

Anna Turczak*

Zachodniopomorska Szkoła Biznesu w Szczecinie

Patrycja Zwiech**

Uniwersytet Szczeciński

CZYNNIKI WPŁYWAJĄCE NA ZAKUPY PRZEZ INTERNET

Streszczenie

Celem artykułu jest określenie, w jakim stopniu wybrane czynniki (wiek, poziom wykształcenia, płeć oraz miejsce zamieszkania) rzutują na fakt, że dana osoba korzysta z Internetu w celach dokonywania zakupów. W drodze przeprowadzonych badań udowodniono, że wiek, poziom wykształcenia, płeć oraz miejsce zamieszkania są zmiennymi, które mają statystycznie istotny wpływ na dokonywanie zakupów przez Internet przez Polaków. Na podstawie obliczonych wartości współczynnika V Cramera dodatkowo stwierdzono, że z czterech wymienionych zmiennych to wiek i wykształcenie są tymi, które w największym stopniu odpowiadają za zróżnicowanie rozpatrywanej cechy statystycznej.

Słowa kluczowe: zakupy przez Internet, wiek, wykształcenie, miejsce zamieszkania, płeć

* Adres e-mail: aturczak@zpsb.szczecin.pl.

** Adres e-mail: patrycjazwiech@tlen.pl.

Wprowadzenie

W 2013 roku w Polsce 60% osób w wieku 16–74 lat regularnie (to znaczy co najmniej raz w tygodniu) korzystało z Internetu, podczas gdy w całej Unii Europejskiej odsetek regularnych użytkowników Internetu wyniósł wówczas 72%. Wśród 28 krajów UE w 2013 r. największy był udział osób regularnie użytkujących Internet w Luksemburgu (93%), a najmniejszy – w Rumunii (45%) (Eurostat database). Najczęściej użytkownicy Internetu korzystali z poczty elektronicznej. Odsetek użytkowników poczty elektronicznej w ogólnej liczbie ludności w Polsce w wieku 16–74 lata wyniósł 51,5%, a wśród osób korzystających z Internetu przekroczył nawet 81%. Drugą najczęściej wykonywaną za pośrednictwem Internetu czynnością było wyszukiwanie informacji o towarach i usługach. W 2013 r. informacje takie wyszukiwało ponad 45,5% osób z przedziału wiekowego 16–74 lat oraz 72,4% tych osób, które w ogóle korzystały z Internetu (*Spoleczeństwo...*, 2013, s. 137).

W Polsce z roku na rok coraz więcej osób zaopatruje się w sieci. W 2004 r. było to 5% wszystkich mieszkańców kraju w wieku 16–74 lat, w 2005 r. – 7%, w 2006 r. – 12%, w 2007 r. – 16%, w 2008 r. – 18%, w 2009 r. – 23%, w 2010 r. – 29%, w 2011 r. – 30%, w 2012 r. – 30% i w 2013 r. – blisko 32% (*Spoleczeństwo...*, 2010, s. 89; *Spoleczeństwo...*, 2013, s. 147). Pomimo przedstawionych wskaźników, przeciętne miesięczne wydatki na osobę na towary i usługi nabyte przez Internet były w gospodarstwach domowych w Polsce na bardzo niskim poziomie. W 2010 r. wydatki te wyniosły jedynie 4,61 zł, w 2011 r. – 5,34 zł, a w 2012 r. – 5,83 zł (*Spoleczeństwo...*, 2013, s. 147). Udział środków przeznaczonych na zakupy przez Internet w ogólnych wydatkach na towary i usługi konsumpcyjne stopniowo wzrastał. Miesięczna wartość wszystkich nabytych towarów i usług konsumpcyjnych *per capita* w 2010 r. to 945,80 zł, w 2011 r. – 971,83 zł, a w 2012 r. – 1.005,19 zł (*Budżety...*, 2011, s. 112; *Budżety...*, 2012, s. 115; *Budżety...*, 2013, s. 108). Z tego na zakupy przez Internet w 2010 r. poniesiono zaledwie 0,49% tych wydatków, w 2011 r. – 0,55%, a w 2012 r. – 0,58%.

Udział osób robiących zakupy przez Internet jest różny w różnych grupach. Jest on zatem niejednorodny i zależy od wielu czynników. Stąd właśnie udzielenie odpowiedzi na pytanie, w jakim stopniu wiek, poziom wykształcenia, płeć i miejsce zamieszkania rzutują na fakt, że dana osoba korzysta z Internetu w celach zakupowych, stało się celem niniejszego opracowania. W związku z tak

zdefiniowanym celem postawiono cztery odrębne hipotezy stanowiące, że istnieje statystycznie istotny związek między:

- wiekiem,
- poziomem wykształcenia,
- płcią,
- miejscem zamieszkania,

a kupowaniem przez Internet towarów lub usług konsumpcyjnych do użytku prywatnego. Do weryfikacji postawionych hipotez posłużył test niezależności chi-kwadrat¹. W ramach artykułu określona została również siła i charakter zależności między wymienionymi czterema cechami statystycznymi a faktem dokonywania zakupów przez Internet. Do określenia siły występujących zależności użyty został współczynnik V Cramera². Okres badawczy obejmował lata 2004–2013.

