Jerzy Marzec, Katedra Ekonometrii i Badań Operacyjnych, Uniwersytet Ekonomiczny w Krakowie

JERZY MARZEC

BAYESOWSKA ANALIZA MODELI DYSKRETNEGO WYBORU (DWUMIANOWYCH) [1]

1. WSTĘP

W latach siedemdziesiątych ubiegłego stulecia nastąpił gwałtowany rozwój ekonometrycznych modeli dla zmiennych dyskretnych (jakościowych), które w skrócie nazywa się modelami dyskretnego wyboru (ang. *quantal response or discrete choice models*). Podstawową cechą tych modeli jest to, że zmienna endogeniczna przyjmuje tylko skończoną liczbę wartości. W chwili obecnej ekonometryczne modelowanie zmiennych jakościowych w ujęciu klasycznym stanowi standardową treść średnio zaawansowanych anglojęzycznych podręczników ekonometrii, np. Greene [7].

W polskiej literaturze jednym z pierwszych monograficznych ujęć metod ekonometrycznej analizy zjawisk jakościowych była praca Wiśniewskiego [19], natomiast monografia Gruszczyńskiego [8] stanowi prezentację aktualnego dorobku mikroekonometrii stosowanej w finansach i bankowości, interesującą z punktu widzenia badań prowadzonych przez autora niniejszego opracowania. Z drugiej strony, w ciągu ostatnich dwudziestu lat na gruncie ekonometrii bayesowskiej pojawiło się wiele metodologicznych propozycji dotyczących bayesowskiej specyfikacji, estymacji i uogólnień klasycznych modeli dyskretnego wyboru. Jednymi z pierwszych autorów byli Zellner [21] oraz Zellner i Rossi [22].

Głównych celem niniejszego opracowania jest prezentacja nowych (bayesowskich) podejść do modelowania zmiennych jakościowych i ich zastosowanie w analizie rzeczywistych danych finansowych. W szczególności, w niniejszej pracy przedstawimy specyfikację bayesowskich modeli dwumianowych, tzn. dla zmiennych binarnych (dychotomicznych), zaproponowaną przez Alberta i Chiba [1]. Zaprezentujemy przypadek modelu opartego na rozkładzie t-Studenta z nieznaną (podlegającą estymacji) liczbą stopni swobody, który stanowi uogólnienie najczęściej stosowanego w praktyce modelu probitowego. Następnie zastosujemy to podejście do badania ryzyka kredytowego pojedynczej umowy kredytowej dla klientów detalicznych pewnego polskiego banku komercyjnego. W celu uzyskania momentów i gęstości brzegowych rozkładów a posteriori

[1] Autor pragnie wyrazić podziękowania Profesorowi Jackowi Osiewalskiemu za cenne uwagi merytoryczne w trakcie powstawania niniejszego opracowania. Praca wykonana w ramach badań statutowych finansowanych przez Akademię Ekonomiczną w Krakowie w roku 2003 r.

parametrów modeli wykorzystamy losowanie Gibbsa. Wyniki a posteriori pokazują, że zaproponowane uogólnienie (związane z rozkładem t-Studenta) lepiej opisuje dane empiryczne, więc jego wykorzystanie w badaniach jest w pełni uzasadnione.

2. MODEL DYSKRETNEGO WYBORU

Niech $y_1,\ldots,y_T$ będzie ciągiem niezależnych zmiennych losowych o rozkładzie zero-jedynkowym z prawdopodobieństwem sukcesu $\Pr(y_t=1)$ równym p_t. W literaturze zależność między p_t a wektorem zmiennych egzogenicznych x_t (lub znanych ich funkcji, jak w części 5) opisuje się za pomocą modelu dyskretnego wyboru (ang. *binary choice models*). Model ten, nazywany także modelem dwumianowym lub dychotomicznym, ma postać

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

gdzie β to wektor $k\times 1$ nieznanych parametrów, zaś $x_t=(x_{t1}\ldots\ x_{tk})$ oznacza wektor ustalonych wartości k zmiennych egzogenicznych, $F(\cdot)$ jest znaną funkcją wiążącą prawdopodobieństwo p_t z x_t i β, określającą klasę modelu. Jeżeli $F(\cdot)$ jest dystrybuantą zmiennej losowej o standaryzowanym rozkładzie normalnym, to mamy do czynienia z modelem probitowym[2]. Teoretyczne podstawy innych definicji modeli dyskretnego wyboru przedstawia m.in. Amemiya w pracy [2].

Równoważną specyfikację modelu dwumianowego otrzymamy wprowadzając T niezależnych zmiennych ukrytych (nieobserwowalnych) $z_1,\ldots,z_T$. Jeżeli założymy, że obserwujemy $y_t=1$, gdy $z_t\geq 0$ i $y_t=0$ w przeciwnym przypadku, to model dyskretnego wyboru zapiszemy za pomocą dwóch równań

$$\begin{aligned} z_t &= x_t \cdot \beta + \varepsilon_t \\ y_t &= I_{\{0,\infty)}(z_t), \end{aligned} \tag{2}$$

gdzie $I_\Omega(w)=1$, gdy $w\in\Omega$ i $I_\Omega(w)=0$, jeżeli $w\notin\Omega$; zob. np. [1] lub [17]. Powyższa konstrukcja modelu ma szersze znaczenie, ponieważ wykorzystuje się ją także do konstrukcji modelu dla danych wielomianowych; zob. [3], [9] i [13].

Najczęściej stosowany model probitowy otrzymamy przyjmując dla składników losowych ε_t standaryzowane (niezależne) rozkłady normalne, $\varepsilon_t \sim N(0, 1)$. Wówczas zmienne ukryte z_t posiadają rozkład normalny o wartości oczekiwanej $x_t\beta$ i jednostkowej wariancji, $z_t\sim N(x_t\beta, 1)$. Zauważmy, że znamy jedynie znak nieobserwowalnych zmiennych z_t, zatem wariancja składnika losowego jest nieidentyfikowalnym parametrem, stąd przyjmuje się, że wynosi ona jeden. Korzystając z definicji dystrybuanty możemy łatwo sprawdzić równoważność specyfikacji (1) i (2):

---

[2] Drugim najczęściej stosowanym modelem w przypadku analizy danych jakościowych jest model logitowy, który uzyskujemy przyjmując za $F(\cdot)$ dystrybuantę rozkładu logistycznego.

$$p_t = 1 - \Pr(y_t = 0) = 1 - \int_{-\infty}^{0} f_N(w|x_t\beta,\, 1)dw = 1 - \int_{-\infty}^{-x_t\beta} f_N(w'|0,\, 1)dw' = 1 - F(-x_t\beta) = F(x_t\beta), \quad (3)$$

gdzie $f_N(\cdot|a,b)$ jest funkcją gęstości rozkładu normalnego o wartości oczekiwanej a i wariancji b. Powyższa zależność jest prawdziwa nie tylko w przypadku rozkładu normalnego, ale dla zmiennej z_t o dowolnym ciągłym rozkładzie prawdopodobieństwa z symetryczną funkcją gęstości.

W niniejszej pracy w celu estymacji modelu (2) wykorzystamy podejście bayesowskie, które w przypadku analizy danych jakościowych stosują m.in. autorzy prac: [1], [12], [13], [21] i [22].

