

Modelowanie zachowań konsumentów

Agnieszka Pleśniak

Niektóre cechy zachowań konsumentów na rynku odnoszące się do dokonywanych zakupów można przewidzieć na podstawie modeli teoretycznych. Zwykle różne marki produktu różnią się pod względem popularności. Marki popularne są częściej kupowane przez swoich nabywców, zaś nabywcy marek mniej popularnych w większym stopniu skłonni są kupować również inne marki. Tego typu prawidłowości opisują zjawisko podwójnego ryzyka i prawo współwystępowania zakupów. Zgodnie z teorią zakupów wielokrotnych Ehrenberga wiele aspektów zachowań nabywczych konsumentów może być przewidziane tylko na podstawie wskaźnika penetracji oraz wskaźnika intensywności zakupów, i nawet te zmienne są powiązane.

Teoria Ehrenberga przyczyniła się do powstania modeli, takich jak model $w(1-b)$, LSD czy NBD-Dirichleta, pozwalających przewidzieć zachowanie konsumentów na rynku na podstawie znajomości wskaźnika penetracji i wskaźnika intensywności zakupów.

Model NBD-Dirichleta opiera się na złożeniu czterech rozkładów. Ideą modelu jest dopasowanie wspomnianych rozkładów do danych empirycznych, dzięki czemu staje się możliwe uchwycenie w sposób ilościowy najważniejszych aspektów zachowania konsumentów (jak prawo podwójnego ryzyka czy prawo współwystępowania zakupu) oraz określenie (lub przewidzenie – w przypadku nowej marki) wielkości charakteryzujących zachowania konsumentów.

1. Wstęp

Na temat zachowań konsumentów powstało wiele teorii oraz modeli, pozwalających w różny sposób opisać, przewidzieć bądź wytłumaczyć zachowania konsumentów.

Ze względu na cel modele można podzielić na konceptualne i ilościowe. Pierwsza grupa stara się stworzyć model zachowań, który by w uproszczony sposób wyjaśniał te zachowania, uwzględniając motywy, wpływy otoczenia na podejmowane decyzje, postawy, cechy osobowości konsumentów itd. (więcej informacji na temat różnych modeli a także ich klasyfikacje można znaleźć w: Smyczek 2003a, 2003b, Rudnicki 2002, Jachnis, Terelak 1998).

Modele ilościowe, w odróżnieniu od modeli konceptualnych, nie starają się wyjaśnić zachowania konsumenta, rzadko biorą pod uwagę jego motywy, nie analizują czynników wpływających na decyzję itp. Modele te starają się zmierzyć zachowa-

nie konsumenta na rynku (tzn. przedstawić je w kategoriach ilościowych, np.: 10% konsumentów kupuje daną markę w analizowanym okresie) lub je przewidzieć.

Według Ehrenberga, wiele spośród bardzo dużej ilości zmiennych, które mogą wpływać na zachowanie, tak naprawdę nie ma znaczenia. Zgodnie z jego teorią zakupów wielokrotnych wiele aspektów zachowań nabywczych może być przewidziane tylko na podstawie wskaźnika penetracji i wskaźnika intensywności zakupów, i nawet te zmienne są powiązane (Ehrenberg 1972).

Teoria Ehrenberga przyczyniła się do powstania modeli, które pozwalają nie tylko przewidzieć pewne cechy zachowań konsumentów na rynku (przykładowo: ilu konsumentów będzie kupowało daną markę w określonym czasie czy jak dużo produktu kupują nabywcy danej marki), ale uchwycić w sposób ilościowy obserwowane regularności na rynku takie jak prawo współwystępowania zakupu czy zjawisko podwójnego ryzyka.

Zjawisko podwójnego ryzyka polega na tym, że marki o małym udziale w rynku nie dość, że są rzadziej kupowane przez swoich nabywców, to jeszcze ich nabywcy są mniej wobec nich lojalni. Zjawisko to po raz pierwszy zaobserwował w latach 60. William Mc Phee w odniesieniu do komiksów i prezenterów radiowych. Zauważył on, że nabywcy popularniejszych marek bardziej je lubią (postawy) i częściej kupują (zachowania). Wyjaśnienie Mc Phee w odniesieniu do postaw można w skrócie (w całości wyjaśnienie Mc Phee można znaleźć w: Ehrenberg, Goodhardt, Barwise 1990) przedstawić następująco:

Załóżmy, że w pewnym mieście są dwie restauracje, jedna bardzo znana i druga mało znana. Ponadto osoby, które znają obie restauracje, uważają ich poziom (pod każdym względem) za jednakowy. Jeśli zapytamy, którą restaurację lubią, pojawi się zjawisko podwójnego ryzyka.

Spośród wielu osób, które znają bardziej popularną restaurację, tylko niewielu zna tę drugą, zatem na pytanie, którą restaurację lubią, większość z nich odpowie, że tę dobrze znaną (tej drugiej nie znają). W dodatku „głosy” osób, które znają obie restauracje (i tak będące już w mniejszości) rozłożą się na dwie restauracje i ewentualnie na brak zdecydowania. Biorąc pod uwagę, że większość ludzi znających mniej popularną restaurację zna również tę znaną (ich „głosy” zostaną podzielone), pozostaje niewielu, którzy będą wybierać tylko tę mniej znaną.

Analogicznie można wytłumaczyć zjawisko podwójnego ryzyka w odniesieniu do zachowań, gdyż ludzie będą jadać w restauracji, którą znają (więcej na temat zjawiska podwójnego ryzyka oraz przykłady jego występowania można znaleźć m. in. w Ehrenberg, Goodhardt, Barwise 1990, Keng, Ehrenberg 1984, Lambkin 2001: 84-124).

Inną, zaobserwowaną na rynkach regularnością jest prawo współwystępowania zakupu (ang. *duplication law*), tłumaczone również jako prawo podwojenia zakupu (Lambkin 2001: 96). Mówi ono o tym, że konsumenci zmieniający markę wybierają nową markę z prawdopodobieństwem proporcjonalnym do udziału nowo wybranej marki w rynku (szerzej na temat prawa współwystępowania zakupu zob.: Ehrenberg 1972, Lambkin 2001).