Określenie zależności między wiekiem a faktem kupowania przez Internet

Pierwszym postawionym zadaniem jest odpowiedź na pytanie, czy istnieje statystycznie istotna zależność między tym, w jakim wieku jest dana osoba, a tym, czy dokonuje zakupów w sieci. A zatem, weryfikacji podlega hipoteza H_0 stanowiąca, że badane dwie cechy są stochastycznie niezależne, wobec hipotezy alternatywnej H_1 orzekającej, że występuje stochastyczna zależność między tymi cechami (Aczel, 2000, s. 758). Procedura weryfikacyjna przeprowadzona została

¹ Przy podejmowaniu decyzji w teście niezależności chi-kwadrat bierze się pod uwagę jednostronny (a konkretnie – prawostronny) obszar krytyczny, który określa nierówność $\chi^2 \geq \chi_\alpha^2$. Statystyka χ_α^2 jest wartością krytyczną odczytaną z tablic rozkładu chi-kwadrat dla przyjętego z góry poziomu istotności α i dla $(r-1)(s-1)$ stopni swobody, gdzie r oznacza liczbę wariantów pierwszej cechy, a s – liczbę wariantów drugiej cechy. Obliczoną wartość statystyki testowej χ^2 porównuje się z wartością krytyczną χ_α^2 i jeśli spełniona jest nierówność $\chi^2 \geq \chi_\alpha^2$, to przy przyjętym poziomie istotności hipotezę zerową należy odrzucić na korzyść hipotezy alternatywnej. Orzeka się wówczas, że rozpatrywane cechy są zależne. Gdy natomiast zachodzi nierówność $\chi^2 < \chi_\alpha^2$, nie ma podstaw do odrzucenia hipotezy zerowej o niezależności badanych cech (Zeliaś, 2000, s. 286; Majewski, Walczak, 2013, s. 43; Sulewski, 2013, s. 2–3).

² Współczynnik V Cramera jest wielkością niemianowaną i unormowaną – przyjmuje wartości wyłącznie z przedziału $(0, 1)$. Jeżeli jest równy 0, to między cechami nie występuje zależność. Z kolei im bliższy jest liczby 1, tym zależność jest silniejsza. Do obliczenia tego współczynnika służy wzór:

$$V = \sqrt{\frac{\chi^2}{n \min(r-1, s-1)}}$$

gdzie n jest liczbą przebadanych jednostek statystycznych (Pułaska-Turyńska, 2005, s. 254).

za pomocą testu niezależności chi-kwadrat i objęła następujące kroki obliczeniowe (Onchiri, 2013, s. 19):

- a) na podstawie informacji o liczebności całkowitej, udziale procentowym osób z poszczególnych grup wiekowych (16–24, 25–34, 35–44, 45–54, 55–64 i 65–74 lata) oraz odpowiedziach respondentów z tych grup na temat korzystania przez nich z Internetu w celu dokonywania zakupów, wyznaczono elementy macierzy liczebności empirycznych n_{ij} ;
- b) obliczono odpowiednie prawdopodobieństwa empiryczne p_{ij} ;
- c) następnie obliczone zostały prawdopodobieństwa teoretyczne p_{ij}^* , które występowałyby przy stochastycznej niezależności badanych cech;
- d) z kolei mnożąc wyznaczone prawdopodobieństwa hipotetyczne p_{ij}^* przez całkowitą liczbę osób w wieku 16–74 lat (czyli przez n), otrzymano liczebności hipotetyczne n_{ij}^* ;
- e) dysponując już wszystkimi elementami macierzy liczebności empirycznych n_{ij} oraz macierzy liczebności teoretycznych n_{ij}^* , obliczono poszczególne składniki statystyki testowej χ^2 ³.

Opisaną procedurę obliczeniową powtórzono dziesięć razy – oddzielnie dla poszczególnych lat z okresu 2004–2013. Dla każdego roku objętego analizą hipotezę H_0 o niezależności tego, czy mieszkaniec Polski kupuje przez Internet towary lub usługi do użytku prywatnego, od tego, jaki jest jego wiek, należało odrzucić na rzecz hipotezy alternatywnej orzekając jednocześnie, że rozpatrywane cechy są zależne⁴ (uzyskane odchylenia między liczebnościami empirycznymi i teoretycznymi były wystarczająco duże, aby odrzucić przypuszczenie o niezależności)⁵.

³ Statystykę χ^2 , którą przyjmuje się jako syntetyczną miarę odchylenia liczebności rzeczywistych od liczebności teoretycznych, oblicza się zgodnie z formułą (Ankarali, Cangur, 2013, s. 95):

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^s \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*} \quad (i = 1, 2, \dots, r; j = 1, 2, \dots, s).$$

⁴ Hipotezę zerową zweryfikowano w analogiczny sposób między innymi w następujących pracach: (Carlin, Doyle, 2001, s. 392; Kukowska, 2011, s. 12–20; Dowling Fink, Lunsford, 2010, s. 12; Guty, Mrozowicz, 2013, s. 138–143; B. Fura, M. Fura, 2011 s. 42–46).

⁵ Decyzja dotycząca hipotezy zerowej była taka sama dla wysokiej wartości współczynnika istotności (na przykład 0,1), jak i dla bardzo niskiego poziomu istotności (przykładowo 0,0001). Można więc orzec, że istnieje statystycznie istotny związek między wiekiem danej osoby a faktem dokonywania przez nią zakupów przez Internet i błąd odrzucenia prawdziwej hipotezy zerowej jest w tym przypadku bliski zeru.

W następnym etapie, dla każdego roku objętego analizą, określono relacje między poszczególnymi prawdopodobieństwami empirycznymi oraz odpowiadającymi im prawdopodobieństwami teoretycznymi. Wyznaczono także udział procentowy każdego składnika otrzymanej statystyki χ^2 . Informacje na temat zależności między obliczonymi wartościami p_{ij} (prawdopodobieństwami rzeczywistymi⁶) i wartościami p_{ij}^* (prawdopodobieństwami hipotetycznymi⁷) dla badanego okresu umieszczono w tabeli 1. W tabeli tej ukazano również strukturę statystyki χ^2 .

