

STUDIA METODOLOGICZNE

Henryk GURGUL, Marcin SUDER

Rozkład prawdopodobieństwa dziennych wypłat z bankomatów

Do wczesnych lat 80. ub. wieku system wpłat i wypłat był stosunkowo prosty i składał się z gotówki, czeków, banków oraz przelewów. Wraz z wprowadzeniem kart debetowych i kredytowych ten system zmienił się zasadniczo.

Jednym z istotnych elementów systemu bankowego stały się sieci bankomatów (*Automated Teller Machines* — ATM). Historia bankomatów sięga końca lat 70. XX w., kiedy to pierwszy bankomat udostępniono klientom nowojorskiego oddziału Chemical Bank. Począwszy od lat 80. ub. wieku liczba bankomatów zaczęła stale rosnąć, zaś liczba okienek bankowych odpowiednio spadać.

W ostatnich latach liczba automatów w krajach wysoko rozwiniętych ustabilizowała się, a nawet zaczęła spadać — i to pomimo realnego ciągłego wzrostu gotówki. Przyczyną tego mogą być banki, które są odpowiedzialne za dystrybucję gotówki, ponieważ operują częścią bankomatów. Jednak należy zaznaczyć, że banki bezpośrednio nie korzystają finansowo na dystrybucji gotówki (korzyść odnosi tu bank centralny), wprost przeciwnie — dystrybucja gotówki jest dla banków źródłem poważnych kosztów. Chociaż banki mogą pokryć część tych kosztów poprzez różne opłaty i taryfy, stopa zwrotu netto z gotówki jest ujemna. Co więcej, użycie gotówki zmniejsza depozyty i w ten sposób zmniejszają się zyski banków. Jednak banki nie mogą wstrzymać dostaw gotówki, ponieważ znaczna część klientów ciągle jej jeszcze używa. Jeśliby więc bank nie miał możliwości prowadzenia usług gotówkowych, to mógłby utracić istotną część klientów.

W Polsce pierwszy bankomat pojawił się w drugiej połowie lat 80. ub. wieku w III Oddziale Banku Polska Kasa Opieki SA w Warszawie (dziś Pekao SA). Jednak dopiero w drugiej połowie lat 90. ub. wieku bankomaty zaczęły odgrywać istotną rolę w kształtowaniu całego systemu. Od tego czasu ich liczba zaczęła wyraźnie wzrastać, jednocześnie wzrastała liczba operacji bankomatowych i obroty w bankomatach. Zwiększała się również liczba operatorów bankomatów.

Badania wykazały, że istnieje istotna i odporna zależność między liczbą bankomatów (w relacji do ludności) i liczbą sieci (operatorów) bankomatów. Jedną z nich jest Euronet, który powstał w 1994 r. w Budapeszcie, a już w roku 1996 wprowadził pierwsze bankomaty w Polsce.

Od tego czasu rozwój sieci bankomatowych był bardzo dynamiczny. Na wykresie zobrazowano zmiany liczby bankomatów wszystkich operatorów w Polsce w latach 1999—2011.

W Polsce działa prawie 18 tys. bankomatów, w których np. w trzecim kwartale 2011 r. dokonano prawie 180 mln operacji (wypłaty, depozyty, przelewy, zakup towarów lub usług) na kwotę ponad 70 mld zł. Są to kwoty, które znacznie przewyższają sumy operacji pieniężnych w oddziałach bankowych.

Tak duża ilość operacji oraz ogromny przepływ gotówki powodują, że zarządzanie obsługą sieci bankomatów ma ogromne znaczenie dla operatora tej sieci. Błędne decyzje związane z zarządzaniem bankomatami, którego elementami są m.in.: lokalizacja bankomatu, wielkość i moment doładowania oraz logistyka konwoju mogą nieść za sobą istotne zwiększenie kosztów obsługi. Operatorzy zatem poszukują rozwiązań mogących zoptymalizować obsługę sieci bankomatów, którymi zarządzają.

Jednym z istotniejszych elementów zarządzania bankomatem lub siecią bankomatów jest prognoza wielkości wypłat z danej maszyny. Możliwość dobrej predykcji wielkości wypłaty z bankomatu pozwala na zbudowanie optymalnego systemu zarządzania minimalizującego koszty obsługi poprzez fakt, że w bankomacie nie będzie za dużo pieniędzy, które nie byłyby w obrocie, a w rezultacie nie przynosiłyby zysku. Ponadto trafne prognozowanie wielkości wypłat minimalizuje ryzyko, że dany bankomat będzie pusty, co mogłoby powodować niezadowolenie klientów.

Z punktu widzenia obniżki kosztów obsługi bankomatów istotną rolę mogą odgrywać także ogólne prognozy zapotrzebowania klientów bankomatów na gotówkę. Prognozy te mogą być użyteczne z punktu widzenia zarządzania gotówką, określania poziomu ryzyka, a przede wszystkim mogą być podstawą przyjęcia przez operatora odpowiedniej strategii napełniania bankomatów gotówką. Są to złożone zagadnienia, bowiem wyniki badań (prognoz) nie mogą być oparte tylko na przeszłych zachowaniach klientów, ale powinny brać pod uwagę dodatkowe czynniki, jak np. aktualną liczbę i strukturę ludności obszaru, na którym znajdują się bankomaty. W modelach badawczych należy uwzględnić nie tylko występowanie składnika systematycznego, ale i efektów losowych.

Aby możliwa była predykcja szeregu czasowego, w pierwszym etapie należy poznać własności statystyczne danych empirycznych pochodzących z maszyny bankomatowej.

W artykule podjęto próbę dopasowania rozkładu sumy dziennych wypłat z bankomatów. Autorzy mają na celu zbadanie, czy wybrane przez nich rozkłady mogą skutecznie opisywać strukturę statystyczną szeregu wypłat z bankomatów. Dodatkowo sprawdzimy, czy rodzaj rozkładu, jaki dopasowano do danych ma związek z miejscem, w którym dany bankomat się znajduje (województwo, miejscowość, typ lokalizacji).

Określenie rozkładu teoretycznego dla danych empirycznych ma istotne znaczenie w prognozowaniu. Znajomość prawdopodobieństwa wystąpienia danych o niskich oraz wysokich wartościach może istotnie wpływać na jakość prognozy. Niektóre szeregi czasowe wykazują tzw. „efekt kalendarza”, tzn. ich wartości mogą zależeć od pory roku, miesiąca, dnia tygodnia, pory dnia itd. W sytuacji gdy szereg czasowy jest nim obciążony, to jeśli znany jest teoretyczny rozkład prawdopodobieństwa dla analizowanych danych, jesteśmy w stanie zweryfikować, czy prognozy „odstające” są dopuszczalne czy też należy je skorygować.

Dodatkowo przedstawimy i porównamy wyniki badań podstawowych własności statystycznych wielkości wypłat z bankomatów według różnych typów ich lokalizacji.

PRZEGLĄD LITERATURY

Zagadnienie optymalizacji zarządzania siecią bankomatową jest problemem dość znanym. Istnieje wiele firm oferujących rozwiązania w tym zakresie. Jednak w literaturze przedmiotu trudno napotkać na publikacje naukowe związane z badaniami za pomocą modeli ekonometrycznych wielkości wypłat gotówkowych z bankomatów. Wynika to głównie z braku dostępu badaczy do tego typu danych. Operatorzy bankomatów, szukając optymalnych rozwiązań w zakresie zarządzania, zwracają się zazwyczaj do firm komercyjnych, a nie do pracowników naukowych. W związku z tym wszelkie analizy tych szeregów są utajnione przez te firmy.