Do modeli opartych na teorii Ehrenberga należą:

- model $w(1-b)$ zakładający, że wyrażenie $w(1-b)$ jest stałe dla poszczególnych marek (model ten został opisany w Ehrenberg 1972, Lambkin 2001),
- model LSD wykorzystujący rozkład szeregu logarytmicznego (z ang. *loga-*

rithmic series distribution, więcej na temat modelu zob. Ehrenberg 1972, ponadto bardzo ciekawe zastosowanie modelu LSD można znaleźć w: Geyer-Schulz, Hahsler, Jahn 2001),

- a także model NBD-Dirichleta omówiony poniżej.

2. Model Dirichleta

2. 1. Istota modelu

Model Dirichleta (a ściślej NBD-Dirichleta) pozwala na podstawie danych panelowych określić, np. jaki jest odsetek konsumentów wybierających daną markę, jak często ją oni wybierają (ile przeciętnie jej kupują w określonym czasie), jak dużo produktu kupują przeciętnie nabywcy poszczególnych marek, jaki jest odsetek konsumentów lojalnych wobec marki, ile przeciętnie kupują oni produktu, (czyli de facto marki) itd.

Model Dirichleta pozwala również przewidzieć występowanie takich prawidłowości na rynku jak prawo współwystępowania zakupu czy zjawisko podwójnego ryzyka dla różnych marek konkurujących na rynku. Nie ogranicza się przy tym do stwierdzenia, że np. zjawisko podwójnego ryzyka istnieje, tzn., że marki o mniejszym udziale w rynku będą rzadziej kupowane przez ich nabywców, ale mówi, o ile rzadziej są one kupowane (a więc ujmuje to zjawisko w sposób ilościowy).

Model opiera się na złożeniu czterech rozkładów (Poissona, gamma, wielomianowego i Dirichleta). Ideą modelu jest dopasowanie wspomnianych rozkładów do danych empirycznych, dzięki czemu staje się możliwe uchwycenie w sposób ilościowy najważniejszych aspektów zachowania konsumentów (jak prawo podwójnego ryzyka czy prawo współwystępowania zakupu) oraz określenie (lub przewidzenie – w przypadku nowej marki) wielkości charakteryzujących zachowania konsumentów.

Zgodnie z teorią Ehrenberga, wiele aspektów zachowań konsumentów może być przewidziane tylko na podstawie wskaźnika penetracji i wskaźnika intensywności zakupów (Ehrenberg 1972).

Wskaźnik penetracji, oznaczany jako B , jest to odsetek nabywców produktu, czyli osób, które kupiły przynajmniej raz produkt. Wskaźnik intensywności zakupów produktu – W – informuje natomiast o tym, ile produktu przeciętnie kupują jego nabywcy.

Łatwo zauważyć, że $M=BW$. Gdzie M jest wskaźnikiem stopy zakupów produktu i informuje o tym, ile kupionego produktu przypada na osobę. Ilość zakupionego produktu może być wyrażona w ujęciu ilościowym (np. 5 paczek), wartościowym (np. za 30zł), a także liczbą dokonanych zakupów w określonym czasie (np. 3 razy). W przypadku analizy danych panelowych wygodniej jest posługiwać się liczbą dokonanych w analizowanym okresie zakupów, akcentując raczej sam fakt zakupu niż ilości zakupionych towarów czy też ich wartość. Fakt dokonania zakupu produktu kodowany jest jako TAK, jeśli konsument zakupił jedną lub więcej jednostek produktu w sytuacji zakupowej. Ilość zakupionego jednorazowo produktu lub zapłacona kwota nie mają w tym ujęciu znaczenia.

W odniesieniu do marki oznaczenia są analogiczne: b – wskaźnik penetracji marki, w – wskaźnik intensywności zakupów marki, m – wskaźnik stopy zakupów marki.

Kupując produkt, nabywca dokonuje dwóch wyborów: kiedy dokona zakupu (lub zakupów) i jaką markę wybierze. Zgodnie z przyjętą konwencją, prawdopodobieństwa dotyczące zakupu produktu oznaczono wielkimi literami z subskryp-tem oznaczającym liczbę dokonanych zakupów w badanym okresie, zaś prawdopodobieństwa odnoszące się do wyboru marki oznaczone zostały małymi literami. Przykładowo: P_n – oznacza prawdopodobieństwo dokonania zakupu produktu n -razy, gdzie $n = 0, 1, \dots, v$, $p(r)$ – prawdopodobieństwo wyboru r -razy i -tej marki ($r = 0, 1, 2, \dots, v$; $i = 1, 2, \dots, g$).

Każdy spośród N konsumentów dokonuje n -razy ($n=0, 1, \dots, v$) zakupu produktu w badanym okresie. Konsument za każdym razem spośród marek dostępnych na rynku wybiera i -tą markę ($i=1, 2, \dots, g$), zgodnie z własnymi preferencjami. Preferencje konsumentów określają ich indywidualne wektory prawdopodobieństw wyboru marek ($\tilde{p} = [\tilde{p}_1, \tilde{p}_2, \dots, \tilde{p}_g]$). Przy kolejnym zakupie konsument może wybrać ponownie tę samą markę lub też inną, zakładamy jednak, że fakt wyboru konkretnej marki nie wpływa na wybory przy kolejnej okazji zakupowej. Ponadto indywidualny wektor prawdopodobieństw wyboru marek jest dla danego konsumenta stały w analizowanym czasie. W sekwencji n zakupów konsument wybiera r_i – razy ($r_i = 0, 1, 2, \dots, \rho$) markę i ($i = 1, 2, \dots, g$), czyli r_1 -razy markę 1, r_2 – razy markę 2, ... r_g -razy markę g , przy czym suma wszystkich r_i wynosi n .