Tabela 1

Wyniki uzyskane w badaniu zależności między wiekiem i zakupami przez Internet

Wiek	Kupują przez Internet		Nie kupują przez Internet		Razem
1	2		3		4
2013 r.					
od 16 do 24 lat	$p_{ij} > p_{ij}^*$	5,9%	$p_{ij} < p_{ij}^*$	3,1%	9,0%
od 25 do 34 lat	$p_{ij} > p_{ij}^*$	21,3%	$p_{ij} < p_{ij}^*$	11,0%	32,3%
od 35 do 44 lat	$p_{ij} > p_{ij}^*$	3,0%	$p_{ij} < p_{ij}^*$	1,5%	4,5%
od 45 do 54 lat	$p_{ij} < p_{ij}^*$	4,1%	$p_{ij} > p_{ij}^*$	2,1%	6,2%
od 55 do 64 lat	$p_{ij} < p_{ij}^*$	16,1%	$p_{ij} > p_{ij}^*$	8,3%	24,4%
od 65 do 74 lat	$p_{ij} < p_{ij}^*$	15,6%	$p_{ij} > p_{ij}^*$	8,1%	23,7%
2012 r.					
od 16 do 24 lat	$p_{ij} > p_{ij}^*$	7,7%	$p_{ij} < p_{ij}^*$	3,7%	11,3%
od 25 do 34 lat	$p_{ij} > p_{ij}^*$	21,0%	$p_{ij} < p_{ij}^*$	10,1%	31,1%
od 35 do 44 lat	$p_{ij} > p_{ij}^*$	2,3%	$p_{ij} < p_{ij}^*$	1,1%	3,4%
od 45 do 54 lat	$p_{ij} < p_{ij}^*$	4,0%	$p_{ij} > p_{ij}^*$	1,9%	6,0%
od 55 do 64 lat	$p_{ij} < p_{ij}^*$	17,6%	$p_{ij} > p_{ij}^*$	8,5%	26,1%
od 65 do 74 lat	$p_{ij} < p_{ij}^*$	14,9%	$p_{ij} > p_{ij}^*$	7,2%	22,1%
2011 r.					
od 16 do 24 lat	$p_{ij} > p_{ij}^*$	10,6%	$p_{ij} < p_{ij}^*$	4,8%	15,4%
od 25 do 34 lat	$p_{ij} > p_{ij}^*$	18,9%	$p_{ij} < p_{ij}^*$	8,5%	27,4%

⁶ Są to prawdopodobieństwa empiryczne, że obserwacja przyjmuje i -ty wariant cechy X i j -y wariant cechy Y .

⁷ Są to prawdopodobieństwa teoretyczne, które wystąpiłyby, gdyby hipoteza zerowa była prawdziwa, czyli gdyby rozpatrywane cechy X i Y były niezależne.

1	2		3		4
od 35 do 44 lat	$P_{ij} > P_{ij}^*$	1,7%	$P_{ij} < P_{ij}^*$	0,8%	2,5%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	4,3%	$P_{ij} > P_{ij}^*$	1,9%	6,2%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	16,9%	$P_{ij} > P_{ij}^*$	7,6%	24,5%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	16,6%	$P_{ij} > P_{ij}^*$	7,4%	24,0%
2010 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	12,7%	$P_{ij} < P_{ij}^*$	5,5%	18,3%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	17,7%	$P_{ij} < P_{ij}^*$	7,7%	25,4%
od 35 do 44 lat	$P_{ij} > P_{ij}^*$	1,5%	$P_{ij} < P_{ij}^*$	0,7%	2,2%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	5,5%	$P_{ij} > P_{ij}^*$	2,4%	8,0%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	17,0%	$P_{ij} > P_{ij}^*$	7,4%	24,3%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	15,2%	$P_{ij} > P_{ij}^*$	6,6%	21,8%
2009 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	14,5%	$P_{ij} < P_{ij}^*$	4,7%	19,2%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	19,9%	$P_{ij} < P_{ij}^*$	6,4%	26,4%
od 35 do 44 lat	$P_{ij} > P_{ij}^*$	1,0%	$P_{ij} < P_{ij}^*$	0,3%	1,3%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	6,7%	$P_{ij} > P_{ij}^*$	2,2%	8,9%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	17,6%	$P_{ij} > P_{ij}^*$	5,7%	23,3%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	15,8%	$P_{ij} > P_{ij}^*$	5,1%	20,9%
2008 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	18,3%	$P_{ij} < P_{ij}^*$	4,4%	22,7%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	23,1%	$P_{ij} < P_{ij}^*$	5,5%	28,6%
od 35 do 44 lat	$P_{ij} \approx P_{ij}^*$	0,0%	$P_{ij} \approx P_{ij}^*$	0,0%	0,0%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	8,3%	$P_{ij} > P_{ij}^*$	2,0%	10,3%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	16,4%	$P_{ij} > P_{ij}^*$	3,9%	20,3%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	14,6%	$P_{ij} > P_{ij}^*$	3,5%	18,0%
2007 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	20,6%	$P_{ij} < P_{ij}^*$	4,1%	24,7%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	24,6%	$P_{ij} < P_{ij}^*$	4,9%	29,5%
od 35 do 44 lat	$P_{ij} \approx P_{ij}^*$	0,2%	$P_{ij} \approx P_{ij}^*$	0,0%	0,3%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	9,3%	$P_{ij} > P_{ij}^*$	1,9%	11,1%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	14,7%	$P_{ij} > P_{ij}^*$	2,9%	17,6%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	14,0%	$P_{ij} > P_{ij}^*$	2,8%	16,7%

