MODEL DWUMIANOWY II RZĘDU I SKOŚNY ROZKŁAD STUDENTA W ANALIZIE RYZYKA KREDYTOWEGO*

Jacek Osiewalski

e-mail: eeosiewa@cyf-kr.edu.pl
Katedra Ekonometrii
Akademia Ekonomiczna
PL 31-150 Kraków, ul. Rakowicka 27

Jerzy Marzec

e-mail: marzecj@ae.krakow.pl
Katedra Ekonometrii
Akademia Ekonomiczna
PL 31-150 Kraków, ul. Rakowicka 27

Praca przedstawiona przez autorów na posiedzeniu Komisji Nauk Ekonomicznych i Komisji Statystyczno-Demograficznej Oddziału PAN w Krakowie dniu 23 czerwca 2004 r.

ABSTRACT

J. Osiewalski, J. Marzec., Binomial model of order 2 and the skewed Student *t* distribution in the analysis of loan risk, Folia Oeconomica Cracoviensia.

In this paper binomial choice models of order 1 and 2 are distinguished. They are all based on $F(x_t\beta)$, where $F(.)$ is some cumulative distribution function. In usual order 1 models, x_t consists of original explanatory variables w_{tj}, while order 2 models also use squares and products of w_{tj}, thus making $x_t\beta$ a second order polynomial in w_{tj}. We use the cumulative distribution function of the two-parameter family of skewed Student *t* distributions as the functional form of *F*. This allows us to test special cases, which are based on a symmetric *t* distribution or on a normal distribution (the probit model). In the (skewed) Student case (with unknown degrees of freedom), the likelihood function does not integrate to a constant and the ML estimator has unknown properties. Also, in order 2 models multicollinearity can be a severe problem. Hence we advocate the Bayesian approach with proper priors for the parameters and propose the Metropolis-Hastings MCMC algorithm to draw from the posterior.

Our example uses the proposed Bayesian model and the data on consumer loans (from 39040 bank accounts) in order to assess risk of an individual loan. Empirical results show that our order 2 model cannnot be reduced to its order 1 submodel. Also, the CDF of a skewed *t* distribution with very low degrees of freedom and strong left skewness is most adequate from the statistical viewpoint.

KEY WORDS - SŁOWA KLUCZOWE
discrete choice model, Bayesian inference, skewed *t* distribution, scoring model, loan risk
model dyskretnego wyboru, wnioskowanie bayesowskie, skośny rozkład *t*, model scoringowy, ryzyko kredytowe

* Praca przygotowana w ramach badań statutowych Akademii Ekonomicznej w Krakowie w roku 2004.

1. WPROWADZENIE

Model dwumianowy (dychotomiczny) jest podstawowym modelem wyjaśniającym jakościową zmienną endogeniczną. Przedstawia on zależność między prawdopodobieństwem wyboru jednej z dwóch możliwości (oznaczanych umownie jako 0 i 1) a egzogenicznymi zmiennymi objaśniającymi, które opisują cechy możliwych alternatyw lub indywidualne charakterystyki podmiotów podejmujących decyzję. Postać tego modelu jest następująca:

$$p_t \equiv \Pr(y_t = 1) = G(x_t \cdot \beta) = 1 - F(-x_t \cdot \beta) \quad \text{dla } t=1,\ldots,T, \tag{1}$$

gdzie β jest wektorem $k\times1$ nieznanych parametrów ($\beta\in R^k$), $x_t = (x_{t1} \ldots\ x_{tk})$ oznacza wektor ustalonych wartości k zmiennych egzogenicznych lub ich znanych funkcji, zaś $G(\cdot)$ i $F(\cdot)$ są funkcjami wiążącymi p_t, czyli prawdopodobieństwo zaobserwowania sukcesu, z x_t i β. Modelem dwumianowym I rzędu nazywamy taką specyfikację (1), w której x_t zawiera tylko oryginalne zmienne egzogeniczne i jedynkę (odpowiadającą wyrazowi wolnemu), a $x_t\beta$ jest liniową funkcją tych wielkości (przy ustalonym β). Model I rzędu jest najczęstszym, powszechnie rozważanym i stosowanym przypadkiem modelu dychotomicznego. Natomiast modelem dwumianowym II rzędu nazywamy taki, w którym $x_t\beta$ jest wielomianem drugiego stopnia względem zmiennych egzogenicznych, czyli x_t zawiera też kwadraty i iloczyny wartości oryginalnych zmiennych. Oczywiście, w obu przypadkach (modele I i II rzędu) $x_t\beta$ jest liniową funkcją parametrów (przy ustalonym x_t), różny jest natomiast wymiar wektora β.

Funkcja $F(\cdot)$ we wzorze (1) ma własności dystrybuanty rozkładu prawdopodobieństwa i określa klasę modelu. Równoważną specyfikację otrzymujemy przez wprowadzenie modelu regresji liniowej (ze względu na β) dla ukrytej (nie obserwowalnej) zmiennej ciągłej z_t, której znak determinuje zaobserwowaną wartość y_t (0 lub 1):

$$z_t = x_t \cdot \beta + \varepsilon_t,$$
$$y_t = I_{[0,\infty)}(z_t) = \begin{cases} 1, & \text{gdy} \quad z_t \geq 0, \\ 0, & \text{gdy} \quad z_t < 0, \end{cases} \tag{2}$$

czyli $I_A(.)$ jest funkcją charakterystyczną zbioru A. O składnikach losowych ε_t zakłada się, że są niezależne i mają ten sam rozkład o zerowym parametrze położenia i jednostkowej skali (zwykle wartości oczekiwanej i wariancji, jeśli istnieją). Dla rozkładu symetrycznego zapis (1) sprowadza się do $p_t \equiv \Pr(y_t = 1) = F(x_t \cdot \beta)$. Szczegóły dotyczące niebayesowskiej estymacji modeli jakościowej zmiennej zależnej oraz wiele ich zastosowań empirycznych z zakresu ekonomii prezentują m.in. Amemiya (1981, 1985), Maddala (1983) i Greene (1993).

Najbardziej znanymi i powszechnie stosowanymi specyfikacjami są modele probitowy i logitowy, które odpowiadają przyjęciu dla ε_t rozkładu normalnego lub logistycznego. Do ich estymacji wykorzystywana jest zwykle metoda największej wiarygodności (MNW), mająca pożądane własności stochastyczne. Naturalne uogólnienie modelu probitowego polega na przyjęciu dla ε_t rozkładu *t* Studenta o nieznanej liczbie stopni swobody $\nu > 0$, co dopuszcza brak wariancji ($\nu \leq 2$) a nawet wartości oczekiwanej zmiennej ε_t ($\nu \leq 1$). Z tych powodów rozważamy rozkłady o zerowej modalnej i jednostkowej precyzji. Klasa rozkładów *t* Studenta zawiera rozkład normalny jako przypadek graniczny ($\nu = +\infty$), zaś – jak podają Albert i Chib (1993) – rozkład logistyczny może być przybliżany przez rozkład *t* Studenta o ok. 7 – 9 stopniach swobody. Uogólnienie to pozwala zatem testować (choćby w przybliżeniu) empiryczną adekwatność dwóch podstawowych modeli dwumianowych. Jednak zastosowanie MNW w tym przypadku jest niewskazane, ponieważ nie są znane własności estymatora MNW dla modeli z nieznanym parametrem ν.

Albert i Chib (1993) zaproponowali specyfikację i estymację bayesowskiego modelu dychotomicznego z rozkładem *t* Studenta. W celu numerycznej aproksymacji brzegowych rozkładów a posteriori interesujących wielkości wykorzystali algorytm Gibbsa, metodę Monte Carlo typu łańcuchów Markowa (ang. *Markov Chain Monte Carlo*, MCMC). Marzec (2003c) wykorzystał to podejście w modelu II rzędu w celu zbadania ryzyka pojedynczych umów kredytowych klientów detalicznych banku komercyjnego. Wyniki empiryczne wskazywały na konieczność zastosowania modelu II rzędu opartego na rozkładzie *t* Studenta, gdyż redukcja do modelu I rzędu okazała się bezzasadna, a rozkład a posteriori parametru ν skupiony był w przedziale (1, 2) – świadcząc o nieadekwatności specyfikacji probitowej czy logitowej.