3. BAYESOWKI MODEL PROBITOWY.

Na gruncie bayesowskim każdą nieznaną wielkość traktujemy jako zmienną losową, a zatem funkcja (gęstości) prawdopodobieństwa jej rozkładu odzwierciedla pełną wiedzę o tej wielkości. Różne elementy modelu statystycznego (obserwacje - y i parametry - θ) traktowane są symetrycznie, więc wstępną wiedzę o wszystkich nieznanych wielkościach (obserwowalnych i nieobserwowalnych) posiadaną przed zaobserwowaniem zjawiska empirycznego odzwierciedla łączny rozkład prawdopodobieństwa o funkcji gęstości p(y, θ) (lub uogólnionej gęstości w przypadku rozkładu mieszanego, dyskretno-ciągłego). W bayesowskim modelu statystycznym ogólne zasady estymacji parametrów sprowadzają się do wyznaczenia z rozkładu łącznego p(y, θ) warunkowej gęstości dla wektora parametrów θ, przy zaobserwowanym wektorze y, czyli funkcji gęstości tzw. rozkładu a posteriori, danej wzorem Bayesa:

$$p(\theta|y) = \frac{p(y,\theta)}{p(y)} = \frac{p(y|\theta)\cdot p(\theta)}{\int_{\Theta} p(y|\theta)\cdot p(\theta)d\theta} \propto l(\theta|y)\cdot p(\theta), \quad (4)$$

gdzie $l(\theta|y)$ jest funkcją wiarygodności, wyznaczoną z rozkładu próbkowego ($l(\theta|y)$=p($y|\theta$)), a $\propto$ jest znakiem proporcjonalności, zob. Osiewalski [14].

Model (2) jest przykładem modelu hierarchicznego, więc łączna funkcja uogólnionej gęstości dla parametrów β i wielkości nieobserwowalnych $z=(z_1,\ldots,z_T)'$ oraz wektora obserwacji $y=(y_1,\ldots,y_T)'$, charakteryzująca jednoznacznie bayesowski model statystyczny, ma postać[3]

$$p(y,z,\beta) = p(y|z,\beta)\cdot p(z,\beta) = p(y|z,\beta)\cdot p(z|\beta)\cdot p(\beta), \quad (5)$$

gdzie $p(\beta)$ to rozkład a priori, a p($z|\beta$) to gęstość warunkowego rozkładu wektora zmiennych ukrytych z, której postać zależy od postaci przyjętego rozkładu dla ε_t w równaniu (2). Korzystając ze wzoru (4), gdzie $\theta=(\beta'\ z')'$ otrzymujemy postać łącznej funkcji gęstości rozkładu a posteriori dla β i z przy danym y:

[3] W zapisie pominięto, że wszystkie funkcje gęstości prawdopodobieństwa są warunkowe względem egzogenicznej macierzy danych X=(x_1',…,x_T')′.

$$p(z,\beta|y)=\frac{p(y,z,\beta)}{p(y)}\propto p(y|z,\beta)\cdot p(z|\beta)\cdot p(\beta),\ \text{gdzie} \tag{6}$$

zdegenerowany (jednopunktowy) rozkład próbkowy dla zaobserwowanego wektora y (warunkowy względem z i β) ma postać

$$p(y|z,\beta)=\prod_{t=1}^{T}\left(y_t\cdot I_{[0,\infty)}(z_t)+(1-y_t)\cdot I_{(-\infty,0)}(z_t)\right). \tag{7}$$

W celu uzyskania z formuły (6) brzegowych rozkładów a posteriori powinniśmy dokonać wielokrotnego całkowania, przy czym uzyskanie np. brzegowej gęstości a posteriori p$(\beta|y)$ wymaga T-krotnego całkowania (względem T zmiennych ukrytych) , gdzie T to liczba obserwacji. Za sprawą niestandardowej postaci rozkładu próbkowego (6) całkowanie analityczne jest wykluczone. W tego typu zagadnieniach wykorzystuje się metody numerycznej aproksymacji brzegowych rozkładów a posteriori stosując losowanie Gibbsa. (ang. *Gibbs Sampling*). W metodzie tej posługujemy się jedynie rozkładami warunkowymi, z których uzyskujemy poprzez wielokrotne generowanie liczb pseudolosowych próbę z rozkładu a posteriori (choć tylko w sensie asymptotycznym), zob. np. [4], [15] i [18].

W przypadku modelu probitowego, gdy wektor składników losowych $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_T)'$ ma wielowymiarowy rozkład normalny o jednostkowej macierzy kowariancji (I_T) scentrowany wokół zera, $\varepsilon\sim N^{(T)}(0,I_T)$, łączna funkcja gęstości dla β i z przy danym y ma postać

$$p(z,\beta|y)\propto p(\beta)\cdot\prod_{t=1}^{T}\left(y_t\cdot I_{[0,\infty)}(z_t)+(1-y_t)\cdot I_{(-\infty,0)}(z_t)\right)\cdot f_N(z_t|x_t\beta,1). \tag{8}$$

W celu zastosowania algorytmu Gibbsa wyznaczamy z gęstości (8) warunkowy względem z rozkład a posteriori dla β, co zapiszemy jako p($\beta|z,y$), oraz rozkład a posteriori dla z warunkowy względem β. Zauważmy, że ze wzoru (8) otrzymujemy

$$p(\beta|z,y)\propto p(\beta)\cdot\prod_{t=1}^{T}f_N(z_t|x_t\beta,1), \tag{9}$$

co odpowiada przypadkowi normalnej regresji liniowej postaci $z=X\beta+\varepsilon$, gdzie $\varepsilon\sim N^{(T)}(0,I_T)$; nie pojawiają się tu zmienne dyskretne y_t.

W zastosowaniach wnioskowania bayesowskiego istotny jest dobór rozkładu a priori. Przy jego doborze próbujemy pogodzić zasadę wiernego odzwierciedlenia wstępnej wiedzy o parametrze (w naszym przypadku jej braku) i unikanie komplikacji obliczeń na etapie wyznaczania rozkładu a posteriori. W omawianym przypadku wybór rozkładu a priori dla β nie stanowi problemu. Jeżeli przyjmiemy jako p(β) niewłaściwy rozkład jednostajny na całej przestrzeni R^k, to warunkowy względem z rozkład a posteriori dla β to k-wymiarowy rozkład normalny o wektorze wartości oczekiwanych $\hat{\beta}_z$ i macierzy kowariancji $(X'X)^{-1}$, co zapisujemy

$$\beta|z,y\sim N^{(k)}\left(\hat{\beta}_z,(X'X)^{-1}\right),\ \text{gdzie}\ \hat{\beta}_z=(X'X)^{-1}X'z. \tag{10}$$

Jeżeli natomiast dla β przyjmiemy informacyjny rozkład a priori, tj. normalny o wartości oczekiwanej β^* i macierzy kowariancji H^*, który należy do tzw. sprzężonej rodziny rozkładów, to poszukiwany rozkład a posteriori dla wektora parametrów β (warunkowy względem z i y) jest także rozkładem normalnym

$$\beta | z, y \sim N^{(k)}\left(\tilde{\beta}, \tilde{H}\right), \tag{11}$$

o następujących parametrach

$$\tilde{\beta} = \tilde{H}\left(H^{*-1}\beta^* + X'z\right)$$
$$\tilde{H} = \left(H^{*-1} + X'X\right)^{-1}.$$

W tym przypadku istotne znaczenie ma dobór stałych β^* i H^*, aby przyjęty rozkład a priori dobrze reprezentował wstępną wiedzę badacza.

Pełny warunkowy rozkład a posteriori dla wektora z wyznaczamy w oparciu o wzór (8). Założyliśmy wcześniej, że zmienne $z=(z_1,\ldots,z_T)'$ są niezależne (przy ustalonym β), więc

$z_t | \beta, z$ ma rozkład $N(x_t\beta, 1)$ ucięty na lewo (na prawo) od zera, gdy $y_t=1$ ($y_t=0$). (12)

Aby uzyskać próbę z rozkładu a posteriori (8) stosując schemat Gibbsa, w pojedynczym cyklu losujemy najpierw z gęstości rozkładu normalnego (12), a następnie (w zależności od przyjętego rozkładu a priori dla β) losujemy z gęstości (10) albo (11). Za wartości inicjujące łańcuch Markowa przyjmiemy dla β oceny uzyskane metodą największej wiarygodności (MNW) lub oceny estymatora metody najmniejszych kwadratów dla liniowego modelu prawdopodobieństwa, czyli korzystając z formuły $(X'X)^{-1}X'y$.