Jedną z najbardziej trafnych prac statystycznych na omawiany temat jest analiza transakcji za pomocą kart kredytowych autorstwa Handa i Blunt (2001). Jednakże w literaturze ekonomicznej można znaleźć też pozycje zajmujące się związkiem pomiędzy liczbą bankomatów i popytem na gotówkę. Amromin i Chakravorti (2007) przeprowadzili analizę za pomocą regresji liniowej danych panelowych w celu porównania krajów, głównie pod kątem wzrostu transakcji przy użyciu kart debetowych w punktach sprzedaży. Boeschoten (1998) oraz Snellman i Viren (2006) mieli podobny cel, ale swoje wnioski sformułowali na podstawie wyników zastosowania modeli opisujących indywidualne zachowania konsumenta.

Boeschoten analizował dane dotyczące konsumentów holenderskich z okresu 1990—1994 stosując deterministyczny model inwentarzowy. Zakładano w nim, że klienci utrzymują stały poziom gotówki i zużywają ją ze stałą stopą, przy czym, jeśli jej ilość spadnie poniżej określonego poziomu, uzupełniają zasoby gotówkowe. Wymienieni badacze wykazali, że klienci bankomatów mają na ogół niższe zapasy gotówki niż osoby niekorzystające z bankomatów, ponieważ koszt uzupełnienia gotówki przez tych pierwszych jest na ogół niższy niż tych drugich.

Związek pomiędzy kosztem uzyskania gotówki i liczbą bankomatów modelowali wspomniani Snellman i Viren w postaci deterministycznego problemu optymalizacyjnego. Zakładali oni, że koszty są proporcjonalne do odległości od bankomatu. Z kolei Simutis i in. (2007) podjęli próbę optymalizacji zarządzania gotówką w sieci bankomatowej poprzez wybór odpowiedniego modelu predykcji wypłat. Autorzy ci wykazują, że zastosowanie sieci neuronowych do prognozowania wypłat z bankomatów pozwala na obniżenie kosztów obsługi sieci bankomatowej o ok. 15—20%. Z kolei Armenise i in. (2010) jako metodę prognozowania wielkości wypłat z bankomatów zaproponowali zastosowanie algorytmów genetycznych.

Pomimo tego, że wymienieni naukowcy stosowali modele wypłat gotówki przez klientów bankomatów z uwzględnieniem czasu, to jednak przyjęte przez nich uproszczenia uniemożliwiały sformułowanie wniosków na poziomie makroekonomicznym, jeśli chodzi o łączne zapotrzebowanie na gotówkę. Autorzy podkreślili, że wnioski mogłyby być uściśnione, gdyby badania były przeprowadzone na poziomie indywidualnych wypłat.

W badaniu zapotrzebowania na gotówkę w bankomatach mogą być istotne takie aspekty, jak ilość i wielkość wypłat, ich rozkłady empiryczne, autokorelacja i efekty czasowe. Istotną rolę w badaniach odgrywają rozkłady. W literaturze dotyczącej rozkładów wypłat występują generalnie dwa typy. Pierwszy rodzaj to rozkłady opisujące sytuację, w której wypłata nie zależy od historii wypłat (rozkład wielomianowy oraz rozkład Dirichleta efektu losowego). W drugim typie modeli nie specyfikuje się rozkładu składnika losowego, lecz aproksymuje się go za pomocą rozkładu empirycznego.

W przypadku różnych danych mających postać szeregu czasowego często zakłada się, że rozkład normalny w przybliżeniu dobrze pasuje do tych danych. Na tym założeniu oparte są m.in. modele ekonometrii finansowej, np. Blacka-Scholesa (Black, Scholes, 1973) służący do wyceny opcji, model Markowitza budowy portfela, model CAPM (*Capital Asset Pricing Model*) zaproponowany przez Sharpe'a (1964) czy też metody analizy technicznej. Jednak w późniejszych pracach, m.in. Mandelbrot (1963) oraz Fama (1965), wykazali, że założenie o normalności rozkładu ekonomicznych szeregów czasowych jest często nieprawdziwe.

Rozwój techniki komputerowej pozwalający na szybką analizę danych przyczynił się do wykazania, że dane mające postać szeregu czasowego odznaczają się większą kurtozą niż to jest w przypadku rozkładu normalnego, a ponadto są

asymetryczne. Wykazano też, że prawdopodobieństwo występowania wartości ekstremalnych jest dużo większe niż w rozkładzie gaussowskim (tzw. „grube ogony”). Odrzucenie normalności rozkładu wielu finansowych szeregów czasowych skłoniło badaczy do poszukiwania nowych rozkładów teoretycznych, które dobrze opisywałyby zachowanie się danych pochodzących z rynków finansowych.

Pomimo dość obszernej literatury dotyczącej tego tematu (zwłaszcza obcojęzycznej), badania opisywane w pracach naukowych zazwyczaj dotyczyły głównie indeksów lub spółek notowanych przede wszystkim na różnych giełdach światowych. Ponadto w większości przypadków autorzy poprzestawali na stwierdzeniu braku normalności tychże rozkładów lub na zbadaniu ich zgodności tylko z pewną grupą wspomnianych rozkładów.

W Polsce nie zostały dotąd przeprowadzone badania własności statystycznych wielkości wypłat z bankomatów. Autorom nie są również znane prace na ten temat w literaturze zagranicznej.

STRUKTURA DANYCH I ICH WŁASNOŚCI STATYSTYCZNE

W badaniach wykorzystano dane z 293 bankomatów firmy Euronet. Jest to największy spośród 28 działających operatorów sieci bankomatów w Polsce. Firma obsługuje prawie 25% bankomatów w Polsce, a w jej posiadaniu jest ok. 3 tys. bankomatów.

Do analizy uzyskano dane w postaci sumy dziennych wypłat ze wszystkich bankomatów znajdujących się w województwach małopolskim i podkarpackim. W przypadku bankomatów, które funkcjonowały przed 2008 r. wykorzystano dane od 1 stycznia 2008 r. Z kolei dla pozostałych bankomatów uwzględniono w badaniu dane od początku funkcjonowania bankomatu. Ostatnia obserwacja dotycząca wszystkich bankomatów pochodzi z 30 listopada 2011 r. Długość analizowanych szeregów czasowych kształtuje się od 280 do 1429, w zależności od tego, jak długo dany bankomat funkcjonował.

Analizowane dane są dobowymi sumami wypłat z danej maszyny, przy czym usunięto dane o wartości 0, czyli takie, które oznaczają, że bankomat był niedostępny (zepsuty lub w miejscu niedostępnym w danym dniu, np. nieczynny hipermarket).