Aby można było zastosować model Dirichleta, konsumenci muszą mieć okazję, żeby dokonać powtórnie zakupu, zatem rozważany okres powinien być na tyle długi, aby zdążyli oni zużyć produkt i dokonać kolejnych zakupów.

Zastosowanie modelu Dirichleta na rynkach dóbr trwałych jest więc ograniczone ze względu na zbyt długi okres użytkowania produktów. Warto dodać, że analizowany okres nie może być nadmiernie długi, gdyż w takim wypadku preferencje konsumentów wobec marek mogą ulegać zmianie, co wiąże się ze zmianą indywidualnych wektorów prawdopodobieństw wyboru marek. Również inne wielkości charakteryzujące dany rynek mogą ulegać zmianie.

2.2. Założenia

Założenia modelu dotyczące rynku: (por. Goodhardt, Ehrenberg, Chatfield 1984: 621, Ehrenberg 2001: 100):

- rynek jest stabilny, tzn. udziały poszczególnych marek lub inne zbiorcze miary (wskaźnik penetracji, wskaźnik intensywności zakupów) są stałe w analizowanym okresie czasu lub zmieniają się w niewielkim stopniu (dyskusję na temat stabilności można znaleźć w Ehrenberg 1972: 12);
- rynek nie jest podzielony na segmenty, jeśli nawet istnieje tendencja, by pewna marka (lub marki) była adresowana do konkretnej grupy odbiorców, to jest ona słaba.

Jak wspomniano, model Dirichleta opiera się na czterech rozkładach. Założenia dotyczące rozkładów są następujące (zob. Goodhardt, Ehrenberg, Chatfield 1984: 624-626, Ehrenberg 2001: 101, Lambkin 2001: 100):

Pierwsze dwa założenia dotyczą zakupów produktu:

1. Liczba zakupów dokonanych przez każdego konsumenta ma rozkład Poissona z wartością oczekiwaną μ_i dla zakupów l -tego konsumenta.

Warto dodać, że zakupy konsumenta są zgodne z rozkładem Poissona, jeśli są

one na tyle nieregularne, że mogą być traktowane jako losowe (co często jest obserwowane) i nie zależą od wcześniejszych zakupów. W związku z tym rozpatrywany okres czasu musi być odpowiednio długi, tak aby zakupy dokonane w jednym okresie nie wpływały na zakupy dokonywane w kolejnym (np. jeśli analizowany okres będzie zbyt krótki, żeby zużyć zakupiony produkt, to fakt zakupu wyrobu w jednym okresie będzie przesądzał o braku zakupu w kolejnym). Z drugiej strony analizowany okres czasu nie może być zbyt długi, tak aby założenie dotyczące stabilności rynku było spełnione.

2. Wartości oczekiwane μ_l różnią się dla poszczególnych konsumentów i mają rozkład gamma. Ze złożenia rozkładu Poissona i gamma otrzymamy rozkład identyczny z rozkładem dwumianowym ujemnym (NBD, stąd nazwa modelu NBD-Dirichleta). Zgodnie z rozkładem dwumianowym ujemnym prawdopodobieństwo dokonania zakupu produktu n -razy dane jest wzorem:

$$P_n = \left(1 + \frac{M}{K}\right)^{-K} \frac{\Gamma(K+n)}{n! \Gamma(K)} \left(\frac{M}{M+K}\right)^n, \quad (1)$$

gdzie: M - wskaźnik stopy zakupów produktu, K - parametr, Γ - funkcja gamma.

Kolejne dwa założenia odnoszą się do wyboru marki:

3. Liczba zakupów każdej marki dokonywanych przez l -tego konsumenta w sekwencji n zakupów produktu jest zgodna z rozkładem wielomianowym z parametrami n, p_1, p_2, \dots, p_g , dla l -tego konsumenta, gdzie n liczba dokonanych zakupów produktu, p_1, p_2, \dots, p_g - prawdopodobieństwa wyboru poszczególnych marek ($i=1, 2, \dots, g$).
4. Prawdopodobieństwa wyboru poszczególnych marek podczas robienia zakupów są różne dla różnych konsumentów (ale są stałe w analizowanym czasie) i wyraża je rodzaj wielowymiarowego rozkładu beta¹ - rozkład Dirichleta. Wartość oczekiwana prawdopodobieństwa wyboru marki i jest

zarazem udziałem marki i w rynku i wynosi: $\frac{\alpha_i}{S}$ (dla $i=1, 2, \dots, g$), gdzie:

α_i i $S = \sum_{i=1}^g \alpha_i$ są parametrami rozkładu Dirichleta.

Wygodnym uproszczeniem rozkładu Dirichleta jest potraktowanie rynku dychotomicznie (biorąc pod uwagę markę i oraz resztę rynku). Korzystając z własności addytywności rozkładu Dirichleta, pozostałe marki sumuje się w jedną zbiorczą

markę z udziałem w rynku $\frac{(S - \alpha_i)}{S}$. W ten sposób prawdopodobieństwo

dokonania r_i ($r_i=0, 1, \dots, p_i$) zakupów marki i pod warunkiem dokonania n zakupów produktu ($r_i \leq n$) można wyrazić rozkładem dwumianowym beta:

$$p(r_i | n) = \binom{n}{r_i} \frac{B(\alpha_i + r_i, S - \alpha_i + n - r_i)}{B(\alpha_i, S - \alpha_i)}, \quad (2)$$

gdzie: B jest funkcją beta.

5. Ostatnie założenie jest założeniem o niezależności, zgodnie z którym: prawdopodobieństwa wyboru marki i zakupu wyrobu rozkładają się w populacji niezależnie od siebie.

