Kazimierz Czerwiński
Uniwersytet Kazimierza Wielkiego w Bydgoszczy

Wybrane problemy analizy związków między zmiennymi w studenckich badaniach pedagogicznych

1. Uwagi wprowadzające

Analiza danych ilościowych, zgromadzonych w wyniku własnych badań prowadzonych przez studentów (głównie dla potrzeb prac licencjackich i magisterskich) bywa przez nich postrzegana jako niezwykle trudna i nierzadko będąca poza zasięgiem ich możliwości – zwłaszcza jeśli są to studenci kierunków humanistycznych. Dotyczy to szczególnie mocno studentów pedagogiki, ponieważ metodologia tej dyscypliny w wielu aspektach silnie zbliża się do metodologii socjologii i innych dyscyplin społecznych, a z drugiej strony – do metodologii badań psychologicznych. Zaś te dyscypliny obficie korzystają z instrumentarium metod i technik badawczych przynależnych do orientacji ilościowej – mimo ogłaszanego tu i ówdzie jej zmierzchu. Istota poruszanego tu problemu tkwi w tym, że pełna i pogłębiona analiza danych masowych nie jest możliwa bez użycia technik statystycznych – w wielu przypadkach wysoce zaawansowanych.

Wśród pytań badawczych tradycyjnie rozróżnia się pytania o stan (poziom) zmiennych oraz pytania o relacje (zależności, związki) między zmiennymi. Jeśli spojrzeć na ten podział z punktu widzenia najczęściej wymienianych funkcji nauki, jakimi są opis, wyjaśnianie i przewidywanie (niektórzy autorzy wymieniają eksplorację, opis i wyjaśnianie, np. Babbie 2003), to jasne jest, że pytania pierwszego typu (i odpowiedzi na nie) umożliwiają jedynie dokonywanie opisu, są to bowiem pytania pozwalające ustalić „jak jest". Wyjaśnianie zaś (i w perspektywie ewentualne przewidywanie) możliwe jest po postawieniu pytań, w których badacza interesują relacje między zmiennymi. Dopiero wtedy możliwe jest ustalenie „dlaczego tak jest", inaczej: jakie są przyczyny wyjaśnianego zjawiska.

Studenci, motywowani po trosze zdrową ambicją naukową, w większym zaś stopniu naciskami opiekunów swych prac, formułują często pokaźną liczbę pytań badawczych o zależności między zmiennymi. Jednakże późniejsza lektura analizy wyników studenckich badań i wypływających z nich wniosków pozostawia zazwyczaj ogromny niedosyt co do rozwiązania postawionych problemów. Niedosyt ów polega głównie na tym, że dane – zgromadzone najczęściej przy dużym nakładzie pracy i zaangażowaniu badacza, a także osób badanych – nie zostają należycie wykorzystane. Tym samym pokaźna ilość ludzkiej energii idzie na marne.

Dostępne na rynku wydawniczym liczne podręczniki czy poradniki dotyczące użycia statystyki dla potrzeb badań pedagogicznych są oczywiście potencjalnym remedium na poruszany tu problem – byle tylko studentom „chciało się chcieć" z nich skorzystać. Dlaczego zatem nie jest z tym najlepiej? Wydaje się, że można zaryzykować próbę dwojakiego typu wyjaśnień. Po pierwsze – niektóre z tych pozycji literaturowych wymagają całkiem sporego przygotowania matematycznego (np. Góralski 1987; Guilford 1960). Studenci dość otwarcie buntują się przeciwko temu, by zgłębiać elementy wyższej matematyki, używając argumentu, że są wszak studentami kierunku humanistycznego. Pozostawiając na uboczu dyskusję, czy pedagogika to dyscyplina raczej humanistyczna, czy raczej społeczna, należy tak czy owak uznać, że studenci mają sporo racji. Jako drugą przyczynę niedostatku statystycznych analiz zależności między zmiennymi w studenckich badaniach można wskazać tę, że ta problematyka bywa we wzmiankowanych podręcznikach niewystarczająco reprezentowana – przeważa zaś statystyczna analiza jednozmiennowa. Często też bywa i tak, że dany autor, pisząc rzecz o statystyce dla pedagogów, wkłada wiele wysiłku (udanego), by maksymalnie „uprzystępnić" humanistom procedury statystyczne – lecz owa przystępność kończy się wraz z zakończeniem rozdziałów poświęconych analizie jednej zmiennej; przy przejściu do analizy wielu (zwykle dwu) zmiennych, poziom przystępności obniża się.