Wszystkie trzy rozważane rozkłady prawdopodobieństwa (normalny, logistyczny, *t* Studenta) charakteryzują się symetrią, różniąc się grubością ogonów (szybkością zbieżności dystrybuanty do wartości granicznych 0 i 1). Proponujemy więc dalsze uogólnienie modelu probitowego, które polega na przyjęciu dla ε_t klasy skośnych rozkładów *t* Studenta; zob. też Osiewalski i Marzec (2004). Klasa ta jest charakteryzowana przez dwa parametry: stopnie swobody ν i współczynnik asymetrii γ. Estymacja parametrów β, ν, γ i ich funkcji możliwa jest na gruncie bayesowskim przy wykorzystaniu metod MCMC.

Asymetryczne rozkłady wielowymiarowe (w tym typu *t* Studenta) rozważali Fernández, Osiewalski i Steel (1995), natomiast szczegółową definicję i formalne własności skośnego rozkładu *t* w przypadku jednowymiarowym podali Fernández i Steel (1998), którzy rozkład ten zastosowali dla składnika losowego modelu regresji liniowej. Z kolei Osiewalski i Pipień (1999, 2000) wykorzystali go jako rozkład warunkowy w modelach GARCH dla finansowych szeregów czasowych, wskazując na jego użyteczność w ekonometrii finansowej.

W tej pracy omówimy genezę i własności modeli rzędu II (część 2) oraz bayesowski model dychotomiczny wykorzystujący dystrybuantę skośnego rozkładu *t* (część 3). Głównym celem pracy jest empiryczna weryfikacja użyteczności obu uogólnień w badaniu spłacalności kredytu;

zagadnieniu temu poświęcono część 4. W części 5 przedstawiamy podstawy i wyniki formalnego bayesowskiego porównania konkurencyjnych modeli dychotomicznych spłacania kredytu.

2. MODELE DWUMIANOWE I I II RZĘDU

Głównymi charakterystykami wyznaczanymi na podstawie modelu dwumianowego są efekty krańcowe. Jeżeli oryginalne zmienne objaśniające w_{tj} (j=1,...m) mogą przyjmować dowolne wartości rzeczywiste i nie są powiązane zależnościami funkcyjnymi, to h-ty efekt krańcowy (tj. zmiana prawdopodobieństwa p_t na skutek wzrostu w_{th} o małą jednostkę) jest równy pochodnej cząstkowej $\partial p_t / \partial w_{th}$. Dla modeli I rzędu, tj. gdy $k=m+1$ i $x_{tj}=w_{tj}$ (j=1,...,m), efekt krańcowy ma postać $\partial p_t / \partial w_{th} = \beta_h f(-x_t \beta)$, gdzie f jest gęstością odpowiadającą dystrybuancie F, definiującej konkretny model dychotomiczny. W tym przypadku ilorazy efektów krańcowych są niezależne od zmiennych objaśniających (stałe) i równe β_h / β_i, co jest bardzo mocnym założeniem. Aby uzmiennić względne efekty krańcowe, Marzec (2003c) wykorzystał wielomian stopnia 2 względem zmiennych w_{tj}, czyli przyjął

$$x_t \beta = \beta_1 + \sum_j w_{tj} \beta_j + \sum_j \sum_{i \geq j} w_{tj} w_{ti} \beta_{ij} \ .$$

Ponieważ efekty krańcowe wynoszą $\partial p_t / \partial w_{th} = f(-x_t \beta) \times \partial (x_t \beta) / \partial w_{th}$, więc w modelu II rzędu ich ilorazy są ilorazami liniowych funkcji zmiennych w_{tj}. Oczywiście, taka postać efektu krańcowego może wydać się szczególna, powstaje więc pytanie o uzasadnienie modeli II rzędu i o możliwość ich uogólniania. Rozważmy model dwumianowy postaci

$$p_t \equiv \Pr(y_t = 1) = G(a(w_{t1},..., w_{tm})) = 1 - F(-a(w_{t1},..., w_{tm})),$$

odpowiadający zależności $z_t = a(w_{t1},...,w_{tm}) + \varepsilon_t$ dla zmiennej ukrytej z_t. Jeśli a jest funkcją różniczkowalną, to efekty krańcowe mają postać $\partial p_t / \partial w_{th} = f(-a(w_{t1},..., w_{tm})) \times \partial a / \partial w_{th}$. Niech a będzie funkcją posiadającą rozwinięcie w szereg Taylora w otoczeniu ustalonego punktu z dziedziny. Wówczas model I rzędu można interpretować jako stosujący – zamiast funkcji a – jej lokalną aproksymacją rzędu I, czyli funkcję liniową, zaś model rzędu II jako oparty na lokalnym przybliżeniu za pomocą wielomianu stopnia II. Uwzględniając kolejne wyrazy rozwinięcia funkcji a w szereg Taylora, można definiować modele dwumianowe wyższych rzędów; wszystkie one mają ogólną postać (1), liniową względem parametrów, różniąc się wymiarem k wektorów x_t i β. Ponieważ k w modelu rzędu s jest wielomianem stopnia s względem liczby m oryginalnych zmiennych w_{tj}, więc modele rzędu III i wyższych posiadają zbyt wiele swobodnych parametrów, by stosować je w praktyce. Specyfikacje rzędu II są kompromisem między jakością aproksymacji funkcji a i efektów krańcowych oraz oszczędnością parametryzacji. Już w modelu rzędu II problemem może być przybliżona współliniowość składowych wektora x_t, powodująca słabą identyfikowalność (niską precyzję szacunku) parametrów β_t i problemy numeryczne przy ich estymacji.

Powyższe rozumowanie zakłada konkretną postać dystrybuanty F (jak w modelu probitowym czy logitowym). Jednak w następnej części pracy proponujemy stosowanie dwuparametrowej rodziny dystrybuant skośnych rozkładów Studenta. Znacznie zwiększa to giętkość modelu za cenę estymacji tylko dwóch nowych parametrów; ma wpływ na postać efektów krańcowych, ale nie ich ilorazów, zatem nie zastąpi rozszerzenia modeli rzędu I do modeli rzędu II. Oba proponowane uogólnienia wydają się więc komplementarne.

Dotychczasowe rozważania zakładały, że zmienne egzogeniczne w_{tj} mogą przyjmować dowolne wartości rzeczywiste. Uzasadnienie modelu II rzędu (poprzez odwołanie się do lokalnej aproksymacji nieznanej funkcji a) nawiązuje wtedy do koncepcji giętkich form funkcyjnych Diewerta (1971), znanej z empirycznej mikroekonomii; zob. też Wróbel-Rotter (2001). Często buduje się jednak modele wyboru z dyskretnymi zmiennymi objaśniającymi. Również wówczas jest sens rozważać uogólnienie specyfikacji podstawowej (I rzędu) do modelu II rzędu, choć nie można odwołać się do wyżej podanej argumentacji. Wielkości odpowiadające różniczkowym efektom krańcowym – szacowane dla dyskretnych zmiennych objaśniających – tracą interpretację. Dla zero-jedynkowej zmiennej w_{tj} interpretowalnym odpowiednikiem efektu krańcowego jest

$$\eta_{tj} = \Pr(y_t = 1 \mid w_{t,j} = 1) - \Pr(y_t = 1 \mid w_{t,j} = 0).$$

3. BAYESOWSKI MODEL DWUMIANOWY Z DYSTRYBUANTĄ SKOŚNEGO ROZKŁADU STUDENTA

Przyjmijmy, że składnik ε_t we wzorze (2) ma skośny rozkład t Studenta o zerowej modalnej, jednostkowej precyzji, ν stopniach swobody ($\nu>0$) i parametrze asymetrii $\gamma > 0$; funkcja gęstości tego rozkładu ma postać:

$$p(\varepsilon_t|\theta) = f_{skS}(\varepsilon_t|\nu,\gamma) = \frac{2}{\gamma+\gamma^{-1}}\left\{f_\nu(\gamma\varepsilon_t)\cdot I_{(-\infty,0)}(\varepsilon_t) + f_\nu(\varepsilon_t\gamma^{-1})\cdot I_{[0,+\infty)}(\varepsilon_t)\right\}, \tag{3}$$

gdzie $\theta = (\beta'\ \nu\ \gamma)'$, zaś $f_\nu(\cdot)$ jest funkcją gęstości symetrycznego rozkładu t Studenta o modalnej 0, precyzji 1 i ν stopniach swobody; zob. Fernández i Steel (1998).