2.3. Zastosowania

Model Dirichleta dostarcza punktu odniesienia przy opracowaniu strategii marketingowej, pozwala bowiem określić, jakie cele są możliwe do zrealizowania, a jakie nie. Przykładowo wzrost sprzedaży można osiągnąć poprzez skłonienie istniejących nabywców marki, aby kupowali jej więcej, przez pozyskanie nowych nabywców lub przez kombinację tych elementów. Z modelu Dirichleta wynika, że samo skłonienie nabywców marki, aby kupowali więcej danej marki, może okazać się mało realne, zwłaszcza jeśli mamy do czynienia z marką o małym udziale w rynku. Z drugiej strony pozyskanie nowych nabywców będzie się wiązało równocześnie ze zwiększeniem intensywności zakupów, zatem najbardziej możliwa jest kombinacja wspomnianych sposobów. Należy jednak zwrócić uwagę, że nawet jeżeli marki różnią się znacznie między sobą pod względem udziału w rynku, to jeśli chodzi o intensywność zakupów, różnice te są dużo mniejsze, w związku z czym zwiększanie intensywności zakupów danej marki jest ograniczone.

W przypadku nowej marki można na podstawie modelu (znając docelowy udział marki w rynku) przewidzieć, ile osób będzie po ustabilizowaniu się rynku kupować nową markę, jak często będą ją kupować, jaki odsetek nabywców marki będzie wobec niej lojalny, czy też ile produktu będą przeciętnie kupować nabywcy nowej marki.

Model dostarcza pewnych norm (dla rynków ustabilizowanych), dzięki temu można stwierdzić, które marki są nietypowe. Teoretyczne normy pozwalają ocenić, w jakim stopniu np. wzrost udziału marki w rynku spowodowany jest przez przyciągnięcie nowych nabywców, a w jakim przez zwiększenie intensywności zakupów przez istniejących nabywców. Zatem można wykorzystać model Dirichleta w sytuacji zmian do interpretacji zaistniałej zmiany.

3. Przykład

Obliczenia pozwalające dopasować model do danych empirycznych można przedstawić na poniższym przykładzie². Rozważmy trzy marki (X , Y , Z) o udziałach w rynku odpowiednio: 0,48; 0,35; 0,17, dane zawiera tablica 1.

Na podstawie tablicy 1. można obliczyć zaobserwowane wskaźniki penetracji, intensywności zakupów i stopy zakupów dla produktu oraz poszczególnych marek:

Wskaźnik penetracji produktu $B=20/100=0,20$ oznacza, że w analizowanym okresie 20% osób biorących udział w panelu kupiło przynajmniej raz produkt. Ponadto wskaźnik intensywności zakupów produktu $W=58/20=2,9$ oznacza, że w analizowanym okresie czasu nabywcy produktu dokonywali średnio 2,9 razy zakupu produktu, wskaźnik stopy zakupów produktu $M=58/100=0,58$, jak wspomniano, informuje o tym, ile kupionego produktu przypada na osobę. Dla poszczególnych marek można przeprowadzić analogiczne obliczenia:

wskaźnik penetracji:

$$\text{marka X: } b_1 = 14/100 = 0,14,$$

$$\text{marka Y: } b_2 = 11/100 = 0,11,$$

$$\text{marka Z: } b_3 = 7/100 = 0,07.$$

wskaźnik intensywności zakupów:

$$\text{marka X: } w_1 = 28/14 = 2,00,$$

$$\text{marka Y: } w_2 = 20/11 = 1,82,$$

$$\text{marka Z: } w_3 = 10/7 = 1,43.$$

wskaźnik stopy zakupów:

$$\text{marka X: } m_1 = 28/100 = 0,28,$$

$$\text{marka Y: } m_2 = 20/100 = 0,2,$$

$$\text{marka Z: } m_3 = 10/100 = 0,1.$$

członkowie panelu	zakupy w tygodniu												zakupy produktu		zakupy marki						zakupy produktu przez nabywców marki			
	1	2	3	4	5	6	7	8	9	10	11	12	przynajmniej raz		łącznie		przynajmniej raz		łącznie		X	Y	Z	
													X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
1	X	X	Y	X	X	X	Y	X		Y	X	Y	1	1		7	4		11	11				
2		X	X	Z	X	X				Z	X	X	1	1	1	4		2	6		6			6
3					Z		X	X		X	Z		1	1	1	2		2	4		4			4
4				Y			Y	Y		X			1	1	1	1	3		4		4			4
5	Y	X	X						X			Y	1	1	1	3	2		5		5			
6	X						Z	Z	Y				1	1	1	1	1	2	4		4			4
7		Z	X									X	1	1	1	2		3		3				3
8		Y		Y				X					1	1	1	1	2		3		3			3
9	Z						X				Y		1	1	1	1	1		3		3			3
10						Y							1	1	1	1		1		1				1
11				Y			X						1	1	1	1	1		2		2			2
12						X			X				1	1	1	2		2		2				2
13	Y				Y								1	1	1	1		2		2				2
14												X	1	1	1	1		1		1				1
15				Z									1	1	1			1		1				1
16					Z								1	1	1			1		1				1
17							X						1	1	1	1		1		1				1
18			Y			Y							1	1	1		2		2		2			2
19					Y								1	1	1			1		1				1
20				X									1	1	1	1		1		1				1
21-100																								
liczba kupujących	5	4	6	4	5	6	5	5	4	5	5	4	20	14	11	7								
liczba dokonanych zakupów	5	4	6	4	5	6	5	5	4	5	5	4					28	20	10		50	38		22

Tab. 1. Uproszczony przykład zachowań nabywczych dla trzech marek w okresie 12 tygodni.

3.1. Prawdopodobieństwa dotyczące zakupów produktu

Przystępując do szacowania modelu, w pierwszej kolejności trzeba oszacować prawdopodobieństwa dokonania n -zakupów w analizowanym okresie. Należy w tym celu oszacować parametr K . Łatwo zauważyć, że podstawiając do wzoru (1) $n=0$, otrzymamy prawdopodobieństwo niedokonania zakupu produktu:

$$P_0 = \left(1 + \frac{M}{K}\right)^{-K}, \quad (3)$$

gdzie: M – wskaźnik stopy zakupów produktu, K – parametr.

