

*Krzysztof Błoński*  
Katedra Marketingu  
Uniwersytet Szczeciński

*Ewa Putek-Szeląg*  
Katedra Ekonometrii  
Uniwersytet Szczeciński

# Wykorzystanie modeli ze zmiennymi porządkowymi w badaniach satysfakcji klienta

## Streszczenie

Pomiar satysfakcji klienta możliwy jest z wykorzystaniem skal ilościowych, jak również jakościowych. Jednak stosowanie każdej z nich związane jest z problemami i ograniczeniami. Niektórzy autorzy, np. N. Hill i J. Alexander w publikacji *Pomiar satysfakcji i lojalności klientów* (2003) jednoznacznie wskazują najlepszy według nich rodzaj skal do pomiaru satysfakcji – skalę liczbową. Mimo takich wskazówek najczęściej stosowana jest skala porządkowa. Jej użycie powoduje konieczność wykorzystania metod statystycznych pozwalających na analizowanie danych jakościowych.

Artykuł przedstawia zastosowanie analizy klas ukrytych w badaniach satysfakcji. Do budowy modelu wykorzystano dane z pomiaru dokonanego na skalach porządkowych.

**Słowa kluczowe:** analiza klas ukrytych, satysfakcja, samorząd terytorialny, sektor publiczny.

## 1. Wprowadzenie

Pomiar satysfakcji klienta jest możliwy na wszystkich spośród skal zaproponowanych przez S.S. Stevensona. Jednak użycie każdej skali pomiaru łączy się

z ograniczeniami metodologicznymi. Stosunkowo często wykorzystywana jest skala porządkowa. Jej wadą jest ograniczona liczba możliwych odpowiedzi – jednakże jeśli skala ma być czytelna, nie powinno się wykorzystywać zbyt wielu opcji wyboru. Dodatkowo trzeba zadbać, aby odpowiedzi respondentów ściśle korespondowały ze stwierdzeniami na skali. Stosując ją, można użyć tylko podstawowych miar z obszaru statystyki opisowej. Często spotykanym zabiegiem jest przypisanie odpowiedziom *a priori* arbitralnie wartości liczbowych. Takie podejście jest szeroko krytykowane, ponieważ założenie o liniowej relacji poszczególnych stopni satysfakcji nie zawsze jest zgodne z rzeczywistymi warunkami rynkowymi. Ponadto przejście z jednego poziomu satysfakcji na kolejny nie dostarcza tej samej wartości klientowi oraz nie jest proporcjonalne do wysiłku organizacji. Należy również pamiętać, że poziomy wymagań klientów mogą być różne dla różnych towarów lub usług [Grigoroudis i Siskos 2010, s. 24].

Sposobem na przewyżczenie niedogodności związanych z użyciem podstawowych miar z obszaru statystyki opisowej może być wykorzystanie metod statystycznej analizy danych jakościowych. Przegląd i opis tych metod zawierają m.in. prace: [Analiza danych... 2011], [Kenett i Salini 2012], [Zaawansowane metody... 2012]. Celem artykułu jest przedstawienie wykorzystania modeli klas ukrytych dla danych jakościowych na podstawie wyników badań klientów korzystających z usług świadczonych przez jednostki samorządu terytorialnego.

## 2. Modele klas ukrytych

Wśród metod pozwalających na modelowanie związków pomiędzy zmiennymi oprócz analizy czynnikowej i analizy kowariancji można wykorzystać analizę zmiennych ukrytych. Wyróżnia się kilka rodzajów modelowania związków ukrytych pomiędzy zmiennymi [Zaawansowane metody... 2012, s. 140]:

- analizę z ukrytymi charakterystykami (*latent trait analysis*), która pozwala na modelowanie ciągłych zmiennych ukrytych na podstawie cech skokowych,
- analizę z ukrytymi profilami (*latent profile analysis*), pozwalającą na odkrywanie ukrytych skokowych zależności na podstawie obserwowalnych zmiennych ciągłych,
- analizę z ukrytymi klasami (*latent class analysis*), umożliwiającą modelowanie zależności o charakterze jakościowym, bez różnicowania, na jakiej skali następował pomiar oraz w jaki sposób były zakodowane poziomy cech objaśniających.