Ze specyfikacji (2) wynika, że prawdopodobieństwo zaobserwowania $y_t = 1$ wynosi

$$\Pr(y_t = 1|\theta) = \Pr(z_t \geq 0|\theta) = \Pr(\varepsilon_t \geq -x_t\beta|\theta) = 1 - \Pr(\varepsilon_t < -x_t\beta|\theta) = 1 - F_{skS}(-x_t\beta|\nu,\gamma), \tag{4}$$

gdzie dystrybuanta skośnego rozkładu *t* Studenta o modalnej 0, precyzji 1, ν stopniach swobody i parametrze asymetrii γ (obliczona w punkcie *a*) wyraża się formułą

$$F_{skS}\left(a|\nu,\gamma\right)=\frac{2}{\gamma+\gamma^{-1}}\left[\gamma^{-1}F_{\nu}\left(a\gamma\right)I_{(-\infty,0)}\left(a\right)+\left(\frac{\gamma^{-1}-\gamma}{2}+\gamma F_{\nu}\left(a\gamma^{-1}\right)\right)I_{[0,+\infty)}\left(a\right)\right], \qquad (5)$$

przy czym $F_{\nu}(\cdot)$ jest dystrybuantą symetrycznego rozkładu *t* Studenta o modalnej 0, precyzji 1 i ν stopniach swobody. Łatwo sprawdzić, że funkcja we wzorze (3) jest pochodną funkcji (5).

Stopień asymetrii rozkładu zmiennej ε_t określony jest przez iloraz prawdopodobieństw na prawo i na lewo od modalnej, równy kwadratowi parametru γ (i niezależny od ν):

$$\frac{\Pr\left(\varepsilon_t\geq 0\mid\nu,\gamma\right)}{\Pr\left(\varepsilon_t<0\mid\nu,\gamma\right)}=\gamma^2 . \qquad (6)$$

Innymi słowy, γ parametryzuje wartość dystrybuanty w zerze: $F_{skS}(0|\nu,\gamma)=1/(\gamma^2+1)$. Jeżeli parametr asymetrii γ równy jest jedności, to rozkład jest symetryczny i $F_{skS}(0|\nu, 1)=1/2$.

Wzór (4) określa rozkład pojedynczej obserwacji (przy ustalonych parametrach) jako rozkład dwupunktowy o funkcji prawdopodobieństwa:

$$p\left(y_t|\theta\right)=F_{skS}\left(-x_t\beta|\nu,\gamma\right)I_{\{0\}}\left(y_t\right)+\left[1-F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right]I_{\{1\}}\left(y_t\right).$$

W przypadku *T* niezależnych obserwacji ich łączne prawdopodobieństwo można zapisać jako:

$$p\left(y|\theta\right)=p\left(y_1,\ldots,y_T|\theta\right)=\prod_{t=1}^{T}p\left(y_t|\theta\right)=\left[\prod_{t:y_t=0}F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right]\cdot\left[\prod_{t:y_t=1}\left(1-F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right)\right].$$

Przy ustalonych obserwacjach, powyższa formuła określa funkcję wiarygodności dla modelu dychotomicznego rozważanego w tej pracy. Funkcja ta, traktowana jako funkcja argumentu ν (przy pozostałych ustalonych) bardzo szybko zmierza do dodatniej stałej równej wartości wiarygodności przy (skośnym) rozkładzie normalnym ($\nu=+\infty$). Ta stałość wiarygodności dla dużych ν może być poważną przeszkodą w klasycznej estymacji parametrów modelu. Oczywiście, tę samą własność ma już funkcja wiarygodności w modelach z symetrycznym rozkładem Studenta. Autorzy nie znają żadnej pracy określającej własności estymatora MNW w takich przypadkach. Własne badania symulacyjne ukazują jego systematyczne obciążenie.

Podstawowym elementem analizy bayesowskiej jest statystyczny model bayesowski, czyli łączny rozkład obserwacji i parametrów, określony przez dyskretny warunkowy rozkład wektora obserwacji *y*, $p\left(y|\theta\right)$, i ciągły brzegowy rozkład wektora parametrów (tzw. rozkład a priori), $p(\theta)$. W tej pracy zakładamy stochastyczną niezależność wektora β oraz parametrów ν i γ, przyjmując dla β *k*–wymiarowy normalny rozkład a priori o wartościach oczekiwanych 0 i diagonalnej macierzy kowariancji. Dla ν przyjmujemy wykładniczy rozkład a priori o wartości oczekiwanej *r* (*r* = 10), zaś dla γ standardowy rozkład logarytmiczno-normalny.

Z uwagi na to, że zbiory dopuszczalnych wartości parametrów ν i γ są równe R_+, warto dokonać reparametryzacji $\theta_{k+1} = \ln(\nu/r)$, $\theta_{k+2} = \ln(\gamma)$ i redefiniować wektor wszystkich parametrów jako $\theta = [\beta' \ \theta_{k+1} \ \theta_{k+2}]'$. Wówczas przestrzeń parametrów jest całym zbiorem R^{k+2}, co bardzo upraszcza stronę numeryczną analizy bayesowskiej. Dla tak określonego θ mamy następującą strukturę a priori:

$$p(\theta) = p(\beta) \cdot p(\theta_{k+1}) \cdot p(\theta_{k+2}),$$

$$p(\beta) = f_N^k\left(\beta \mid 0, s_0^2 I\right), \quad p(\theta_{k+1}) = \exp(\theta_{k+1})\exp(-\exp(\theta_{k+1})), \quad p(\theta_{k+2}) = f_N\left(\theta_{k+2} \mid 0, 1\right), \qquad (7)$$

gdzie s_0^2 jest wariancją a priori dla składowych wektora β (u nas: s_0^2=100). Warto zauważyć, że wykładniczy rozkład a priori dla ν (o wartości oczekiwanej r) prowadzi do rozkładu wartości ekstremalnych (Gumbela) dla θ_{k+1}=ln(ν/r), zaś informacja o parametrze skośności jest reprezentowana przez standaryzowany rozkład normalny dla ln(γ). Określona powyżej struktura a priori reprezentuje nikłą wstępną wiedzę obserwatora o parametrach. Prowadzi ona do pełnego, dyskretno-ciągłego modelu bayesowskiego określonego przez uogólnioną gęstość postaci

$$p(y,\theta) = p(y|\theta)p(\theta) = p(\theta)\left[\prod_{t:y_t=0} F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right] \cdot \left[\prod_{t:y_t=1}\left(1 - F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right)\right], \qquad (8)$$

gdzie $\nu = r\exp(\theta_{k+1})$, $\gamma = \exp(\theta_{k+2})$. Wnioskowanie bayesowskie wykorzystuje dekompozycję tego rozkładu łącznego na rozkład a posteriori, tj. warunkowy rozkład ciągły o gęstości:

$$p(\theta|y) = \frac{p(y|\theta)\,p(\theta)}{p(y)} \propto p(\theta)\left[\prod_{t:y_t=0} F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right] \cdot \left[\prod_{t:y_t=1}\left(1 - F_{skS}\left(-x_t\beta|\nu,\gamma\right)\right)\right],$$

oraz brzegowy rozkład obserwacji (dyskretny) o funkcji prawdopodobieństwa postaci:

$$p(y) = \int_{\Theta} p(y|\theta)\,p(\theta)\,d\theta .$$

Aby dekompozycja ta była możliwa (tj. by istniał rozkład a posteriori), powyższa całka musi być skończona. Z tego powodu przyjęto właściwy rozkład a priori (miarę probabilistyczną na przestrzeni parametrów). Jest to oczywiste w przypadku stopni swobody. Zbieżność funkcji wiarygodności przy $\nu \to +\infty$ do dodatniej stałej oznacza bowiem, że całka tej funkcji (po całej przestrzeni parametrów) jest nieskończona, więc niewłaściwy jednostajny rozkład a priori dla ν (tj. miara σ–skończona, ale nie probabilistyczna) prowadziłby do braku rozkładu a posteriori.

