



Anna Pajor

Uniwersytet Ekonomiczny w Krakowie
Wydział Zarządzania
Katedra Ekonometrii i Badań Operacyjnych
pajora@uek.krakow.pl

STANDARDOWY BŁĄD NUMERYCZNY DLA ESTYMATORÓW CHM I CAM¹

Streszczenie: W pracy zaproponowano sposób obliczania standardowego błędu numerycznego dla estymatorów wartości brzegowej gęstości wektora obserwacji, opartych na skorygowanej średniej harmonicznej oraz skorygowanej średniej arytmetycznej. W części empirycznej porównano numeryczne własności tych estymatorów w kontekście modeli Copula-AR-GARCH. Dodatkowo zastosowano metodę Chiba i Jeliaskova. Wyniki jednoznacznie pokazały, że estymator oparty na skorygowanej średniej arytmetycznej charakteryzuje się najmniejszym standardowym błędem numerycznym.

Słowa kluczowe: standardowy błąd numeryczny, skorygowana średnia arytmetyczna, skorygowana średnia harmoniczna.

Wprowadzenie

Bayesowskie podejście do wnioskowania statystycznego daje możliwość formalnego porównywania modeli, poprzez czynniki Bayesa, a także łączenia wiedzy pochodzącej z różnych modeli. Stosowanie zasad Bayesowskiego porównywania modeli wymaga estymacji wartości brzegowej gęstości wektora obserwacji. Estymacja ta jest poważnym wyzwaniem obliczeniowym².

Jednym z mierników dokładności szacowania wartości brzegowej gęstości wektora obserwacji jest standardowy błąd numeryczny (ang. *Numerical Stan-*

¹ Praca została sfinansowana ze środków przyznanych Wydziałowi Zarządzania Uniwersytetu Ekonomicznego w Krakowie, w ramach dotacji na utrzymanie potencjału badawczego.

² Dlatego też w ramach wnioskowania bayesowskiego nierzadko wykorzystywane są nieformalne metody, np. kryterium informacyjne DIC (ang. *Deviance Information Criterion*, zob. Spiegelhalter, Besta, Carlina, Van der Linde, 2002).

dard Error, NSE). W pracy zaproponowano sposób obliczania tego błędu dla estymatora opartego na skorygowanej średniej harmonicznej (ang. *Corrected Harmonic Mean*, CHM; zob. Lenk, 2009; Pajor, Osiewalski, 2013) oraz estymatora opartego na skorygowanej średniej arytmetycznej (ang. *Corrected Arithmetic Mean*, CAM; zob. Pajor, 2017). W części empirycznej, dla porównania, rozważono również estymator Chiba i Jeliaskova [2001]. Metoda Chiba i Jeliaskova jest jedną z najpopularniejszych metod szacowania wartości brzegowej gęstości wektora obserwacji, należy do zbioru „złoty standardów” (ang. *gold-standard method*). Pozostałe dwie metody są stosunkowo nowe, choć oparte na znanych średnich. Przy pewnych założeniach estymatory CHM, CAM oraz Chiba i Jeliaskova są symulacyjnie zgodne (ang. *simulation-consistent*), tzn. że wraz ze wzrostem liczby losowań realizacje estymatorów zbiegają do szacowanej wartości [zob. Chib, Jeliazkov, 2001; Lenk, 2009; Pajor, 2017], jednak szybkość tej zbieżności jest różna. O rozproszeniu możliwych (w kontekście powtarzanych wielokrotnie symulacji) realizacji tych estymatorów informuje m.in. standardowy błąd numeryczny. Sposób obliczenia NSE dla estymatora Chiba i Jeliaskova można znaleźć w pracy [Chib, Jeliazkov 2001], natomiast dla pozostałych dwóch estymatorów, CHM i CAM, dokonano stosownych wyprowadzeń, które zaprezentowano w niniejszej pracy.

Artykuł składa się z trzech części. W pierwszej przedstawiono ogólne zasady formalnego porównywania modeli Bayesowskich. Część druga poświęcona jest standardowym błędom numerycznym dla estymatorów CHM i CAM. Ilustrację empiryczną, związaną z porównaniem modeli Copula-GARCH, zawarto w części trzeciej. Część ostatnią stanowi podsumowanie.

1. Porównywania modeli Bayesowskich

W ujęciu Bayesowskim podstawowym kryterium porównawczym modeli są ich prawdopodobieństwa *a posteriori*, obliczone ze wzoru Bayesa [zob. Osiewalski, Steel, 1993; Zellner, 1971]. Załóżmy, że mamy kompletny zbiór parami wykluczających się modeli Bayesowskich, $\{M_1, M_2, \dots, M_l\}$:

$$M_i: p_i(y, \eta_i) = p_i(y|\eta_i) p_i(\eta_i), \quad i = 1, \dots, l, \quad (1)$$

gdzie:

$\eta_i \in H_i$ – wektor parametrów modelu M_i ;

$p_i(\eta_i)$ – gęstość rozkładu *a priori* wektora η_i w i -tym modelu;

$p_i(y|\eta_i)$ – próbkowa gęstość wektora obserwacji, y , w i -tym modelu.