Modele klas ukrytych mają podobne zastosowanie do metod klasyfikacji jak analiza skupień. Dlatego analiza klas ukrytych jest wykorzystywana jako technika analizy tablic kontyngencji umożliwiającą zidentyfikowanie wzajemnie rozłącz-

nych klas. Jednak zmienne ukryte, na podstawie których dokonuje się podziału, nie wnoszą żadnych dodatkowych informacji ponad te, które są zawarte w zmien-nych obserwowanych. Służą natomiast do syntezy lub agregacji właściwości zawartych w zmiennych obserwowanych [Analiza danych... 2011, s. 206]. Wyodrębniona zmienna ukryta w dalszych analizach może występować jako zmienna zarówno niezależna, jak i zależna, co wynika z rozróżnienia głównych obszarów zastosowań modeli zmiennych ukrytych. W tym podziale można wyróżnić umieszczenie analizowanych przypadków w segmentach (*latent class cluster models*), redukcję zmiennych (*latent class factor models*), jak również konstrukcję skali oraz predykcje zmiennej zależnej (*latent class regression and choice models*) [Zaawansowane metody... 2012, s. 142].

Przewagą analizy klas ukrytych nad innymi metodami wielowymiarowej analizy zmiennych jakościowych jest brak ograniczeń takich, jak: liniowość, normalność rozkładu zmiennych wskaźnikowych, jednorodność wariancji czy brak skorelowania zmiennych objaśniających. Dodatkowo efekt finalny analizy jest przedstawiany w postaci parametrów będących prawdopodobieństwami. Są to:

- prawdopodobieństwo bezwarunkowe –  $\gamma_c$  – czyli prawdopodobieństwo przynależności jednostki do klasy ukrytej  $c$ , gdzie:  $P(L = c) = \gamma_c$  oraz  $\sum_{c=1}^C \gamma_c = 1$ ,
- prawdopodobieństwo warunkowe –  $\rho_{j, r_{jc}}$  – prawdopodobieństwo odpowiedzi  $r_j$  na pytanie wskaźnikowe  $j$  pod warunkiem przynależności jednostki do klasy ukrytej  $c$ .

Podstawowy model można przedstawić następującym równaniem:

$$P(Y = y) = \sum_{c=1}^C P(L = c) P(Y = y | L = c), \quad (1)$$

gdzie:

$P(L = c)$  – proporcja osób należących do klasy  $c$ ;

$L$  – zmienna ukryta o  $c$  kategoriach, gdzie  $c = 1, \dots, C$  oznacza liczbę wyodrębnionych klas ukrytych.

Wykorzystanie wzoru (1) wymaga spełnienia założenia o lokalnej niezależności zmiennych  $J$  zmiennych wskaźnikowych w poszczególnych klasach ukrytych, gdzie zmienne wskaźnikowe mają  $r_j$  poziomów, co przedstawia następujący wzór:

$$P(Y = y, L = c) = P(L = c)P(Y = y | L = c) = \gamma_c \prod_{j=1}^J \prod_{r_j=1}^{R_j} \rho_{j, r_{jc}}^{I(y_j = r_j)}. \quad (2)$$

Jeżeli założymy, że zmienne reprezentują pytania zawarte w kwestionariuszu, a kategorie zmiennych – możliwe odpowiedzi, to wynikiem przeprowadzonej analizy klas ukrytych będzie tablica przedstawiająca rozkład procentowy liczebności zawartych w początkowej tablicy kontyngencji z podziałem na klasy.