Te przyczyny nie usprawiedliwiają oczywiście studentów w pełni, lecz jedynie po części. Zdolni czy ambitni studenci, albo mający predyspozycje do zgłębiania materii nauk ścisłych, albo bez takich predyspozycji, ale wkładający w to ogromną ilość wysiłku – jakoś sobie radzą i potrafią całkiem przyzwoicie (a czasem nawet na wysokim poziomie) wykorzystać zgromadzone dane i prawidłowo przeprowadzić wszystkie sensowne analizy między różnymi parami zmiennych. Ale tacy studenci wśród pedagogów są nieliczni, prawda zaś jest taka, że większość ma z tym problem.

Celem tego tekstu jest zachęcenie studentów – badaczy do szerszego korzystania z dostępnego instrumentarium statystycznych procedur umożliwiających analizę związków między zmiennymi. Autor jako pedagog z przygotowaniem matematycznym wyraża nadzieję, że potrafi uczynić to w sposób przystępny.

W dobie powszechnego wykorzystania elektronicznej techniki obliczanie jakichkolwiek wielkości statystycznych przestaje zresztą być problemem. Prawie wszystkie analizy, dokonane w oparciu o zaprezentowany w dalszych partiach tekstu przykład, są możliwe do wykonania już w dowolnym arkuszu kalkulacyjnym, nie wspominając o wyspecjalizowanych komputerowych programach statystycznych. Ale również kalkulatory programowalne, a nawet te nie mające takiej możliwości, lecz posiadające tzw. tryb statystyczny, znakomicie ułatwiają dokonywanie obliczeń („same” obliczają m.in. średnie czy odchylenia standardowe). Nie należy zresztą ignorować faktu, że jeśli przeciętny student (przeważnie: studentka) pedagogiki potrzebuje dla celów swojej pracy dokonać statystycznej analizy wyników swych badań, to i tak najczęściej robi to przy pomocy zaprzyjaźnionego studenta ekonomii, informatyki czy innego „bardziej ścisłego” kierunku. Sądzę, że nie należy z tego powodu rozpaczać. Problem nie leży w samej technicznej realizacji obliczeń statystycznych – te obliczenia zawsze ktoś lub coś wykona. Chodzi o to, by student pedagogiki umiał „dogadać się” – zarówno z owym zaprzyjaźnionym kolegą, jak i z maszyną.

Chodzi więc przede wszystkim o: 1) prawidłowe postawienie problemu (co obliczyć?, jakie procedury w danej sytuacji zastosować?), 2) prawidłowe odczytanie i interpretację wyniku lub wyników (jaka jest odpowiedź?, co ten wynik oznacza?), 3) rozumienie istotności lub jej braku (czy korelacja, zbieżność, różnica – jest statystycznie istotna?, na jakim poziomie istotności?, co to oznacza – głównie w aspekcie relacji: próba – populacja?), 4) poprawne wyciąganie wniosków (co uzyskane wyniki oznaczają z punktu widzenia weryfikacji sformułowanych wcześniej hipotez?).

Nie zależy nam więc na takich sytuacjach, w których student samodzielnie umie na przykład „obliczyć chi kwadrat”, lecz nie wie, w jakim celu. Dopuśćmy, by współczynnika chi-kwadrat samodzielnie nie umiał obliczać, niech zrobi to za niego maszyna, także z czyjąś pomocą. Ważne, by umiał przeprowadzić całe rozumowanie, które jest do wykonania „wokół” obliczanego współczynnika, by w efekcie tego mógł kompetentnie ustalić, czy należy postawioną hipotezę przyjąć, czy też odrzucić, oraz z jakim prawdopodobieństwem.