Każdy bankomat opisano trzema cechami: województwo, liczba mieszkańców miejscowości (kategoria miejscowości) oraz typ lokalizacji. Jako typ lokalizacji Euronet wprowadza 15 kategorii: lotnisko, oddział bankowy, dworzec autobusowy, centrum rozrywki, zakład przemysłowy, hipermarket, hotel, stacja benzynowa, dworzec kolejowy, osiedle, restauracja, sklep, centrum handlowe, węzeł komunikacyjny, pozostałe. Miało też miejsce łączenie grup lokalizacyjnych w tych przypadkach, gdy liczebność bankomatów była zbyt mała, aby miało sens porównywanie z innymi lokalizacjami. Bankomaty, które znajdują się na lotnisku, dworcach kolejowym oraz autobusowym, a także w węźle komunika-

cyjnym zgrupowano w kategorię „transport osób”. Niska liczba bankomatów zlokalizowanych w zakładach przemysłowych oraz na osiedlach była przyczyną zaliczenia tych bankomatów do kategorii „pozostałe”. Dodatkowo każda z miejscowości, w której znajdują się bankomaty, została przydzielona do jednej z kategorii w zależności od liczby mieszkańców miejscowości, w której bankomat się znajdował:

- 0— 20 tys. — wsie oraz małe miasteczka — kategoria I,
- 21— 50 tys. — miasta średnie — kategoria II,
- 51—100 tys. — miasta duże — kategoria III,
- 101—200 tys. — miasta bardzo duże (Rzeszów, Tarnów) — kategoria IV,
- powyżej 200 tys. — metropolia Kraków — kategoria V.

Analizowane cechy pozwalają na zweryfikowanie hipotez związanych z tym, że wielkości wypłat z bankomatów mogą zależeć od ogólnie pojętego umiejscowienia bankomatu (województwo, miejscowość, lokalizacja).

TABL. 1. LICZBA BADANYCH BANKOMATÓW WEDŁUG LOKALIZACJI ORAZ WOJEWÓDZTW (stan na 1 marca 2011 r.)

Typ lokalizacji	Małopolskie		Podkarpackie		Razem	
	liczba bankomatów	w % ogólnej liczby bankomatów w województwie	liczba bankomatów	w % ogólnej liczby bankomatów w województwie	liczba bankomatów	w % ogólnej liczby bankomatów badanych województw
R a z e m	232	79,18	61	20,82	293	100,00
Oddział bankowy	57	24,57	27	44,26	84	28,67
Centrum rozrywki	6	2,59	—	—	6	2,05
Hipermarket	32	13,79	8	13,11	40	13,65
Hotel	6	2,59	—	—	6	2,05
Stacja paliw	23	9,91	2	3,28	25	8,53
Sklep	41	17,67	8	13,11	49	16,72
Centrum handlowe	30	12,93	12	19,67	42	14,33
Transport osób	4	1,72	—	—	4	1,37
Pozostałe	33	14,22	4	6,56	37	12,63

Źródło: opracowanie własne.

Z tabl. 1 wynika, że 232 bankomaty zlokalizowano w Małopolsce, natomiast 61 w woj. podkarpackim. Rozkład ze względu na typ lokalizacji również nie jest równomierny. Ponad 28% bankomatów było w oddziałach bankowych, przy czym w badanych województwach występuje zróżnicowanie w tym rozkładzie. W woj. podkarpackim ponad 44% analizowanych bankomatów było usytuowanych w oddziale bankowym, a w woj. małopolskim prawie co czwarty bankomat znajdował się w banku. Na kolejnych miejscach — jeśli chodzi o liczbę bankomatów — uplasowały się sklepy, centra handlowe, hipermarkety. Jednak w przypadku tych lokalizacji nie ma aż tak dużych dysproporcji między województwami.

Podobnego podziału dokonano ze względu na liczbę mieszkańców (tabl. 2).

**TABL. 2. LICZBA BADANYCH BANKOMATÓW WEDŁUG LOKALIZACJI
ORAZ KATEGORII MIEJSCOWOŚCI (stan na 1 marca 2011 r.)**

Typ lokalizacji	Kategoria miejscowości				
	I	II	III	IV	V
Liczba bankomatów					
R a z e m	15	47	25	39	167
Oddział bankowy	2	17	13	14	38
Centrum rozrywki	1	2	1	—	2
Hipermarket	3	2	5	5	25
Hotel	—	1	—	—	5
Stacja paliw	2	4	1	1	17
Sklep	5	10	—	4	30
Centrum handlowe	1	6	5	10	20
Transport osób	—	1	—	1	2
Pozostałe	1	4	—	4	28
W % ogólnej liczby bankomatów w miejscowości					
R a z e m	5,12	16,04	8,53	13,31	57,00
Oddział bankowy	13,33	36,17	52,00	35,90	22,75
Centrum rozrywki	6,67	4,26	4,00	—	1,20
Hipermarket	20,00	4,26	20,00	12,82	14,97
Hotel	—	2,13	—	—	2,99
Stacja paliw	13,33	8,51	4,00	2,56	10,18
Sklep	33,33	21,28	—	10,26	17,96
Centrum handlowe	6,67	12,77	20,00	25,64	11,98
Transport osób	—	2,13	—	2,56	1,20
Pozostałe	6,67	8,51	—	10,26	16,77

Ź r ó d ł o: jak przy tabl. 1.

Można zauważyć, że większość badanych bankomatów znajdowała się w Krakowie (kategoria V), natomiast najmniejszy odsetek bankomatów odnotowano w miejscowościach o liczbie mieszkańców poniżej 20 tys. Również podział bankomatów w poszczególnych grupach kategorii miejscowości był różny. W najmniejszych miejscowościach przeważały bankomaty w sklepach, natomiast w pozostałych miejscowościach najwięcej bankomatów znajdowało się w oddziałach bankowych.

W celu lepszego poznania struktury danych przeanalizowano podstawową statystykę dotyczącą wszystkich bankomatów łącznie, a następnie podobnej analizy dokonano dla bankomatów należących do tych samych grup, porównując jednocześnie wyniki w poszczególnych grupach.

**ZESTAWIENIE (1) STATYSTYKI ŚREDNICH DZIENNYCH WYPŁAT
ZE WSZYSTKICH BANKOMATÓW**

Średnia	Odchylenie standardowe	Współczynnik zmienności w %	Minimum	Maksimum
46857,7	27673,9	59,06	2755	161600

Ź r ó d ł o: jak przy tabl. 1.

W zestawieniu (1) przedstawiono podstawową statystykę dotyczącą średnich dziennych wypłat ze wszystkich bankomatów. Szczególną uwagę zwraca umiarkowana zmienność średnich. Współczynnik zmienności na poziomie ok. 60% świadczy o tym, że zróżnicowanie średnich wypłat z bankomatów jest dosyć duże.

W zestawieniu (2) przedstawiono analogiczną analizę współczynników zmienności ze wszystkich bankomatów wyrażonych w ułamkach dziesiętnych.

**ZESTAWIENIE (2) STATYSTYKI WSPÓŁCZYNNIKA ZMIENNOŚCI WYPŁAT
ZE WSZYSTKICH BANKOMATÓW**

Średnia w %	Odchylenie standardowe	Współczynnik zmienności	Minimum	Maksimum
		w %		
201,91	0,471456	23,35	76	335

Źródło: jak przy tabl. 1.

Wyniki tej analizy wskazują na dużą zmienność w strukturach danych poszczególnych bankomatów. Średni współczynnik zmienności kształtował się na poziomie 200%. Jednak jego odchylenie standardowe jest nieduże, co świadczy o tym, że generalnie zmienność w przypadku wszystkich bankomatów jest zbliżona.

W analizie skośności poszczególnych bankomatów istotna jest liczba bankomatów o asymetrii lewo- i prawostronnej. W przypadku kurtozy podano liczbę bankomatów, dla których kurtoza jest wyższa (niższa) od tej dla rozkładu normalnego.