1	2		3		4
2006 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	19,5%	$P_{ij} < P_{ij}^*$	2,9%	22,4%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	26,9%	$P_{ij} < P_{ij}^*$	4,0%	30,8%
od 35 do 44 lat	$P_{ij} \approx P_{ij}^*$	0,1%	$P_{ij} \approx P_{ij}^*$	0,0%	0,2%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	9,0%	$P_{ij} > P_{ij}^*$	1,3%	10,3%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	16,3%	$P_{ij} > P_{ij}^*$	2,4%	18,8%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	15,3%	$P_{ij} > P_{ij}^*$	2,3%	17,5%
2005 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	31,3%	$P_{ij} < P_{ij}^*$	2,4%	33,8%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	13,6%	$P_{ij} < P_{ij}^*$	1,1%	14,7%
od 35 do 44 lat	$P_{ij} \approx P_{ij}^*$	0,5%	$P_{ij} \approx P_{ij}^*$	0,0%	0,5%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	10,5%	$P_{ij} > P_{ij}^*$	0,8%	11,3%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	17,6%	$P_{ij} > P_{ij}^*$	1,4%	18,9%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	19,2%	$P_{ij} > P_{ij}^*$	1,5%	20,7%
2004 r.					
od 16 do 24 lat	$P_{ij} > P_{ij}^*$	18,1%	$P_{ij} < P_{ij}^*$	0,9%	19,0%
od 25 do 34 lat	$P_{ij} > P_{ij}^*$	30,9%	$P_{ij} < P_{ij}^*$	1,6%	32,5%
od 35 do 44 lat	$P_{ij} \approx P_{ij}^*$	0,0%	$P_{ij} \approx P_{ij}^*$	0,0%	0,0%
od 45 do 54 lat	$P_{ij} < P_{ij}^*$	15,2%	$P_{ij} > P_{ij}^*$	0,8%	16,0%
od 55 do 64 lat	$P_{ij} < P_{ij}^*$	17,2%	$P_{ij} > P_{ij}^*$	0,9%	18,0%
od 65 do 74 lat	$P_{ij} < P_{ij}^*$	13,8%	$P_{ij} > P_{ij}^*$	0,7%	14,5%

Źródło: obliczenia własne na podstawie (*Spoleczeństwo...*, 2010a, s. 89; *Spoleczeństwo...*, 2010b, s. 99; *Spoleczeństwo...*, 2013, s. 141–142; *Mały...*, 2005, s. 112–113; *Mały...*, 2006, s. 112–113; *Rocznik...*, 2007, s. 140–141; *Rocznik...*, 2008, s. 140–141; *Rocznik...*, 2009, s. 134–135; *Rocznik...*, 2010, s. 134–135; *Rocznik...*, 2011, s. 130–131; *Rocznik...*, 2012, s. 132–133; *Rocznik...*, 2013, s. 136–137; *Rocznik...*, 2014, s. 140–141).

Analiza informacji zawartych w sporządzonej w ten sposób tablicy pozwoliła na wyciągnięcie wniosku, że osobami relatywnie najczęściej zaopatrującymi się w sieci są te, które mają od 25 do 34 lat. Z kolei względnie najrzadziej kupują przez Internet osoby mające co najmniej 55 lat.

Określenie zależności między wykształceniem a faktem kupowania przez Internet

Drugim zadaniem jest odpowiedź na pytanie, czy istnieje statystycznie istotna zależność między tym, jakie jest wykształcenie danej osoby a tym, czy dokonuje zakupów przez Internet? Procedurę weryfikacyjną dla hipotezy zerowej stanowiącej, że badane dwie cechy są stochastycznie niezależne, zrealizowano za pomocą testu niezależności chi-kwadrat analogicznie jak w podpunkcie 1, przy czym w tym przypadku zastosowano podział na trzy grupy – trzy poziomy wykształcenia. Dla każdego roku objętego analizą hipotezę H_0 trzeba było odrzucić na rzecz hipotezy alternatywnej orzekając jednocześnie, że rozpatrywane cechy są zależne⁸.

W tabeli 2 umieszczono informacje na temat relacji między poszczególnymi prawdopodobieństwami empirycznymi i odpowiadającymi im prawdopodobieństwami teoretycznymi oraz o udziale procentowym każdego elementu składającego się na otrzymaną wartość statystyki χ^2 .

Tabela 2

Wyniki uzyskane w badaniu zależności między wykształceniem i zakupami przez Internet

Wykształcenie	Kupują przez Internet	Nie kupują przez Internet	Razem		
1	2	3	4		
2013 r.					
podstawowe oraz gimnazjalne ¹	$p_{ij} < p_{ij}^*$	13,3%	$p_{ij} > p_{ij}^*$	6,1%	19,4%
zasadnicze zawodowe lub średnie ²	$p_{ij} < p_{ij}^*$	5,1%	$p_{ij} > p_{ij}^*$	2,3%	7,4%
wyższe ³	$p_{ij} > p_{ij}^*$	50,2%	$p_{ij} < p_{ij}^*$	23,0%	73,2%
2012 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	14,0%	$p_{ij} > p_{ij}^*$	6,1%	20,1%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	4,4%	$p_{ij} > p_{ij}^*$	1,9%	6,3%
wyższe	$p_{ij} > p_{ij}^*$	51,3%	$p_{ij} < p_{ij}^*$	22,4%	73,6%

⁸ Analogicznie jak w przypisie 5.