2. Wybrane pojęcia teorii weryfikacji hipotez statystycznych

Analiza zależności między zmiennymi prowadzona jest oczywiście w oparciu o dane empiryczne zebrane od wybranej do badania próbki osób (uczniów, wychowanków, nauczycieli itp.). Najbardziej wartościowe naukowo są zaś wnioski z przeprowadzonych analiz możliwie najbardziej ogólne. W badaniach studenckich zwykle nie jest osiągalny zbyt wysoki poziom uogólnień, tym niemniej każdy – nawet początkujący – badacz powinien mieć ambicje, by jego wnioski dotyczyły obszerniejszego zbioru ludzi, niż tylko zbiór ludzi objętych badaniami. Zakładam, że w tym kontekście znane są studentowi pojęcia: „próba” i „populacja”, a także problematyka takiego doboru próby do badań, by była ona reprezentatywna dla populacji. Wypowiadanie się o populacji na podstawie wyników uzyskanych z próby wymaga jednak posługiwania się (nieco sformalizowanym) językiem teorii weryfikacji hipotez statystycznych. Należy jednak podkreślić, że zaprezentowany niżej zwięzły przegląd podstawowych pojęć tej teorii dotyczy tylko hipotez odnoszących się do związków między zmiennymi.

Hipoteza statystyczna to przypuszczenie dotyczące związku między zmiennymi w populacji. Hipotezy formułuje się w dwóch postaciach:

– hipoteza zerowa (H_0) – założenie, że zmienne w populacji nie są zależne, inaczej: że stwierdzona na podstawie próby zależność nie jest statystycznie istotna, że jest zdarzeniem incydentalnym, przypadkowym, nie mającym charakteru powtarzalnej tendencji;
– hipoteza alternatywna (H_1) – założenie, że zmienne w populacji są zależne, inaczej: że stwierdzona na podstawie próby zależność jest statystycznie istotna, że nie jest zdarzeniem przypadkowym, przeciwnie – jest egzemplifikacją pewnej trwałej tendencji.

Hipotezy formułuje się, a następnie weryfikuje, na podstawie danych z próby, dotyczą zaś populacji. Nie można więc mieć pewności, że weryfikacja (przyjęcie bądź odrzucenie) zostaną wykonane bezbłędnie. Przeciwnie – należy się liczyć z błędami; statystyka wyróżnia tu dwa ich rodzaje.

Tab. 1. Rodzaje błędów występujących podczas weryfikacji hipotez statystycznych

Stan faktyczny (nieznany)	Decyzja podjęta w wyniku próby	
	przyjęcie H_0 (odrzucenie H_1)	odrzucenie H_0 (przyjęcie H_1)
H_0 prawdziwa (H_1 fałszywa)	decyzja prawidłowa	**błąd pierwszego rodzaju**
H_0 fałszywa (H_1 prawdziwa)	**błąd drugiego rodzaju**	decyzja prawidłowa

Zerowe prawdopodobieństwa obu błędów nie są możliwe, chodzi przynajmniej o to, aby je zminimalizować. Prawdopodobieństwo błędu pierwszego rodzaju oznaczane jest symbolem α i nosi nazwę: **poziom istotności**. Prawdopodobieństwo błędu drugiego rodzaju oznaczane jest symbolem β, zaś liczba $1 - \beta$ nosi nazwę: moc testu.

Test statystyczny to – w uproszczeniu – sposób weryfikacji hipotezy statystycznej, a bardziej dokładnie: reguła postępowania, która każdemu wynikowi próby pozwala – z określonym prawdopodobieństwem – przyporządkować decyzję przyjęcia bądź odrzucenia hipotezy zerowej. Testy statystyczne są tak skonstruowane, że pozwalają użytkownikowi kontrolować poziom istotności (α), jednocześnie (jakby „w tle”) zapewniają minimalizację prawdopodobieństwa błędu drugiego rodzaju (β).