**ZESTAWIENIE (3) WARTOŚCI WSPÓŁCZYNNIKÓW ASYMETRII
ORAZ KURTOZY WYPŁAT ZE WSZYSTKICH BANKOMATÓW**

Asymetria		Kurtoza	
lewostronna	prawostronna	ujemna	dodatnia
31 (10,6%)	262 (89,4%)	112 (38,2%)	181 (61,8%)

Źródło: jak przy tabl. 1.

W przypadku większości bankomatów uzyskane wartości współczynnika asymetrii wskazują na umiarkowaną lub dużą siłę asymetrii w badanych szeregach, co stanowi podstawy do testowania zgodności z rozkładami asymetrycznymi. Również wartości kurtozy prawie wszystkich bankomatów istotnie odbiegają od wartości charakterystycznej dla rozkładu normalnego, co również skłania do zastosowania rozkładów teoretycznych, których funkcja gęstości ma inny kształt niż w przypadku rozkładu gaussowskiego. Wyniki analizy przedstawiono w zestawieniu (3) wskazując, że znaczna większość danych bankomatowych charakteryzuje się prawostronną asymetrią. Większa część bankomatów charakteryzuje się wyższą kurtozą niż w przypadku danych z rozkładu normalnego.

W kolejności przedstawimy analogiczne analizy z uwzględnieniem odpowiednich grup związanych z lokalizacją.

W tabl. 3 zamieszczono podstawowe dane dotyczące przeciętnych dziennych wypłat. Wyniki wskazują, że średnia przeciętnych dziennych wypłat z bankomatów w woj. małopolskim była znacznie wyższa (o ponad 12 tys. zł) niż na Podkarpaciu. Jednocześnie zmienność tych średnich jest praktycznie identyczna. Analiza wariancji ($p\text{-value}=0,005$ dla testu F) wykazała istotność w zróżnicowaniu średniej z przeciętnych wypłat w tych województwach.

**TABL. 3. STATYSTYKA ŚREDNICH WYPŁAT Z BANKOMATÓW
WEDŁUG WOJEWÓDZTW**

Województwa	Średnia	Odchylenie standardowe	Współczynnik zmienności w %	Minimum	Maksimum
Małopolskie	49370,9	28615,3	57,96	2755	161600
Podkarpackie	37299,1	21384,0	57,33	6083	104000

Źródło: jak przy tabl. 1.

Informacje zawarte w tabl. 4 wskazują, że rodzaj miejscowości, w jakiej znajduje się bankomat, określony przez liczbę mieszkańców, ma związek ze średnią wielkością wypłat z bankomatów ($p\text{-value}=0$ dla testu F). Nie jest to jednak zależność proporcjonalna. Można bowiem zauważyć, że najwyższą średnią wypłat była w najmniejszych miejscowościach. Nieco mniejszą średnią uzyskano w Krakowie, a najmniejszą w miejscowościach o liczbie mieszkańców od 50 tys. do 100 tys. Aby wyniki te można było poprawnie zinterpretować, należałoby zbadać dodatkowo liczbę bankomatów przypadających na mieszkańca w poszczególnych miejscowościach. Niestety nie jest to możliwe z powodu zbyt dużej zmienności liczby bankomatów na danym terenie.

**TABL. 4. STATYSTYKA ŚREDNICH WYPŁAT Z BANKOMATÓW
WEDŁUG KATEGORII MIEJSCOWOŚCI**

Kategoria miejscowości	Średnia	Odchylenie standardowe	Współczynnik zmienności w %	Minimum	Maksimum
I	56480,0	19174,5	33,95	15260	87390
II	39671,3	22249,0	56,08	9623	124600
III	34225,6	17726,2	51,79	7011	89800
IV	42640,6	22711,6	53,26	6083	104000
V	50891,7	30787,8	60,50	2755	161600

Źródło: jak przy tabl. 1.

Na podstawie danych zawartych w tabl. 8 można stwierdzić, że najczęściej wypłacono z bankomatów znajdujących się w hipermarketach, centrach handlowych oraz na stacjach paliw. Przeciętnie najmniejsze wypłaty odnotowano w hotelach oraz oddziałach bankowych. Również w tym przypadku wpływ lokalizacji na wielkość dziennych wypłat z bankomatów jest istotny ($p\text{-value}=0,0023$ dla testu F).

**TABL. 5. STATYSTYKA ŚREDNICH WYPŁAT Z BANKOMATÓW
WEDŁUG TYPU LOKALIZACJI**

Typ lokalizacji	Średnia	Odchylenie standardowe	Współczynnik zmienności w %	Minimum	Maksimum
Oddział bankowy	30085,8	13460,7	44,74	2755	62100
Centrum rozrywki	47528,3	26451,9	55,66	10550	74730
Hipermarket	69484,8	32145,0	46,26	13570	155900
Hotel	26792,5	25503,8	95,19	5855	71060
Stacja paliw	60170,4	33343,4	55,41	18940	161600
Sklep	44952,4	19649,8	43,71	19310	100300
Centrum handlowe	61748,6	32066,9	51,93	7011	142100
Transport osób	38072,5	18725,0	49,18	17080	54790
Pozostałe	41192,2	17611,9	42,76	14990	90970

Źródło: jak przy tabl. 1.

Podsumowując tę część analizy można wskazać, że średnia wielkość wypłat jest istotnie uzależniona od tego, w którym województwie, w jakiej miejscowości oraz w jakiej lokalizacji znajduje się bankomat.

Przeprowadzono analogiczną analizę w zakresie współczynników zmienności, skośności oraz asymetrii. Wyniki wskazują, że również one przyjmują zróżnicowane wartości w zależności od zlokalizowania bankomatu.

Uzyskane wyniki mogą stać się podstawą do postawienia hipotezy, że własności statystyczne sumy dziennych wypłat z bankomatów są uzależnione od lokalizacji bankomatu. Istnieją powody, aby przypuszczać, że rodzaj rozkładu, który można dopasować do danych, również może być uzależniony od tej lokalizacji.

PRZEGLĄD ROZKŁADÓW ORAZ METODOLOGIA BADAŃ

W tej części artykułu szcharakteryzujemy rozkłady teoretyczne, które zostały użyte w celu dopasowania do danych. W szczególności przedstawimy funkcje gęstości rozkładów wraz z opisem ich parametrów.

Rozkład logistyczny. Ten rozkład teoretyczny zaproponował w 1981 r. Smith (1981). Testowali go też m.in. Gray i French (1990), Peiró (1994), Aparicio i Estrada (2001) oraz Suder i Wójtowicz (2005). Jest to rozkład podobny do rozkładu normalnego, a główną różnicę stanowią „grube ogony”. Funkcja gęstości rozkładu logistycznego dana jest wzorem:

$$f(x) = \frac{\exp\left(-\frac{x-\mu}{\alpha}\right)}{\alpha \cdot \left[1 + \exp\left(-\frac{x-\mu}{\alpha}\right)\right]^2}$$

gdzie:

μ — średnia,

α — parametr skali,

przy czym dla wariancji prawdziwa jest formuła $\sigma^2 = \frac{\pi^2 \alpha^2}{3}$.

Rozkład skalowany t -Studenta. W pracach (Praetz, 1972; Blattberg, Gonedes, 1974; Peiró, 1994; Aparicio, Estrada, 2001; Broca, 2002; Suder, Wójtowicz, 2005) pokazano, że ten rozkład dobrze aproksymuje np. rozkłady empiryczne dziennych stóp zwrotu:

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sigma \cdot \Gamma\left(\frac{v}{2}\right) \sqrt{\pi v}} \cdot \left[1 + \frac{\left(\frac{x-\mu}{\sigma}\right)^2}{v} \right]^{-\frac{v+1}{2}}$$

gdzie:

$\Gamma(\cdot)$ — funkcja gammy Eulera,

$\mu \in (-\infty, \infty)$ — parametr położenia,

$\sigma^2 \neq 0$ — parametr rozproszenia,

$v > 0$ — parametr odpowiedzialny za kształt.