4. BAYESOWSKI MODEL t-STUDENTA.

Specyfikacja modelu dwumianowego poprzez wprowadzenie w (2) zmiennych ukrytych z_t pozwala na uogólnienie modelu probitowego. Rozszerzenie klasy modelu można uzyskać poprzez wprowadzenie dla ε_t rozkładu z rodziny t-Studenta jak proponują Albert i Chib [1]. Jeżeli przyjmiemy, że parametrem podlegającym estymacji jest również liczba stopni swobody $\nu>0$, to wówczas pozwolimy danym empirycznym wyznaczyć najbardziej prawdopodobną a posteriori wartość tego parametru, a wartości zakładane a priori w innych modelach mogą podlegać testowaniu. Stosowanie tej rodziny rozkładów jest tym bardziej uzasadnione, że (jak zauważyli autorzy pracy [1]) często stosowany model logistyczny w przybliżeniu odpowiada przyjęciu założenia o około 8 stopniach swobody dla rozkładu t-Studenta.[4] Zatem w ramach modelu z

[4] Albert i Chib [1] uzyskali m.in. wyniki a posteriori dla modelu t-Studenta z 8 stopniami swobody zbliżone do wyników dla modelu logitowego. Ponadto G. Mudholkar i E. George (”A Remark on the Shape of the Logistic

rozkładem t-Studenta można dokonać statystycznej weryfikacji obu najbardziej znanych modeli dla danych dychotomicznych: logistycznego $\nu \approx 8$ i probitowego $\nu = +\infty$. W szczególności, gdyby uzyskane wyniki wskazywały na dużą liczbę stopni swobody, to wówczas model probitowy byłby wystarczająco adekwatny.

Zatem przyjmijmy, że zmienne z_t w równaniu (2) mają niezależne rozkłady t-Studenta z ν stopniami swobody, parametrem położenia (niecentralności) $x_t\beta$ i jednostkowym parametrem skali. Równoważną specyfikację tego modelu uzyskujemy budując model hierarchiczny, który uwzględnia dodatkowe zmienne $\lambda_1, \ldots, \lambda_T$ o rozkładzie gamma z parametrami $\nu/2$ i $\nu/2$ i przyjmując dla z_t warunkowy (względem λ_t) rozkład normalny o parametrach $x_t\beta$ i λ_t^{-1}:

$$\begin{aligned} z_t \mid \beta, \lambda &\sim \mathrm{N}\left(x_t\beta, \lambda_t^{-1}\right) \\ \lambda_t \mid \nu &\sim \mathrm{Gamma}\left(\nu/2, \nu/2\right) \end{aligned} \tag{13}$$

Powyższa dekompozycja wynika z tożsamości

$$p(z_t|\beta,\nu) = \int_0^\infty p(z_t, \lambda_t|\beta,\nu)d\lambda_t = \int_0^\infty f_N\left(z_t|x_t\beta,\lambda_t^{-1}\right) f_G\left(\lambda_t|\nu/2,\nu/2\right)d\lambda_t = f_{St}\left(z_t|\nu, x_t\beta, 1\right),$$

gdzie $f_G(\cdot|a,b)$ jest funkcją gęstości rozkładu gamma o wartości oczekiwanej a/b i wariancji a/b^2, natomiast $f_{St}(\cdot)$ jest funkcją gęstości rozkładu t-Studenta (zob. [14]).

Przyjmując za autorami [1] niewłaściwy rozkład a priori dla β, otrzymamy łączną gęstość rozkładu a posteriori dla z, β, λ i ν (przy zaobserwowanym y) postaci

$$\begin{aligned} & p(z, \beta, \lambda, \nu|y) \propto \\ & p(\nu) \cdot \prod_{t=1}^T \left(y_t \cdot I_{[0,\infty)}(z_t) + (1-y_t)\cdot I_{(-\infty,0)}(z_t)\right) \cdot f_N\left(z_t|x_t\beta, \lambda_t^{-1}\right) \cdot f_G\left(\lambda_t|\tfrac{\nu}{2}, \tfrac{\nu}{2}\right), \end{aligned} \tag{14}$$

gdzie $p(\nu)$ to rozkład a priori dla stopni swobody. Warunkowe rozkłady a posteriori dla z, β i λ, będące podstawą do wykorzystania algorytmu Gibbsa, wyprowadzamy z łącznej gęstości (14) uzyskując

$$z_t|\beta, \lambda, \nu, y \sim N\left(x_t\beta, \lambda_t^{-1}\right) \text{ ucięty na lewo (na prawo) od zera, gdy } y_t = 1 \ (y_t = 0). \tag{15}$$

$$\beta|z, \lambda, \nu, y \sim N^{(k)}\left(\hat{\beta}_{z,\lambda}, (X'WX)^{-1}\right), \text{ gdzie } \hat{\beta}_{z,\lambda} = (X'WX)^{-1}X'Wz, \tag{16}$$

$$\lambda_t|z, \beta, \nu, y \sim Gamma\left(\tfrac{\nu+1}{2}, \tfrac{1}{2}\left(\nu + (z_t - x_t\beta)^2\right)\right), \tag{17}$$

gdzie W jest diagonalną macierzą o wymiarach T na T zawierającą na przekątnej elementy λ_t, czyli $W = \mathrm{Diag}(\lambda_1, \ldots, \lambda_T)$. Natomiast warunkowy rozkład a posteriori dla ν nie jest gęstością żadnego ze znanych rozkładów prawdopodobieństwa, przy czym jądro tego rozkładu ma postać

Distribution", Biometrica, 1978) pokazali, że rozkład logistyczny ma identyczną kurtozę, jak rozkład t-Studenta z 9 stopniami swobody.

$$p(\nu|z,\beta,\lambda,y) \propto p(\nu)\cdot \prod_{t=1}^{T} \Gamma\left(\tfrac{\nu}{2}\right)^{-1} \cdot \left(\tfrac{\nu}{2}\right)^{\nu/2} \cdot \lambda_t^{\nu/2-1} \exp\left(-\frac{\nu}{2}\lambda_t\right). \tag{18}$$

Jako rozkład a priori dla parametru swobody przyjmujemy, jak w pracy [6], rozkład wykładniczy z parametrem γ, o funkcji gęstości $p(\nu)=\gamma\exp(-\gamma\cdot\nu)$. Wartość parametru γ równa np. 10^{-1} implikuje rozkład a priori o wartości oczekiwanej 10 i wariancji 100, który możemy uznać za prawie nieinformacyjny rozkład a priori. Uwzględniając wykładniczy rozkład a priori, postać warunkowego rozkładu a posteriori dla ν ma postać

$$p(\nu|z,\beta,\lambda,y) = c_0\cdot c_1\cdot\Gamma\left(\tfrac{\nu}{2}\right)^{-T}\cdot(\nu/2)^{T\nu/2}\cdot\exp(-\nu\cdot c_2), \text{ gdzie} \tag{19}$$

c_0 – stała normująca, $c_1=\gamma\cdot\prod_{t=1}^{T}\lambda_t$ oraz $c_2=\gamma+\frac{1}{2}\left(\sum_t^T\lambda_t-\sum_t^T\ln\lambda_t\right)$, przy czym wielkości c_0 i c_1 nie mają znaczenia w dalszej, gdyż nie zależą od ν.