Poziom istotności jest bardzo ważnym parametrem każdego testu, warto uświadomić sobie dokładniej jego znaczenie. Z poczynionych ustaleń wynika, że poziom istotności można zdefiniować jako prawdopodobieństwo odrzucenia hipotezy zerowej pomimo tego, że jest ona prawdziwa. Przy rozpatrywanych tutaj rodzajach hipotez i testów oznaczać to będzie, że gdy odrzucimy hipotezę zerową, to uznamy, że w populacji ma miejsce związek między zmiennymi. Jeśli nasza decyzja została podjęta – dajmy na to – na poziomie istotności 0,02, to oznacza, że dopuszczamy prawdopodobieństwo co najwyżej 2%, że nasza decyzja jest błędna, czyli że tak naprawdę ten związek nie zachodzi. Oczywistym wydaje się, że im poziom istotności jest mniejszy, tym lepiej (dokładnie biorąc, sprawa jest nieco bardziej skomplikowana – w pewnych sytuacjach przyjęcie bardzo niskiej wartości poziomu istotności może skutkować podwyższeniem prawdopodobieństwa błędu drugiego rodzaju – w naszych rozważaniach możemy jednak to niebezpieczeństwo pominąć). Powszechnie uznawana konwencja dopuszcza, by poziom istotności mógł wynosić co najwyżej 0,05. W ten sposób godzimy się na co najwyżej 5-procentowy „margines błędu” w naszych decyzjach weryfikacyjnych – lecz nie większy. Liczba 0,05 staje się tym samym pewną wielkością graniczną, wyznacza „linię podziału” między sytuacjami, w których uznajemy, że korelacja (zbieżność itp.) jest statystycznie istotna, a tymi, w których uznajemy, że jest statystycznie nieistotna. Pierwszy przypadek ma miejsce, gdy poziom istotności nie przekracza 0,05 (w formalnym zapisie: $\alpha \leq 0{,}05$), zaś drugi przypadek – gdy przekracza ($\alpha > 0{,}05$). W pierwszym przypadku nie należy ograniczać się do samego stwierdzenia, że zależność (korelacja, zbieżność itp.) jest statystycznie istotna, lecz należy podać konkretną wartość α. W przypadku zaś braku istotności domyślnie przyjmujemy, że α przekracza 0,05 i dokładna wartość już nas nie interesuje.

Drugim (pomocniczym) parametrem obecnym w testach statystycznych jest **liczba stopni swobody**, oznaczana zazwyczaj *df*. Definiuje się ją jako liczbę niezależnych porównań możliwych do wykonania na elementach próby – ale zagłębianie się w istotę tego parametru nie jest nam tutaj potrzebne. Każdy test statystyczny zawiera – jako integralny element – wzór na obliczanie liczby stopni swobody.

Test statystyczny przeprowadza się w czterech etapach:

1. Sformułowanie hipotez: zerowej i alternatywnej.
2. Obliczenie tzw. **sprawdzianu testowego**.

3. Ustalenie poziomu istotności (α), obliczenie liczby stopni swobody (*df*), a następnie – na podstawie tych dwu wielkości – odczytanie z odpowiedniej tabeli tzw. **wartości krytycznej**.
4. Podjęcie decyzji weryfikacyjnej.

Sprawdzian testowy to wielkość liczbowa obliczana według określonego wzoru – będącego integralną częścią danego testu. Tabele, z których odczytuje się wartość krytyczną, to – w referowanych sytuacjach – albo tabela rozkładu Studenta, albo tabela rozkładu chi-kwadrat; obie z reguły są zamieszczane jako załączniki w każdym podręczniku statystyki. Decyzja weryfikacyjna zaś przebiega według schematu: jeśli wartość sprawdzianu testowego (ewentualnie wartość bezwzględna – w przypadku otrzymania wartości ujemnej) jest większa lub równa niż wartość krytyczna, to należy odrzucić hipotezę zerową (czyli uznać, że zależność między zmiennymi jest statystycznie istotna – na poziomie istotności α), w przeciwnym przypadku brak jest podstaw do odrzucenia hipotezy zerowej (wówczas należy uznać, że zależność między zmiennymi jest statystycznie nieistotna).

Ten skrótowy przegląd najważniejszych ustaleń teorii weryfikacji hipotez statystycznych – będącej częścią teorii wnioskowania statystycznego – jest potrzebny dlatego, że prawidłowo prowadzona analiza zależności między zmiennymi zawiera w sobie zarówno elementy opisu, jak i wnioskowania. Ustalamy, jakiego typu zależność ma miejsce między elementami badanej próby (opis), by następnie dążyć do uogólnienia tych ustaleń na dużo większy niż zbadana próba zbiór osób (wnioskowanie).