Jeśli zmienna losowa X ma rozkład skalowany t -Studenta o parametrach μ , σ oraz v , wówczas zmienna $\frac{x-\mu}{\sigma}$ ma rozkład t -Studenta z liczbą stopni swobody wynoszącą v .

Rozkład potęgowo-wykładniczy. Z uwagi na „grube ogony” i występujące w skalowanym rozkładzie t -Studenta wysokie prawdopodobieństwo małych zmian, Hsu (1982), Gray i French (1990) oraz Broca (2002) zaproponowali stosowanie w badaniu rozkładów stóp zwrotu rozkładu potęgowo-wykładniczego. Funkcja gęstości rozkładu potęgowo-wykładniczego rzędu β dana jest wzorem:

$$f(x) = \frac{1}{2\beta^{\frac{1}{\beta}} \cdot \sigma \cdot \Gamma\left(1 + \frac{1}{\beta}\right)} \exp\left(-\left|\frac{x-\mu}{\beta\sigma^\beta}\right|^\beta\right)$$

gdzie $\mu \in (-\infty, \infty)$, $\sigma > 0$ oraz $\beta \in (0, \infty)$ są parametrami odpowiednio: położenia, rozrzutu i kształtu.

Jeśli $\beta < 2$, to rozkład posiada „grube ogony”, gdy $\beta = 2$, to rozkład jest normalny, a dla $\beta > 2$ mamy do czynienia z rozkładem o „wąskich ogonach”. W przypadku gdy $\beta = 1$ otrzymujemy rozkład Laplace’a, natomiast przy $\beta \rightarrow \infty$ otrzymujemy rozkład jednostajny.

Mieszanka dwóch rozkładów normalnych. Innym, używanym w literaturze rozkładem teoretycznym (Press, 1967; Kon, 1984; Peiró, 1994; Aparicio, Estrada, 2001) jest mieszanka rozkładów normalnych. Funkcja gęstości mieszanki dwóch rozkładów normalnych ma postać:

$$f(x) = \lambda \cdot \frac{1}{\sigma_1^2 \sqrt{2\pi}} \exp\left(\frac{-(x-\mu_1)}{2\sigma_1^2}\right) + (1-\lambda) \frac{1}{\sigma_2^2 \sqrt{2\pi}} \exp\left(\frac{-(x-\mu_2)}{2\sigma_2^2}\right)$$

Parametry $-\infty < \mu_i < \infty$ oraz $\sigma_i > 0$ (dla $i=1, 2$) opisują średnią i odchylenie standardowe każdego z dwóch rozkładów mieszanki. Parametry λ oraz $1-\lambda$ są wagami odpowiednich rozkładów w mieszance ($\lambda \in (0,1)$).

Mieszanka trzech rozkładów normalnych. Jest to rozkład znacznie rzadziej stosowany w badaniach niż mieszanka dwóch rozkładów normalnych. Jednak zastosowanie mieszanki więcej niż dwóch rozkładów można znaleźć m.in. w pracach: Baringhaus i Henze (1988), Henze (2002), Rencher (1995). Funkcja gęstości mieszanki trzech rozkładów normalnych ma postać:

$$f(x) = \lambda_1 \cdot \frac{1}{\sigma_1^2 \sqrt{2\pi}} \exp\left(\frac{-(x-\mu_1)}{2\sigma_1^2}\right) + \lambda_2 \frac{1}{\sigma_2^2 \sqrt{2\pi}} \exp\left(\frac{-(x-\mu_2)}{2\sigma_2^2}\right) + (1-\lambda_1-\lambda_2) \frac{1}{\sigma_3^2 \sqrt{2\pi}} \exp\left(\frac{-(x-\mu_3)}{2\sigma_3^2}\right)$$

Parametry $-\infty < \mu_i < \infty$ oraz $\sigma_i > 0$ (dla $i=1, 2, 3$) opisują średnią i odchylenie standardowe każdego z trzech rozkładów mieszanki. Parametry λ_1, λ_2 oraz $1-\lambda_1-\lambda_2$ są wagami odpowiednich rozkładów w mieszance ($\lambda_1, \lambda_2 \in (0,1)$, $\lambda_1 + \lambda_2 < 1$).

Rozkład hiperboliczny. Funkcja gęstości tego rozkładu dana jest wzorem:

$$f(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha \cdot \delta \cdot K_1(\delta \sqrt{\alpha^2 - \beta^2})} \exp\left\{-\alpha \sqrt{\delta^2 + (x-\mu)^2} + \beta(x-\mu)\right\}$$

gdzie:

$\alpha(\alpha > 0)$ oraz $\beta(0 \leq |\beta| < \alpha)$ — parametry kształtu,

$\delta(\delta > 0)$ — parametry skali,

$\mu(\mu \in R)$ — odpowiada za położenie,

$K_1(\cdot)$ — zmodyfikowana funkcja Bessela trzeciego rodzaju z indeksem 1.

Rozkład hiperboliczny zawiera jako przypadki graniczne m.in. rozkłady gaussowski (gdy $\alpha \rightarrow \infty, \delta \rightarrow \infty$) oraz Laplace'a (dla $\delta = 0$).

Rozkład normalny odwrotny gaussowski (NIG). Funkcja gęstości tego rozkładu dana jest wzorem:

$$f(x) = \frac{\alpha \cdot \delta}{\pi} \exp\left\{\delta \sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right\} \frac{K_1\left(\alpha \cdot \sqrt{\delta^2 + (x - \mu)^2}\right)}{\sqrt{\delta^2 + (x - \mu)^2}}$$

Parametry tego rozkładu mają analogiczną interpretację, jak parametry rozkładu hiperbolicznego. Rozkład hiperboliczny oraz rozkład normalny odwrotny gaussowski należą do grupy tzw. uogólnionych rozkładów hiperbolicznych. Ich zastosowanie pokazują m.in. Eberlein i Keller (1995), Küchler i in. (1999) oraz Suder i Wójtowicz (2005).

Rozkład α -stabilny. W przeciwieństwie do omówionych rozkładów jest on zdefiniowany za pomocą funkcji charakterystycznej danej wzorem:

$$\Phi(t) = \begin{cases} \exp\left\{-\gamma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sgn}(t) \operatorname{tg} \frac{\pi\alpha}{2}\right) + i\delta t\right\} & \alpha \neq 1 \\ \exp\left\{-\gamma |t| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \ln|t|\right) + i\delta t\right\} & \alpha = 1 \end{cases}$$

Rozkład ma cztery parametry, które są związane z czterema pierwszymi momentami rozkładu. Parametr α ($0 < \alpha \leq 2$) jest nazywany indeksem stabilności rozkładu i określa „grubość ogona”. Jeśli $\alpha = 2$, to zmienna ma rozkład normalny, gdy $\alpha < 2$, to rozkład ma „ogony grubsze” niż występujące w rozkładzie normalnym. $\beta(0 \leq |\beta| \leq 1)$ jest parametrem skośności, natomiast $\delta(-\infty < \delta < \infty)$ i $\gamma(\gamma > 0)$ są parametrami odpowiednio położenia i skali. Dla $\alpha = 2$ i $\beta = 0$ rozkład α -stabilny staje się rozkładem normalnym, dla $\alpha = 1$ i $\beta = 0$ jest rozkładem Cauchy'ego, natomiast przy $\alpha = \frac{1}{2}$ i $\beta = 0$ równy jest rozkładowi Levy'ego.