Zatem w przypadku parametru ν musimy zastosowań inną technikę uzyskiwania realizacji z rozkładu a posteriori, mianowicie wykorzystujemy losowanie z odrzucaniem (ang. *acceptance – rejection sampling*) opisane w omawianym przypadku m.in. w pracy [6] lub [16]. Schemat losowania z odrzucaniem polega na tym, że chcąc uzyskać realizację z ciągłego rozkładu (19), który oznaczymy $p(\nu;\, T,\, c_2)$, w pierwszym kroku losujemy realizację innej zmiennej losowej - ν^* - z rozkładu pomocniczego $g(\nu^*;\, \alpha)$, gdzie α jest parametrem tego rozkładu. Zakładamy, że istnieje kres górny ilorazu obu gęstości: $\sup_\nu\, [p(\nu;\, T, c_2)/g(\nu;\, \alpha)] = c(\alpha)<\infty$. Następnie jeżeli zachodzi nierówność $[c(\alpha)]^{-1}\cdot p(\nu^*;\, T, c_2)/g(\nu^*;\, \alpha)>u$ przyjmujemy, że $\nu=\nu^*$, gdzie u jest realizacją z rozkładu jednostajnego na przedziale (0;1). Oznacza to, że losując z gęstości pomocniczej $g(\nu^*;\, \alpha)$, z prawdopodobieństwem akceptacji równym $[c(\alpha)]^{-1}\cdot p(\nu^*;\, T, c_2)/g(\nu^*;\, \alpha)$, uzyskaliśmy realizację zmiennej ν z rozkładu $p(\nu;\, T, c_2)$. W przeciwnym razie powtarzamy losowanie z $g(\nu^*;\, \alpha)$, aż do spełnienia powyższej nierówności. Bezwarunkowe prawdopodobieństwo spełnienia powyższej nierówności jest równe stałej $[c(\alpha)]^{-1}$, więc tak dobieramy parametr funkcji pomocniczej α, aby to prawdopodobieństwo było maksymalne.

W omawianym przypadku niech $g(\nu^*;\, \alpha)$ będzie gęstością rozkładu wykładniczego z parametrem α. Następnie konstruujemy funkcje pomocniczą:

$$\begin{aligned} &Q(\nu,\alpha;T,c_2)= \\ &\ln\frac{p(\nu;T,c_0,c_1,c_2)}{g(\nu;\alpha)} = \ln(c_0c_1) - T\ln\Gamma\left(\tfrac{\nu}{2}\right) + T\nu/2\cdot\ln\left(\tfrac{\nu}{2}\right) + \nu(\alpha-c_2) - \ln\alpha, \end{aligned} \tag{20}$$

aby znaleźć optymalne α i stałą $c(\alpha)$ rozwiązując następujący problem

$$\inf_\alpha\left(\sup_\nu(Q(\nu,\alpha;T,c_2))\right).$$

Zauważmy, że $\partial Q/\partial\alpha=\nu-1/\alpha$, więc optymalna wartość parametru α gęstości pomocniczej wynosi $1/\nu$. Zatem warunek konieczny (i wystarczający) na istnieje ekstremum funkcji, przy czym $\alpha=1/\nu$, ma postać:

$$\frac{\partial Q}{\partial \nu}=\frac{T}{2}\left(\ln\frac{\nu}{2}+1-\frac{\partial \ln\Gamma\left(\frac{\nu}{2}\right)}{\partial \frac{\nu}{2}}\right)-c_2+\alpha=0\,. \tag{21}$$

Następnie numerycznie rozwiązujemy równanie (21), uzyskując rozwiązanie optymalne ν_{opt}. Podsumowując, losowanie z odrzucaniem w omawianym przypadku składa się z następujących kroków:

1. losujemy ν^* z rozkładu wykładniczego z parametrem $1/\nu_{opt}$, czyli o wartości oczekiwanej ν_{opt} i wariancji $(\nu_{opt})^2$,
2. losujemy u z rozkładu jednostajnego (0;1),
3. jeżeli zachodzi poniższa nierówność

$$\Gamma\left(\tfrac{\nu^*}{2}\right)^{-T}\Gamma\left(\tfrac{\nu_{opt}}{2}\right)^{T}\left(\frac{\nu^*}{2}\right)^{0.5\cdot T\cdot\nu^*}\left(\frac{\nu_{opt}}{2}\right)^{-0.5\cdot T\cdot\nu^*} e^{\left(\nu_{opt}-\nu^*\right)\left(c_2-1/\nu_{opt}\right)}\leq u\,,$$

to powtarzamy losowanie z punktu 1,

4. w przeciwnym przypadku przyjmujemy, że ν^* pochodzi z rozkładu o gęstości (19).

Przedstawione powyżej losowanie z odrzucaniem (dla uzyskania realizacji z warunkowego rozkładu a posteriori parametru ν) będziemy stosować w ramach losowania Gibbsa. Mając dane $z^{(q)}$, $\beta^{(q)}$, $\lambda^{(q)}$, $\nu^{(q)}$ jako rezultat q-tego losowania, w kolejnym cyklu dokonujemy następujących kroków:

- $z^{(q+1)}$ jest losowane z rozkładu o gęstości (15) dla $\beta=\beta^{(q)}$, $\lambda=\lambda^{(q)}$,
- $\beta^{(q+1)}$ jest losowane z rozkładu o gęstości (16) dla $z=z^{(q+1)}$, $\lambda=\lambda^{(q)}$,
- $\lambda_t^{(q+1)}$ (t=1,…,T) jest losowane z rozkładu o gęstości (17) dla $z=z^{(q+1)}$, $\beta=\beta^{(q+1)}$, $\nu=\nu^{(q)}$,
- $\nu^{(q+1)}$ jest losowane z rozkładu o gęstości (19) dla $\lambda=\lambda^{(q+1)}$ w sposób opisany wcześniej,

przy czym przyjmujemy takie same wartości początkowe dla $\beta^{(0)}$ jak w przypadku bayesowskiego modelu probitowego, natomiast $\lambda_t^{(0)}=1$ dla t=1,…,T oraz $\nu^{(0)}=10$.

Zasadniczą kwestią w sytuacji, gdy stosujemy schemat Gibbsa jako numeryczną metodę aproksymacji brzegowych rozkładów a posteriori, jest zbadanie zbieżności tego algorytmu, którą w praktyce uzyskuje się po wykonaniu odpowiednio dużej liczby cykli – losowań Gibbsa. Możemy spodziewać się, że stabilizacja w kolejnych cyklach wartości oczekiwanych i odchyleń standardowych a posteriori po odrzuceniu odpowiedniej liczby początkowych losowań, tzw. cykli spalonych, oznacza osiągnięcie zbieżności tej metody. Jedną z graficznych metod badania zbieżności algorytmu jest metoda CumSum, zaproponowana przez Yu i Myklanda i opisana m.in. w pracy [5].

Chcąc zastosować w praktyce wnioskowanie bayesowskie dla rozważanego modelu dychotomicznego, musimy w pierwszej kolejności ustalić parametry rozkładów a priori dla wektora β i parametru ν. W przypadku rozkładu a priori dla β przyjęliśmy, że β^* jest wektorem zerowym, a macierz kowariancji $H^*=\sigma_\varepsilon^2\cdot I$ - macierzą diagonalną. Wartości stałej σ_ε^2 ustaliliśmy w trzech wariantach na poziomie odpowiednio 49, 100 i 256. Ponadto przyjęliśmy, że wartość oczekiwana i wariancja wykładniczego rozkładu a priori dla stopni swobody ν wynosi odpowiednio 10 i 100, a zatem $\gamma=10^{-1}$. Uzyskane wyniki empiryczne (w postaci rozkładów a posteriori oraz odpowiednich charakterystyk tych rozkładów dla parametrów β i wielkości prawdopodobieństwa p_t) nie były wrażliwe przyjęte na wartości parametrów rozkładów a priori.