Niech ponadto $p(M_1), p(M_2), \dots, p(M_l)$ będą prawdopodobieństwami *a priori* tych modeli. Wówczas prawdopodobieństwo *a posteriori* modelu M_i wynosi:

$$p(M_i | y) = \frac{p(M_i)p(y | M_i)}{\sum_{j=1}^l p(M_j)p(y | M_j)}, \quad i = 1, 2, \dots, l \quad (2)$$

gdzie:

$p(y | M_i)$ – brzegowa gęstość wektora obserwacji w modelu M_i :

$$p(y | M_i) = p_i(y) = \int_{H_i} p_i(y | \eta_i) p_i(\eta_i) d\eta_i. \quad (3)$$

Prawdopodobieństwo *a posteriori* danego modelu zależy więc od prawdopodobieństw *a priori* wszystkich modeli oraz wartości brzegowych gęstości wektora obserwacji. Co do prawdopodobieństw *a priori* modeli, to w literaturze proponuje się przyjąć, iż są one jednakowe (czyli $p(M_i) = 1/l, i = 1, \dots, l$), lub dobrać je tak, aby modele o mniejszej liczbie parametrów były bardziej prawdopodobne, zgodnie z tzw. zasadą brzytwy Ockhama (np. $p(M_i) = a^{-k_i} / \sum_{i=1}^l a^{-k_i}, i = 1, \dots, l$,

gdzie a jest liczbą rzeczywistą większą od jedności, zaś k_i oznacza liczbę parametrów modelu M_i , por. Jeffreys, 1961; Osiewalski, Steel, 1993). Natomiast obliczenie wartości brzegowej gęstości wektora obserwacji dla zdecydowanej większości modeli nie jest możliwe w sposób analityczny i stanowi ogromne wyzwanie numeryczne. W literaturze proponuje się różne metody Monte Carlo, a w szczególności te oparte na łańcuchach Markowa (ang. *Markov – chain Monte Carlo*, MCMC), na przykład losowanie Gibbsa lub algorytm Metropolisa i Hastingsa [zob. Newton, Raftery, 1994; Kass, Raftery, 1995; Chib, 1995; Chib, Jeliazkov, 2001; Raftery, 1996; Lenk, 2009; Pajor, Osiewalski, 2013; Pajor, 2017].

W niniejszej pracy porównamy własności numeryczne trzech metod szacowania wartości brzegowej gęstości wektora obserwacji: Chiba i Jeliazkova oraz dwóch opartych na odpowiednio skorygowanej średniej harmonicznej oraz skorygowanej średniej arytmetycznej. Ostatnie dwie metody wymagają przyjęcia zbioru $A_i \subseteq H_i$, którego zarówno prawdopodobieństwo *a priori*, jak i *a posteriori* jest skończone i niezerowe³. Zbiór ten został więc ustalony na podstawie próby pseudolosowej z rozkładu *a posteriori* jako przekrój (iloczyn mnogościowy) kostki wyznaczonej przez skrajne stany łańcucha Markowa i zbioru punktów spełniających warunek, że wartość funkcji wiarygodności jest niemniejsza od najmniejszej z uzyskanych w ramach symulacji MCMC:

³ Prawdopodobieństwo *a priori* nie musi być miarą unormowaną, wystarczy, że jest miarą σ -skończoną [zob. Pajor, 2017].

$$A_i = \times_j [\min \{\eta_{i,j}^{(k)}\}, \max \{\eta_{i,j}^{(k)}\}] \cap \{\eta_i : p_i(y | \eta_i) \geq L_i\},$$

gdzie:

$$L_i = \min \{p_i(y | \eta_i^{(k)}), k = 1, \dots, K\},$$

$\{\eta_i^{(k)}\}_{k=1}^K$ – próba pseudolosowa z rozkładu *a posteriori* w *i*-tym modelu;

$\eta_{i,j}^{(k)}$ – *j*-tą składową wektora $\eta_i^{(k)}$.

Do estymacji $P_i(A_i)$ oraz obliczenia skorygowanej średniej arytmetycznej wykorzystano dodatkowo metodę Monte Carlo z funkcją ważności [por. Pajor, 2017]. Jako funkcję ważności przyjęto gęstość wielowymiarowego rozkładu normalnego, scentrowanego wokół wektora wartości oczekiwanych *a posteriori*, z macierzą kowariancji równą w przybliżeniu macierzy kowariancji rozkładu *a posteriori* (obliczoną numerycznie na podstawie symulacji MCMC).

2. Standardowy błąd numeryczny dla CHM i CAM

Miernikiem dokładności szacowania wartości brzegowej gęstości wektora obserwacji jest standardowy błąd numeryczny.