1	2	3	4		
2011 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	15,3%	$p_{ij} > p_{ij}^*$	6,5%	21,7%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	3,4%	$p_{ij} > p_{ij}^*$	1,5%	4,9%
wyższe	$p_{ij} > p_{ij}^*$	51,5%	$p_{ij} < p_{ij}^*$	21,9%	73,4%
2010 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	13,1%	$p_{ij} > p_{ij}^*$	5,4%	18,5%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	4,4%	$p_{ij} > p_{ij}^*$	1,8%	6,2%
wyższe	$p_{ij} > p_{ij}^*$	53,3%	$p_{ij} < p_{ij}^*$	22,0%	75,3%
2009 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	12,1%	$p_{ij} > p_{ij}^*$	3,7%	15,8%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	5,1%	$p_{ij} > p_{ij}^*$	1,6%	6,6%
wyższe	$p_{ij} > p_{ij}^*$	59,4%	$p_{ij} < p_{ij}^*$	18,2%	77,6%
2008 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	12,0%	$p_{ij} > p_{ij}^*$	2,7%	14,7%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	5,2%	$p_{ij} > p_{ij}^*$	1,2%	6,3%
wyższe	$p_{ij} > p_{ij}^*$	64,4%	$p_{ij} < p_{ij}^*$	14,6%	79,0%
2007 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	10,7%	$p_{ij} > p_{ij}^*$	2,1%	12,8%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	5,4%	$p_{ij} > p_{ij}^*$	1,1%	6,5%
wyższe	$p_{ij} > p_{ij}^*$	67,6%	$p_{ij} < p_{ij}^*$	13,2%	80,8%
2006 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	12,9%	$p_{ij} > p_{ij}^*$	1,9%	14,8%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	4,1%	$p_{ij} > p_{ij}^*$	0,6%	4,7%
wyższe	$p_{ij} > p_{ij}^*$	70,3%	$p_{ij} < p_{ij}^*$	10,2%	80,5%
2005 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	15,9%	$p_{ij} > p_{ij}^*$	1,2%	17,2%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	2,8%	$p_{ij} > p_{ij}^*$	0,2%	3,1%
wyższe	$p_{ij} > p_{ij}^*$	74,0%	$p_{ij} < p_{ij}^*$	5,7%	79,8%

1	2		3		4
2004 r.					
podstawowe oraz gimnazjalne	$p_{ij} < p_{ij}^*$	4,2%	$p_{ij} > p_{ij}^*$	0,2%	4,4%
zasadnicze zawodowe lub średnie	$p_{ij} < p_{ij}^*$	9,1%	$p_{ij} > p_{ij}^*$	0,5%	9,6%
wyższe	$p_{ij} > p_{ij}^*$	81,9%	$p_{ij} < p_{ij}^*$	4,1%	86,0%

¹ Zaliczono tutaj również osoby bez formalnego wykształcenia i z wykształceniem niepełnym podstawowym.

² Zaliczono tutaj osoby z wykształceniem zasadniczym zawodowym, z wykształceniem średnim technicznym lub ogólnokształcącym oraz z wykształceniem pomaturalnym.

³ Zaliczono tutaj osoby, które ukończyły studia z tytułem zawodowym licencjata, inżyniera lub magistra, ukończyły kolegium nauczycielskie albo uzyskały stopień tudzież tytuł naukowy (doktora, doktora habilitowanego czy profesora).

Źródło: obliczenia własne na podstawie: (*Spoleczeństwo...*, 2010a, s. 89; *Spoleczeństwo...*, 2010b, s. 99; *Spoleczeństwo...*, 2013, s. 142; *Mały...*, 2006, s. 115; *Mały...*, 2007, s. 118; *Mały...*, 2008, s. 118; *Mały...*, 2009, s. 118; *Mały...*, 2010, s. 118; *Mały...*, 2011, s. 121; *Mały...*, 2012, s. 124; *Mały...*, 2013, s. 124; *Mały...*, 2014, s. 120).

Transakcje przez Internet zdecydowanie najczęściej przeprowadzane były w grupie osób z wykształceniem wyższym, a najrzadziej w grupie osób bez żadnego wykształcenia szkolnego, z wykształceniem niepełnym podstawowym, pełnym podstawowym czy gimnazjalnym.

Określenie zależności między płcią a faktem kupowania przez Internet

Trzecim zadaniem jest odpowiedź na pytanie, czy istnieje statystycznie istotna zależność między tym, jaka jest płeć danej osoby a tym, czy dokonuje zakupów przez Internet? Dla każdego roku badanego okresu dziesięciu lat hipotezę o niezależności należało odrzucić na rzecz hipotezy alternatywnej. Pozwoliło to na konstatację, że rozpatrywane cechy są zależne⁹.

W tabeli 3 zebrano informacje na temat relacji między obliczonymi wartościami p_{ij} i p_{ij}^* oraz podano strukturę statystyki χ^2 .

⁹ Analogicznie jak w przypisie 5.