W celu zbadania zgodności rozkładu empirycznego sumy dziennych wypłat z poszczególnych bankomatów z analizowanymi rozkładami teoretycznymi za-

stosowano najbardziej popularne testy zgodności, tj. chi-kwadrat, Kołmogorowa-Smirnowa oraz Andersona-Darlinga.

Dla testu chi-kwadrat dane zostały podzielone na k grup. Wykorzystano statystykę:

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

gdzie:

O_i — częstość grupy o indeksie i ,

E_i — oczekiwana wartość teoretyczna częstości obliczana na podstawie dystrybuanty empirycznej.

Tak wyznaczona statystyka ma rozkład χ^2 z $k-p-1$ stopniami swobody (p określa liczbę estymowanych parametrów rozkładu). Test chi-kwadrat jest wrażliwy na sposób grupowania danych przy obliczaniu statystyki.

Wartość statystyki testowej testu chi-kwadrat jest uzależniona od liczności próby. Identyczne proporcje w tabeli w przypadku większej próby skutkują wyższą wartością statystyki testowej. Dlatego test ten nie może być wykorzystywany do porównania siły związku między zmiennymi w próbach różniących się licznością. Zależność wartości krytycznych testu od liczby stopni swobody nie pozwala na porównanie wyników uzyskanych w tabelach różniących się liczbą kolumn lub wierszy. Aby takie porównania były zasadne, trzeba przeliczyć wartość chi-kwadrat na odpowiedni do rodzaju danych współczynnik siły związku. Często jest to współczynnik V -Cramera. Wielkości te są miarami korelacji między zmiennymi nominalnymi i pozwalają wyrobić sobie zdanie na temat możliwości prognozowania jednej zmiennej ze względu na drugą zmienną. Im są one bliższe zera, tym możliwość ta jest mniejsza.

Test Kołmogorowa-Smirnowa sprawdza, czy rozkład analizowanych danych dla pewnej zmiennej losowej różni się od rozkładu teoretycznego przy danej skończonej liczbie obserwacji tej zmiennej. Test wykorzystuje statystykę postaci:

$$D_n = \sup_{1 \leq i \leq n} |F(x_i) - F_n(x_i)|$$

gdzie:

F — zadana dystrybuanta teoretyczna,

F_n — dystrybuanta empiryczna,

x_i — znane wartości zmiennej losowej.

Test Kołmogorowa-Smirnowa jest dobry do oceny zgodności rozkładu w otoczeniu średniej.

Jednym z głównych symptomów braku normalności danych jest występowanie „grubych ogonów”. Do badania zgodności rozkładu wykorzystuje się tu statystykę Andersona-Darlinga, która jest zdefiniowana następująco:

$$A^2 = -n - S$$

gdzie:

$$S = \sum_{i=1}^n \frac{2i-1}{n} [\ln F(x_i) + \ln(1 - F(x_{n+1-i}))],$$

F — zadana dystrybucja teoretyczna,

n — liczność próbek.

Statystyka Andersona-Darlinga jest pewną modyfikacją testu Craméra-von Misesa pozwalającą na poprawę jego czułości w „ogonach” testowanego rozkładu. Ze względu na wysoką czułość w „ogonach”, czyli dla wartości ekstremalnych, w wielu pracach, także w naszej, statystyka ta jest rozstrzygającym kryterium w dopasowywaniu rozkładów.

Warto zwrócić uwagę na niedokładności związane z zastosowaniem opisanych testów. Najważniejszą z nich jest fakt, że testuje się zgodność z rozkładami, których parametry zostały wcześniej wyestymowane metodą największej wiarygodności. Pociąga to za sobą konieczność ostrożnej interpretacji uzyskanych wyników, w szczególności w przypadku, gdy hipoteza o zgodności z badanym rozkładem zostaje odrzucona.

WYNIKI DOPASOWANIA ROZKŁADÓW

Zwykle w statystyce odrzucenie hipotezy zerowej przez jeden z testów jest powodem ostatecznego odrzucenia tej hipotezy. Zdając sobie sprawę z różnej wrażliwości wymienionych trzech testów autorzy oszacowali „stopień niedopasowania” rozkładów, biorąc pod uwagę liczbę testów odrzucających hipotezę zerową.

Współczynniki rozważanych rozkładów wyestymowano metodą największej wiarygodności, z użyciem oprogramowania komputerowego R oraz Matlab, dla sumy dziennych wypłat każdego z bankomatów. Następnie, stosując testy: chi-kwadrat, Kołmogorowa-Smirnowa i Andersona-Darlinga, zweryfikowano hipotezy o dopasowaniu danych do wybranych rozkładów teoretycznych. W przeprowadzanych testach poziom istotności wynosił 0,05.

W dalszej części pracy przedstawiono wyniki analizy związanej z dopasowaniem każdego z rozkładów całej grupy bankomatów. Zgodność dopasowania ocenimy w skali 0—3, przy czym wartość ta określa, w przypadku ilu testów

hipoteza zerowa nie została odrzucona, np. 0 oznacza, że wszystkie testy odrzuciły hipotezę o zgodności danych z rozkładem teoretycznym, natomiast 3 oznacza, że żaden z testów nie odrzucił badanego rozkładu.

TABL. 6. SIŁA DOPASOWANIA POSZCZEGÓLNYCH ROZKŁADÓW TEORETYCZNYCH DO DANYCH EMPIRYCZNYCH

Typ rozkładu	0		1		2		3	
	liczba	w %	liczba	w %	liczba	w %	liczba	w %
Normalny	180	61,43	28	9,56	71	24,23	14	4,78
Logistyczny	213	72,70	37	12,63	38	12,97	5	1,71
Skalowany <i>t</i> -Studenta	162	55,29	30	10,24	87	29,69	14	4,78
Potęgowo-wykładniczy	160	54,61	26	8,87	89	30,38	18	6,14
Mieszaną dwóch rozkładów normalnych ...	46	15,70	28	9,56	178	60,75	41	13,99
Mieszaną trzech rozkładów normalnych ...	1	0,34	1	0,34	190	64,85	101	34,47
NIG	120	40,96	20	6,83	110	37,54	43	14,68
Hiperboliczny	129	44,03	19	6,48	103	35,15	42	14,33
α -stabilny	204	69,62	11	3,75	60	20,48	18	6,14

Źródło: jak przy tabl. 1.

W tabl. 6 zamieszczono wyniki dopasowania wybranych rozkładów według wszystkich danych bez ich kategoryzowania. Łatwo zauważyć, że rozkład normalny został odrzucony przez wszystkie trzy testy dotyczące dziennych wypłat z ponad 60% badanych bankomatów. Jeśli chodzi o dopasowanie, to gorzej procentowo wypadają rozkłady logistyczny oraz α -stabilny. Również dane z ponad połowy bankomatów nie zostały dopasowane do rozkładów skalowanego *t*-Studenta oraz potęgowo-wykładniczego. O niewiele lepszym dopasowaniu można mówić w przypadku rozkładów hiperbolicznego oraz odwróconego gaussowskiego. Dla danych z ponad 40% bankomatów wszystkie testy odrzuciły zgodność z tymi rozkładami. Ewidentnymi liderami, jeśli chodzi o zgodność z analizowanymi danymi empirycznymi są mieszaniki rozkładów normalnych, zwłaszcza trzech rozkładów normalnych. Okazało się, że tylko w przypadku jednego bankomatu wszystkie testy odrzuciły mieszaną trzech rozkładów jako rozkład opisujący dzienne wypłaty z bankomatów. Dla 291 z 293 bankomatów przynajmniej dwa z trzech testów nie odrzuciły mieszaniki trzech rozkładów normalnych.