Niech $\hat{h}_i = G^{-1} \sum_{g=1}^G h_i^{(g)}$ będzie oszacowaniem całki

$E(h_i | y) = \int_{H_i} h(\eta_i) p_i(\eta_i | y) d\eta_i$, gdzie $h_i^{(g)} = h(\eta_i^{(g)})$ jest wektorem wy-

miaru $d \times 1$, $\{\eta_i^{(g)}\}_{g=1}^G$ są realizacjami ergodycznego łańcucha Markowa o rozkładzie stacjonarnym zadanym przez gęstość $p_i(\eta_i | y)$. Z twierdzeń ergodycznych dla łańcuchów Markowa wynika, że $\hat{h}_i \xrightarrow{p.w.} E(h_i | y)$, gdy $G \rightarrow +\infty$ [Tierney, 1994].

Estymatorem asymptotycznej macierzy kowariancji \hat{h}_i jest [Hannan, 1970; Newey, West, 1987; Chib, 1995]:

$$\hat{V}(\hat{h}_i) = G^{-1} \left[\Omega_{i,0} + \sum_{s=1}^b (1 - s/(b+1)) (\Omega_{i,s} + \Omega_{i,s}') \right], \quad (4)$$

gdzie: $\Omega_{i,s} = G^{-1} \sum_{g=s+1}^G (h_i^{(g)} - \hat{h}_i)(h_i^{(g-s)} - \hat{h}_i)'$, $s = 0, 1, \dots, b$.

Niech teraz $\psi_i = g(\hat{h}_i)$, gdzie $g: \mathbb{R}^d \rightarrow \mathbb{R}$ jest funkcją różniczkowalną. Wykorzystując tzw. metodę delta, można pokazać, że estymator asymptotycznej wariancji ψ_i wynosi:

$$\widehat{\text{var}} g(\hat{h}_i) \approx \frac{\partial g}{\partial \hat{h}_i} \hat{V}(\hat{h}_i) \frac{\partial g}{\partial \hat{h}_i}, \quad (5)$$

gdzie: $\frac{\partial g}{\partial \hat{h}_i}$ jest wektorem pochodnych cząstkowych funkcji g .

Pierwiastek kwadratowy z $\widehat{\text{var}} g(\hat{h}_i)$ jest standardowym błędem numerycznym dla ψ_i [por. Chib, 1995].

Sposób obliczenia NSE dla estymatora Chiba i Jeliazkova można znaleźć w pracy [Chib, Jeliazkov, 2001].

Estymator wartości brzegowej gęstości wektora obserwacji, oparty na skorygowanej średniej harmonicznej, jest następujący:

$$\hat{p}_{i,CHME}(y) = \hat{P}_i(A_i) \left(\frac{1}{K} \sum_{k=1}^K \frac{I_{A_i}(\eta_i^{(k)})}{p_i(y | \eta_i^{(k)})} \right)^{-1}, \quad (6)$$

gdzie: $\{\eta_i^{(k)}\}_{k=1}^K$ jest próbą pseudolosową z rozkładu *a posteriori* w i -tym modelu [zob. Pajor, Osiewalski, 2013]. Zatem:

$$\ln \hat{p}_{i,CHME}(y) = \ln \hat{P}_i(A_i) - \ln \left(K^{-1} \sum_{k=1}^K \frac{I_{A_i}(\eta_i^{(k)})}{p_i(y | \eta_i^{(k)})} \right), \quad (7)$$

ponadto:

$$\ln \hat{P}_i(A_i) = \ln \left(J^{-1} \sum_{j=1}^J \frac{I_{A_i}(\eta_i^{s(j)}) p_i(\eta_i^{s(j)})}{s_i(\eta_i^{s(j)})} \right), \quad (8)$$

gdzie: $\{\eta_i^{s(j)}\}_{k=1}^J$ jest próbą pseudolosową z rozkładu o gęstości s_i (zwanej funkcją ważności w i -tym modelu). Analogicznie jak w [Chib, Jeliazkov, 2001], przyjęto $K = J$ oraz:

$$h_i^{(k)} = \begin{bmatrix} h_{i,1}^{(k)} \\ h_{i,2}^{(k)} \end{bmatrix} = \begin{bmatrix} \frac{I_{A_i}(\eta_i^{s(k)}) p_i(\eta_i^{s(k)})}{s_i(\eta_i^{s(k)})} \\ \frac{I_{A_i}(\eta_i^{(k)})}{p_i(y | \eta_i^{(k)})} \end{bmatrix}, \quad \hat{h}_i = K^{-1} \sum_{k=1}^K h_i^{(k)} = (\hat{h}_{i,1}, \hat{h}_{i,2})'.$$