Nie przedstawiono jednak nigdzie dowodu, że właściwy rozkład a priori parametru ν jest wystarczający dla istnienia rozkładu a posteriori przy niewłaściwym rozkładzie a priori wektora β (jednostajnym na całej przestrzeni R^k), stosowanym we wszystkich pracach ze swobodnym parametrem ν; zob. Albert i Chib (1993), Marzec (2003b,c), Osiewalski i Marzec (2004). Wstępne symulacje sugerują możliwość braku rozkładu a posteriori, gdy zakłada się niewłaściwy rozkład a priori dla β i właściwy rozkład a priori dla ν. Ponadto, w przypadku modeli II rzędu, duża liczba zmiennych objaśniających – silnie powiązanych między sobą – prowadzi do współliniowości, powodującej istotne problemy numeryczne. Właściwy, normalny rozkład a priori dla β przyczynił się do uzyskania w tej pracy wyników stabilnych numerycznie i poprawnych z punktu widzenia teorii wnioskowania bayesowskiego (niezależnie od przyszłych rozstrzygnięć w zakresie istnienia rozkładu a posteriori przy niewłaściwym rozkładzie a priori dla β).

Rozkład a posteriori parametrów modelu dwumianowego – o gęstości proporcjonalnej do (8) – jest skomplikowanym, niestandardowym rozkładem wielowymiarowym. Ponadto przedmiotem wnioskowania są nie tylko oryginalne parametry, ale przede wszystkim ich skomplikowane funkcje nieliniowe takie jak:

- prawdopodobieństwo $p_* \equiv \Pr(y_* = 1 \mid \theta) = 1 - F_{skS}(-x_* \cdot \beta)$ podjęcia określonej decyzji przez nie obserwowaną jednostkę o ustalonych charakterystykach,
- efekty krańcowe, omówione w poprzedniej części pracy.

Uzyskanie brzegowej funkcji gęstości a posteriori dla wielkości będącej przedmiotem analizy jest złożonym problemem całkowania w przestrzeni (k+2)–wymiarowej. Proponujemy w tym celu zastosowanie metod Monte Carlo typu łańcuchów Markowa (MCMC), a w szczególności losowania Metropolisa i Hastingsa.

Metody MCMC polegają na tym, że ciąg kolejnych losowań w przestrzeni parametrów $(\theta^{(1)},...,\theta^{(n)},...)$ tworzy łańcuch Markowa (o nieprzeliczalnej liczbie stanów) z rozkładem stacjonarnym równym rozkładowi a posteriori $p(\theta \mid y)$. W efekcie, po osiągnięciu zbieżności łańcucha do rozkładu stacjonarnego, generujemy realizacje (otrzymujemy próbę) z rozkładu a posteriori; zob. np. O'Hagan (1994), Gamerman (1997). Algorytm Metropolisa i Hastingsa buduje łańcuch Markowa poprzez zadanie $\theta^{(0)}$ – arbitralnego punktu startowego oraz gęstości $q(\theta^*;\theta^{(m-1)})$ rozkładu losowań kandydackich θ^* (m=1,2,...); dla danego $\theta^{(m-1)}$ przyjmujemy

$\theta^{(m)} = \theta^*$ z prawdopodobieństwem $P(\theta^*,\theta^{(m-1)})$,

$\theta^{(m)} = \theta^{(m-1)}$ z prawdopodobieństwem $1 - P(\theta^*,\theta^{(m-1)})$,

przy czym prawdopodobieństwo akceptacji wylosowanego wstępnie θ* dane jest wzorem

$$P(\theta^*,\theta^{(m-1)}) = \min\left\{\frac{h_y(\theta^*)q(\theta^{(m-1)};\theta^*)}{h_y(\theta^{(m-1)})q(\theta^*;\theta^{(m-1)})}, 1\right\},$$

gdzie $h_y(\theta)$ to jądro gęstości a posteriori $p(\theta \mid y)$, u nas – prawa strona wzoru (8). Dogodny mechanizm losowań wstępnych to $q(\theta^*;\theta^{(m-1)}) = f_S(\theta^* \mid 3, \theta^{(m-1)}, 3C^{-1})$, wielowymiarowy rozkład t Studenta o 3 stopniach swobody, modalnej równej poprzedniemu stanowi łańcucha oraz macierzy precyzji takiej, że C jest macierzą kowariancji (równą wstępnej ocenie macierzy kowariancji rozkładu a posteriori). W tym przypadku gęstość rozkładu losowań wstępnych $q(\theta^*;\theta^{(m-1)})$ jest symetryczna względem argumentów, więc prawdopodobieństwo akceptacji zależy tylko od ilorazu gęstości a posteriori:

$$P(\theta^*,\theta^{(m-1)}) = \min\left\{\frac{h_y(\theta^*)}{h_y(\theta^{(m-1)})}, 1\right\}.$$

W praktyce początkowe S stanów łańcucha Markowa służy uzyskaniu zbieżności (cykle spalone), a następne M stanów – generowaniu próby z rozkładu stacjonarnego i aproksymacji jego charakterystyk zgodnie z ogólnym wzorem:

$$E[g(\theta)|y] \approx \frac{1}{M}\sum_{q=S+1}^{S+M} g\left(\theta^{(q)}\right).$$

Wyniki badań empirycznych prezentowanych w następnej części uzyskano na podstawie S=10^5 cykli spalonych i M=10^6 realizacji tworzących próbę z rozkładu a posteriori. Kilka krótszych przebiegów wstępnych pozwoliło wcześniej wykalibrować macierz C oraz ocenić zbieżność algorytmu Metropolisa i Hastingsa w przypadku naszych modeli i danych.

4. MODELOWANIE PRAWDOPODOBIEŃSTWA NIE SPŁACANIA KREDYTU

Zaprezentowany w poprzednich częściach pracy bayesowski model dwumianowy rzędu II ze skośnym rozkładem *t* Studenta stosujemy do badania spłacalności kredytów detalicznych w oparciu o dane, które wykorzystali wcześniej Marzec (2003a,b,c) oraz Osiewalski i Marzec (2004). Zmienna objaśniana y_t przyjmuje dwie wartości, tzn. y_t=1, gdy kredytobiorca na dzień 30.09.2001 miał zaległości w spłacie rat kapitałowo-odsetkowych (opóźnienie w spłacie ostatniej raty wynosiło więcej niż miesiąc), natomiast y_t=0 w przeciwnym przypadku. Badana liczba (T) indywidualnych rachunków kredytowych wynosi 39040. Jako potencjalne zmienne wyjaśniające ryzyko pojedynczej umowy kredytowej wprowadziliśmy:

- płeć (zmienna jest równa 1, jeżeli klientem jest mężczyzna, 0 w przypadku kobiety),
- wiek kredytobiorcy (w setkach lat),
- wpływy, tzn. wielkość kwartalnych wpływów w latach 2000-2001 (w setkach tys. zł) na rachunki typu ROR kredytobiorcy w badanym banku,
- posiadanie ROR w analizowanym banku (1 – posiada, 0 – nie posiada),
- informację o tym, czy kredytobiorca posiada karty płatnicze lub kredytowe wydane przez ten bank (1 – posiada choć jedną kartę płatniczą, 0 – nie posiada),
- sposób udzielenia kredytu (1 – przez pośrednika, 0 – bezpośrednio przez bank),

- typ kredytu (1 – kredyt konsumpcyjny, 0 – kredyt hipoteczny),
- okres na jaki został udzielony kredyt (w dziesiątkach lat),
- kwota przyznanego kredytu (w setkach tysięcy złotych),
- waluta kredytu (1 – DEM, EUR lub USD; 0 – PLN),
- podstawowe źródło dochodu uzyskiwanego przez kredytobiorcę (zmienne *zrdoch*), tj. umowa o pracę, albo renta lub emerytura, albo własna działalność, umowa o dzieło lub umowa zlecenie, albo inne źródło (np. stypendium).