Tabela 3

Wyniki uzyskane w badaniu zależności między płcią i zakupami przez Internet

Płeć	Kupują przez Internet		Nie kupują przez Internet		Razem
1	2		3		4
2013 r.					
kobieta	$p_{ij} < p_{ij}^*$	33,5%	$p_{ij} > p_{ij}^*$	15,5%	49,0%
mężczyzna	$p_{ij} > p_{ij}^*$	34,8%	$p_{ij} < p_{ij}^*$	16,2%	51,0%
2012 r.					
kobieta	$p_{ij} < p_{ij}^*$	34,2%	$p_{ij} > p_{ij}^*$	14,9%	49,1%
mężczyzna	$p_{ij} > p_{ij}^*$	35,4%	$p_{ij} < p_{ij}^*$	15,5%	50,9%
2011 r.					
kobieta	$p_{ij} < p_{ij}^*$	34,5%	$p_{ij} > p_{ij}^*$	14,6%	49,1%
mężczyzna	$p_{ij} > p_{ij}^*$	35,7%	$p_{ij} < p_{ij}^*$	15,2%	50,9%
2010 r.					
kobieta	$p_{ij} < p_{ij}^*$	34,7%	$p_{ij} > p_{ij}^*$	14,2%	48,9%
mężczyzna	$p_{ij} > p_{ij}^*$	36,3%	$p_{ij} < p_{ij}^*$	14,8%	51,1%
2009 r.					
kobieta	$p_{ij} < p_{ij}^*$	37,5%	$p_{ij} > p_{ij}^*$	11,4%	48,9%
mężczyzna	$p_{ij} > p_{ij}^*$	39,2%	$p_{ij} < p_{ij}^*$	11,9%	51,1%
2008 r.					
kobieta	$p_{ij} < p_{ij}^*$	40,0%	$p_{ij} > p_{ij}^*$	8,8%	48,9%
mężczyzna	$p_{ij} > p_{ij}^*$	41,9%	$p_{ij} < p_{ij}^*$	9,2%	51,1%
2007 r.					
kobieta	$p_{ij} < p_{ij}^*$	41,2%	$p_{ij} > p_{ij}^*$	7,7%	48,9%
mężczyzna	$p_{ij} > p_{ij}^*$	43,1%	$p_{ij} < p_{ij}^*$	8,0%	51,1%
2006 r.					
kobieta	$p_{ij} < p_{ij}^*$	42,9%	$p_{ij} > p_{ij}^*$	6,0%	48,8%
mężczyzna	$p_{ij} > p_{ij}^*$	44,9%	$p_{ij} < p_{ij}^*$	6,3%	51,2%
2005 r.					
kobieta	$p_{ij} < p_{ij}^*$	45,4%	$p_{ij} > p_{ij}^*$	3,4%	48,8%
mężczyzna	$p_{ij} > p_{ij}^*$	47,6%	$p_{ij} < p_{ij}^*$	3,6%	51,2%

1	2		3		4
2004 r.					
kobieta	$p_{ij} < p_{ij}^*$	46,6%	$p_{ij} > p_{ij}^*$	2,2%	48,8%
mężczyzna	$p_{ij} > p_{ij}^*$	48,9%	$p_{ij} < p_{ij}^*$	2,3%	51,2%

Źródło: obliczenia własne na podstawie (*Spoleczeństwo...*, 2010a, s. 89; *Spoleczeństwo...*, 2010b, s. 99; *Spoleczeństwo...*, 2013, s. 141; *Mały...*, 2005, s. 112–113; *Mały...*, 2006, s. 112–113; *Rocznik...*, 2007, s. 140–141; *Rocznik...*, 2008, s. 140–141; *Rocznik...*, 2009, s. 134–135; *Rocznik...*, 2010, s. 134–135; *Rocznik...*, 2011, s. 130–131; *Rocznik...*, 2012, s. 132–133; *Rocznik...*, 2013, s. 136–137; *Rocznik...*, 2014, s. 140–141).

Na podstawie zawartości tak zbudowanej tablicy dotyczącej poszczególnych lat okresu 2004–2013 można stwierdzić, że odsetek mężczyzn kupujących przez Internet towary lub usługi konsumpcyjne jest większy niż odsetek kobiet.

Określenie zależności między miejscem zamieszkania a faktem kupowania przez Internet

Z kolei czwartym zadaniem jest odpowiedź na pytanie, czy istnieje statystycznie istotna zależność między tym, gdzie mieszka dana osoba, a tym, czy dokonuje zakupów przez Internet? W celu wykonania tego zadania badane osoby podzielono na trzy rozłączne i wyczerpujące grupy – mieszkańców dużych miast, mieszkańców mniejszych miast oraz mieszkańców wsi¹⁰.

W przypadku każdego roku z badanych dziesięciu lat sprawdzono – z wykorzystaniem testu chi-kwadrat – prawdziwość hipotezy zerowej o niezależności rozpatrywanych cech. Ponieważ obliczona wartość statystyki χ^2 znalazła się w prawostronnym obszarze krytycznym, toteż hipotezę H_0 odrzucono na korzyść hipotezy H_1 . Pozwoliło to na sformułowanie stwierdzenia, że analizowane cechy są zależne¹¹.

¹⁰ Duże miasta zdefiniowano jako te o liczbie mieszkańców przekraczającej 100 tys., natomiast mniejsze miasta to miasta mające do 100 tys. mieszkańców.

¹¹ Analogicznie jak w przypisie 5.

Tabela 4

Wyniki uzyskane w badaniu zależności między miejscem zamieszkania i zakupami przez Internet