5. WYNIKI EMPIRYCZNE

W celu empirycznej prezentacji bayesowskich modeli dychotomicznych: probitowego i t-Studenta wykorzystamy dane o kredytach detalicznych, tj. kredytach konsumpcyjnych i hipotecznych, które zostały udzielone klientom indywidualnym przez jeden z dużych, polskich banków komercyjnych w okresie 01.01.2000 - 30.09.2001 r. Wcześniej wykorzystano te dane w pracy [10] i [11] do estymacji m.in. modelu logitowego i probitowego metodą największej wiarygodności.

Niech objaśniana zmienna dychotomiczna y_t przyjmuje następujące wartości:

- $y_t=1$ w przypadku, gdy kredytobiorca na dzień 30.09.2001 ma zaległości w spłacie rat kapitałowo-odsetkowych, tzn. opóźnienie w spłacie ostatniej raty wynosi więcej niż jeden miesiąc. W tym przypadku bank ma obowiązek odprowadzić rezerwy celowe w wysokości 20%, albo 50%, albo 100% wartości zadłużenia w zależności od okresu niespłacania rat przez klienta[5].
- $y_t=0$ w przypadku, gdy kredytobiorca na dzień 30.09.2001 w terminie spłaca raty kapitałowo-odsetkowe od zaciągniętego kredytu.

Dla uproszczenia możemy zatem przyjąć, że z punku widzenia banku jeżeli $y_t=1$, to kredytobiorca jest „złym” klientem, a w przeciwnym przypadku „dobrym”.

W niniejszej analizie wykorzystaliśmy prawie 40 tysięcy rachunków kredytowych, a jako potencjalne zmienne egzogeniczne wyjaśniające ryzyko pojedynczej umowy kredytowej przyjęliśmy (jak w pracy [10] i [11]]):[6]

- płeć (zmienna przyjmuje wartość 1, jeżeli klientem jest mężczyzna, 0 w przypadku kobiety),

---

[5] Uchwała nr 8/1999 Komisji Nadzoru Bankowego z 22 grudnia 1999 r. stanowi zasady tworzenia przez banki rezerw celowych od należności zagrożonych.

[6] Podstawowe charakterystyki tego zbioru danych zostały przedstawione w pracy [10].

- wiek klienta (w setkach lat, aby odpowiednio wyskalować dane),
- wpływy, tzn. wielkość miesięcznych wpływów w latach 2000-2001 (w setkach tys. zł) na rachunki *a vista* kredytobiorcy w badanym banku (przede wszystkim rachunki oszczędnościowo-rozliczeniowe ROR); jeżeli nie posiada rachunku ROR w tym banku przyjęto, że wpływy wynoszą zero,
- posiadanie przez kredytobiorcę rachunku ROR w analizowanym banku (1 - posiada, 0 – nie posiada),
- informację o tym, czy kredytobiorca posiada karty płatnicze lub kredytowe wydane przez rozważany bank (1 - posiada choć jedną kartę, 0 - nie posiada),
- sposób udzielenia klientowi kredytu (1 – udzielono go poprzez pośrednika kredytowego, 0 – bezpośrednio przez bank),
- typ kredytu (1 - kredyt konsumpcyjny, 0 – kredyt hipoteczny),
- podstawowe źródło dochodu uzyskiwanego przez kredytobiorcę (zmienna *zrdoch*), tj. umowa o pracę, albo renta lub emerytura, albo własna działalność, umowa o dzieło lub umowa zlecenie, albo inne źródło (np. stypendium).

Zmienna *zrdoch* przyjmuje cztery różne wartości. Chcąc ją uwzględnić w równaniu regresji z wyrazem wolnym, za referencyjną wartość tej zmiennej przyjęliśmy „umowę o pracę” (dla 75% kredytobiorców stanowi podstawowe źródło dochodu). Typ źródła dochodu określony jest przez trzy następujące zmienne zerojedynkowe *zrdoch1*, *zrdoch2*, *zrdoch3*, przy czym źródłem dochodu kredytobiorcy jest umowa o pracę, jeżeli wszystkie te zmienne przyjmują wartość jeden. W pozostałych przypadkach, gdy

- źródłem dochodu kredytobiorcy jest renta lub emerytura, to *zrdoch1* = 0 i *zrdoch2*=*zrdoch3* = 1,
- źródłem dochodu kredytobiorcy jest własna działalność, umowa o dzieło lub umowa zlecenie, to *zrdoch2* = 0 i *zrdoch1*=*zrdoch3* = 1,
- źródło dochodu jest inne niż wcześniej wymienione, np. stypendium, to *zrdoch3* = 0 i *zrdoch1*=*zrdoch2* = 1.

Podstawowe charakterystyki wykorzystywanego zbioru danych przedstawiamy w pracy [10]. W niniejszej pracy - w odróżnieniu do prac [10] i [11] - założyliśmy, że w modelu (1) prawdopodobieństwo niedotrzymania umowy przez kredytobiorcę (p_t) może zależeć liniowo (poprzez funkcję F) nie tylko od zmiennych egzogenicznych w_{tj}, a także od iloczynów tych zmiennych oraz kwadratów zmiennych ciągłych (wiek, wpływy), co prowadzi do następującego, bardziej ogólnego niż w [10] i [11], modelu:

$$\begin{aligned} z_t &= \beta_1 + \sum_j w_{tj}\beta_j + \sum_j \sum_{i \geq j} w_{tj} w_{ti} \beta_{ij} + \varepsilon_t \\ y_t &= I_{\{0,\infty)}(z_t). \end{aligned} \tag{22}$$

W efekcie wymiar wektora parametrów β zwiększył się z 11 do 54 (po uwzględnieniu faktu, że zmienna *pośrednik* determinuje *typ kredytu*). Z punktu widzenia omówionych wcześniej metod wnioskowania, sposób wprowadzenia zmiennych egzogenicznych w formule (22) nie wnosi żadnych komplikacji, gdyż z_t jest nadal liniowo zależne od parametrów β. Powyższa modyfikacja pozwoli na lepsze oszacowanie p_t i umożliwi poszukiwanie optymalnych wartości zmiennych ciągłych (ze względu na minimalizację wielkości p_t). Zauważmy, że jeżeli w modelu (22) założymy β_{ij}=0, wówczas iloraz pochodnych cząstkowych prawdopodobieństwa p_t względem zmiennych w_{ti} i w_{tj}, tzw. efektów krańcowych, jest równy ilorazowi parametrów, tj. β_i/β_j, a zatem nie zależy od wartości tych zmiennych. Zatem uwzględnienie w (22) iloczynów i kwadratów zmiennych egzogenicznych w_{tj} powoduje, że iloraz efektów krańcowych zależy od wszystkich zmiennych egzogenicznych, czyli dla każdej obserwacji może być inny. Powyższe rozszerzenie liczby czynników wyjaśniających p_t może być przedmiotem statystycznej weryfikacji, co pokazujemy w dalszej części.

Wyniki przeprowadzonych badań wskazują (zgodnie z intuicją), że w przypadku bardzo dużej liczby obserwacji, wyniki dla modelu probitowego uzyskane metodą największej wiarygodności są identyczne z rezultatami dla bayesowskiego modelu o normalnym rozkładzie zmiennej ukrytej z_t (por. [11]). Zatem w dalszej części przedstawiamy przede wszystkim wyniki estymacji bayesowskiego modelu t-Studenta o nieznanej liczbie stopni swobody na tle klasycznego modelu probitowego, który jest najczęściej wykorzystywany w praktyce. Tabela 1 zawiera wartości oczekiwane i odchylenia standardowe a posteriori dla parametrów modelu t-Studenta, czyli dla wektora β oraz stopni swobody ν.

