HZ), Tintnera (T), trzech sum (S) oraz Muchy (M) – metody te wymieniono w kolumnie 1 tabeli 1. W celu oceny jakości uzyskanych modeli wyznaczono następujące wielkości: odchylenie standardowe reszt S , współczynnik zmienności v oraz współczynnik zbieżności φ^2 .

Rysunek 1. Wyposażenie gospodarstw domowych w telefony komórkowe [w %]
 TK_m – linia kropkowana z kółkami. Wyniki obliczeń: metodą Hotellinga – fH_m linia przerywana, zmodyfikowaną metodą Hotellinga – fHZ_m linia ciągła



Źródło: opracowanie własne.

Estymację wykonano dla danych dotyczących lat 2004–2013, natomiast obserwacja z roku 2014 posłużyła do wyznaczenia błędu względnego prognozy *ex post wp*. Kierując się najmniejszą wartością współczynnika zmienności, rozpatrzone metody można uszeregować następująco: HZ, M, S, T oraz H. Kryterium wartości współczynnika zbieżności powoduje zamianę miejsc metod M i S. Na rysunku 1 zamieszczono wyniki obliczeń najłabszej metody (H) – fH_m linia przerywana oraz najlepszej metody (HZ) – fHZ_m linia ciągła.

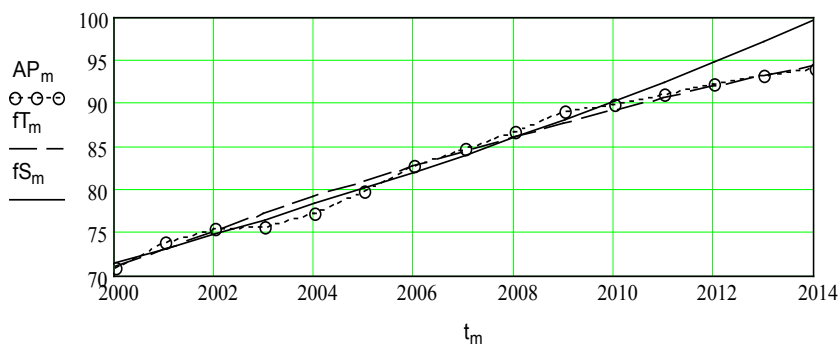
Tabela 1. Wartości mierników zgodności funkcji logistycznej opisującej wyposażenie gospodarstw domowych w telefony komórkowe

| Lp. | Metoda | Odchylenie standardowe S | Współczynnik zmienności v | Współczynnik zbieżności φ^2 | Błąd względny <i>ex post</i> w_p |
|-----|--------|----------------------------|-----------------------------|-------------------------------------|------------------------------------|
| 1 | H | 3,500 | 0,0453 | 0,03561 | 0,0351 |
| 2 | HZ | 0,531 | 0,0069 | 0,00082 | 0,0258 |
| 3 | T | 0,938 | 0,0121 | 0,00256 | 0,0191 |
| 4 | S | 0,746 | 0,0096 | 0,00090 | 0,0311 |
| 5 | M | 0,659 | 0,0085 | 0,00126 | 0,0214 |

Źródło: opracowanie własne.

Na rysunku 2 zamieszczono dane dotyczące wyposażenia gospodarstw domowych w automaty pralnicze (w %) – AP_m – linia kropkowana z kółkami, a także wyniki obliczeń: metodą Tintnera – fT_m linia przerywana oraz metodą trzech sum – fS_m linia ciągła.

Rysunek 2. Wyposażenie gospodarstw domowych w automaty pralnicze [w %] AP_m – linia kropkowana z kółkami. Wyniki obliczeń: metodą Tintnera – fT_m linia przerywana, metodą trzech sum – fS_m linia ciągła



Źródło: opracowanie własne.

Na podstawie wartości współczynnika zbieżności (kolumna 4 tabeli 2) można stwierdzić, że cztery metody (M, T, HZ, H) wykazują zbliżone wartości φ^2 , natomiast najslabiej wypada metoda S. Prowadzi ona do błędnego oszacowania funkcji logistycznej: $\hat{a} = -144,47$, $\hat{b} = -3,067$, $\hat{c} = 0,015$ wbrew założeniu [wzór (1)] $a > 0$ oraz $b > 1$. W rezultacie krzywa fS_m przedstawiona na rysunku 2 nie wykazuje tendencji malejących przyrostów.