W tym przypadku $\psi_i = \ln \hat{h}_{i,1} - \ln \hat{h}_{i,2}$. Dodatkowo, ponieważ losowania z rozkładu *a posteriori* oraz z rozkładu o gęstości s_i są niezależne, to estymator przybliżonej wariancji logarytmu naturalnego estymatora wartości brzegowej gęstości wektora obserwacji wynosi:

$$\widehat{\text{var}}(\ln \hat{p}_{i,CHME}(y)) = \widehat{\text{var}}(\ln \hat{h}_{i,1}) + \widehat{\text{var}}(\ln \hat{h}_{i,2}), \quad (9)$$

gdzie: $\widehat{\text{var}}(\ln \hat{h}_{i,1}) \approx \widehat{\text{var}}(\hat{h}_{i,1}) / \hat{h}_{i,1}^2$, $\widehat{\text{var}}(\ln \hat{h}_{i,2}) \approx \widehat{\text{var}}(\hat{h}_{i,2}) / \hat{h}_{i,2}^2$. Standardowy błąd numeryczny oszacowania $\ln p_i(y)$ obliczamy ze wzoru:

$$NSE_{CHME} \approx \sqrt{\widehat{\text{var}}(\hat{h}_{i,1}) / \hat{h}_{i,1}^2 + \widehat{\text{var}}(\hat{h}_{i,2}) / \hat{h}_{i,2}^2}. \quad (10)$$

Estymator oparty na skorygowanej średniej arytmetycznej (z uwzględnieniem losowania z funkcją ważności s_i) można zapisać:

$$\hat{p}_{i,CAME}(y) = \frac{1}{\hat{P}_i(A_i | y)} \frac{1}{K} \sum_{k=1}^K \frac{p_i(y | \eta_i^{(k)}) p_i(\eta_i^{(k)}) I_{A_i}(\eta_i^{(k)})}{s_i(\eta_i^{(k)})}, \quad (11)$$

gdzie: $\{\eta_i^{(k)}\}_{k=1}^K$ jest próbą pseudolosową z rozkładu o gęstości s_i [zob. Pajor, 2017]. A zatem:

$$\ln \hat{p}_{i,CAME}(y) = \ln \left(K^{-1} \sum_{k=1}^K \frac{p_i(y | \eta_i^{(k)}) p_i(\eta_i^{(k)}) I_{A_i}(\eta_i^{(k)})}{s_i(\eta_i^{(k)})} \right) - \ln \hat{P}_i(A_i | y). \quad (12)$$

Ponadto: $\ln \hat{P}_i(A_i | y) = \ln \left(J^{-1} \sum_{j=1}^J I_{A_i}(\eta_i^{p(j)}) \right)$, gdzie: $\{\eta_i^{p(j)}\}_{j=1}^J$ jest próbą pseudolosową z rozkładu *a posteriori*. Zbiór A_i został jednak tak dobrany, że $\hat{P}_i(A_i | y) = 1$. Zatem standardowy błąd numeryczny dla estymatora CAM jest związany tylko z pierwszym składnikiem, czyli:

$$NSE_{CAME} \approx \sqrt{\widehat{\text{var}}(\hat{h}_i) / \hat{h}_i^2}, \quad (13)$$

gdzie: $\hat{h}_i = K^{-1} \sum_{k=1}^K h_i^{(k)}$, natomiast $h_i^{(k)} = \frac{p_i(y | \eta_i^{(k)}) p_i(\eta_i^{(k)}) I_{A_i}(\eta_i^{(k)})}{s_i(\eta_i^{(k)})}$.

3. Ilustracja empiryczna

Własności numeryczne wybranych metod szacowania wartości brzegowej gęstości wektora obserwacji przedstawimy w kontekście porównania modeli Copula-AR-GARCH.

3.1. Bayesowskie modele Copula-AR(1)- t -GARCH(1,1)

Niech $\{x_t = (x_{1,t}, x_{2,t})', t = -1, 0, 1, \dots, T\}$ oznacza szereg czasowy cen aktywów finansowych w chwili t . Dla logarymicznych stóp zwrotu $\{y_t = (y_{1,t}, y_{2,t})', t = 0, 1, 2, \dots, T\}$, obliczonych według formuły $y_{i,t} = 100 \ln(x_{i,t}/x_{i,t-1})$ $t = 1, \dots, T$, $i = 1, 2$, przyjęto następującą strukturę Copula-AR(1)- t -GARCH(1,1) [zob. Mokrzycka, Pajor, 2016]:

$$y_{i,t} = \beta_{i,0} + \beta_{i,1}y_{i,t-1} + z_{i,t}, \quad (14)$$

$$z_{i,t} = \varepsilon_{i,t} \sqrt{h_{i,t}}, \quad (15)$$

$$h_{i,t} = \alpha_{i,0} + \alpha_{i,1}z_{i,t-1}^2 + \gamma_{i,1}h_{i,t-1}, \quad i = 1, 2, t = 1, 2, \dots, T, \quad (16)$$

gdzie: $\alpha_{i,0} > 0$, $\alpha_{i,1} \geq 0$, $\gamma_{i,1} \geq 0$, $\{\varepsilon_{i,t}\}_{t=1}^T \sim iit(0,1, v_i)$ dla każdego $i \in \{1, 2\}$ ⁴.