Tabela 1

Wartość oczekiwana a posteriori dla stopni swobody wynosi około 1,3 przy niewielkim odchyleniu standardowym, a zatem założenie normalności składnika losowego w równaniu (2) jest bezzasadne[7]. Rozkład próbkowy dla zmiennych z_t charakteryzuje się rozkładem o tak bardzo grubych ogonach, że nie posiada wariancji. Innymi słowy, dystrybuanta F w (1) jest znacznie spłaszczona w stosunku do modelu probitowego. W świetle wyników dla ν ogólniejszy model t-Studenta jest zdecydowanie bardziej preferowany przez dane niż model ze składnikiem losowym o rozkładzie normalnym.

Ogólna postać (22) wprowadzająca iloczyny i kwadraty zmiennych w_{tj}, może być przedmiotem testowania. Przy weryfikacji tej specyfikacji wykorzystaliśmy bayesowski odpowiednik klasycznego testu F na redukcję modelu, posługując się formą kwadratową:

[7] W przypadku bayesowskiego modelu t-Studenta z 11 parametrami wartość oczekiwana dla parametru stopni swobody wynosi około 2, por. Marzec [2003b].

$$u(\beta_{(2)}) = (\beta_{(2)} - b_{(2)})^T H_{22}^{-1}(\beta_{(2)} - b_{(2)})/ k_2, \text{ gdzie } \beta_{(2)} = [\beta_{12} \dots \beta_{54}]'. \quad (23)$$

Jeżeli brzegowy rozkład a posteriori wektora $\beta_{(2)}$ jest k_2-wymiarowym rozkładem t-Studenta o T-k stopniach swobody ($1 \le k_2 \le k$), wektorze niecentralności $b_{(2)}$ i macierzy precyzji H_{22}^{-1}, to rozkład a posteriori u($\beta_{(2)}$) jest rozkładem F-Snedecora o (k_2, T-k) stopniach swobody; por. Zellner [20]. Dla restrykcji $\beta_{(2)} = [\beta_{12} \dots \beta_{54}]' = [0 \dots 0]'$, która odpowiada redukcji modelu (22) do modelu liniowego względem w_{tj}, wartość u([0…0]′) wynosi około 279. Jest to wartość z ogona rozkładu F(k_2,T-k), który jest dobrą aproksymacją rozkładu a posteriori tej formy kwadratowej. Wobec tego wybrany element podprzestrzeni parametrów $\beta_{(2)} = [0 \dots 0]'$ znajduje się w podzbiorze wartości parametrów mało prawdopodobnych a posteriori. W takim wypadku nie jesteśmy skłonni przyjąć, że $\beta_{(2)} = [0 \dots 0]'$ i nie dokonujemy redukcji modelu. Testy wyraźnie wskazują na przewagę modelu z 54 parametrami (ogólniejszego) nad modelem z 11 parametrami (zagnieżdżonym)[8]. Posługując się najprostszymi skalarnymi miernikami dopasowania modelu do danych empirycznych nie uzyskujemy tak jednoznacznych wyników (faworyzujących model lepiej sparametryzowany). Współczynnik determinacji Efrona - $R^2 = 1 - \sum(y_t - \hat{p}_t)^2 / \sum(y_t - \bar{y})^2$ (zob. np. [1]) - dla obu tych modeli przyjmuje zbliżone wartości rzędu 0.28 – 0.27, przy czym tak niskie wartości są typowe z uwagi na dychotomiczny charakter zmiennej y_t. Oba modele - ogólniejszy i zagnieżdżony – równie dobrze prognozują Pr(y_t=1), tzn. udział złych prognoz zmiennej y_t w obu przypadkach wynosi około 19%, przy czym w modelu (22) jest on nieznacznie niższy. Przez złą prognozę rozumiemy sytuację, gdy obserwujemy $y_t = 1$, a oszacowane prawdopodobieństwo jest mniejsze od 0.5 oraz gdy $y_t = 0$, a $\hat{p}_t$ jest co najmniej 0.5 (zob. np. [2], [7]).

Pojedyncze parametry β_j badanego modelu (22) nie mają bezpośredniej interpretacji. W celu porównania wyników obu modeli – z 11 i 54 parametrami, obliczyliśmy średnie arytmetyczne po wszystkich obserwacjach dla efektów krańcowych względem zmiennych egzogenicznych, tj. dla $\partial \mathrm{Pr}(y_t=1)/\partial w_{tj}$. Wyniki te przedstawia Tabela 2.

Tabela 2

Należy zwrócić uwagę na zgodność znaków badanych charakterystyk w przedstawionych powyżej modelach, z wyłączeniem znaku przy zmiennej *ROR*. Z uwagi na to, że spośród przedstawionych modeli dane zdecydowanie preferują ogólniejszy model t-Studenta z 54 parametrami, zatem interpretację wyników przedstawimy przede wszystkich dla tego modelu. Dodatni znak efektu krańcowego informuje nas, że wzrost w_{tj} powoduje wzrost szans, że $y_t = 1$. Zatem, jeżeli hipotetycznym klientem okazuje się być mężczyzna, to ryzyko niedotrzymania przez niego umowy jest wyższe niż w przypadku kobiety. Analogicznie posiadanie przez kredytobiorcę rachunku oszczędnościowo-rozliczeniowego, w przeciwieństwie do posiadania choć jednej karty płatniczej lub kredytowej, wiąże się z wyższym prawdopodobieństwem niespłacenia przez niego kredytu. Udzielenie kredytu poprzez pośrednika, podobnie jak udzielenie kredytu konsumpcyjnego zamiast hipotecznego, zwiększa ryzyko kredytowe. W przypadku zmiennych określających źródło dochodu wszystkie trzy modele zgodnie informują, że studenci korzystający z kredytu studenckiego (*zrdoch3*) oraz emeryci i renciści (*zrdoch1*) są mniej ryzykownymi kredytobiorcami niż klienci zatrudnieni na umowę o pracę. Największe ryzyko kredytowe wiąże się z udzieleniem kredytu detalicznego klientom prowadzącym własną działalność gospodarczą (*zrdoch2*). Podsumowując, spośród zmiennych zero-jedynkowych w_{tj} największy wpływ na ryzyko kredytowe ma typ udzielonego kredytu oraz fakt, czy klient posiada kartę płatniczą lub kredytową czy nie. Ponadto wraz z wiekiem kredytobiorcy i wielkością jego wpływów na rachunek ROR maleje prawdopodobieństwo niedotrzymania umowy kredytowej.

Iloraz efektów krańcowych względem pary zmiennych np. w_{ti} i w_{tj} informuje, ile razy większa jest reakcja $\Pr(y_t=1)$ na jednostkowy przyrost w_{ti}, w porównaniu z reakcją $\Pr(y_t=1)$ na jednostkowy przyrost w_{tj}. Przykładowo, w bayesowskim modelu t-Studenta z 54 parametrami, reakcja prawdopodobieństwa niedotrzymania umowy przez kredytobiorcę ze względu na sposób udzielenia kredytu jest około 28,5 razy większa niż reakcja $\Pr(y_t=1)$ na to, czy kredytobiorcą jest kobietą, czy mężczyzną. Natomiast w klasycznym modelu probitowym oraz w drugim modelu t-Studenta (k=11) parametrami iloraz ten wynosi odpowiednio około 31 i 90. Przedstawione modele różnią się znacząco ze względu na wartość i ranking efektów krańcowych oraz ilorazy tych wielkości.

Ponadto zbadaliśmy, czy istnieją optymalne ze względu na p_t wartości zmiennych ciągłych, tj. wieku i wpływów. Dla czterech hipotetycznych kredytobiorców (por. Tabela 3) obliczyliśmy optymalne wartości obu zmiennych, które minimalizują p_t, jednakże wykraczają one poza zakres obserwowanych wartości tych zmiennych. Ponadto ze statystycznego punktu widzenia duże odchylenie standardowe a posteriori dla parametru β_{21} (przy kwadracie zmiennej *wiek*) w stosunku do wartości oczekiwanej wskazuje, że zerowa wartość tego parametru jest wysoce prawdopodobna a posteriori. Zatem rozproszenie rozkładu a posteriori dla optymalnego wieku kredytobiorcy byłoby relatywnie duże, a wnioskowanie o tej wielkości byłoby słabe.