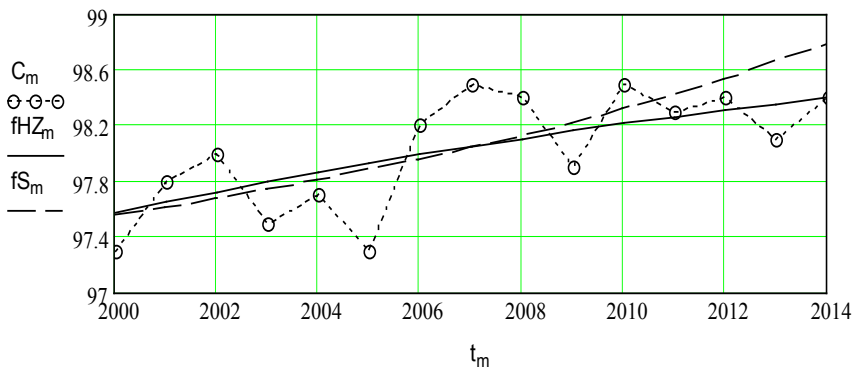
Tabela 2. Wartości mierników zgodności funkcji logistycznej opisującej wyposażenie gospodarstw domowych w automaty pralnicze

| Lp. | Metoda | Odchylenie standardowe S | Współczynnik zmienności v | Współczynnik zbieżności φ^2 | Błąd względny <i>ex post</i> w_p |
|-----|--------|----------------------------|-----------------------------|-------------------------------------|------------------------------------|
| 1 | H | 1,033 | 0,0124 | 0,0159 | -0,0022 |
| 2 | HZ | 0,968 | 0,0117 | 0,0140 | -0,0084 |
| 3 | T | 0,966 | 0,0116 | 0,0139 | -0,0050 |
| 4 | S | 1,288 | 0,0155 | 0,0411 | -0,0604 |
| 5 | M | 0,916 | 0,0110 | 0,0125 | -0,0092 |

Źródło: opracowanie własne.

Na uwagę zasługuje duża wartość błędu względnego *ex post* dla metody trzech sum. Znajduje to potwierdzenie na rysunku 2, gdzie dane za okres 2000–2013 posłużyły do budowy modelu, natomiast obserwacja dla 2014 roku została wykorzystana do wyznaczenia błędu *ex post*.

Rysunek 3. Wyposażenie gospodarstw domowych w chłodziarki [w %] C_m – linia kropkowana z kółkami. Wyniki obliczeń: metodą trzech sum – fS_m linia przerywana, zmodyfikowaną metodą Hotellinga – fHZ_m linia ciągła



Źródło: opracowanie własne.

Jako kolejny przykład rozpatrzono dane dotyczące wyposażenia gospodarstw domowych w chłodziarki (w %) – na rysunku 3 oznaczono C_m – linia kropkowana z kółkami. Ponadto zamieszczono wyniki obliczeń: metodą trzech sum – fS_m linia przerywana oraz zmodyfikowaną metodą Hotellinga – fHZ_m linia ciągła.

Z danych przedstawionych w tabeli 3 wynika, że trzy metody prowadzą do ujemnej wartości oszacowania parametru b (H, T oraz S). Natomiast metoda M estymuje zespoloną wartość parametru b . Przyczyny tego zdarzenia należy upatrywać w dużych amplitudach wahań wartości liczby chłodziarek – krzywa C_m na rysunku 3. Jedynie metoda HZ zapewniła poprawne wyniki uzyskanego modelu. Na podstawie rysunku 3 stwierdza się analogiczne zachowanie funkcji fS_m jak na rysunku 2 – rosnące przyrosty trendu.

Z tabeli 3 wynika niebezpieczeństwo, jakie niesie wyciąganie wniosków na podstawie wartości współczynnika zmienności oraz współczynnika zbieżności bez uwzględnienia wartości oszacowanych parametrów trendu. Mianowicie na podstawie wartości v (φ^2) można stwierdzić równoważność metody trzech sum oraz zmodyfikowanej metody Hotellinga, co stoi w sprzeczności z wartościami wyestymowanych parametrów funkcji logistycznej (tabela 3).

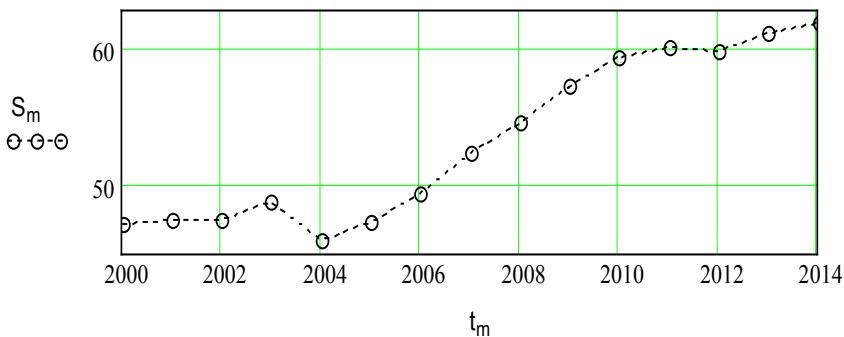
Tabela 3. Wartości mierników zgodności funkcji logistycznej opisującej wyposażenie gospodarstw domowych w chłodziarki