P_0 można oszacować odsetkiem niekupujących produktu, fakt ten wykorzystamy, aby oszacować parametr K , podstawiając do równania (2) za P_0 : $1-B=0,8$, a następnie rozwiązując je po K :

$$0,8 = \left(1 + \frac{0,58}{K}\right)^{-K}, \quad \text{stad } K = 0,13.$$

Mając oszacowany parametr K , najprościej przekształcić wzór (1) na P_n na równania rekurencyjne (co znacznie ułatwia obliczenia):

$$P_n = P_{n-1} \left(\frac{M}{M+K}\right) \frac{K+n-1}{n}, \quad (4)$$

dla $n=1, 2, \dots, v$.

Poczynając od P_1 obliczamy P_n dla kolejnych $n=1, 2, \dots$:

$$P_1 = 0,8 \cdot \left(\frac{0,58}{0,58+0,13}\right) \frac{0,13+1-1}{1} = 0,08496$$

$$P_2 = 0,08496 \cdot \left(\frac{0,58}{0,58+0,13}\right) \frac{0,13+2-1}{2} = 0,03921, \quad \text{itd.}$$

Obliczenia dla pozostałych P_n zestawiono w tabelicy 2.

n	P_n	ΣP_n	nP_n	ΣnP_n
0	0.80	0.8	0	0
1	0.08496	0.88496	0.0850	0.0850
2	0.03921	0.92417	0.0784	0.1634
3	0.02274	0.94691	0.0682	0.2316
4	0.01454	0.96145	0.0582	0.2898
5	0.00981	0.97126	0.0490	0.3388
6	0.00685	0.97811	0.0411	0.37992
7+	0.02189	1.00000	0.2001	0.58000
Σ	1.0000		0.5800	=M

Tab. 2. Prawdopodobieństwa dokonania zakupu produktu n -razy.

Kwestię maksymalnej liczby zakupów można rozwiązać na dwa sposoby. Pierwszy z nich polega na przyjęciu n na pewnym poziomie v i obliczeniu

dla niego P_v , jako $1 - \sum_{n=0}^{v'-1} P_n$. Drugi z nich dzieli między $v'-1$ i v' pozostałą część prawdopodobieństwa w ten sposób, żeby spełniona była równość:

$$(v'-1)P_{v'-1} + v'P_v = M - \sum_{n=0}^{v'-2} nP_n \quad (\text{por. Ehrenberg 1988}).$$

W przykładzie zastosowano pierwszy z wymienionych sposobów (przyjęto $v'=7$), w przypadku zastosowania sposobu drugiego łatwo sprawdzić, że $v'-1=9$, zaś $v'=10$.

3.2. Prawdopodobieństwa dotyczące wyboru marki

Po oszacowaniu prawdopodobieństw zakupu produktu n -razy dla kolejnych wartości n przechodzimy do obliczenia prawdopodobieństw $p(r_i|n)$. W tym celu wykorzystamy równanie (2).

Wyznaczając α_i należy zauważyć, że m_i/M jest w istocie udziałem marki i w rynku. W związku z tym α_i można wyznaczyć, porównując zaobserwowany udział i -tej marki m_i/M z jej udziałem teoretycznym α_i/S . Wcześniej należy jednak oszacować parametr S .

W celu oszacowania parametru S wykorzystuje się odsetek konsumentów, którzy nie kupili marki i w analizowanym okresie. Stąd w pierwszej kolejności skoncentrujemy się na konsumentach niekupujących marki i .

Poczynając od:

$$p(0|0) = 1, \quad (5)$$

należy wyznaczyć prawdopodobieństwa niekupienia marki i pod warunkiem dokonania n -razy zakupu produktu (dla $n=1,2,\dots,v$). W tym celu przekształcamy wzór (2) na równania rekurencyjne:

$$p(0|n) = p(0|n-1) \frac{S - \alpha_i + n - 1}{S + n - 1}, \quad \text{dla } n > 0. \quad (6)$$

Obliczeń dla poszczególnych marek dokonujemy, podstawiając na początku za parametr S wartość S' . Wyznaczamy prawdopodobieństwa warunkowe $p(0|n)$ na podstawie wzoru (6), dla kolejnych $n = 1, 2, \dots, v$, a następnie, korzystając z wzoru na prawdopodobieństwo całkowite, obliczamy prawdopodobieństwo niekupienia marki i :

$$p(0) = \sum_{n=0}^v p(0|n)P_n. \quad (7)$$

W sposób iteracyjny szukamy takiej wartości parametru S , aby różnica między $p(0)$, obliczonym na podstawie wzoru (7) a zaobserwowanym odsetkiem konsumentów niekupujących i -tej marki (czyli $1-b$) była jak najmniejsza. W praktyce można wykorzystać dodatek *Solver* lub opcję *szukaj wyniku...* dostępne w arkuszu Excel. Dla marki X obliczenia dla wybranych wartości parametru S przedstawia tabela 3. W przypadku pozostałych marek obliczenia są analogiczne, a wartości parametru S dla marki Y i Z wynoszą odpowiednio 8,1 oraz 46,11.

produkt		marka					
		obliczenia dla S'=2		obliczenia dla S''=35,25		obliczenia dla S=50,75	
n	P _n	p(0 n)	p(0,n)= =p(0 n)P _n	p(0 n)	p(0,n)= =p(0 n)P _n	p(0 n)	p(0,n)= =p(0 n)P _n
0	0.8	1	0.8	1	0.8	1	0.8
1	0.08496	0.51724	0.04394	0.51724	0.04394	0.51724	0.04394
2	0.03921	0.35077	0.01375	0.27443	0.01076	0.27236	0.01068
3	0.02274	0.26610	0.00605	0.14906	0.00339	0.14586	0.00332
4	0.01454	0.21472	0.00312	0.08274	0.00120	0.07938	0.00115
5	0.00981	0.18017	0.00177	0.04687	0.00046	0.04386	0.00043
6	0.00685	0.15532	0.00106	0.02705	0.00019	0.02458	0.00017
7+	0.02189	0.13657	0.00299	0.01589	0.00035	0.01397	0.00031
Σ	1.00000	x	0.87269	x	0.86029	x	0.86000

Tab. 3. Prawdopodobieństwa warunkowe i łączne nie kupienia marki X wyznaczone dla wybranych wartości parametru S.