3. Przykład danych do analizy

Zaprezentowany poniżej przykład będzie stanowił pewnego rodzaju pokaz możliwości dostępnych dla studenta analiz związków między różnymi parami zmiennych. W tabeli 2 zawarte są dane (umowne) zawierające fragment wyników z pewnego badania, którym objęto 20 uczniów gimnazjum. Analizie zostanie poddanych 5 zmiennych. Oto ich charakterystyka:

- Z.1 – **płeć** (1 – dziewczynka, 2 – chłopiec);
- Z.2 – **typ szkoły ponadgimnazjalnej** wskazywanej w dalszych planach edukacyjnych (1 – liceum ogólnokształcące, 2 – inna szkoła);
- Z.3 – **motywacja do nauki** (1 – bardzo wysoka, 2 – wysoka, 3 – przeciętna, 4 – niska, 5 – bardzo niska);
- Z.4 – **wynik testu humanistycznego** (w punktach);

– Z.5 – **wynik testu matematyczno-przyrodniczego** (w punktach).

Tab. 2. Dane (umowne) do analizy

Numer osoby badanej	Wartość zmiennej:				
	Z.1	Z.2	Z.3	Z.4	Z.5
1	1	1	1	45	37
2	2	1	1	42	50
3	1	2	3	40	35
4	1	1	2	36	26
5	1	1	1	34	42
6	1	1	2	32	20
7	2	1	3	31	50
8	1	1	3	30	22
9	2	1	3	28	35
10	1	1	4	27	13
11	2	2	4	26	44
12	1	2	3	25	34
13	2	2	3	25	28
14	1	1	4	25	17
15	1	1	4	23	6
16	2	1	2	22	48
17	2	2	2	19	17
18	1	2	5	19	9
19	2	2	5	14	37
20	2	2	5	11	22

Jest sprawą oczywistą, że w odniesieniu do trzech pierwszych zmiennych zawarte w tabeli liczby oznaczające ich „wartości” mają jedynie sens umowny – nie mogą być na nich wykonywane żadne działania matematyczne. Zamiast tych liczb można było oczywiście użyć oznaczeń literowych, na przykład „dz.” oraz „chł.” – w stosunku do płci, zaś dla zmiennej Z.3 mogły być użyte na przykład symbole: „bw”, „w”, „p” itd., lub jakiekolwiek inne. Świadomie zostały tu jednak zastosowane oznaczenia numeryczne, by uczulić Czytelników na czyhającą w tym kontekście pułapkę. Polega ona na tym, że na jednym z dalszych etapów przetwarzania danych badacz może zapomnieć o czysto konwencjonalnym charakterze użytych oznaczeń i zacznie te symbole traktować, jak gdyby miały one sens liczbowy – i policzy na przykład średnią dla zmiennej Z.1, po czym zacznie się zastanawiać, co ona właściwie oznacza. Tego typu błędy zdarzają się zresztą nie tylko studentom...

Warto jeszcze dodać, że zaprezentowany w tabeli układ danych jest typowy dla arkuszy kalkulacyjnych – wartości wszystkich zmiennych dotyczące określonej osoby badanej są umieszczone w jednym wierszu (noszącym ten sam numer, co ta osoba), natomiast wartości określonej zmiennej dla wszystkich osób badanych są umieszczone w określonej kolumnie. Dostępne w arkuszu kalkulacyjnym funkcje i procedury statystyczne domyślnie traktują bowiem każdą kolumnę danych jako zbiór wartości kolejnej zmiennej.