Struktura zależności wektora $(\varepsilon_{1,t}, \varepsilon_{2,t})'$ zadana jest poprzez kopulę o gęstości $c(u_1, u_2)$, czyli gęstość łącznego rozkładu tego wektora jest postaci:

$$p_{\varepsilon}(\varepsilon_{1,t}, \varepsilon_{2,t}) = c(F_{St, v_1}(\varepsilon_{1,t}), F_{St, v_2}(\varepsilon_{2,t})) f_{St}(\varepsilon_{1,t}; 0, 1, v_1) f_{St}(\varepsilon_{2,t}; 0, 1, v_2), \quad (17)$$

gdzie $F_{St, v}(\cdot)$ oznacza dystrybuantę jednowymiarowego rozkładu t Studenta o zerowej modalnej, jednostkowej precyzji i v stopniach swobody, zaś $f_{St}(\cdot; 0, 1, v)$ jest gęstością tego rozkładu.

W badaniu wykorzystano jedenaście postaci funkcji gęstości $c(u_1, u_2)$. Rozważono dwie kopule eliptyczne: gaussowską i t Studenta; pięć kopuli archimedeesowych: Franka, Clayтона, Gumbela, Clayтона-Gumbela (BB1) i Joego-Claytona (BB7); symetryzowaną kopulę Joego-Claytona; dwie kopule przeżycia: obrócone (o 180 stopni) kopule Clayтона i Gumbela [zob. Joe, 1987; Doman, 2011]. Ponadto wzięto pod uwagę kopulę odpowiadającą niezależności rozkładów warunkowych, $c(u_1, u_2) = 1$. Podstawowe charakterystyki tych kopuli oraz postaci funkcji gęstości zestawiono w pracy Mokrzyckiej i Pajor [2016]. Również rozkłady *a priori* parametrów rozważanych modeli przyjęto jak w tej pracy.

3.2. Wyniki empiryczne

Przedmiotem modelowania jest zmienność i struktura zależności dziennych, procentowych logarymicznych stóp zwrotu, wyliczonych na podstawie kursu zamknięcia subindeksów sektorowych indeksu WIG: WIG-Banki, WIG-Informatyka, WIG-Spożywczy i WIG-Budownictwo⁵. Analizowane szeregi zawierają po 2539 obserwacji pochodzących z okresu od 1 sierpnia 2005 r. do

⁴ Zapis $\{\varepsilon_{i,t}\}_{t=1}^T \sim iit(0,1, v_i)$ oznacza, że zmienne losowe $\{\varepsilon_{i,t}, t = 1, 2, \dots, T\}$ są niezależne i mają rozkład t Studenta z zerową modalną, jednostkową precyzją i v_i stopniami swobody.

⁵ Dane pochodzą ze strony www.stooq.pl i były przedmiotem analizy w pracy [Mokrzycka, Pajor, 2016].

21 września 2015 r. Dwie obserwacje z każdego szeregu zostały wykorzystane jako warunki początkowe, stąd liczba obserwacji modelowanych wyniosła $T = 2537$ dla każdego szeregu czasowego.

Do symulacji z rozkładu *a posteriori* wykorzystano algorytm Metropolis i Hastings. Wstępne stany łańcucha generowano z rozkładu *t* Studenta z trzema stopniami swobody, scentrowanego wokół ostatniego stanu łańcucha Markowa, i macierzą kowariancji ustaloną na podstawie wstępnych przebiegów algorytmu [por. Osiewalski, Pipień, 2004]. Uzyskane w ten sposób wektory pseudolosowe (z rozkładu *a posteriori* jako rozkładu stacjonarnego) wykorzystano do obliczenia ocen wartości brzegowej gęstości wektora obserwacji.

Tabela 1. Oszacowanie wartości brzegowej gęstości macierzy obserwacji oraz standardowy błąd numeryczny (dla par subindeksów: WIG-Informatyka – WIG-Spożywczy, WIG-Informatyka – WIG-Budownictwo, WIG-Informatyka – WIG-Banki)*