W przypadku wielu marek parametr S można wyznaczyć jako średnią z wartości parametru wyznaczonych dla poszczególnych marek ważoną ich udziałem w rynku (m_i/M). W rozważanym przykładzie byłoby to: $50,77 \cdot 0,483 + 8,1 \cdot 0,345 + 46,11 \cdot 0,172 = 35,25$. Warto dodać, że Ehrenberg w przypadku odbiegającego od pozostałych szacunku S dopuszcza pominięcie nietypowej wartości i obliczenie średniej na podstawie pozostałych wartości S (szerzej na ten temat zob. Ehrenberg 1988, dodatek C: 345). Ze względu na niewielką liczbę marek w analizowanym przykładzie w dalszych obliczeniach wykorzystana zostanie wartość parametru S obliczona na podstawie wszystkich marek, czyli 35,25 (tym bardziej, że różnice w szacowanych prawdopodobieństwach byłyby niewielkie – rzędu 2-4 dziesięciotysięcznych, por. tablica 3.).

Mając oszacowaną wartość parametru S, wyznaczamy dla poszczególnych marek:

$$\alpha_i = S \cdot \frac{m_i}{M}, \tag{8}$$

zatem:

dla marki X: $\alpha_1 = 35,25 \cdot 0,483 = 17,0186$;

dla marki Y: $\alpha_2 = 35,25 \cdot 0,345 = 12,1561$;

dla marki Z: $\alpha_3 = 35,25 \cdot 0,172 = 6,0781$.

Aby obliczyć prawdopodobieństwa kupienia marki i r_i -razy dla $r_i > 0$ pod warunkiem zakupu produktu n-razy, przekształcamy wzór (2) następująco:

$$p(r_i | n) = p(r_i - 1 | n) \frac{n - r_i + 1}{r_i} \cdot \frac{\alpha_i + r - 1}{S - \alpha_i + n - r_i}, \tag{9}$$

dla $0 < r_i \leq n$.

$$p(1 | 1) = p(0 | 1) \frac{1 - 1 + 1}{1} \cdot \frac{17,0186 + 1 - 1}{35,25 - 17,0186 + 1 - 1} = 0,5172 \cdot 0,9335 = 0,4828,$$

$$p(1|2) = p(0|2) \frac{2-1+1}{1} \cdot \frac{17,0186+1-1}{35,25-17,0186+2-1} = 0,2744 \cdot 1,7699 = 0,4857$$

...

$$p(1|7) = p(0|7) \frac{7-1+1}{1} \cdot \frac{17,0186+7-1}{35,25-17,0186+7-1} = 0,01589 \cdot 4,9164 = 0,0781$$

$$p(2|2) = p(1|2) \frac{2-2+1}{2} \cdot \frac{17,0186+2-1}{35,25-17,0186+2-2} = 0,4857 \cdot 0,4942 = 0,2400$$

itd.

Odsetek konsumentów kupujących produkt n -razy i kupujących markę i r_i -razy dany jest przez iloczyn $p(r_i|n)$ i P_n (z NBD):

$$p(r_i|n) = p(r_i|n)P_n, \tag{10}$$

$$p(1,1) = p(1|1)P_1 = 0,4828 \cdot 0,08496 = 0,40101,$$

$$p(1,2) = p(1|1)P_2 = 0,4857 \cdot 0,03921 = 0,01904,$$

...

$$p(1,7) = p(1|7)P_1 = 0,0781 \cdot 0,0219 = 0,00171,$$

$$p(2,2) = p(2|2)P_2 = 0,2400 \cdot 0,03921 = 0,00941,$$

itd.

Obliczenie zawiera tablica 4.

zakupy	n	0	1	2	3	4	5	6	7+	razem
Produktu	P_n	0.8	0.08496	0.03921	0.02274	0.01454	0.00981	0.00685	0.02189	1.00000
zakupy marki	r									P_i
	0	0.8	0.04394	0.01076	0.00339	0.00120	0.00046	0.00019	0.00035	0.86029
	1		0.04101	0.01904	0.00855	0.00386	0.00176	0.00081	0.00171	0.07675
	2			0.00941	0.00801	0.00515	0.00299	0.00165	0.00398	0.03119
	3				0.00279	0.00340	0.00281	0.00197	0.00567	0.01663
	4					0.00093	0.00146	0.00146	0.00535	0.00920
	5						0.00034	0.00064	0.00333	0.00431
	6							0.00013	0.00127	0.00140
	7+								0.00023	0.00023
P_n	- prawdopodobieństwo zakupu produktu n-razy									1.00000
$p_n = p(r_i, n)$	- prawdopodobieństwo zakupu marki r-razy i produktu n-razy									
$p_i = p(r_i)$	- prawdopodobieństwo zakupu marki r-razy									

Tab. 4. Prawdopodobieństwa dotyczące zakupów produktu i marki - obliczenia dla marki X.

Teoretyczne prawdopodobieństwo zakupu r_i -razy marki i otrzymujemy, sumując $p(r_i, n)$ po wszystkich n (por. tablica 4, ostatnia kolumna):

$$p(r_i) = \sum_{n=0}^{\nu} p(r_i, n) \quad (11)$$

Sumując $p(r_i, n)$ po odpowiednich wartościach n i r_i , można obliczyć wszystkie statystyki dla konkretnej marki.

3.3. Estymacje dla marki

Korzystając z prawdopodobieństw zestawionych w tabelicy 1., obliczmy charakterystyki dla marki X. Obliczenia dla pozostałych marek są analogiczne.