Weryfikacji uzyskanych modeli dokonuje się za pomocą analizy absolutnego oraz względnego dopasowania modelu. Absolutne dopasowanie modelu ma na celu określenie, jak dobrze model odwzorowuje empirycznie zaobserwowaną zmienność analizowanych cech. Najczęściej stosowanym podejściem w ocenie są statystyki oparte na ilorazie wiarygodności  $G^2$ , który wyrażony jest wzorem:

$$G^2 = 2 \sum_{i=1}^W n_i \ln \left( \frac{n_i}{\hat{n}_i} \right), \quad (3)$$

gdzie:

$n_i$  – rzeczywista liczebność w komórce w tablicy kontyngencji,

$\hat{n}_i$  – teoretyczna liczebność w komórce w tablicy kontyngencji.

Przyjmuje się, że model jest dopasowany do danych, jeżeli wartość  $G^2$  jest dostatecznie niska dla wskazanego poziomu istotności. Najlepszym modelem spośród identyfikowalnych z różną liczbą klas ukrytych jest ten, w przypadku którego nastąpi duże obniżenie wartości  $G^2$ , a dalszy spadek będzie nieznaczny. Względne dopasowanie modelu pozwala zweryfikować, który z modeli wśród tych z różną liczbą klas ukrytych jest lepszy. Ocena jest dokonywana według kryteriów informacyjnych pozwalających zoptymalizować równowagę pomiędzy dopasowaniem modelu do danych a jego prostotą. Najczęściej stosowanym kryterium informacyjnym jest kryterium AIC (*Akaike information criterion*) wraz z kryterium BIC (*Bayesian information criterion*). Miary AIC oraz BIC można skonstruować na podstawie wartości  $G^2$ :

$$AIC = G^2 + 2P \quad (4)$$

$$BIC = G^2 + [\log(N)]P, \quad (5)$$

gdzie:

$P$  – liczba parametrów estymowanych w modelu,

$N$  – wielkość próby.

### 3. Analiza klas ukrytych na podstawie badań satysfakcji odbiorców usług świadczonych przez jednostki samorządu terytorialnego

Celem badań ankietowych przeprowadzonych na terenie 16 jednostek samorządu terytorialnego (JST) województwa zachodniopomorskiego był pomiar jakości i zadowolenia z usług świadczonych przez urzędy gminne i (lub) powiatowe. Badania przeprowadzono dwa razy – pierwsze w okresie listopad–grudzień

2009 r., drugie w czerwcu i lipcu 2010 r. Podmiotami badania byli m.in. mieszkańcy danej gminy lub powiatu<sup>1</sup>. Łącznie przebadano 6053 respondentów (z czego 3264 w pierwszym badaniu, a w drugim 2789). Wielkość próby badawczej do badań ankietowych wyznaczono na podstawie liczby ludności zamieszkującej na terenie działania poszczególnych JST. Badania zostały przeprowadzone za pomocą ankiety bezpośredniej wśród interesantów odwiedzających siedzibę danej JST (50% badanych) oraz na jej terenie (50% badanych). Pozwoliło to na zebranie informacji na temat:

- oceny warunków zamieszkiwania na terenie gminy lub powiatu,
- oceny wizerunku gminy i możliwości jej rozwoju,
- istotności dla mieszkańców atrybutów zapewniających właściwy poziom świadczonych usług,
- subiektywnych ocen bieżącej sytuacji w zakresie składowych świadczonych usług.

Badani zostali również poproszeni o ogólną ocenę jakości obsługi w danej jednostce samorządu terytorialnego. W artykule przeanalizowano wyniki uzyskane podczas drugiego badania. Na ich podstawie oszacowano sześć różnych modeli różniących się między sobą liczbą zmiennych oraz klas ukrytych:

1. Model A – zawierający 28 zmiennych szacowanych za pomocą pięciostopniowej skali porządkowej. Wybrane zmienne przedstawiają ocenę bieżącego stanu świadczenia usług.