Wyniki wskazują więc na mieszaną trzech rozkładów normalnych jako ten rozkład, który powinien zostać wzięty pod uwagę w dalszych badaniach nad tworzeniem modelu prognostycznego.

Z tabl. 6 można odczytać, jak każdy z rozkładów dopasowuje się do analizowanych danych empirycznych. Jednak nie ma tu możliwości porównania dopasowania poszczególnych rozkładów do konkretnych bankomatów. W tym celu dla każdego bankomatu wybrano rozkład, który najlepiej charakteryzuje sumę

dziennych wypłat. Za najlepiej dopasowany rozkład uznano ten, który nie został odrzucony przez największą liczbę testów. W przypadku gdy dwa rozkłady miały identyczny wynik wybierano ten, dla którego statystyka Andersona-Darlinga była największa. Jak już podkreślono, test ten charakteryzuje się wysoką czułością, bada zgodność danych empirycznych z rozkładem teoretycznym w „ogonach” rozkładu. Znajomość rozkładu w „ogonach” jest ważna ze względu na wybór odpowiedniego modelu prognostycznego. Wyniki opisanej analizy zamieszczono w tabl. 7.

TABL. 7. PORÓWNANIE WYNIKÓW DOPASOWANIA POSZCZEGÓLNYCH ROZKŁADÓW

Typ rozkładu	Liczba	W %
Brak	1	0,34
Potęgowo-wykładniczy	1	0,34
Skalowany <i>t</i> -Studenta	1	0,34
Mieszanka dwóch rozkładów normalnych	23	7,85
Mieszanka trzech rozkładów normalnych	238	81,23
NIG	13	4,44
Hiperboliczny	16	5,46

Źródło: jak przy tabl. 1.

Wyniki potwierdzają wcześniejsze analizy, bowiem w przypadku aż 238 bankomatów, czyli w ponad 80%, najlepszym dopasowaniem charakteryzowała się mieszanka trzech rozkładów normalnych. W tym rankingu kolejne miejsca zajęły: mieszanka dwóch rozkładów normalnych (7,85%), rozkład hiperboliczny (5,46%) i odwrócony gaussowski (4,44%). Tylko dla jednego bankomatu żaden rozkład nie został dopasowany.

Wyniki potwierdzają, że mieszanka trzech rozkładów normalnych najlepiej opisuje strukturę dziennych wypłat z bankomatów. Z tabl. 7 można dodatkowo odczytać, że w przypadku prawie 90% bankomatów mieszanki rozkładów normalnych najlepiej aproksymują rozkłady empiryczne analizowanych danych.

W kolejnej części pracy zbadano, czy istnieje zależność pomiędzy dopasowaniem rozkładu do danych a miejscem, w jakim dany bankomat się znajduje. Aby weryfikacja miała sens, pod uwagę wzięto tylko mieszanki rozkładów normalnych oraz rozkłady NIG i hiperboliczny.

TABL. 8. WYNIKI PORÓWNANIA DOPASOWANIA ROZKŁADÓW WEDŁUG WOJEWÓDZTW

Województwa	Hiperboliczny		Mieszanka dwóch rozkładów normalnych		Mieszanka trzech rozkładów normalnych		NIG	
	liczba	w %	liczba	w %	liczba	w %	liczba	w %
Małopolskie	11	4,80	19	8,30	188	82,10	11	4,80
Podkarpackie	5	8,20	4	6,56	50	81,97	2	3,28

Źródło: jak przy tabl. 1.

Wyniki zamieszczone w tabl. 8 wskazują na brak zróżnicowania w rozkładzie zgodności poszczególnych rozkładów w badanych województwach. Potwierdza to wynik testu niezależności, w którym $p\text{-value}=0,6964$.

**TABL. 9. WYNIKI PORÓWNIANIA DOPASOWANIA ROZKŁADÓW
WEDŁUG KATEGORII MIEJSCOWOŚCI**

Kategoria miejscowości	Hiperboliczny		Mieszanka dwóch rozkładów normalnych		Mieszanka trzech rozkładów normalnych		NIG	
	liczba	w %	liczba	w %	liczba	w %	liczba	w %
I	—	—	—	—	14	93,33	1	6,67
II	4	8,70	1	2,17	41	89,13	—	—
III	3	12,00	1	4,00	20	80,00	1	4,00
IV	1	2,56	4	10,26	34	87,18	—	—
V	8	4,85	17	10,30	129	78,18	11	6,67

Źródło: jak przy tabl. 1.

Dopasowanie rozkładów do danych dotyczących bankomatów znajdujących się w miejscowościach o różnym zaludnieniu wykazało lekkie dysproporcje (tabl. 9). Można zauważyć, że dla małych miejscowości mieszanka trzech rozkładów normalnych dopasowuje się najlepiej w ponad 90% przypadków, jednocześnie dla Krakowa (kategoria V) udział ten był poniżej 80%. Również wyniki pozostałych rozkładów wskazują na niewielkie zróżnicowanie zgodności dopasowania w zależności od liczby mieszkańców. Jednak przeprowadzony test niezależności nie wskazuje na istotność tego zróżnicowania. Wartość $p\text{-value}$ w tym przypadku wynosi 0,1889.

**TABL. 10. WYNIKI PORÓWNIANIA DOPASOWANIA ROZKŁADÓW
WEDŁUG TYPÓW LOKALIZACJI**

Typ lokalizacji	Hiperboliczny		Mieszanka dwóch rozkładów normalnych		Mieszanka trzech rozkładów normalnych		NIG	
	liczba	w %	liczba	w %	liczba	w %	liczba	w %
Oddział bankowy	3	3,61	5	6,02	73	87,95	2	2,41
Centrum rozrywki	1	16,67	—	—	5	83,33	—	—
Hipermarket	3	7,89	6	15,79	27	71,05	2	5,26
Hotel	—	—	—	—	3	50,00	3	50,00
Stacja paliw	1	4,00	1	4,00	23	92,00	—	—
Sklep	1	2,04	4	8,16	43	87,76	1	2,04
Centrum handlowe	5	11,90	3	7,14	31	73,81	3	7,14
Pozostałe	1	2,70	3	8,11	33	89,19	—	—

Źródło: jak przy tabl. 1.

Wyniki zamieszczone w tabl. 10 wskazują, że rodzaj lokalizacji bankomatu najbardziej różnicuje typy rozkładów, które najlepiej opisują strukturę danych.

Według wszystkich typów lokalizacji mieszanka trzech rozkładów normalnych najtrafniej przybliża rozkład empiryczny danych, z wyjątkiem bankomatów w hotelach, gdzie równie dobre dopasowanie uzyskuje rozkład NIG. W przypadku bankomatów zlokalizowanych na stacjach paliw, w oddziałach bankowych i w sklepach mieszanka trzech rozkładów daje najlepsze dopasowanie w blisko 90% przypadków. Jednak w przypadku bankomatów w hipermarketach i centrach handlowych jest to niewiele ponad 70%. Jeśli chodzi o bankomaty zlokalizowane w hipermarketach, dobry wynik dopasowania uzyskuje również mieszanka dwóch rozkładów. W centrach handlowych dla niektórych bankomatów pozostałe rozkłady dają lepsze dopasowanie. Opisane zróżnicowanie zgodności rozkładów w zależności od typu lokalizacji okazuje się istotne statystycznie, gdyż wartość p -value dla testu niezależności wynosi 0.