Kopula	WIG-Informatyka – WIG-Spożywczy			WIG-Informatyka – WIG-Budownictwo			WIG-Informatyka – WIG-Banki		
	ChibJ	CHME	CAME	ChibJ	CHME	CAME	ChibJ	CHME	CAME
Franka	-8332,6 (0,972)	-8344,5 (0,368)	-8331,6 (0,004)	-8145,7 (0,352)	-8147,8 (0,655)	-8146,4 (0,005)	-8448,4 (0,954)	-8461,6 (0,463)	-8443,6 (0,006)
Claytona	-8316,4 (0,994)	-8304,0 (0,506)	-8304,5 (0,005)	-8136,7 (0,985)	-8129,5 (0,479)	-8129,1 (0,005)	-8473,5 (0,937)	-8463,5 (0,216)	-8463,9 (0,005)
Obrócona Claytona	-8410,6 (0,989)	-8398,6 (0,335)	-8400,3 (0,005)	-8230,3 (0,98)	-8221,3 (0,232)	-8221,7 (0,005)	-8551,6 (0,979)	-8539,0 (0,414)	-8541,7 (0,005)
Gumbela	-8362,2 (0,989)	-8350,2 (0,360)	-8352,2 (0,005)	-8158,7 (0,988)	-8150,2 (0,318)	-8150,3 (0,005)	-8468,4 (0,984)	-8453,0 (0,716)	-8457,3 (0,005)
Obrócona Gumbela	-8302,4 (0,984)	-8287,2 (0,291)	-8290,4 (0,005)	-8094,6 (0,328)	-8093,4 (0,528)	-8093,0 (0,005)	-8424,0 (0,971)	-8409,1 (0,421)	-8413,3 (0,005)
Gumbela- Claytona	-8305,3 (0,989)	-8290,3 (0,702)	-8292,2 (0,005)	-8092,9 (0,999)	-8084,7 (0,535)	-8084,2 (0,006)	-8395,4 (0,367)	-8396,9 (0,783)	-8394,5 (0,005)
Joego- Claytona	-8306,4 (0,994)	-8292,0 (0,497)	-8293,7 (0,005)	-8096,4 (0,984)	-8087,9 (0,532)	-8086,6 (0,005)	-8411,6 (0,979)	-8398,7 (0,660)	-8400,4 (0,005)
Sym. Joego- Claytona	-8305,5 (0,994)	-8294,8 (0,566)	-8292,7 (0,005)	-8094,9 (0,989)	-8086,5 (0,688)	-8085,4 (0,006)	-8413,4 (0,963)	-8398,3 (0,385)	-8400,4 (0,006)
Normalna	-8322,9 (0,989)	-8310,5 (0,523)	-8312,6 (0,005)	-8106,6 (0,275)	-8106,2 (0,333)	-8105,4 (0,005)	-8414,0 (0,989)	-8402,4 (0,626)	-8403,7 (0,006)
<i>t</i> Studenta	-8310,1 (0,989)	-8296,2 (0,487)	-8297,4 (0,006)	-8102,2 (0,98)	-8093,8 (0,86)	-8089,7 (0,006)	-8395,3 (0,623)	-8395,0 (0,460)	-8394,1 (0,007)
Niezależność	-8529,8 (0,994)	-8520,1 (0,716)	-8519,0 (0,005)	-8461,1 (0,984)	-8453,3 (0,42)	-8452,9 (0,006)	-8872,1 (0,883)	-8856,6 (0,641)	-8859,0 (0,006)

* ChibJ – metoda Chiba i Jeliaskova, CHME – skorygowana średnia harmoniczna, CAME – skorygowana średnia arytmetyczna. NSE podano w nawiasach.

Źródło: Wyniki dotyczące ChibJ i CHME pochodzą z pracy [Mokrzycka, Pajor, 2016]. Pozostałe to obliczenia własne.

W tabelach 1 i 2 zamieszczono logarytmy naturalne oszacowań wartości brzegowej gęstości wektora obserwacji wraz z standardowymi błędami numerycznymi (w nawiasach)⁶. Jak widać, najmniejszymi wartościami NSE charakteryzuje się estymator oparty na skorygowanej średniej arytmetycznej. Dla wszystkich par szeregów czasowych są one ponad dwa rzędy wielkości mniejsze niż te uzyskane dla estymatora Chiba i Jeliaskova. Standardowe błędy numeryczne dla estymatora opartego na skorygowanej średniej harmonicznnej są około 1,5 rzędu wielkości większe niż te uzyskane dla estymatora CAM. Precyzja estymatora CAM jest więc dużo większa niż pozostałych dwóch estymatorów. Najgorzej w tym porównaniu wypadł estymator Chiba i Jeliaskova.

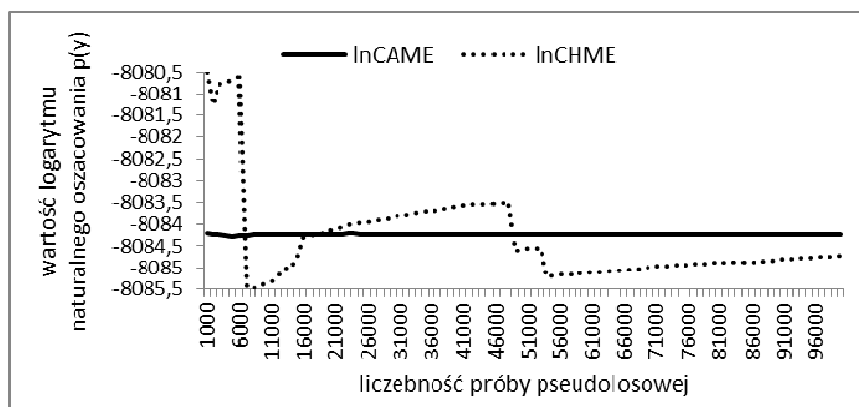
Tabela 2. Oszacowanie wartości brzegowej gęstości macierzy obserwacji oraz standardowy błąd numeryczny (dla par subindeksów: WIG-Spożywczy – WIG-Budownictwo, WIG-Banki – WIG-Spożywczy, WIG-Banki – WIG-Budownictwo)*