Oszacowane modele możemy wykorzystać do celów prognostycznych, czyli prognozowania prawdopodobieństwa „złego kredytu” w przypadku potencjalnego kredytobiorcy. Dla uproszczenia rozważmy cztery hipotetyczne sylwetki potencjalnych klientów starających się o kredyt, które przedstawia Tabela 3 (zob. także [10] i [11]). Wszystkie modele zgodnie prognozują, że największe ryzyko kredytowe związane jest z klientem będącym młodym mężczyzną, który utrzymuje się z

---

[8] Podobne wyniki otrzymaliśmy także w przypadku klasycznego modelu probitowego, w którym wartość ilorazu wiarygodności modelu ogólniejszego i zagnieżdżonego wynosi około 377.

prowadzenia własnej działalności i nie korzysta z jakichkolwiek innych usług badanego banku oprócz kredytu, który został mu udzielony poprzez pośrednika. Preferowany przez dane model t-Studenta z 54 parametrami wskazuje, że prawdopodobieństwo niedotrzymania umowy kredytowej (p_t) przez tego „młodego biznesmena" jest bardzo wysokie i wynosi 0,55 (±0,05). Jednocześnie we wszystkich modelach precyzja wnioskowania o p_t dla „młodego biznesmena" jest najmniejsza w porównaniu do pozostałych klientów. Najmniejsze ryzyko kredytowe, spośród czterech rozważanych kredytobiorców, związane jest ze starszą panią utrzymującą się z emerytury, której udzielono kredyt hipoteczny. Model t-Studenta (k=54) wskazuje, że praktycznie brak jest jakiegokolwiek ryzyka w przypadku tego klienta. Największe różnice w oszacowaniu p_t dotyczą „najczęstszego klienta", tzn. o cechach najczęstszych w próbie (dotyczy zmiennych jakościowych) i przeciętnych (dla zmiennych ciągłych) w badanej zbiorowości, który uzyskał kredyt poprzez pośrednika. W przypadku tego klienta, w modelu t-Studenta (k=54) prawdopodobieństwo „złego kredytu" wynosi 0,044, to jest istotnie mniej niż w przypadku modelu probitowego, który szacuje tę wielkość na poziomie 0,3-0,19. W przypadku „najczęstszego klienta", który uzyskał kredyt konsumpcyjny bezpośrednio z banku, model t-Studenta (k=54) szacuje p_t na poziomie 0,015, co z praktycznego punktu widzenia oznacza brak ryzyka. Także pozostałe modele prognozują to prawdopodobieństwo na podobnym poziomie. Podsumowując, widoczne są istotne różnice w wielkości oszacowanego prawdopodobieństwa „złego kredytu" między preferowanym przez dane modelem t-Studenta (k=54), a często stosowanym modelem probitowym, a także modelem logistycznym; por. [9]. Natomiast różnice między dwoma bayesowskimi modelami t-Studenta z 54 i 11 parametrami nie są już tak wyraźne.

Tabela 3

6. PODSUMOWANIE

W niniejszym opracowaniu przedstawiliśmy, odwołując się do literatury przedmiotu, specyfikację bayesowskich modeli dla danych dwumianowych, zarówno z normalnym składnikiem losowym, jak i z rozkładem t-Studenta o nieznanej liczbie stopni swobody. Dla tej klasy modeli omówiliśmy szczegółowo wykorzystanie losowania Gibbsa jako metody numerycznej aproksymacji brzegowych rozkładów a posteriori. Przeprowadzone badania empiryczne wskazują, że podejście bayesowskie pozwoliło na uzyskanie nowych wyników. Warto pamiętać, że na gruncie klasycznym metoda największej wiarygodności, wystarczająca w przypadku tak dużej liczbie obserwacji, nie ma zbadanych własności w przypadku modelu t-Studenta, stąd potrzeba zastosowania podejścia bayesowskiego. Ponadto w przypadku małej próby podejście bayesowskie jest polecane z uwagi na nieasymptotyczne (małopróbkowe) własności, o czym w przypadku analizy modeli dychotomicznych pisze np. Zellner [21]. Wyniki empiryczne wyraźnie preferują model t-Studenta z około jednym stopniem swobody, tj. model z rozkładem Cauchy'ego. Zatem zastosowanie w tym przypadku modelu probitowego czy logitowego jest nieuzasadnione ze statystycznego punktu widzenia, zwłaszcza, że uzyskane prognozy prawdopodobieństwa niespłacenia kredytu czy efekty krańcowe są bardzo rozbieżne w modelach: probitowym i t-Studenta.

Akademia Ekonomiczna w Krakowie

LITERATURA

[1] Albert J. Chib S., 1993, *Bayesian Analysis of Binary and Polychotomous Response Data*, Journal of the American Statistical Association, 88, s. 669-679.

[2] Amemiya T., 1981, *Qualitative Response Models: A Survey*, Journal of Economic Literature, 19.

[3] Amemiya T., 1985, *Advanced Econometrics*, Harvard University Press, Cambrige Massachusetts.

[4] Casella G., E. George, 1992, *Explaining the Gibbs Sampler*, The American Statistician, 46.

[5] Cowles M.K., B.P. Carlin, 1996, *Markov Chain Monte Carlo Covergence Diagnostic: A Comparative Review*, Journal of the American Statistical Association, 91, s. 883-904.

[6] Geweke J., 1996, *Monte Carlo Simulation and Numerical Integration* in H. Amman, D. Kendrick and J. Rust (eds.), Handbook of Computational Economics, Amsterdam: North-Holland.

[7] Greene W.H., 1993, *Econometric Analysis*, Macmillan Publishing Company, New York.

[8] Gruszczyński M., 2001, *Modele i prognozy zmiennych jakościowych w finansach i bankowości*, Monografie i Opracowania SGH, Warszawa, nr 6.

[9] Maddala G.S., 1983, *Limited dependent and qualitative variables in econometrics*, Cambrigde University Press, Cambrigde.

[10] Marzec J., 2003a, *Badanie niewypłacalności kredytobiorcy na podstawie modeli logitowych i probitowych*, Zeszyty Naukowe Akademii Ekonomicznej w Krakowie nr 628 (w druku).

[11] Marzec J., 2003b, *Badanie niespłacalności kredytów za pomocą bayesowskich modeli dychotomicznych - założenia i wyniki*, Metody ilościowe w naukach ekonomicznych (red. A. Welfe), Wydawnictwo SGH w Warszawie (w druku).

[12] McCulloch R.E, N.G. Polson, P. E. Rossi, 2000, *A Bayesian Analysis of the Multinomial Probit Model with Fully Identified Parameters*, Journal of Econometrics, 99, s. 173-193.

[13] McCulloch R.E., P. E. Rossi, 1993, *An exact Likelihood Analysis of the Multinomial Probit Model*, Journal of Econometrics, 64, s. 207-240.

[14] Osiewalski J., 1991, *Bayesowska estymacja i predykcja dla jednorównaniowych modeli ekonometrycznych*, Akademia Ekonomiczna w Krakowie, Zeszyty Naukowe, Seria specjalna: Monografie, nr 100, Kraków.

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

[16] Pajor A., 2002, *Bayesowska estymacja i prognozowanie w modelu stochastycznej zmienności z błędem t-Studenta*, Dynamiczne Modele Ekonometryczne (VII Ogólnopolskie Seminarium Naukowe 4-6 września 2001), Wydawnictwo Uniwersytetu M. Kopernika, Toruń, s. 256-274.