Na początku zazwyczaj przeprowadza się analizę każdej zmiennej z osobna. W tym więc przypadku można stwierdzić, że: a) badaniem objęto 11 dziewczynek i 9 chłopców, b) 12 uczniów zamierza uczyć się w liceum ogólnokształcącym, zaś 8 – w szkole innego typu, c) troje uczniów prezentuje bardzo wysoki poziom motywacji, czworo – poziom wysoki, sześcioro – przeciętny, czworo – niski oraz troje – bardzo niski (ze względu na to, że ta zmienna przyjmuje aż 5 kategorii, warto jej rozkład przedstawić w formie tabeli – co pozostawiam Czytelnikowi), d) wyniki testu humanistycznego zawarte są w zakresie od 11 do 45 punktów, wynik średni to 27,7, natomiast odchylenie standardowe wynosi 8,874 – co Czytelnik zechce sprawdzić (możliwe jest oczywiście również obliczenie bardziej zaawansowanych parametrów, jak współczynnik asymetrii czy współczynnik spłaszczenia, zwany kurtozą, można wskazać, że mediana ma wartość 26,5, można wreszcie zbudować – i to na kilka sposobów – przedziałowy szereg rozdzielczy, a następnie wskazać przedział dominujący oraz obliczyć częstości względne dla wszystkich przedziałów – te elementy analizy jednej zmiennej nie będą tu jednak rozwijane, przedmiotem tego tekstu jest bowiem analiza dwuzmiennowa), e) wyniki testu matematyczno-przyrodniczego zawarte są w zakresie od 6 do 50 punktów, wynik średni to 29,6, natomiast odchylenie standardowe wynosi 13,605 (powyższa uwaga na temat dalszych analiz pozostaje w mocy).

Zarówno w przypadku powyżej prezentowanych danych, jak i w przypadku obliczeń, które będą prezentowane poniżej, przyjmujemy konwencję zaokrąglania wyników z dokładnością do 3 miejsc po przecinku.

Przed przystąpieniem do badania związków między zmiennymi koniecznie należy ustalić poziom możliwości pomiarowych poszczególnych zmiennych – to będzie bowiem zasadniczo rzutować na wybór odpowiedniej procedury czy współczynnika. Zakładając, że Czytelnikowi jest znany dokonany przez Stevensa podział zmiennych na cztery kategorie, stwierdźmy, że zmienne Z.1 oraz Z.2 są rejestrowane na poziomie nominalnym (mówimy w uproszczeniu, że są to zmienne nominalne, także: cechy nominalne, cechy jakościowe itp.), zmienna Z.3 jest porządkowa (inaczej: rangowa), zaś zmienne Z.4 i Z.5 mają poziom ilorazowy (inaczej: stosunkowy).

4. Badanie zależności dwu zmiennych nominalnych

Na początek podnieśmy problem, czy plany dalszej edukacji badanych uczniów gimnazjum mają coś wspólnego z ich płcią. Pytanie badawcze – w sposób bardziej formalny – można sformułować na kilka sposobów, np.: „Jaki jest związek między płcią uczniów a wskazywanym przez nich typem szkoły ponadgimnazjalnej?” albo „Czy płeć uczniów różnicuje ich wybory dotyczące dalszego kształcenia?” albo wreszcie – korzystając z wprowadzonych oznaczeń – „Jaki jest związek między zmiennymi Z.1 oraz Z.2?”. Obie zmienne są nominalne, więc właściwą do analizy procedurą jest **test niezależności chi-kwadrat**.

Pierwszym krokiem jest zbudowanie **tablicy korelacyjnej**, zwanej też tablicą krzyżową, tablicą kombinowaną lub jeszcze w inny sposób. Jej istota sprowadza się do zaprezentowania podziału osób badanych na kategorie powstałe z jednoczesnego uwzględnienia wyróżnionych już wcześniej kategorii obu zmiennych. W naszym przykładzie obie zmienne są dychotomiczne (dwukategorialne), powstanie więc tablica czteropolowa.

Tab. 3. Zależność typu wybieranej szkoły od płci

Płeć	Typ wybieranej szkoły		Razem
	liceum ogólnokształcące	inna szkoła	
dziewczynka	**8**	**3**	11
chłopiec	**4**	**5**	9
Razem	12	8	20

Warto podkreślić, że zasadnicza tablica ma dwa wiersze i dwie kolumny danych, wiersz i kolumna oznaczone „Razem” są elementami dodatkowymi, chociaż do analizy bardzo przydatnymi. Dane zawarte w tych elementach dodatkowych noszą nazwę **rozkładów brzegowych** – są to po prostu opisane wcześniej rozkłady każdej ze zmiennych oddzielnie. Jeśli w parze analizowanych zmiennych potrafimy wyróżnić niezależną oraz zależną, to regułą jest, że kategorie tej pierwszej umieszczamy w tzw. boczku tabeli, natomiast kategorie zmiennej zależnej – w tzw. główce.