| Metoda | Wartości oszacowania parametrów | | | Współczynnik zmienności | Współczynnik zbieżności | Błąd względny <i>ex post</i> |
|--------|---------------------------------|------------------|------------------|-------------------------|-------------------------|------------------------------|
| | a | b | c | | | |
| H | 98,086 | -1,311 | 0,619 | 0,756 | 25886 | -0,0436 |
| HZ | 99,357 | 0,019 | 0,045 | 0,0034 | 0,532 | -0,0468 |
| T | 98,083 | -97,413 | 0,961 | 1,454 | 95830 | -0,0509 |
| S | 96,675 | -0,00854 | -0,061 | 0,0037 | 0,618 | -0,0509 |
| M | 98,131 | Liczba zespolona | Liczba zespolona | Liczba zespolona | Liczba zespolona | Liczba zespolona |

Źródło: opracowanie własne.

Ostatni przykład dotyczy wyposażenia gospodarstw domowych w samochód osobowy, co ilustruje rysunek 4. Analogicznie jak w poprzednich przykładach dane za okres 2000–2013 wykorzystano do estymacji parametrów modelu, natomiast obserwacja dotycząca 2014 roku posłużyła do wyznaczenia błędu *ex post*. W wyniku obliczeń dla czterech metod (H, HZ, T, S) uzyskano ujemne wartości oszacowań parametrów, natomiast w przypadku metody M były to wartości zespolone.

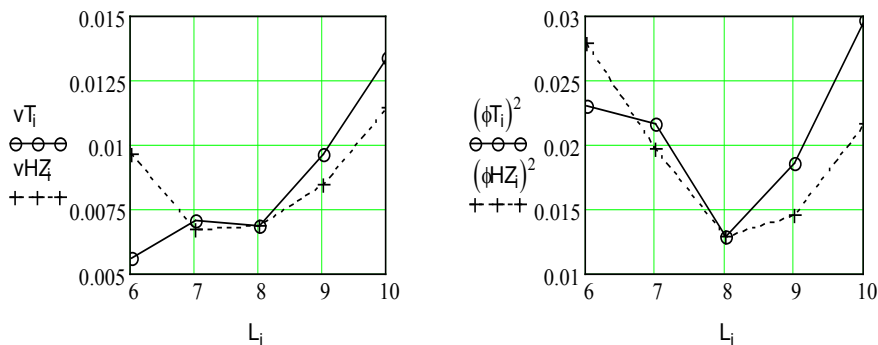
Rysunek 4. Wyposażenie gospodarstw domowych w samochody osobowe [w %] S_m – linia kropkowana z kółkami



Źródło: Rocznik statystyczny.

Obliczenia powtórzono, zmniejszając liczbę ostatnio uwzględnionych obserwacji – kolejno odrzucano obserwacje dotyczące roku 2000, 2001 itd. Gdy uwzględniono $L = 10$ obserwacji (lata 2004–2013), oszacowania parametrów dla wszystkich metod przyjęły wartości dodatnie. Na rysunku 5 przedstawiono wyniki obliczeń dla liczby ostatnio uwzględnionych obserwacji $L = 6 - 10$ dla metody T i HZ, które zapewniały najlepsze dopasowanie modelu do danych empirycznych. Rysunek po lewej stronie nie rozstrzyga, jaką wartość L należy wybrać. Z rysunku po prawej stronie jednoznacznie wynika optymalna wartość $L = 8$.

Rysunek 5. Wartości współczynnika zmienności vT_i (linia ciągła z kółkami) oraz vHZ_i (linia kropkowana z plusami). Po prawej stronie: wartości współczynnika zbieżności $(\phi T_i)^2$ (linia ciągła z kółkami) oraz $(\phi HZ_i)^2$ (linia kropkowana z plusami)



Źródło: opracowanie własne.

W tabeli 4 zamieszczono wartości mierników zgodności wyznaczonych dla $L = 8$ ostatnich obserwacji.

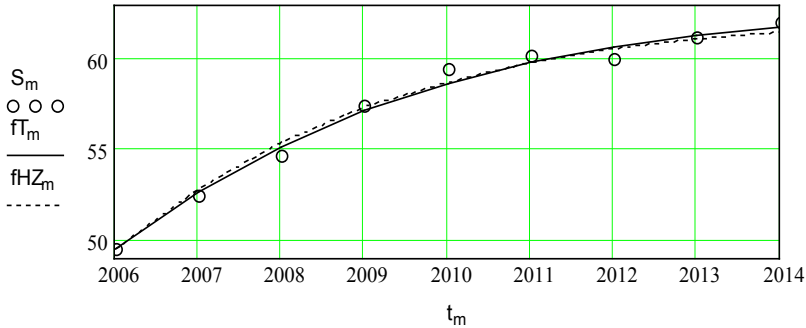
Tabela 4. Wartości mierników zgodności funkcji logistycznej opisującej wyposażenie gospodarstw domowych w samochody osobowe