Miejsce zamieszkania	Kupują przez Internet		Nie kupują przez Internet		Razem
1	2		3		4
2013 r.					
duże miasta	$p_{ij} > p_{ij}^*$	42,9%	$p_{ij} < p_{ij}^*$	19,7%	62,6%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,3%	$p_{ij} \approx p_{ij}^*$	0,1%	0,4%
obszary wiejskie	$p_{ij} < p_{ij}^*$	25,3%	$p_{ij} > p_{ij}^*$	11,6%	36,9%
2012 r.					
duże miasta	$p_{ij} > p_{ij}^*$	42,3%	$p_{ij} < p_{ij}^*$	18,1%	60,4%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,1%	$p_{ij} \approx p_{ij}^*$	0,0%	0,1%
obszary wiejskie	$p_{ij} < p_{ij}^*$	27,6%	$p_{ij} > p_{ij}^*$	11,8%	39,4%
2011 r.					
duże miasta	$p_{ij} > p_{ij}^*$	33,6%	$p_{ij} < p_{ij}^*$	14,0%	47,6%
mniejsze miasta	$p_{ij} > p_{ij}^*$	1,3%	$p_{ij} < p_{ij}^*$	0,6%	1,9%
obszary wiejskie	$p_{ij} < p_{ij}^*$	35,6%	$p_{ij} > p_{ij}^*$	14,9%	50,5%
2010 r.					
duże miasta	$p_{ij} > p_{ij}^*$	41,5%	$p_{ij} < p_{ij}^*$	16,6%	58,1%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,0%	$p_{ij} \approx p_{ij}^*$	0,0%	0,0%
obszary wiejskie	$p_{ij} < p_{ij}^*$	29,9%	$p_{ij} > p_{ij}^*$	12,0%	41,9%
2009 r.					
duże miasta	$p_{ij} > p_{ij}^*$	43,8%	$p_{ij} < p_{ij}^*$	12,9%	56,7%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,0%	$p_{ij} \approx p_{ij}^*$	0,0%	0,0%
obszary wiejskie	$p_{ij} < p_{ij}^*$	33,5%	$p_{ij} > p_{ij}^*$	9,9%	43,3%
2008 r.					
duże miasta	$p_{ij} > p_{ij}^*$	32,2%	$p_{ij} < p_{ij}^*$	6,9%	39,2%
mniejsze miasta	$p_{ij} > p_{ij}^*$	4,3%	$p_{ij} < p_{ij}^*$	0,9%	5,2%
obszary wiejskie	$p_{ij} < p_{ij}^*$	45,8%	$p_{ij} > p_{ij}^*$	9,8%	55,7%
2007 r.					
duże miasta	$p_{ij} > p_{ij}^*$	37,5%	$p_{ij} < p_{ij}^*$	6,8%	44,3%
mniejsze miasta	$p_{ij} > p_{ij}^*$	2,4%	$p_{ij} < p_{ij}^*$	0,4%	2,8%
obszary wiejskie	$p_{ij} < p_{ij}^*$	44,7%	$p_{ij} > p_{ij}^*$	8,2%	52,9%

1	2	3	4		
2006 r.					
duże miasta	$p_{ij} > p_{ij}^*$	49,9%	$p_{ij} < p_{ij}^*$	6,8%	56,7%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,0%	$p_{ij} \approx p_{ij}^*$	0,0%	0,0%
obszary wiejskie	$p_{ij} < p_{ij}^*$	38,2%	$p_{ij} > p_{ij}^*$	5,2%	43,3%
2005 r.					
duże miasta	$p_{ij} > p_{ij}^*$	51,6%	$p_{ij} < p_{ij}^*$	3,8%	55,4%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,1%	$p_{ij} \approx p_{ij}^*$	0,0%	0,1%
obszary wiejskie	$p_{ij} < p_{ij}^*$	41,5%	$p_{ij} > p_{ij}^*$	3,0%	44,5%
2004 r.					
duże miasta	$p_{ij} > p_{ij}^*$	49,7%	$p_{ij} < p_{ij}^*$	2,5%	52,2%
mniejsze miasta	$p_{ij} \approx p_{ij}^*$	0,4%	$p_{ij} \approx p_{ij}^*$	0,0%	0,4%
obszary wiejskie	$p_{ij} < p_{ij}^*$	45,1%	$p_{ij} > p_{ij}^*$	2,2%	47,4%

Źródło: obliczenia własne na podstawie: (*Spoleczeństwo...*, 2010a, s. 89; *Spoleczeństwo...*, 2010b, s. 99; *Spoleczeństwo...*, 2013, s. 142; *Rocznik...*, 2007, s. 90 i 102; *Rocznik...*, 2008, s. 90 i 102; *Rocznik...*, 2009, s. 84 i 96; *Rocznik...*, 2010, s. 84 i 96; *Rocznik...*, 2011, s. 80 i 92; *Rocznik...*, 2012, s. 82 i 94; *Rocznik...*, 2013, s. 86 i 98; *Rocznik...*, 2014, s. 90 i 102).

Na podstawie zawartości tak sporządzonej tablicy dla lat 2004–2013 można orzec, że zakupy przez Internet stosunkowo najczęściej realizowali mieszkańcy dużych miast. Z kolei względnie najrzadziej transakcje takie przeprowadzali mieszkańcy wsi.

Podsumowanie

Celem artykułu była odpowiedź na pytanie, jaki jest wpływ czterech wybranych czynników na fakt, że mieszkaniec Polski zamawia przez Internet towary lub usługi do użytku prywatnego. Aby osiągnąć postawiony cel zrealizowano cztery zadania badawcze. W wyniku przeprowadzenia odpowiednich testów istotności poszczególne hipotezy o niezależności zostały odrzucone. Stwierdzono zatem, że wiek, poziom wykształcenia, płeć i miejsce zamieszkania wpływają na zróżnicowanie Polaków pod względem tego, czy dokonują zakupów przez Internet. Dysponując już taką informacją warto więc dodatkowo sprawdzić, które z tych czynników mają największy wpływ na rozpatrywaną cechę, a w przypadku których wpływ jest najmniejszy. W tym celu dla każdej ze zmiennych oraz dla

każdego z dziesięciu lat 2004–2013 obliczono wartość statystyki empirycznej V . Otrzymane wyniki zaprezentowano w tabeli 5.

Tabela 5

Otrzymane wartości statystyki V

Rok	Statystyka V dotycząca zależności między dokonywaniem zakupów przez Internet oraz			
	wiekem	wykształceniem	plcią	miejscem zamieszkania
2004	0,15	0,21	0,07	0,12
2005	0,17	0,20	0,04	0,18
2006	0,27	0,26	0,06	0,19
2007	0,30	0,28	0,07	0,18
2008	0,32	0,30	0,05	0,16
2009	0,35	0,33	0,06	0,16
2010	0,39	0,32	0,06	0,17
2011	0,37	0,30	0,06	0,16
2012	0,40	0,34	0,04	0,17
2013	0,41	0,37	0,05	0,17