Wnioski

W artykule autorzy podjęli próbę zbadania empirycznych własności dziennych wypłat z bankomatów ze szczególnym uwzględnieniem dopasowania rozkładów teoretycznych.

Uzyskane wyniki wskazują, że badane szeregi cechują się niejednorodnością. W znacznie większości przypadków rozkładem, który najlepiej aproksymuje dane empiryczne jest mieszanka trzech rozkładów normalnych.

Dodatkowo wykazano, że największy wpływ na rodzaj rozkładu ma typ lokalizacji, natomiast rodzaj dopasowania rozkładu nie jest powiązany z województwem oraz miejscowością, w których znajduje się bankomat.

Uzyskane wyniki mogą mieć duże znaczenie przy wyborze metody prognozowania oraz modelu prognostycznego. Dobre dopasowanie mieszanki trzech rozkładów normalnych może wskazywać na możliwość zastosowania modelu przełącznikowego do opisu podszeregów dających się wyróżnić z danego szeregu czasowego wypłat w bankomatach na podstawie odpowiednio zdefiniowanych kryteriów.

Chodzi o to, że współczynniki kombinacji wypukłej (proporcje mieszanki) trzech rozkładów normalnych zapewne zmieniają się na skutek m.in. tzw. efektu kalendarza, obserwowanego w przypadku wypłat z bankomatów. Uwzględnienie tego faktu może nastąpić właśnie poprzez zastosowanie modelu przełącznikowego.

W związku z zaobserwowanym zahamowaniem rozwoju sieci bankomatowych w krajach wysoko rozwiniętych jest rzeczą ważną z praktycznego punktu widzenia określenie perspektywy rozwoju sieci bankomatów w tych krajach, a także w Polsce. W szczególności chodzi o rozpoznanie, w jaki sposób wpłynie to na posługiwanie się gotówką, bowiem obecnie przewiduje się raczej spadek popytu na gotówkę.

Chociaż jest to prawdopodobny scenariusz, to jednak rosnące koszty transakcji mogą sprzyjać pobieraniu gotówki. Otwierają się zatem perspektywy badań w zakresie związków pomiędzy koncentracją w sektorze bankowym i podażą bankomatów z jednej strony oraz popytem na gotówkę i dostępnością usług gotówkowych w postaci głównie bankomatów, ale także alternatywnych sposobów płatności, z drugiej strony.

prof. dr hab. Henryk Gurgul, mgr Marcin Suder — AGH w Krakowie

LITERATURA

- Amromin E., Chakravorti S. (2007), *Debit card and cash usage: a cross-country analysis*, Technical report, Federal Reserve Bank of Chicago
- Aparicio F. M., Estrada J. (2001), *Empirical Distribution of Stock Returns: European Securities Markets 1990—1995*, „The European Journal of Finance”, No. 7
- Armenise R., Birtolo C., Sangianantoni E., Troiano L. (2010), *A generative solution for ATM cash management*, International Conference of Soft Computing and Pattern Recognition 7—10 December
- Baringhaus L., Henze N. (1988), *A consistent test for multivariate normality based on the empirical characteristic function*, „Metrika”, No. 35
- Black F., Scholes M. (1973), *The pricing of options and corporate liabilities*, „Journal of Political Economy”, No. 81
- Blattberg R., Gonedes N. (1974), *A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices*, „Journal of Business”, No. 47
- Boeschoten W. C. (1998), *Cash management, payment patterns and the demand for money*, „De Economist”, No. 146
- Broca D. S. (2002), *The distribution of Indian Stock Returns: A Tale of Two Tails*, „Decision”, No. 29
- Eberlein E., Keller U. (1995), *Hyperbolic distributions in finance*, „Bernoulli”, No. 1
- Fama E. F. (1965), *The behavior of stock market prices*, „Journal of Business”, No. 38
- Gray J. B., French D. W. (1990), *Empirical comparisons of distributional model for stock index prices*, „Journal of Business Finance and Accounting”, No. 17
- Hand D. J., Blunt G. (2001), *Prospecting for gems in credit card data*, „IMA Journal of Management Mathematics”, No. 12
- Henze N. (2002), *Invariant tests for multivariate normality: a critical review*, „Statistical papers”, No. 43
- Hsu D. (1982), *A Bayesian robust detection of shift in the risk structure of stock market returns*, „Journal of the American Statistical Association”, March
- Kon S. J. (1984), *Models of stock returns: a comparison*, „Journal of Finance”, No. 39
- Kühler U., Neumann K., Sørensen M., Streller A. (1999), *Stock returns and hyperbolic distributions*, „Mathematical and Computer Modelling”, No. 29
- Mandelbrot B. (1963), *The variation of certain speculative prices*, „Journal of Business”, No. 36

- Peiró A. (1994), *International evidence on the distribution of stock return*, „Applied Financial Economics”, No. 4
- Praetz P. D. (1972), *The distribution of share price changes*, „Journal of Business”, No. 45
- Press S. J. (1967), *A compound events model for security prices*, „Journal of Business”, No. 45
- Rencher A. C. (1995), *Methods of multivariate analysis*, Wiley, New York
- Sharpe W. F. (1964), *Capital asset price: a theory of market equilibrium under conditions of risk*, „Journal of Finance”, No. 19
- Simutis R., Dilijonas D., Bastina L., Friman J., Drobinov P. (2007), *Optimization of cash management for ATM network*, „Information Technology and Control”, No. 36
- Smith J. B. (1981), *The probability distribution of market returns: a logistic hypothesis*, PhD dissertation, University of Utah
- Snellman H., Viren M. (2006), *ATM networks and cash usage. Technical Report*, „Bank of Finland Research Discussion Papers”, No. 21
- Suder M., Wójtowicz T. (2005), *Rozkłady dziennych stop zwrotu wybranych indeksów GPW w Warszawie*, Zeszyty Naukowe Wyższej Szkoły Ekonomii i Informatyki w Krakowie, nr 3

SUMMARY

ATMs require a continuous connection with the performance management of payments, and distributing cash acceptance, attitude appropriate use of appropriate devices and software, as well as protection against fraud. This paper tries to fit the theoretical distributions to the empirical distribution of withdrawals by ATM installed in Małopolskie and Podkarpackie voivodships. Probability distribution may be useful as a signal to take a risk protection strategy or as a basis for filling ATMs with cash. Knowledge of the probability distribution is essential that the data from the past can be used to draw money management forecasts.

РЕЗЮМЕ

Банкоматы требуют непрерывного управления связанного с проведением выплат, приемом и распределением денег, обеспечением соответствующего оборудования и использованием программного обеспечения, а также защитой от мошенничества. Статья представляет методы приспособливания теоретических и эмпирических распределений выплат денег через банкоматы установленные в малопольском и подкарпатском воеводствах. Распределение вероятностей может быть полезным как сигнал для предпринимания защиты от рисков или же в качестве основы стратегии заполнения банкоматов наличными деньгами. Знание распределения вероятностей является необходимым для использования данных из прошлого в составлении прогнозов управления денежными средствами.