Teoretyczny wskaźnik penetracji b marki i szacowany jest jako (opuszczając w tym wzorze oraz w następnych subskrypt i):

$$b = 1 - p(0). \quad (12)$$

Teoretyczny wskaźnik penetracji dla marki X wyniesie zatem:

$$b = 1 - 0,86029 = 0,1397 \approx 0,14.$$

Wskaźnik intensywności zakupów marki i przez jej nabywcę oblicza się w następujący sposób:

$$w = \frac{\sum_{n=1}^{\nu} \left(P_n \sum_{r=1}^{\rho} r p(r | n) \right)}{1 - p(0)} = \frac{\sum_{r=1}^{\rho} r p(r)}{b}, \quad (13)$$

dla marki X wyniesie on: $w = \frac{0,2574}{0,1397} = 1,84$

Wskaźnik intensywności zakupów produktu dokonanych przez nabywcę marki i informuje o tym, ile zakupów produktu przeciętnie dokonują nabywcy tej marki w danym okresie i jest obliczamy jako:

$$w_p = \frac{\sum_{n=1}^{\nu} \{nP_n(1 - p(0 | n))\}}{1 - p(0)} = \frac{\sum_{n=1}^{\nu} \{n(P_n - p(0, n))\}}{b} \quad (14)$$

Dla nabywców marki X wskaźnik intensywności zakupów produktu wynosi:

$$w_p = \frac{0,4468}{0,1397} = 3,20.$$

Dla konsumentów lojalnych wskaźnik penetracji można wyznaczyć, sumując prawdopodobieństwa $p(n, n)$ dla n większych od zera i nie większych niż ρ :

$$b_s = \sum_{n=1}^{\nu} P_n p(n | n) = \sum_{n=1}^{\nu} p(n, n), \quad 0 < n \leq \rho. \quad (15)$$

Nabywca lojalny, kupujący n -razy wyrób, kupuje również n -razy daną markę, zatem wskaźnik intensywności zakupów dla konsumentów lojalnych:

$$w_s = \frac{\left\{ \sum_{n=1}^v (nP_n p(n|n)) \right\}}{\left\{ \sum_{n=1}^v (P_n p(n|n)) \right\}} \quad (16)$$

Sumując prawdopodobieństwa leżące na przekątnej w tablicy 4. (pomijając $p(0,0)$ z oczywistych względów), otrzymujemy wskaźnik penetracji dla nabywców lojalnych wobec marki X równy 0,05484. Warto zwrócić uwagę, że dla dłuższego okresu odsetek nabywców lojalnych wobec danej marki będzie niższy.

Natomiast wskaźnik intensywności zakupów dla konsumentów lojalnych wobec marki X wyniesie $0,0760/0,0548=1,39$.

W okresie o długości T wszystkie powyższe formuły w modelu Dirichleta nie zmienia się, za wyjątkiem M w równaniu na P_n , (wzór 1), które przejdzie w MT (por. Goodhardt, Ehrenberg, Chatfield 1984: 629).

Zaobserwowane i teoretyczne wartości wskaźników penetracji i intensywności zakupów dla wszystkich trzech marek przedstawia tablica 5.

marka \ wskaźnik	X		Y		Z	
	zaobs.	teoret.	zaobs.	teoret.	zaobs.	teoret.
penetracji	0.14	0.14	0.11	0.11	0.07	0.07
intensywności zakupu	2.00	1.84	1.82	1.61	1.43	1.32

Tab. 5. Zaobserwowane i teoretyczne wskaźniki penetracji i intensywności zakupów dla poszczególnych marek.

Podsumowując, zgodnie z przyjętym modelem możemy się spodziewać, że w ciągu 12 tygodni 14% osób kupi markę X przynajmniej raz. Nabywcy marki X kupią przeciętnie 1,84 razy markę X i 3,2 razy produkt. Konsumenti lojalni wobec marki X będą stanowili prawie 5,5% wszystkich konsumentów (oraz 39% nabywców marki X) i będą oni kupować przeciętnie 1,39 razy produkt (a tym samym markę X). Analogiczne analizy można przeprowadzić dla pozostałych marek.

Jak widać w tablicy 5., zaobserwowane wskaźniki intensywności zakupu są nieco wyższe niż zaobserwowane. Wynika to częściowo z faktu, że w modelu ujęto tylko 3 marki i siłą rzeczy mają one duży udział w rynku (np. marka X 48%, marka Y 35%). W przypadku marek o dużym udziale w rynku zdarza się, że w rzeczywistości mają one wyższy wskaźnik intensywności zakupu od oczekiwanego (por. Lambkin 2001: 122).

Wspólni nabywcy marek i i j

Aby oszacować liczbę wspólnych nabywców dla marek i i j , skorzystać należy z własności addytywności. Tworząc wspólną markę $(i+j)$ i szacując $b_{(i+j)}$ można obliczyć b_{ij} - teoretyczny odsetek nabywców obu marek (czyli kupujących obie marki w analizowanym okresie czasu przynajmniej raz):

$$b_{ij} = b_i + b_j - b_{(i+j)} \quad ()$$

oraz warunkowe odsetki: $b_{i|j} = \frac{b_{ij}}{b_j}$.

marki	X i Y		X i Z		Y i Z	
	zaobs.	teoret.	zaobs.	teoret.	zaobs.	teoret.
$b_{(i+j)}$	0.18	0.18	0.16	0.16	0.16	0.15
$b_i + b_j - b_{(i+j)}$	0.07	0.07	0.05	0.05	0.02	0.03

Tab. 6. Wskaźniki penetracji dla par marek.