[17] Poirier D.J., P.A. Ruud, 1988, *Probit with dependent observations*, Review of Economics Studies, 55, s. 593-614.

[18] Tierney L., 1994, *Markov chains for exploring posterior distributions (with discussion)*, Annals of Statistics, 22, s. 1701-1762.

[19] Wiśniewski J., 1986, *Ekonometryczne badanie zjawisk jakościowych (studium metodologiczne)*, Uniwersytet M. Kopernika, Toruń

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

[21] Zellner A., 1983, *Bayesian Analysis of Simple Multinomial Logit Model*, Economics Letters, 11, s. 133-136.

[22] Zellner A., P. Rossi, 1984, *Bayesian Analysis of Dichotomous Quantal Response Models*, Journal of Econometrics, 25, s. 365-393.

Tabela 1.

Wartości oczekiwane i odchylenia standardowe a posteriori parametrów bayesowskiego modelu t-Studenta o nieznanej liczbie stopni swobody ν.

Zmienna	parametr	E(|*y*)	D(|*y*)	Zmienna	Parametr	E(|*y*)	D(|*y*)
Stała	β_1	-73.376	13.139	$w_2 \cdot w_{10}$	β_{29}	-1.276	1.367
Płeć (w_1)	β_2	2.445	1.990	$(w_3)^2$	β_{30}	0.221	0.008
Wiek (w_2)	β_3	-3.293	7.399	$w_3 \cdot w_4$	β_{31}	5.592	27.477
Wpływy (w_3)	β_4	-425.893	35.095	$w_3 \cdot w_5$	β_{32}	-4.468	2.652
ROR (w_4)	β_5	52.060	9.691	$w_3 \cdot w_6$	β_{33}	54.125	3.917
Karty (w_5)	β_6	-62.388	31.318	$w_3 \cdot w_7$	β_{34}	-47.521	4.392
Pośrednik (w_6)	β_7	9.807	3.949	$w_3 \cdot w_8$	β_{35}	69.560	17.420
Typ kredytu (w_7)	β_8	64.462	12.374	$w_3 \cdot w_9$	β_{36}	-55.882	3.825
Zrdoch1 (w_8)	β_9	12.283	10.314	$w_3 \cdot w_{10}$	β_{37}	341.724	14.861
Zrdoch2 (w_9)	β_{10}	-3.798	1.678	$w_4 \cdot w_5$	β_{38}	45.302	30.235
Zrdoch3 (w_{10})	β_{11}	10.570	14.561	$w_4 \cdot w_6$	β_{39}	-1.627	0.259
$w_1 \cdot w_2$	β_{12}	-0.562	0.257	$w_4 \cdot w_7$	β_{40}	-45.269	9.006
$w_1 \cdot w_3$	β_{13}	-0.294	2.148	$w_4 \cdot w_8$	β_{41}	-0.204	0.490
$w_1 \cdot w_4$	β_{14}	0.395	0.240	$w_4 \cdot w_9$	β_{42}	1.536	0.327
$w_1 \cdot w_5$	β_{15}	0.256	0.305	$w_4 \cdot w_{10}$	β_{43}	-6.385	3.812
$w_1 \cdot w_6$	β_{16}	-0.338	0.203	$w_5 \cdot w_6$	β_{44}	0.856	0.300
$w_1 \cdot w_7$	β_{17}	-2.444	1.963	$w_5 \cdot w_7$	β_{45}	0.081	1.449
$w_1 \cdot w_8$	β_{18}	-0.030	0.083	$w_5 \cdot w_8$	β_{46}	-0.742	0.667
$w_1 \cdot w_9$	β_{19}	0.323	0.166	$w_5 \cdot w_9$	β_{47}	-0.863	0.371
$w_1 \cdot w_{10}$	β_{20}	0.262	0.313	$w_5 \cdot w_{10}$	β_{48}	17.953	6.037
$(w_2)^2$	β_{21}	0.064	1.006	$w_6 \cdot w_8$	β_{49}	-0.808	0.451
$w_2 \cdot w_3$	β_{22}	11.445	15.377	$w_6 \cdot w_9$	β_{50}	1.599	0.263
$w_2 \cdot w_4$	β_{23}	0.469	1.110	$w_6 \cdot w_{10}$	β_{51}	-8.167	3.818
$w_2 \cdot w_5$	β_{24}	-0.715	1.388	$w_7 \cdot w_8$	β_{52}	-11.904	10.285
$w_2 \cdot w_6$	β_{25}	0.568	0.964	$w_7 \cdot w_9$	β_{53}	1.528	1.632
$w_2 \cdot w_7$	β_{26}	0.473	7.149	$w_7 \cdot w_{10}$	β_{54}	-1.947	14.147
$w_2 \cdot w_8$	β_{27}	0.975	0.483	-	ν	1.302	0.067
$w_2 \cdot w_9$	β_{28}	1.715	0.810				

Źródło: obliczenia własne.

Tabela 2.

Oceny MNW oraz wartości oczekiwane i odchylenia standardowe a posteriori dla uśrednionych efektów krańcowych.

	Model probitowy - MMW (11 parametrów)			Bayesowski model t—Studenta (11 parametrów)		Bayesowski model t—Studenta (54 parametrów)	
Zmienna	Oceny	Błędy szacunku	Stat. t	E($\lvert y$)	D($\lvert y$)	E($\lvert y$)	D($\lvert y$)
płeć	0,008	0,003	2,4	0,004	0,005	0,019	0,006
wiek	-0,170	0,017	-10,0	-0,221	0,023	-0,266	0,031
wpływy	-0,333	0,044	-7,6	-8,297	0,227	-15,733	5,204
ROR	-0,056	0,008	-7,5	0,128	0,017	0,662	0,329
karty	-0,034	0,007	-5,2	-0,053	0,019	-7,907	5,274
pośrednik	0,251	0,006	43,3	0,358	0,011	0,542	0,036
typ kredytu	0,036	0,013	2,8	0,237	0,069	9,412	1,511
Zrdoch1	0,018	0,006	3,1	0,021	0,007	0,076	0,031
Zrdoch2	-0,062	0,008	-7,6	-0,018	0,014	-0,030	0,019
Zrdoch3	0,045	0,015	3,1	0,101	0,027	0,592	0,085

Źródło: obliczenia własne.

Tabela 3.

Wartości oczekiwane i odchylenia standardowe a posteriori prawdopodobieństwa niespłacenia kredytu - $\Pr(y_t=1)=F(x_t\beta)$.

	Najczęstszy Klient		„Młody Biznesmen”	„Starsza Pani”
Zmienna	Pośrednik=1	Pośrednik=0		
Stała	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
Zrdoch1	1	1	**1**	**0**
Zrdoch2	1	1	**0**	**1**
Zrdoch3	1	1	**1**	**1**
Model probitowy (MNW, 11 parametrów)				
Ocena p_t	0,306	0,038	0,664	0,011
Błąd szacunku	(0,014)	(0,002)	(0,016)	(0,002)
Model probitowy (MNW, 54 parametrów)				
Ocena p_t	0,193	0,028	0,551	$4{,}16\cdot10^{-6}$
Błąd szacunku	(0,047)	(0,003)	(0,042)	$(4{,}84\cdot10^{-4})$
Bayesowski model t-Studenta (ν estymowane, 11 parametrów)				
$E(p_t \mid y)$	0,034	0,016	0,584	0,039
$D(p_t \mid y)$	(0,002)	(0,001)	(0,025)	(0,007)
Bayesowski model t-Studenta (ν estymowane, 54 parametrów)				
$E(p_t \mid y)$	0,044	0,015	0,553	0,009
$D(p_t \mid y)$	(0,006)	(0,002)	(0,053)	(0,009)

Źródło: obliczenia własne.