Ostatnia zmienna może przyjmować cztery różne wartości. Aby uwzględnić ją obok wyrazu wolnego, wprowadziliśmy trzy zmienne zero-jedynkowe, przyjmując za punkt odniesienia umowę o pracę ($zrdoch1 = zrdoch2 = zrdoch3 = 0$); w pozostałych przypadkach:

- $zrdoch1 = 1$, gdy źródłem dochodu kredytobiorcy jest renta lub emerytura,
- $zrdoch2 = 1$, gdy źródłem dochodu kredytobiorcy jest własna działalność, umowa o dzieło lub umowa zlecenie,
- $zrdoch3 = 1$ w przypadku innego źródła dochodu, np. stypendium.

Ostatecznie mamy m=13 zmiennych objaśniających, w tym 9 zero-jedynkowych.

Marzec (2003c) badał uwarunkowania nie spłacania kredytu wykorzystując model II rzędu, ale rozważał tylko specyfikację ze zwykłym (symetrycznym) rozkładem t. Osiewalski i Marzec (2004) zaproponowali zastosowanie dystrybuanty skośnego rozkładu t i wykazali jej empiryczną przydatność, ale tylko w modelu I rzędu. W obu tych pracach nie uwzględniano kwoty i waluty przyznanego kredytu, przyjmując $m = 11$ zmiennych. Model stosowany w tej pracy jest ogólniejszy zarówno od strony teoretycznej, jak empirycznej (poprzez włączenie dwóch ważnych charakterystyk kredytu).

Zauważmy, że przy $m = 13$ zmiennych w_{tj} wymiar k wektorów x_t i β wynosi k=1+m=14 dla modelu I rzędu, a maksymalny wymiar w przypadku modelu II rzędu nie przekracza $1+m+m(m+1)/2=105$. Ponieważ dla zmiennych zero-jedynkowych zachodzi $(w_{tj})^2 = w_{tj}$, ten wymiar wynosi w naszych badaniach co najwyżej 96. Eliminując iloczyny $w_{tj}w_{th}$ równe 0 (dla wszystkich t) uzyskano $k = 89$. Jednak model o 89 parametrach charakteryzowała przybliżona współliniowość składowych wektora x_t, zwłaszcza niektórych iloczynów zmiennych zero-jedynkowych. Indeks uwarunkowania macierzy X wartości zmiennych x_{ti} (t=1,...,T; i=1,...,k) wyniósł 663 dla modelu II rzędu z $k = 89$, ale tylko 25 dla modelu I rzędu (k=14).[1] Eliminacja dziesięciu silnie powiązanych iloczynów $w_{tj}w_{th}$ obniżyła wartość indeksu z 663 dla $k = 89$ do 154 dla $k = 79$, co wystarczyło by w

---

[1] Indeks uwarunkowania macierzy X podają i wykorzystują do pomiaru współliniowości Belsley, Kuh i Welsh (1980). Indeks ten można policzyć jako pierwiastek kwadratowy ilorazu największej i najmniejszej wartości własnej macierzy niescentrowanych współczynników korelacji $R_N = W^{-1} X'X\ W^{-1}$, gdzie W jest macierzą diagonalną zawierającą na przekątnej długości kolumn macierzy X.

sposób stabilny numerycznie przeprowadzić estymację bayesowską z wykorzystaniem algorytmu Metropolisa i Hastingsa. Przyjętą listę zmiennych modelu II rzędu podano w Tabeli 1, zawierającej charakterystyki a posteriori parametrów.

Tabela 1

Wartości oczekiwane i odchylenia standardowe a posteriori wielu parametrów stojących przy kwadratach i iloczynach zmiennych w_{tj} wyraźnie świadczą o bezzasadności redukcji modelu II rzędu do modelu I rzędu. Zerowe wartości tych parametrów znajdują się o kilka odchyleń standardowych od wartości oczekiwanych a posteriori, więc są nieprawdopodobne a posteriori. Daleko od zera skupiają się też brzegowe rozkłady a posteriori parametrów przy zmiennych: w_{t9} (kwota kredytu), $(w_{t9})^2$ i kilku iloczynach zawierających tę zmienną lub w_{t10} (waluta kredytu), np. $w_{t1}w_{t10}$, $w_{t4}w_{t9}$, $w_{t7}w_{t9}$, $w_{t8}w_{t9}$, i $w_{t9}w_{t10}$. Włączenie zmiennych w_{t9} i w_{t10} jest więc ważne w modelowaniu prawdopodobieństwa nieterminowego spłacania kredytu.

Brzegowy rozkład a posteriori parametru stopni swobody jest skupiony w przedziale [0,40; 0,80], przy czym $E(\nu|y) = 0{,}582$ i $D(\nu|y) = 0{,}062$, więc zwykle przyjmowana specyfikacja probitowa powinna być odrzucona przy modelowaniu rozważanego zbioru obserwacji. Jeśli zaś chodzi o zastosowanie dystrybuanty z klasy skośnych rozkładów *t* (zamiast podklasy symetrycznych rozkładów Studenta), to i takie uogólnienie wydaje się empirycznie zasadne, gdyż wartość oczekiwana a posteriori $E(\gamma | y) = 0{,}31$ jest o około 30 odchyleń standardowych $D(\gamma | y)$ mniejsza od 1. Warto zaznaczyć, że przy $k = 14$ uzyskujemy $E(\gamma| y) = 0{,}344$ i $D(\gamma| y) = 0{,}018$ oraz $E(\nu| y) = 0{,}593$ i $D(\nu| y) = 0{,}049$, czyli wyniki podobne jak w modelu rzędu II. Przy ograniczeniu się do modelu I rzędu dane wymagają dystrybuanty $F(.)$ o tych samych własnościach, co w modelu I rzędu, czyli bardzo różnych od własności dystrybuanty rozkładu $N(0; 1)$.

Wartości parametrów β_i nie są bezpośrednio interpretowalne, zaś informacje o sile i kierunku wpływu zmiennych egzogenicznych w_{tj} na p_t (prawdopodobieństwo niespłacania) uzyskujemy na podstawie efektów krańcowych, których charakterystyki a posteriori podano w Tabeli 2. Przedstawiono w niej nie tylko wyniki uzyskane w modelu II rzędu, ale także – w celu porównania – wyniki dla modeli I rzędu, wykorzystujących skośny rozkład Studenta i rozkład normalny (tradycyjny model probitowy); uzyskano je stosując algorytm Metropolisa i Hastingsa. Rezultaty bayesowskie dla modelu probitowego I rzędu są prawie identyczne z wynikami uzyskanymi za pomocą MNW, ponieważ (zgodnie z teorią) w modelu probitowym oceny MNW w przypadku dużej liczby obserwacji można interpretować jako wartości oczekiwane rozkładu a posteriori (przy dość dowolnym ciągłym rozkładzie a priori). Efekty krańcowe różnią się między modelami, nie zawsze zachowując ten sam rząd wielkości czy nawet znak. Interpretujemy tylko wyniki uzyskane dla modelu najogólniejszego.

Tabela 2

Spośród rozważanych jakościowych zmiennych egzogenicznych największy wpływ na prawdopodobieństwo nieterminowego spłacania kredytu ma sposób jego udzielenia; jeśli udzielono go przez pośrednika, zamiast bezpośrednio przez bank, to wzrost p_t jest najbardziej znaczący, średnio o 0,267 (±0,014). Różnica między prawdopodobieństwem nieterminowego spłacania kredytu konsumpcyjnego i hipotecznego wynosi średnio aż 0,189 (±0,003). Posiadanie ROR też zwiększa p_t (co jest wynikiem zaskakującym), ale znacznie słabiej bo o 0,035 (±0,013). Posiadanie kart płatniczych lub kredytowych, jako przejaw aktywnego korzystanie z usług badanego banku, zmniejsza ryzyko złego kredytu średnio o 0,103 (±0,073), więc szacunek nie jest precyzyjny. Wraz ze wzrostem o 1 rok okresu na jaki został udzielony kredyt, p_t maleje aż o 0,225 (±0,037). Obcej walucie kredytu odpowiada mniejsza wartość p_t – o 0,115 (±0,072), zatem różnica ta nie wydaje się istotna. Wzrost wpływów kwartalnych kredytobiorcy o 1 tys. zł zmniejsza wartość p_t średnio o 0,0015 (±0,0002). Rola zmiennej *wiek klienta* jest nieistotna, wpływ kwoty kredytu na p_t jest istotny, lecz mały, zaś płeć klienta jest zupełnie bez znaczenia.