W zasadniczej części tablicy korelacyjnej dane zawarte w każdym wierszu (a także w każdej kolumnie) tworzą tzw. **rozkład warunkowy**. W tym przypadku liczby 8 i 3 informują o preferencjach co do szkoły wśród dziewczynek, natomiast 4 i 5 – u chłopców. Już samo porównanie ze sobą tych rozkładów warunkowych może pozwolić na (wstępną) odpowiedź na pytanie badawcze. Wśród dziewczynek stosunek wybierających liceum do wybierających inną szkołę wynosi 8 do 3, podczas gdy u chłopców ten stosunek jest zgoła inny: 4 do 5. Odpowiedź może brzmieć: „Tak, płeć różnicuje preferencje dotyczące dalszego kształcenia – dziewczynki stosunkowo częściej niż chłopcy wybierają liceum ogólnokształcące".

Jeśli chcemy znać siłę i istotność tej zależności, nie możemy na tym poprzestać. Należy obliczyć współczynnik chi-kwadrat – jako sprawdzian testu o tej samej nazwie. W przypadku tablicy czteropolowej – przy założeniu że kolejne komórki tabeli oznaczone są jako *a*, *b*, *c*, *d* – wzór przyjmuje postać:

$$\chi^2 = \frac{n(ad-bc)^2}{(a+b)(a+c)(b+d)(c+d)}.$$

W naszym przykładzie będzie to:

$$\chi^2 = \frac{20\cdot(8\cdot 5 - 3\cdot 4)^2}{11\cdot 12\cdot 8\cdot 9} = 1{,}650.$$

Liczbę stopni swobody w tym teście oblicza się według wzoru:

$$df = (w-1)(k-1),$$

gdzie: w – liczba wierszy, k – liczba kolumn. W naszym przykładzie $w = 2$, również $k = 2$, więc $df = 1$.

Dla liczby stopni swobody równej 1 oraz dla poziomu istotności α równego 0,05 odczytujemy z tablicy rozkładu chi-kwadrat wartość krytyczną: $\chi^2_{\alpha,df} = 3{,}841$. Stwierdzamy, że obliczona wartość sprawdzianu testowego jest mniejsza od wartości krytycznej. Na tej podstawie wnioskujemy, że brak jest podstaw do odrzucenia hipotezy zerowej, co – jak wcześniej zostało to objaśnione – oznacza, że zależność między tymi zmiennymi nie jest statystycznie istotna. Nie mamy więc żadnych podstaw by wypowiadać się o zależności tych dwu zmiennych w populacji, czy choćby w jakichś większych niż przebadany zbiorach uczniów. Stwierdziliśmy istnienie zależności wśród zbadanej próby 20 uczniów – i nic ponadto.

Można jeszcze ustalić siłę tej zależności, obliczając np. współczynnik zbieżności Cramera (oznaczany zazwyczaj *V*). W wersji dla tabeli czteropolowej wzór na jego obliczanie ma postać:

$$V = \sqrt{\frac{\chi^2}{n}}.$$

W naszym przykładzie:

$$V = \sqrt{\frac{1{,}65}{20}} = 0{,}287.$$

Oznacza to, że zależność jest słaba (według: Góralski 1987, s. 38). Ogólnie, współczynnik ten może przyjmować wartości od 0 do 1; im jego wartość jest bliższa jedności, tym siła zależności większa. Warto przy tym wspomnieć, że liczne podręczniki podają również inne, konkurencyjne współczynniki zbieżności, na przykład współczynnik zbieżności Czuprowa (*T*) lub współczynnik kontyngencji (*C*), jednak te współczynniki nie osiągają – nawet teoretycznie – wartości 1, co może powodować pewne trudności podczas interpretacji otrzymanej wartości.

Należy tutaj mocno podkreślić, że przypadek tablicy czteropolowej jest przypadkiem szczególnym. Ogólnie zaś obliczanie współczynnika chi-kwadrat przebiega w inny sposób, konieczne jest m.in. obliczenie tzw. liczebności oczekiwanych. Również wzór do obliczania współczynnika zbieżności Cramera w przypadku ogólnym ma nieco inną postać. Literatura statystyczna omawia tę kwestię szczegółowo (np. Ferguson, Takane 1997, s. 240–243; Góralski 1987, s. 35–38 i inne pozycje). Bez względu jednak na sposób obliczania, zaprezentowany powyżej sposób rozumowania pozostaje bez istotnych zmian.