Kopuła	WIG-Spożywczy – WIG-Budownictwo			WIG-Banki – WIG-Spożywczy			WIG-Banki – WIG-Budownictwo		
	ChibJ	CHME	CAME	ChibJ	CHME	CAME	ChibJ	CHME	CAME
Franka	-8463,1 (0,108)	-8464,5 (0,745)	-8464,4 (0,005)	-8850,3 (0,318)	-8852,1 (0,248)	-8846,9 (0,004)	-8593,4 (0,993)	-8602,1 (0,485)	-8594,4 (0,005)
Claytona	-8438,1 (0,999)	-8430,9 (0,679)	-8430,1 (0,004)	-8851,4 (0,994)	-8838,4 (0,655)	-8839,7 (0,005)	-8626,5 (0,976)	-8618,1 (0,665)	-8617,4 (0,004)
Obrócona Claytona	-8552,5 (0,980)	-8544,9 (0,919)	-8542,9 (0,005)	-8945,5 (0,957)	-8933,5 (0,187)	-8935,8 (0,005)	-8708,7 (0,983)	-8695,1 (0,306)	-8700,0 (0,005)
Gumbela	-8501,4 (0,979)	-8493,2 (0,555)	-8492,2 (0,005)	-8894,4 (0,987)	-8880,0 (0,269)	-8883,9 (0,005)	-8625,5 (0,971)	-8617,9 (0,63)	-8616,1 (0,005)
Obrócona Gumbela	-8409,4 (0,198)	-8408,3 (0,473)	-8408,1 (0,004)	-8830,8 (0,988)	-8815,0 (0,334)	-8818,5 (0,005)	-8572,5 (0,987)	-8560,4 (0,624)	-8564,6 (0,005)
Gumbela-Claytona	-8420,4 (0,178)	-8419,5 (0,277)	-8419,5 (0,005)	-8836,2 (0,994)	-8817,1 (0,656)	-8823,0 (0,005)	-8561,8 (0,989)	-8551,3 (0,349)	-8551,5 (0,005)
Joego-Claytona	-8433,1 (0,967)	-8422,0 (0,352)	-8421,9 (0,005)	-8840,5 (0,989)	-8823,7 (0,549)	-8827,1 (0,005)	-8571,2 (0,967)	-8560,0 (0,749)	-8558,9 (0,005)
Sym. Joego-Claytona	-8425,7 (0,989)	-8415,0 (0,376)	-8415,4 (0,005)	-8837,3 (0,984)	-8820,8 (0,504)	-8824,3 (0,005)	-8567,7 (0,993)	-8553,6 (0,537)	-8557,4 (0,005)
Normalna	-8449,3 (0,988)	-8433,6 (0,265)	-8439,8 (0,005)	-8846,5 (0,984)	-8831,7 (0,465)	-8835,6 (0,004)	-8566,7 (0,971)	-8552,1 (0,334)	-8556,9 (0,005)
t Studenta	-8441,6 (0,967)	-8427,8 (1,014)	-8427,8 (0,006)	-8837,8 (0,994)	-8825,2 (0,952)	-8825,4 (0,006)	-8562,0 (0,988)	-8544,2 (0,384)	-8548,8 (0,006)
Niezależność	-8687,8 (0,993)	-8680,3 (0,550)	-8679,6 (0,005)	-9097,7 (0,984)	-9087,6 (0,844)	-9085,7 (0,005)	-9028,0 (0,999)	-9020,4 (0,817)	-9019,6 (0,006)

* ChibJ – metoda Chiba i Jeliaskova, CHME – skorygowana średnia harmonicznna, CAME – skorygowana średnia arytmetyczna, NSE podano w nawiasach.

Źródło: Wyniki dotyczące ChibJ i CHME pochodzą z pracy [Mokrzycka, Pajor, 2016]. Pozostałe to obliczenia własne.

⁶ Prezentowane wyniki uzyskano na podstawie 110 000 iteracji algorytmu MH, w tym 10 000 cykli spalonych oraz 100 000 dodatkowych losowań wykonanych w ramach metody Chiba i Jeliaskova oraz w ramach MCIS dla korekty Lenka i skorygowanej średniej arytmetycznej. Obliczenia wykonano w programie GAUSS 9.0.



Rys. 1. Wartości logarytmów naturalnych skorygowanej średniej arytmetycznej oraz skorygowanej średniej harmonicznej w modelu z kopułą Claytona-Gumbela dla pary WIG-Informatyka – WIG-Budownictwo

Dla przykładu, na rysunku 1 przedstawiono, jak wartości skorygowanej średniej harmonicznej oraz arytmetycznej, w modelu najlepiej dopasowanym do pary WIG-Informatyka i WIG-Budownictwo, zmieniały się w zależności od liczby punktów wylosowanych z przestrzeni parametrów. Wyraźnie widać, że estymator CAM generuje bardziej stabilne oszacowania $\ln p_i(y)$ niż estymator CHM. Obserwowane wahania $\ln \hat{p}_{i,CAME}(y)$ szybko wygasają wraz ze wzrostem liczby losowań. Również wahania $\ln \hat{p}_{i,CHME}(y)$ wygasają, ale znacznie wolniej.