Z tablicy 6 wynika, że w analizowanym okresie konsumenci, którzy kupują zarówno markę X, jak i markę Y stanowią 7% wszystkich konsumentów, natomiast 5% konsumentów to wspólni nabywcy marki X i Z. Biorąc pod uwagę, że odsetek nabywców marki X wynosi 0,14, łatwo obliczyć, że 50% (0,07/0,14) nabywców marki X kupuje również markę Y, zaś 35% nabywców marki X kupuje również markę Z (por. prawo współwystępowania zakupów), z kolei 39% nabywców marki X jest wobec niej lojalnych. Warto dodać, że aż 71% nabywców marki Z i 64% nabywców marki Y kupuje również markę X.

Model Dirichleta potwierdza również występowanie zjawiska podwójnego ryzyka. Marki, które mają mniejszy udział w rynku, są rzadziej kupowane przez swoich nabywców (por. wskaźniki intensywności zakupów marki w tablicy 5.). Ponadto nabywcy marek o mniejszym udziale w rynku są mniej wobec nich lojalni, co przejawia się niższym odsetkiem nabywców lojalnych (zgodnie z rozkładem Dirichleta odsetek konsumentów lojalnych wobec marki X wynosi 5,5 %, wobec marki Y: 3,6% i wobec marki Z: 1,6%, empiryczne odsetki wynoszą odpowiednio 0,04; 0,04; 0,02).

4. Podsumowanie

Model Dirichleta ma niewielkie wymagania odnośnie informacji potrzebnych do oszacowania modelu. Wymaga znajomości udziałów w rynku, poszczególnych marek oraz oszacowania parametrów strukturalnych modelu S i K , które są łatwe w interpretacji (mogą być interpretowane jako odzwierciedlenie niejednorodności konsumentów pod względem preferencji wyboru marek – S oraz ilości dokonywanych zakupów produktu – K) i pozwalają dopasować rozkłady do danych empirycznych.

Zaletą modelu jest duża elastyczność, nie opiera się on na z góry ustalonych parametrach, ale na rozkładach, co pozwala lepiej dopasować model do konkretnego rynku.

Poważnym ograniczeniem tego modelu jest założenie o braku segmentów. Na każdym rynku można wyodrębnić segmenty, gdyż występują różnice w formule produktu, dystrybucji, cenie, jakości, możliwościach zastosowania, sposobach użytkowania itp. Często bardziej będziemy zainteresowani znajomością pozycji marki w segmencie niż w odniesieniu do całego rynku, lub też wzorcami konkurencji w poszczególnych segmentach.

Z drugiej strony przeprowadzane badania dostarczyły argumentów, że pomimo

tego ograniczenia w większości przypadków model dobrze opisuje obserwowane wzorce zachowań konsumentów (model był stosowany m. in. do takich produktów, jak: benzyna, paliwo do samolotów, zupy, kawa, płatki śniadaniowe, słodycze, kosmetyki, mydło, a nawet do sieci sklepów, przykłady zastosowań można znaleźć w: Lambkin 2001: 103, Ehrenberg 1972, 1988, Keng, Ehrenberg 1984). Warto jednak pamiętać, że istnieją sytuacje, gdy model Dirichleta nie do końca jest dobrze dopasowany (por. Ehrenberg 1988), a także wyrób (papierosy), wobec którego model Dirichleta się nie sprawdza.

Informacje o autorce

Mgr Agnieszka Pleśniak, asystent, Zakład Metod Badań Marketingowych, Instytut Statystyki i Demografii, Szkoła Główna Handlowa w Warszawie, e-mail: agnieszka.plesniak@sgh.waw.pl

Przypisy

- 1 Więcej na temat rozkładu wielowymiarowego beta i wielomianowego można znaleźć w: Mosimann 1962; 1963.
- 2 Opisany przykład nawiązuje do przykładu zamieszczonego w: Ehrenberg 1988, dodatek C. Przyjęto taki sam okres analizy 12 tygodni, aby umożliwić ewentualne porównanie.

BIBLIOGRAFIA

- Ehrenberg, A.S.C. 1972. *Repeat Buying- Theory And Applications*, Amsterdam: North-Holland. wyd.2, 1988 London: Charles Griffin and Company LTD, New York: Oxford University Press.
- Ehrenberg, A.S.C. 2001. Nowe gatunki a istniejący rynek. w: Lambkin, M. (red) *Zachowanie Konsumenta. Koncepcje i Badania Europejskie*, s. 86-104. Warszawa: PWN.
- Ehrenberg, A.S.C., Goodhardt, G.J., Barwise, T.P. 1990. Double Jeopardy Revisited. *Journal of Marketing Research*, vol.54 (February), s. 82-91.
- Geyer-Schulz, A., Hahsler, M., Jahn, M. 2001. A Customer Purchase Incidence Model Applied to Recommender Services. <http://citeseer.nj.nec.com/geyer-schulz01customer.html>.
- Grover, R., Srinivasan, V. 1987. A Simultaneous Approach to Market Segmentation and Market Structuring. *Journal of Marketing Research*, nr 24 (May), s. 139-153.
- Jachnis, A., Terelak, J.F. 1998. *Psychologia konsumenta i reklamy*, Bydgoszcz: Oficyna wydawnicza Branta.
- Keng, K.A., Ehrenberg, A.S.C. 1984. Patterns of Store Choice. *Journal of Marketing Research*, nr 21 (November), s. 399-409.
- Lambkin M. (red). 2001. *Zachowanie Konsumenta: Koncepcje i Badania Europejskie*, Warszawa: PWN.
- Mosimann, J.E. 1963. On the compound negative multinomial distribution, and correlations among inversely sampled pollen counts. *Biometrika*, 47-54.
- Mosimann, J.E. 1962. On the compound multinomial distribution, the multivariate b-distribution and correlations among proportions. *Biometrika*, 49, s. 65-82.
- Rudnicki L. 2000. *Zachowanie konsumentów na rynku*. Warszawa: PWE.
- Smyczek, S. 2003a. Modele zachowań rynkowych konsumentów. *Marketing i Rynek*, nr 1,
- Smyczek, S. 2003b. Modele zachowań rynkowych konsumentów. *Marketing i Rynek*, nr 2,