Spośród źródeł dochodu największe ryzyko kredytowe związane jest z kredytobiorcą prowadzącym własną działalność gospodarczą. Kredytobiorcami o niższym ryzyku są osoby zatrudnione na podstawie umowy o pracę, osoby pobierające emeryturę lub rentę, a także studenci spłacający kredyty studenckie, przy czym udział ilościowy i wartościowy tej grupy kredytów jest znikomy.

Warto jednak zaznaczyć, że rozważamy efekty krańcowe uśrednione po obserwacjach, a zatem powyższe interpretacje mają charakter poglądowy. Bardziej użyteczne mogą być efekty krańcowe dla konkretnego kredytobiorcy (tj. dla ustalonego wektora x_t). Ich przykłady podajemy w Tabeli 4 na końcu tej części pracy.

Głównym sposobem wykorzystania modeli jest prognozowanie prawdopodobieństwa nieterminowej spłaty rat kapitałowo-odsetkowych bądź całkowitego zaniechania ich spłat. W tym celu rozważamy cztery hipotetyczne sylwetki kredytobiorców, których charakterystykę zawiera Tabela 3. Dwaj typowi kredytobiorcy określeni są poprzez najczęstsze w zbiorze danych wartości zero-jedynkowych zmiennych objaśniających (z wyjątkiem w_{t6}) i średnie wartości zmiennych ciągłych. Różnią się oni sposobem uzyskania kredytu: przez pośrednika (w_{t6}=1) lub bezpośrednio w banku (w_{t6}=0). Pozostałe dwie sylwetki prezentują klientów o średniej kwocie i średnim okresie kredytu, ale innych cechach bardzo skontrastowanych – tak, by mogły opisywać młodego mężczyznę prowadzącego własną działalność gospodarczą (nie związanego z bankiem i uzyskującego kredyt przez pośrednika) oraz emerytkę korzystająca także z innych usług banku. Zauważmy, że rozkłady a posteriori uzyskane dla p_t w trzech modelach znacznie różnią się od siebie. Zachowany zostaje jedynie ranking kredytobiorców: wysokie zagrożenie spłacania kredytu pojawia się w przypadku młodego przedsiębiorcy, niższe jest dla typowego kredytobiorcy pozyskanego przez pośrednika, jeszcze niższe dla typowego kredytobiorcy pozyskanego wprost przez bank, a najniższe dla emerytki. Jednak, przy ustalonym profilu klienta, wartości oczekiwane i odchylenia standardowe a posteriori dla p_t otrzymane w trzech modelach są tak różne, że mogą prowadzić do odmiennych wniosków praktycznych. Dotyczy to zwłaszcza typowego kredytobiorcy pozyskanego przez pośrednika; w modelu probitowym I rzędu odpowiada mu p_t na poziomie aż 0,304 (±0,014), w modelach ze skośnym rozkładem *t* jest to prawdopodobieństwo znacznie niższe i szacowane z inną precyzją w każdym modelu. Jeśli redukcje modelu II rzędu do modeli I rzędu (zwłaszcza probitowego) nie są zasadne (jak w przypadku naszych danych), to często stosowana ocena ryzyka kredytowego na podstawie modelu probitowego I rzędu może prowadzić do błędnych wniosków.

Tabela 3

Na zakończenie tej ilustracji empirycznej przedstawiamy w Tabeli 4 oceny punktowe efektów krańcowych dla hipotetycznych kredytobiorców, uzyskane w modelu II rzędu przez wstawienie w miejsce nieznanych parametrów ich wartości oczekiwanych a posteriori (z Tabeli 1). Oceny te silnie zależą od profilu kredytobiorcy; warto je porównać z wartościami oczekiwanymi uśrednionych efektów krańcowych, podanymi w Tabeli 2. Zgodnie z intuicją, efekty krańcowe są bliższe 0 w przypadku klientów, dla których p_t jest mniejsze. Na zmiany zmiennych objaśniających szczególnie silnie reaguje prawdopodobieństwo p_t nieterminowego spłacania kredytu zaciągniętego przez młodego przedsiębiorcę, dla którego to p_t jest bardzo wysokie, równe 0,507 (±0,026) – zob. Tabela 3. Zauważmy, że w modelu II rzędu efekty krańcowe względem ustalonej zmiennej w przypadku dwóch profili kredytobiorców mogą mieć przeciwne znaki, np. dla efektu krańcowego względem waluty kredytu. Ponadto indywidualne efekty krańcowe mogą być różnego znaku od uśrednionych efektów prezentowanych w Tabeli 2. Ma to miejsce dla efektów względem wieku, waluty kredytu oraz *Zrdoch1*.

Tabela 4

5. BAYESOWSKIE PORÓWNANIE KONKURENCYJNYCH SPECYFIKACJI

W poprzedniej części pracy przedstawiliśmy wyniki a posteriori nie tylko dla modelu II rzędu z dystrybuantą skośnego rozkładu Studenta (model 1, M_1), ale również dla modelu I rzędu opartego na tej samej klasie dystrybuant (M_2) i dla modelu probitowego I rzędu (M_3). Brzegowe rozkłady a posteriori parametrów rozszerzających M_3 do M_2 i M_2 do M_1 wskazały, że redukcje modelu najbardziej złożonego (M_1) do prostszych (M_2, M_3) nie są uzasadnione. W tej części przedstawimy podstawy teoretyczne i wyniki formalnego porównania tych trzech modeli, wykorzystującego ich prawdopodobieństw a posteriori.

Rozważmy n modeli próbkowych zdefiniowanych na tej samej przestrzeni Y:

$$M_i : p_i(y|\theta_{(i)}), \quad i = 1,...,n,$$

gdzie $y \in Y$ jest wektorem obserwacji (u nas T-elementowym ciągiem zer i jedynek), a $\theta_{(i)} \in \Theta_i$ jest wektorem parametrów modelu M_i. Definiując n rozkładów a priori o gęstościach $p_i(\theta_{(i)})$ mamy n modeli bayesowskich:

$$p_i(y, \theta_{(i)}) = p_i(y|\theta_{(i)})\, p_i(\theta_{(i)}), \quad i = 1,...,n.$$

Formalny wybór najlepszego modelu próbkowego opiera się na prawdopodobieństwach a posteriori poszczególnych modeli; koncepcję tę referują m.in. Zellner (1971) i Osiewalski (2001). Załóżmy, że M_1, ..., M_n są wzajemnie wykluczającymi (nie zagnieżdżonymi) i łącznie dopełniającymi się modelami o prawdopodobieństwach a priori $p(M_1)$, ..., $p(M_n)$. Przy tych założeniach uzyskujemy ze wzoru Bayesa prawdopodobieństwa a posteriori modeli równe:

$$p(M_i|y) = \frac{p(M_i)\, p(y|M_i)}{\sum_{h=1}^{n} p(M_h)\, p(y|M_h)}.$$

W powyższym wzorze kluczową rolę pełni wielkość

$$p(y|M_i) = \int_{\Theta_i} p_i(y|\theta_{(i)})\, p_i(\theta_{(i)})\, d\theta_{(i)}, \qquad i = 1,...,n,$$

czyli brzegowe prawdopodobieństwo wektora obserwacji w modelu M_i, naturalny bayesowski miernik dopasowania modelu do danych empirycznych (wektora y). Prawdopodobieństwa a posteriori modeli, $p(M_i \mid y)$, łączą informację o dopasowaniu i przekonania a priori o stopniu adekwatności modeli, wyrażane przez $p(M_i)$.

Co do prawdopodobieństw a priori $p(M_i)$, to często przyjmuje się, iż są one równe. Można jednak argumentować, że przy równym dopasowaniu modeli, mierzonym wartością $p(y|M_i)$, ogólne względy metodologiczne przemawiają za preferowaniem – przez odpowiedni dobór prawdopodobieństw a priori – modeli prostszych (o mniejszej liczbie parametrów). Jako $p(M_i)$ można przyjąć malejącą funkcję liczby l_i parametrów modelu, np.: $p(M_i) \propto 2^{-l_i}$.