Źródło: obliczenia własne na podstawie: (*Spoleczeństwo...*, 2010a, s. 89; *Spoleczeństwo...*, 2010b, s. 99; *Spoleczeństwo...*, 2013, s. 141–142; *Mały...*, 2005, s. 112–113; *Mały...*, 2006, s. 112–113 i 115; *Mały...*, 2007, s. 118; *Mały...*, 2008, s. 118; *Mały...*, 2009, s. 118; *Mały...*, 2010, s. 118; *Mały...*, 2011, s. 121; *Mały...*, 2012, s. 124; *Mały...*, 2013, s. 124; *Mały...*, 2014, s. 120; *Rocznik...*, 2007, s. 90, 102, 140–141; *Rocznik...*, 2008, s. 90, 102, 140–141; *Rocznik...*, 2009, s. 84, 96, 134–135; *Rocznik...*, 2010, s. 84, 96, 134–135; *Rocznik...*, 2011, s. 80, 92, 130–131; *Rocznik...*, 2012, s. 82, 94, 132–133; *Rocznik...*, 2013, s. 86, 98, 136–137; *Rocznik...*, 2014, s. 90, 102, 140–141).

Obliczone dla poszczególnych lat wartości statystyki V pozwalają stwierdzić, że to czy mieszkaniec Polski zaopatruje się przez Internet w towary i usługi konsumpcyjne, w największym stopniu zależy od jego wieku i wykształcenia, w mniejszym stopniu zależy od jego miejsca zamieszkania, a wpływ płci jest najmniejszy ze wszystkich czterech cech wziętych pod uwagę w niniejszym badaniu. Na podstawie tabeli 5 można również sformułować kilka wniosków szczegółowych. Mianowicie na przestrzeni lat 2004–2013 siła zależności między badaną zmienną a:

- **wiek** – znacznie wzrosła i w 2013 roku była na poziomie umiarkowanym;
- **poziomem wykształcenia** – również rosła, ale wolniej, i w 2013 roku siła zależności między wykształceniem a faktem robienia zakupów przez Internet była prawie taka sama jak siła zależności między wiekiem a tymi zakupami;

- **plcią** – utrzymywała się na względnie stałym i niskim poziomie świadczącym o bardzo słabym wpływie płci na badaną zmienną;
- **miejszem zamieszkania** – utrzymywała się także na względnie stałym i raczej niskim poziomie świadczącym o dość słabym wpływie miejsca zamieszkania na analizowaną zmienną.

Bibliografia

- Aczel A.D. (2000), *Statystyka w zarządzaniu*, Wydawnictwo Naukowe PWN, Warszawa.
- Ankarali H., Cangur S. (2013), *Comparison of Pearson chi-square and log-likelihood ratio statistics in $R \times C$ tables with regard to type I error*, „International Journal of Basic and Clinical Studies”, no. 2 (2), s. 91–104.
- Budżety gospodarstw domowych w 2010 (2011; 2012) r.*, (2011; 2012; 2013), GUS, Warszawa.
- Carlin J.B., Doyle L.W. (2001), *Statistics for clinicians; 5: Comparing proportions using the chi-squared test*, „Journal of Paediatrics and Child Health”, vol. 37, no. 4, s. 392–394.
- Dowling Fink A.D., Lunsford M.L. (2010), *Water taste test data*, „Journal of Statistics Education”, vol. 18, no. 1, s. 1–18.
- Eurostat database, www.appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do?dvsc=8 (19.01.2015).
- Fura B., Fura M. (2011), *Zróżnicowanie kapitału intelektualnego pracowników w wieku przedemerytalnym*, „Nierówności Społeczne a Wzrost Gospodarczy”, nr 23, s. 40–49.
- Guty E., Mrozowicz A. (2013), *Styl życia a zachowania zdrowotne; poglądy kobiet*, „Zdrowie i Dobrostan”, nr 1, s. 133–147.
- Kukowska K. (2011), *Kapitał społeczny jako determinanta rozwoju przedsiębiorczości*, „Zarządzanie. Zeszyty Naukowe Politechniki Częstochowskiej”, nr 2, s. 9–22.
- Majewski P., Walczak D. (2013), *Świadomość ubezpieczeniowa studentów w aspekcie postrzegania rynku ubezpieczeniowego*, „Studia Oeconomica Posnaniensia”, vol. 1, no. 11 (260), s. 42–50.
- Mały Rocznik Statystyczny 2005 (2006; 2007; 2008; 2009; 2010; 2011; 2012; 2013; 2014), GUS, Warszawa.
- Onchiri S. (2013), *Conceptual model on application of chi-square test in education and social sciences*, „Global Journal of Art and Social Science Education”, vol. 1 (1), s. 16–26.

- Pułaska-Turyna B. (2005), *Statystyka dla ekonomistów*, Difin, Warszawa.
- Rocznik Demograficzny 2007 (2008; 2009; 2010; 2011; 2012; 2013; 2014), GUS, Warszawa.
- Spoleczeństwo informacyjne w Polsce. Wyniki badań statystycznych z lat 2004–2008 (2006–2010; 2009–2013)*, (2010a; 2010b; 2013), GUS, Warszawa.
- Sulewski P. (2013), *Modyfikacja testu niezależności*, „Wiadomości Statystyczne”, nr 10 (629), s. 1–19.
- Zeliaś A. (2000), *Metody statystyczne*, PWE, Warszawa.

FACTORS INFLUENCING PURCHASES VIA INTERNET

Summary

The aim of the article was to determine to what extent the selected factors (age, educational level, sex and domicile) influence the fact that a person uses the Internet for shopping. The research proved that age, educational level, sex and domicile are those variables that have a statistically significant impact on purchasing over the Internet by Poles. On the basis of the values of Cramer's V coefficient it was also found that – of these four variables – age and educational level are those which are mostly responsible for the dispersion of the statistical characteristic examined.

Translated by Anna Turczak

Keywords: shopping via Internet, age, educational level, domicile, sex