Podsumowanie

W pracy zaproponowano sposób obliczania standardowego błędu numerycznego dla estymatorów CHM i CAM oraz porównano numeryczne własności trzech wybranych estymatorów wartości brzegowej gęstości wektora obserwacji w kontekście porównania modeli Copula-GARCH. Rozważono estymator Chiba i Jeliaskova, estymatory oparte na skorygowanej średniej harmonicznej (CHM) oraz skorygowanej arytmetycznej (CAM). Jeśli jako kryterium porównania przyjmujemy standardowy błąd numeryczny, to estymator CAM wypada najlepiej.

Literatura

- Chib S. (1995), *Marginal Likelihood From the Gibbs Output*, "Journal of the American Statistical Association", No. 90, s. 1313-1321.
- Chib S., Jeliaskov I. (2001), *Marginal Likelihood from the Metropolis-Hastings Output*, "Journal of the American Statistical Association", No. 96, s. 270-281.

- Doman R. (2011), *Zastosowania kopuli w modelowaniu dynamiki zależności na rynkach finansowych*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
- Grønneberg S., Hjort N.L. (2014), *The Copula Information Criteria*, "Scandinavian Journal of Statistics", No. 42(2), s. 436-459, <http://dx.doi.org/10.1111/sjos.12042>.
- Hannan H.J. (1970), *Multiple Time Series*, Wiley, New York.
- Jeffreys H. (1961), *Theory of Probability*, Oxford University Press, London.
- Joe H. (1997), *Multivariate Models and Dependence Concepts*, Chapman and Hall, London.
- Kass R.E., Raftery A.E. (1995), *Bayes Factors*, "Journal of the American Statistical Association", No. 90, s. 773-795.
- Lenk P. (2009), *Simulation Pseudo-Bias Correction to the Harmonic Mean Estimator of Integrated Likelihoods*, "Journal of Computational and Graphical Statistics", No. 18, s. 941-960.
- Mokrzycka J., Pajor A. (2016), *Formalne porównanie modeli Copula-AR(1)-t-GARCH(1,1) dla subindeksów indeksu WIG*, „Przegląd Statystyczny” [w druku].
- Newey W.K., West K.D. (1987), *A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix*, "Econometrica", No. 55, s. 703-708.
- Newton M.A., Raftery A.E. (1994), *Approximate Bayesian Inference by the Weighted Likelihood Bootstrap*, "Journal of the Royal Statistical Society B", No. 56(1), s. 3-48.
- Osiewalski J., Pipień M. (2004), *Bayesian Comparison of Bivariate ARCH-type Models for the Main Exchange Rates in Poland*, "Journal of Econometrics", No. 123, 371-391.
- Osiewalski J., Steel M.F.J. (1993), *A Bayesian Perspective on Model Selection* [maszynopis], opublikowano w języku hiszpańskim: *Una perspectiva bayesiana en sección de modelos*, "Cuadernos Economicos ICE", No. 55, s. 327-351.
- Pajor A., Osiewalski J. (2013), *A Note on Lenk's Correction of the Harmonic Mean Estimator*, "Central European Journal of Economic Modelling and Econometrics", No. 5(4), s. 271-275, correction: No. 6(1), s. 69.
- Pajor A. (2016), *Estimating the Marginal Likelihood Using the Arithmetic Mean Identity*, "Bayesian Analysis", No. 12(1), s. 261-287.
- Raftery A.E. (1996), *Hypothesis Testing and Model Selection* [in:] W.R. Gilks, D.J. Spiegelhalter, S. Richardson (eds.), *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, s. 163-188.
- Spiegelhalter D.J., Best N.G., Carlin B.P., Van der Linde A. (2002), *Bayesian Measures of Model Complexity and Fit [with Discussion]*, "Journal of the Royal Statistical Society", Vol. B, No. 64(4), s. 583-616.
- Tierney L. (1994), *Markov Chains for Exploring Posterior Distributions [with Discussion]*, "The Annals of Statistics", No. 22, s. 1701-1762.
- Zellner A. (1971), *An Introduction to Bayesian Inference in Econometrics*, J. Wiley, New York.

NUMERICAL STANDARD ERROR FOR CHM AND CAM ESTIMATORS

Summary: The main aim of the paper is to propose methods for calculating numerical standard errors of the corrected harmonic as well as arithmetic mean estimators of the marginal likelihood. We apply these two estimators in Copula-AR-GARCH models for the daily growth rates of four sub-indices of the stock index WIG, published by the Warsaw Stock Exchange. For the sake of comparison Chib and Jeliazkov estimator (as a gold-standard) is also considered. Empirical results demonstrate that the corrected arithmetic mean estimator performs best. It is characterised by smallest numerical standard errors.

Keywords: numerical standard error, corrected arithmetic mean, corrected arithmetic mean.