Powyższa teoria wymaga, by rozważane modele nie były zagnieżdżone. Zakładamy więc, że nie wszystkie parametry rozszerzające model I rzędu do specyfikacji II rzędu są równe zero. Nie jest to sprzeczne z przyjętym normalnym rozkładem a priori dla wektora β, a wyklucza zagnieżdżenie M_2 w M_1. Formalnie rzecz ujmując, model probitowy I rzędu (M_3), odpowiadający $\nu = +\infty$, nie jest zagnieżdżony w modelu M_2 ani M_1, które zakładają, iż ν jest skończone.

Bayesowskie porównywanie modeli wymaga obliczenia – dla każdego z nich – całki definiującej brzegowe prawdopodobieństwo obserwacji $p(y|M_i)$, co w praktyce okazuje się zwykle poważnym problemem numerycznym. W ramach metod MCMC, stosowanych do symulacji rozkładu a posteriori w każdym z modeli, najprostsza (i dość stabilna numerycznie) jest aproksymacja, którą zaproponowali Newton i Raftery (1994):

$$p(y \mid M_i) = \left\{ \int_{\Theta_i} \left[p(y \mid M_i, \theta_{(i)}) \right]^{-1} dP(\theta_{(i)} \mid M_i, y) \right\}^{-1} \approx \left\{ \frac{1}{M} \sum_{q=S+1}^{S+M} \left[p(y \mid M_i, (\theta_{(i)})^{(q)}) \right]^{-1} \right\}^{-1};$$

polega ona na przybliżeniu $p(y|M_i)$ średnią harmoniczną wartości funkcji wiarygodności obliczonych dla kolejnych stanów łańcucha Markowa (po osiągnięciu zbieżności do rozkładu a posteriori). Zastosowanie tej metody dało następujące szacunki logarytmów naturalnych wartości $p(y|M_i)$: $\ln p(y|M_1) = -13308$; $\ln p(y|M_2) = -13529$; $\ln p(y|M_3) = -13833$. Prowadzą one do tego, że model najbardziej złożony skupia praktycznie całą masę prawdopodobieństwa a posteriori. W przypadku jednakowych prawdopodobieństw a priori uzyskano: $p(M_2|y)=10^{-96}$ i $p(M_3|y)=10^{-228}$; jeśli przyjęto $p(M_i) \propto 2^{-l_i}$ (l_1=81, l_2=16, l_3=14), to $p(M_2|y)=3{,}9*10^{-77}$ i $p(M_3|y)=1{,}6*10^{-208}$. Wstępne konkluzje o bezzasadności redukcji modelu M_1, sformułowane w poprzedniej części, zostały więc z całą mocą potwierdzone poprzez formalne bayesowskie porównywanie trzech modeli dychotomicznych. Model II rzędu oparty na dystrybuancie skośnego rozkładu *t* daje więc nowe możliwości modelowania jakościowej, zero-jedynkowej zmiennej zależnej.

BIBLIOGRAFIA

Albert J., Chib S., 1993, *Bayesian analysis of binary and polychotomous response data*, JASA (Journal of the American Statistical Association) vol. 88, 669-679.

Amemiya T., 1981, *Qualitative response models: A survey*, Journal of Economic Literature vol.19, 1483-1536.

Amemiya T., 1985, *Advanced Econometrics*, Harvard University Press, Cambridge (Massachusetts).

Belsley D.A., Kuh E., Welsh R.E., 1980, *Regression Diagnostics*, Wiley, New York.

Diewert W. E., 1971, *An application of the Shephard duality theorem: a generalized Leontief production function*, Journal of Political Economy vol. 79, 481-507.

Fernández C., Osiewalski J., Steel M., 1995, *Modeling and inference with υ-spherical distributions*, JASA (Journal of the American Statistical Association) vol. 90, 1331-1340.

Fernández C., Steel M., 1998, *On Bayesian modeling of fat tails and skewness*, JASA (Journal of the American Statistical Association) vol. 93, 359-371.

Gamerman D., 1997, *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference*, Chapman and Hall, London.

Greene W.H., 1993, *Econometric Analysis*, Macmillan, New York.

Maddala G.S., 1983, *Limited Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, Cambridge.

Marzec J., 2003a, *Badanie niewypłacalności kredytobiorcy na podstawie modeli logitowych i probitowych*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie nr 628, 103-117.

Marzec J., 2003b, *Badanie niespłacalności kredytów za pomocą bayesowskich modeli dychotomicznych - założenia i wyniki*, Metody ilościowe w naukach ekonomicznych. Trzecie Warsztaty Doktorskie z Zakresu Ekonometrii i Statystyki (red. A. Welfe), Wydawnictwo SGH, Warszawa (73-86).

Marzec J., 2003c, *Bayesowska analiza modeli dyskretnego wyboru (dwumianowych)*, Przegląd Statystyczny t. 50, 129-146.

Newton M. A., Raftery A. E., 1994 *Approximate Bayesian inference by the weighted likelihood bootstrap* (with discussion), Journal of the Royal Statistical Society B vol.56, 3-48.

O’Hagan A., 1994, *Bayesian Inference,* Edward Arnold, London.

Osiewalski J., 2001, *Ekonometria bayesowska w zastosowaniach*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.

Osiewalski J., Marzec J., 2004, *Uogólnienie dychotomicznego modelu probitowego z wykorzystaniem skośnego rozkładu Studenta*, Przegląd Statystyczny t. 51 (w druku).

Osiewalski J., Pipień M., 1999, *Bayesian forecasting of foreign exchange rates using GARCH models with skewed t conditional distributions*, MACROMODELS'98 - Conference Proceedings (red. W. Welfe), vol. 2, Absolwent, Łódź (195-218).

Osiewalski J., Pipień M., 2000, *GARCH-In-Mean through skewed t conditional distributions: Bayesian inference for exchange rates*, MACROMODELS'99 – Conference Proceedings (red. W. Welfe, P. Wdowiński), Absolwent, Łódź (354–369).

Wróbel-Rotter R., 2001, *Giętkie formy funkcyjne w empirycznej analizie kosztu. Podejście Diewerta i wnioskowanie bayesowskie*, Ekonomista nr 5/2001, 687-704.

Zellner A., 1971, *An Introduction to Bayesian Inference in Econometrics*, J.Wiley, New York.

Tabela 1. Wartości oczekiwane i odchylenia standardowe a posteriori parametrów bayesowskiego modelu II rzędu ze skośnym rozkładem *t* Studenta; $\varepsilon_t \sim$ St(0, ν, γ).