2. Model B – zawiera 4 zmienne szacowane z wykorzystaniem pięciostopniowych skal porządkowych. Wybrane zmienne dotyczą subiektywnej opinii na temat wizerunku JST, odczuwania stopnia zadowolenia i dumy z zamieszkiwania na terenie danego JST oraz deklaracji chęci wyjazdu.

3. Model C – zawierający 34 zmienne i będący połączeniem dwóch wcześniejszych modeli.

4. Model D – stanowiący rozszerzenie modelu A i oprócz wymienionych zmiennych zawierający subiektywną ocenę jakości obsługi klienta jako zmienną towarzyszącą.

5. Model E – stanowiący rozbudowaną wersję modelu B zawierającą dodatkowo subiektywną ocenę jakości obsługi klienta jako zmienną towarzyszącą.

6. Model F – który był wersją modelu C rozbudowaną o zmienną towarzyszącą.

---

<sup>1</sup> Badania były częścią zadania „Satysfakcja klientów i pracowników urzędów” realizowanego w ramach projektu „Wdrażanie usprawnień zarządczych w JST na obszarze województwa zachodniopomorskiego”. Kierownikiem projektu była dr hab. prof. US Teresa Lubińska, a kierownikiem zadania dr hab. prof. US Jolanta Witek.

Tabela 1. Wyniki absolutnego i względnego dopasowania modelu E do danych empirycznych

Liczba klas	AIC	BIC	$G^2$	$\chi^2$	Stopnie swobody
3	25 019,94	25 328,05	3106,53	48 032,21	52
4	22 817,27	23 232,04	1339,68	13 735,44	70
5	25 229,24	25 750,66	2975,50	5036,01	88

Źródło: opracowanie własne na podstawie przeprowadzonych badań.

Tabela 2. Prawdopodobieństwa warunkowe dla modeli z trzema i czterema klasami ukrytymi

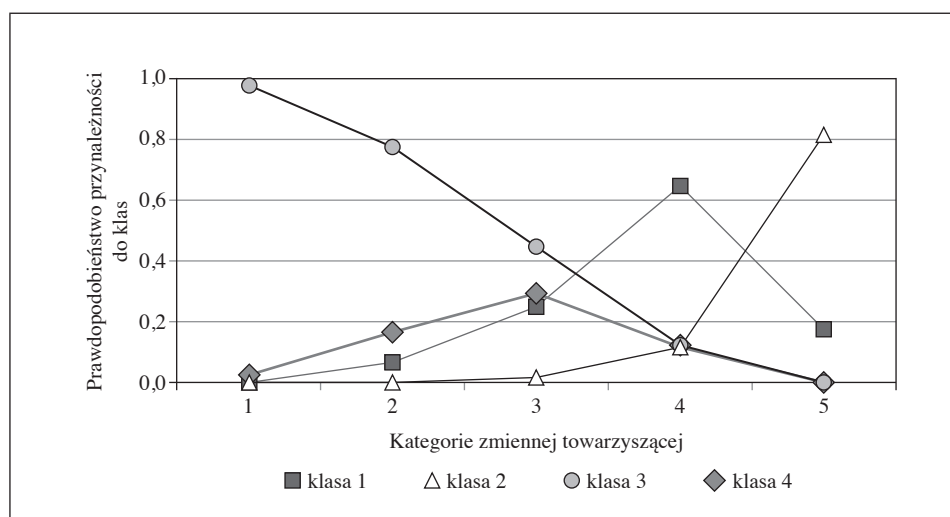
Zmienne	Poziomy odpowiedzi	Model z czterema klasami ukrytymi			
		klasa 1	klasa 2	klasa 3	klasa 4
Opinia o wizerunku danej JST	bardzo negatywny	0,000000	0,001678	0,566849	0,007843
	negatywny	0,025694	0,012343	0,433151	0,147140
	obojętny	0,185548	0,051707	0,000000	0,540430
	pozytywny	0,753157	0,263984	0,000000	0,291534
	bardzo pozytywny	0,035601	0,670287	0,000000	0,013054
Zgodność z opinią dotyczącą zadowolenia z mieszkania na terenie danej JST	całkowicie się nie zgadzam	0,161166	0,018710	0,000000	0,340271
	nie zgadzam się	0,002963	0,003428	0,698193	0,015667
	trudno powiedzieć	0,042891	0,007035	0,301807	0,134964
	zgadzam się	0,758362	0,146202	0,000000	0,466455
	całkowicie się zgadzam	0,034619	0,824625	0,000000	0,042643
Zgodność z opinią dotyczącą dumy z zamieszkiwania na terenie danej JST	całkowicie się nie zgadzam	0,329789	0,055136	0,086798	0,484743
	nie zgadzam się	0,046036	0,003447	0,610918	0,112265
	trudno powiedzieć	0,038656	0,002381	0,302284	0,114069
	zgadzam się	0,566504	0,116805	0,000000	0,240189
	całkowicie się zgadzam	0,019015	0,822231	0,000000	0,048734
Zgodność z opinią dotyczącą chęci wyjazdu, jeśli pojawi się taka możliwość	całkowicie się nie zgadzam	0,430543	0,324353	0,434303	0,383825
	nie zgadzam się	0,062072	0,431316	0,000000	0,067010
	trudno powiedzieć	0,344067	0,188954	0,000000	0,324605
	zgadzam się	0,110548	0,033003	0,085679	0,161900
	całkowicie się zgadzam	0,052770	0,022373	0,480018	0,062660