| Lp. | Metoda | Odchylenie standardowe S | Współczynnik zmienności v | Współczynnik zbieżności ϕ^2 | Błąd względny <i>ex post wp</i> |
|-----|--------|----------------------------|-----------------------------|----------------------------------|---------------------------------|
| 1 | H | 0,996 | 0,0120 | 0,0397 | 0,0098 |
| 2 | HZ | 0,565 | 0,00681 | 0,0128 | 0,0081 |
| 3 | T | 0,568 | 0,00684 | 0,0129 | 0,0045 |
| 4 | S | 1,147 | 0,0138 | 0,0693 | -0,0344 |
| 5 | M | 0,715 | 0,0086 | 0,0205 | 0,0114 |

Źródło: opracowanie własne.

Na rysunku 6 przedstawiono wyniki obliczeń wykonanych dla $L = 8$ obserwacji (lata 2006–2013) dla metody T i metody HZ. Z rysunku wynika dobre dopasowanie modeli do danych empirycznych, co przekłada się na małe wartości współczynnika zmienności: $vT = 0,00684$ i $vHZ = 0,00681$ oraz współczynników zbieżności $\phi T^2 = 0,0129$, $\phi HZ^2 = 0,0128$.

Rysunek 6. Wyposażenie gospodarstw domowych w samochody osobowe w [%] S_m – zaznaczone kółkami. Wyniki obliczeń: metoda Tintnera – fT_m (linia ciągła) oraz zmodyfikowana metoda Hotellinga – fHZ_m linia kropkowana



Źródło: opracowanie własne.

Podsumowanie

Na podstawie wartości mierników zgodności funkcji logistycznej z danymi empirycznymi dokonano podsumowania jakości poszczególnych modeli. Mianowicie dla każdego przykładu uszeregowano metody według rosnących wartości współczynnika zbieżności (współczynnika zmienności). Metoda stojąca na pierwszym miejscu uzyskiwała 1 pkt, a stojąca na ostatnim – 5 pkt. Następnie zsumowano poszczególne punkty. Uzyskano następujące wyniki:

- zmodyfikowana metoda Hotellinga – 6 pkt,
- metoda Tintnera, metoda Muchy – po 12 pkt,
- metoda trzech sum – 14 pkt,
- metoda Hotellinga – 16 pkt.

Z powyższego zestawienia wynika zdecydowana przewaga metody HZ, co znalazło potwierdzenie w wynikach obliczeń zamieszczonych w tabeli 3, kiedy pozostałe metody prowadziły do błędnych rezultatów. Zmodyfikowana metoda Hotellinga różni się tym od metody Hotellinga, że pochodna w metodzie H wyraża się wzorem (4), natomiast pochodna w metodzie HZ – wzorem (11). Ta niewielka różnica we wzorach prowadzi do istotnej różnicy jakości obydwu metod.

W opracowaniu nie rozpatrywano metody najmniejszych kwadratów (MNK), ponieważ ograniczono się do porównania przybliżonych metod estymacji parametrów funkcji logistycznej.

Literatura

- Mucha, Z. (2005). Estymacja parametrów trendu logistycznego. *Wiadomości Statystyczne*, 4, 17–26.
- Purczyński, J. (2003). *Wykorzystanie symulacji komputerowych w estymacji wybranych modeli ekonometrycznych i statystycznych*. Szczecin: Wyd. Naukowe US.
- Rocznik statystyczny Rzeczypospolitej Polskiej, lata 2006–2014.
- Stanisz, T. (1986). *Funkcje jednej zmiennej w badaniach ekonomicznych*. Warszawa: PWN.
- Zeliaś, A (1997). *Teoria prognozy*. Warszawa: PWE.
- Żurowska, J. (2007). Zastosowanie modelu trendu logistycznego do prognozowania wskaźnika motoryzacji w Polsce. *Archiwum Motoryzacji*, 2, 143–148.

SELECTED METHODS OF ESTIMATING LOGISTIC FUNCTION PARAMETERS

Abstract

In this paper the usefulness of the following approximated methods of estimating logistic function parameters was examined: Hotelling's, Tintner's, Mucha's, three sums method and the modified Hotelling's method. As examples, some home appliances popular in Poland were considered: mobile phones, washing machines, refrigerators and cars – data for years 2000–2014. On the basis of goodness-of-fit measures with relation to the logistic function and the empirical data (correlation method, changeability coefficient), evaluations of particular method results were conducted.

Translated by Ewa Stefanowska

Keywords: estimation of logistic function parameters, approximate methods

JEL Codes: C5, C51