Zmienna	Parametr	E(|y)	D(|y)	Zmienna	Parametr	E(|y)	D(|y)
1	β_1	-15,984	3,296	$w_3 \cdot w_{11}$	β_{42}	-5,938	7,814
Płeć (w_1)	β_2	0,374	2,131	$w_3 \cdot w_{12}$	β_{43}	38,580	4,504
Wiek (w_2)	β_3	-2,681	1,772	$w_3 \cdot w_{13}$	β_{44}	-8,203	8,919
Wpływy (w_3)	β_4	-52,817	4,721	$w_4 \cdot w_5$	β_{45}	4,260	4,471
ROR (w_4)	β_5	1,491	0,214	$w_4 \cdot w_6$	β_{46}	-0,058	0,494
Karty (w_5)	β_6	-4,955	4,463	$w_4 \cdot w_8$	β_{47}	-4,444	0,939
Pośrednik (w_6)	β_7	6,451	1,085	$w_4 \cdot w_9$	β_{48}	6,305	1,428
Typ kredytu (w_7)	β_8	16,123	3,287	$w_4 \cdot w_{10}$	β_{49}	5,022	4,789
Okres (w_8)	β_9	-2,163	3,476	$w_4 \cdot w_{11}$	β_{50}	-0,453	0,323
Kwota (w_9)	β_{10}	19,467	3,648	$w_4 \cdot w_{12}$	β_{51}	-0,941	0,383
Waluta (w_{10})	β_{11}	-10,439	5,899	$w_4 \cdot w_{13}$	β_{52}	3,693	2,851
Zrdoch1 (w_{11})	β_{12}	0,600	0,698	$w_5 \cdot w_6$	β_{53}	-0,236	0,565
Zrdoch2 (w_{12})	β_{13}	11,149	2,836	$w_5 \cdot w_8$	β_{54}	0,560	0,982
Zrdoch3 (w_{13})	β_{14}	-2,462	3,025	$w_5 \cdot w_9$	β_{55}	1,011	1,356
$w_1 \cdot w_2$	β_{15}	-0,644	0,544	$w_5 \cdot w_{10}$	β_{56}	-4,421	2,540
$w_1 \cdot w_3$	β_{16}	-9,745	3,748	$w_5 \cdot w_{11}$	β_{57}	0,392	0,389
$w_1 \cdot w_4$	β_{17}	0,533	0,176	$w_5 \cdot w_{12}$	β_{58}	0,424	0,266
$w_1 \cdot w_5$	β_{18}	0,094	0,178	$w_5 \cdot w_{13}$	β_{59}	-0,863	1,319
$w_1 \cdot w_6$	β_{19}	-0,115	0,170	$w_6 \cdot w_8$	β_{60}	-5,194	0,872
$w_1 \cdot w_7$	β_{20}	0,027	2,090	$w_6 \cdot w_9$	β_{61}	3,096	2,614
$w_1 \cdot w_8$	β_{21}	-0,579	0,508	$w_6 \cdot w_{11}$	β_{62}	0,085	0,254
$w_1 \cdot w_9$	β_{22}	-0,236	1,297	$w_6 \cdot w_{12}$	β_{63}	-1,419	0,461
$w_1 \cdot w_{10}$	β_{23}	11,790	3,959	$w_6 \cdot w_{13}$	β_{64}	6,096	3,161
$w_1 \cdot w_{11}$	β_{24}	0,120	0,203	$w_7 \cdot w_8$	β_{65}	1,791	2,885
$w_1 \cdot w_{12}$	β_{25}	-0,220	0,240	$w_7 \cdot w_9$	β_{66}	-13,315	2,909
$w_1 \cdot w_{13}$	β_{26}	-1,731	0,732	$w_7 \cdot w_{12}$	β_{67}	-10,185	2,691
$(w_2)^2$	β_{27}	2,404	2,204	$(w_8)^2$	β_{68}	3,879	1,109
$w_2 \cdot w_5$	β_{28}	0,903	0,697	$w_8 \cdot w_9$	β_{69}	-13,249	2,099
$w_2 \cdot w_6$	β_{29}	-4,925	1,172	$w_8 \cdot w_{10}$	β_{70}	-8,008	3,556
$w_2 \cdot w_8$	β_{30}	4,688	2,447	$w_8 \cdot w_{11}$	β_{71}	-0,472	0,825
$w_2 \cdot w_9$	β_{31}	-2,153	4,388	$w_8 \cdot w_{12}$	β_{72}	2,048	1,101
$w_2 \cdot w_{10}$	β_{32}	-11,305	7,244	$w_8 \cdot w_{13}$	β_{73}	-10,657	3,622
$w_2 \cdot w_{11}$	β_{33}	-2,084	1,178	$(w_9)^2$	β_{74}	-4,221	1,119
$w_2 \cdot w_{12}$	β_{34}	-1,371	1,051	$w_9 \cdot w_{10}$	β_{75}	10,714	3,083
$w_2 \cdot w_{13}$	β_{35}	2,203	2,426	$w_9 \cdot w_{11}$	β_{76}	4,838	2,937
$(w_3)^2$	β_{36}	-3,051	1,400	$w_9 \cdot w_{12}$	β_{77}	-3,488	1,373
$w_3 \cdot w_5$	β_{37}	-1,624	3,027	$w_9 \cdot w_{13}$	β_{78}	-7,591	6,264
$w_3 \cdot w_6$	β_{38}	4,411	6,968	$w_{10} \cdot w_{12}$	β_{79}	-5,153	3,227
$w_3 \cdot w_8$	β_{39}	14,669	3,382	–	ν	0,582	0,062
$w_3 \cdot w_9$	β_{40}	0,591	2,624	–	γ	0,310	0,023
$w_3 \cdot w_{10}$	β_{41}	6,108	5,229				

Źródło: obliczenia własne.

Tabela 2. Wartości oczekiwane i odchylenia standardowe a posteriori uśrednionych efektów krańcowych $\eta_j = \frac{1}{T}\sum_t \partial p_t / \partial w_{tj}$ lub $\eta_j = \frac{1}{T}\sum_t \Pr(y_t = 1 \mid w_{t,j} = 1) - \Pr(y_t = 1 \mid w_{t,j} = 0)$.

	$\varepsilon_t \sim$ St(0, ν, γ) k = 79		$\varepsilon_t \sim$ St(0, ν, γ) k = 14		$\varepsilon_t \sim$St(0, ν=∞, γ=1) k = 14	
Zmienna	E(| y)	D(| y)	E(| y)	D(| y)	E(| y)	D(| y)
Płeć	-7,2E-04	3,4E-03	0,007	0,003	0,008	0,003
Wiek	0,020	0,105	-0,026	0,004	-0,298	0,024
Wpływy	-0,148	0,022	-0,788	0,178	-0,613	0,058
ROR	0,035	0,013	0,048	0,005	-0,059	0,008
Karty	-0,103	0,073	-0,013	0,005	-0,032	0,007
Pośrednik	0,267	0,014	0,247	0,008	0,300	0,008
Typ Kredytu	0,189	0,003	0,173	0,021	0,049	0,020
Okres kredytu	-2,250	0,371	-0,012	0,009	-0,081	0,020
Kwota kredytu	0,074	0,012	0,027	0,009	0,041	0,010
Waluta	-0,115	0,072	0,075	0,067	0,075	0,031
Zrdoch1	-0,006	0,008	-0,034	0,006	-0,019	0,006
Zrdoch2	0,018	0,010	0,035	0,009	0,065	0,009
Zrdoch3	-0,022	0,017	-0,060	0,010	-0,036	0,013

Źródło: obliczenia własne.

Tabela 3. Wartości oczekiwane i odchylenia standardowe a posteriori prawdopodobieństwa nieterminowego płacenia kredytu $p_t = \Pr(y_t=1) = 1 - F(-x_t\beta)$.

	typowy klient		„młody	„starsza
Zmienna	pośrednik=1	pośrednik=0	biznesmen”	pani”
1 (wyraz wolny)	1	1	1	1
Płeć	1	1	1	0
Wiek (w latach)	40,2	40,2	21	60
Wpływy (w tys. zł)	10,2	10,2	0	1
ROR	1	1	0	1
Karty płatnicze	0	0	0	1
Pośrednik	1	0	1	0
Typ kredytu: konsumpcyjny	1	1	1	0
Okres kredytu (w latach)	2,61	2,61	2,61	2,61
Kwota (w tys. Zł)	10,9	10,9	10,9	10,9
Waluta	0	0	0	0
Zrdoch1	0	0	0	1
Zrdoch2	0	0	1	0
Zrdoch3	0	0	0	0
Model probitowy (k=14)				
E(p_t | y)	0,304	0,036	0,668	0,009
D(p_t | y)	0,014	0,002	0,016	0,003
Model skośny t (k=14)				
E(p_t | y)	0,029	0,015	0,555	0,014
D(p_t | y)	0,005	0,001	0,012	0,002
Model skośny t (k=79)				
E(p_t | y)	0,065	0,012	0,507	0,006
D(p_t | y)	0,057	0,001	0,026	0,001

Źródło: obliczenia własne.

Tabela 4. Oceny indywidualnych efektów krańcowych.

	typowy klient		„młody biznesmen”	„starsza pani”
Zmienna	pośrednik=1	pośrednik=0		
Płeć	-0,064	-0,001	-0,014	0,0001
Wiek	-0,108	-0,004	-0,160	-0,001
Wpływy	-2,126	-0,087	-3,143	-0,013
ROR	0,021	0,002	0,030	0,001
Karty	-0,010	-0,0002	-0,467	0,0001
Pośrednik	0,028	0,028	0,274	0,001
Typ Kredytu	0,034	0,007	0,488	0,183
Okres kredytu	-0,087	-0,004	-0,129	-0,001
Kwota kredytu	0,784	0,032	1,158	0,005
Waluta	0,203	0,004	-0,500	-0,002
Zrdoch1	0,062	0,001	-	-
Zrdoch2	0,676	0,627	-	-
Zrdoch3	-0,025	-0,003	-	-