Źródło: opracowanie własne na podstawie przeprowadzonych badań.

Łącznie uzyskano 30 różnych wersji modeli, a do dalszej analizy ze względu na wyniki absolutnego i względnego dopasowania modeli do danych wybrano

model E (tabela 1). Spośród różnych wersji tego modelu najniższe wartości kryterium AIC oraz BIC ma model z 4 klasami ukrytymi. Identyczna sytuacja wynika z analizy absolutnego dopasowania modelu – ze względu na wartość  $G^2$  najlepszym modelem jest wersja z 4 klasami ukrytymi. Przeprowadzona analiza prawdopodobieństw warunkowych wskazuje, że można liczyć w tym przypadku na dominujące wzorce odpowiedzi (tabela 2).

W tym modelu do pierwszej klasy zakwalifikowano 754 respondentów, do drugiej – 457, natomiast do klasy trzeciej i czwartej zaliczono odpowiednio 592 i 232 badanych. Wyniki prawdopodobieństw warunkowych są rozproszone pomiędzy warianty odpowiedzi poszczególnych zmiennych, co jest skutkiem liczby poziomów odpowiedzi w poszczególnych zmiennych (rys. 1).



Uwaga: Wartość 1 oznacza bardzo złą ocenę jakości obsługi klienta, zaś 5 ocenę bardzo dobrą.

Rys. 1. Wpływ zmiennej towarzyszącej na przynależność do klas

Źródło: opracowanie własne na podstawie przeprowadzonych badań.

Pierwszą klasę tworzą osoby przekonane o pozytywnym wizerunku swojej JST, jak również zadowolone z zamieszkiwania na danym terenie. Postawy tej grupy nacechowane są dumą oraz wynikającym z tego całkowitym odrzuceniem możliwości zmiany miejsca zamieszkania. Osoby tworzące drugą grupę są zbliżone w swoich poglądach do osób z pierwszej – są bardziej przekonane o pozytywnym wizerunku JST, cechuje je najwyższy poziom zadowolenia i dumy z miejsca zamieszkania. Od pozostałych odróżnia je mniej zdecydowana postawa w sprawie możliwości wyjazdu.

Trzecia klasa jest przeciwieństwem dwóch poprzednich klas. Osoby wchodzące w skład tej grupy negatywnie oceniają wizerunek JST, są niezadowolone oraz nie odczuwają dumy z zamieszkiwania na danym terenie. Konsekwencją takich poglądów jest chęć wyjazdu z danego miejsca, jeśli pojawi się taka możliwość. Czwartą grupę tworzą osoby, które nie mają wyrobionej opinii na temat wizerunku JST ani nie utożsamiają się z danym miejscem. Jednocześnie są zadowolone i odrzucają możliwość wyjazdu, jeśli pojawi się taka możliwość.

Zmienna towarzysząca w sposób znaczący wpływa na prawdopodobieństwo przynależności do dwóch grup – drugiej i trzeciej. W przypadku drugiej grupy wraz ze wzrostem ogólnego zadowolenia z jakości świadczonych usług prawdopodobieństwo przynależności się zwiększa. Z odwrotną sytuacją mamy do czynienia w przypadku klasy trzeciej, gdzie spadkowi oceny zadowolenia towarzyszy wzrost prawdopodobieństwa. W przypadku dwóch pozostałych grup wpływ zmiennej towarzyszącej jest zróżnicowany i dość dobrze oddaje specyfikę grupy pierwszej i czwartej.

#### 4. Zakończenie

Modele klas ukrytych umożliwiają podział respondentów na homogeniczne grupy na podstawie danych nominalnych i porządkowych zebranych dzięki badaniom ankietowym. Dodatkowo w modelu oprócz zmiennych obserwowanych można uwzględnić zmienne towarzyszące, które istotnie wpływają na prawdopodobieństwo przynależności respondentów do poszczególnych segmentów.

W wyniku zastosowania modeli klas ukrytych ze zmiennymi towarzyszącymi wyodrębniono cztery grupy odbiorców usług świadczonych przez jednostki samorządu terytorialnego. Różnią się one opinią o wizerunku gminy lub powiatu, stosunkiem do miejsca zamieszkania oraz deklaracjami o chęci wyjazdu. Dwie zidentyfikowane grupy (grupa druga i trzecia) mogą być podstawą do dalszych badań. Analiza danych respondentów zaliczanych do drugiej grupy może wskazać, czy podstawą do zadowolenia z usług JST są atrybuty usługi (uwarunkowania), czy również inne czynniki (np. stosunek do miejsca zamieszkania). W przypadku potwierdzenia takiej zależności można poszukiwać kolejnych czynników mających wpływ na poziom satysfakcji, np. ocenę działań urzędu gminy lub powiatu, ocenę wizerunku JST. Natomiast grupa trzecia może podstawą do identyfikacji przyczyn niezadowolenia z usług JST i określenia działań korygujących.



## Literatura

- Analiza danych jakościowych i symbolicznych z wykorzystaniem programu R* [2011], red. E. Gatnar, M. Walesiak, Wydawnictwo C.H. Beck, Warszawa.
- Grigoroudis E., Siskos Y. [2010], *Customer Satisfaction Evaluation: Methods for Measuring and Implementing Service Quality*, Springer, New York.
- Kenett R.N., Salini S. [2012], *Modern Analysis of Customer Surveys: with Applications Using R*, Wiley.
- Zaawansowane metody analiz statystycznych* [2012], red. E. Frątczak, Oficyna Wydawnicza, Szkoła Główna Handlowa w Warszawie, Warszawa.

## The Use of Models with Ordinal Variables in Customer Satisfaction Surveys

Customer satisfaction can be measured using both quantitative and qualitative scales. However, each of them has problems and limitations. Some authors, such as N. Hill and J. Alexander in their *Handbook of Customer Satisfaction and Loyalty Measurement* (2003), clearly believe numerical scales are the best for measuring satisfaction. Nonetheless, ordinal scale, which requires statistical methods for the analysis of qualitative data, remains the most commonly used. This paper presents the use of analysis of latent class in satisfaction surveys. To build a model, data from measurements made with ordinal scales were used.

**Keywords:** latent class analysis, satisfaction, self-government, public sector.