Kolejne dwie pary zmiennych, które można w podobny sposób poddać badaniu z punktu widzenia zależności, to Z.1 i Z.3 oraz Z.2 i Z.3. W obu przypadkach będzie miało miejsce badanie zależności zmiennej nominalnej oraz porządkowej – a więc zmiennych będących na różnych poziomach według klasyfikacji Stevensa. W takiej sytuacji można postąpić tak, jak z dwiema zmiennymi nominalnymi, konstruując tablicę korelacyjną. Dla zmiennych Z.1 i Z.3 taka tablica ma postać:

Tab. 4. Zależność poziomu motywacji od płci

Płeć	Motywacja do nauki					Razem
	bardzo wysoka	wysoka	przeciętna	niska	bardzo niska	
dziewczynka	**2**	**2**	**3**	**3**	**1**	11
chłopiec	**1**	**2**	**3**	**1**	**2**	9
Razem	3	4	6	4	3	20

Porównując rozkłady warunkowe (rozkład motywacji u dziewczynek z rozkładem motywacji u chłopców) nie stwierdzamy znaczących różnic – te występują przede wszystkim w skrajnych kategoriach motywacji. Taka przybliżona analiza danych zawartych w powyższej tabeli jest jednak wszystkim, co można w tej sytuacji uczynić, powodem jest bardzo mała liczebność próby. Nie jest możliwe obliczanie współczynnika chi-kwadrat w oparciu o powyższe dane, ponieważ nie jest spełniony podstawowy do tego warunek – liczebności w poszczególnych komórkach tablicy korelacyjnej są zbyt małe.

Literatura orzeka, że wszystkie liczebności powinny mieć wartość co najmniej 5 (według zaś niektórych autorów: 8). Dokładniej rzecz ujmując, chodzi o liczebności oczekiwane, które należałoby dopiero obliczać przed przystąpieniem do obliczania współczynnika chi-kwadrat, niektórzy autorzy dopuszczają jednak rozpatrywanie w tym aspekcie liczebności empirycznych, czyli zawartych w tabeli. Jeśli warunek minimalnych liczebności nie jest spełniony, należy łączyć ze sobą sąsiednie kategorie, bądź też – gdy odstępstwo od warunku jest niewielkie – można podczas obliczeń stosować pewne udokładniające poprawki, na przykład tzw. poprawkę Yatesa na nieciągłość (patrz np. Krajewska 2001, s. 142). Z tego wynika w każdym razie, że na tabele korelacyjne o dużych liczbach komórek można sobie pozwolić jedynie przy bardzo dużych liczebnościach prób. Uważny i wnikliwy Czytelnik zapewne zresztą zauważył, że również w zbudowanej wcześniej czteropolowej tablicy korelacyjnej ów warunek nie był tak do końca spełniony.

5. Badanie zależności dwu zmiennych interwałowych lub ilorazowych

Kolejno poddajmy analizie parę zmiennych Z.4 oraz Z.5. W przypadku zmiennych mających charakter liczbowy (interwałowych lub ilorazowych) zasadne jest użycie słowa korelacja. Pytanie badawcze może więc zostać sformułowane tradycyjnie: „Jaki jest związek między wynikami testów: humanistycznego i matematyczno-przyrodniczego?", albo też: „W jaki sposób korelują ze sobą zmienne Z.4 oraz Z.5?". Jeśli korelacja między tymi zmiennymi ma charakter liniowy (w odróżnieniu od krzywoliniowego), właściwą jej miarą jest **współczynnik korelacji liniowej Pearsona** (zwany też współczynnikiem korelacji według momentu iloczynowego Pearsona, albo też krótko współczynnikiem korelacji Pearsona).

Sprawdzanie liniowego charakteru związku między zmiennymi jest tym krytycznym momentem, który zazwyczaj bywa pomijany przez pedagogów (i nie tylko przez nich). Tymczasem sprawa wcale nie jest trudna. Dopuszczalne jest przyjęcie założenia, że korelacja jest liniowa, gdy obie zmienne mają rozkład normalny lub bardzo zbliżony (jednomodalny i w miarę symetryczny). Jeśli w warunkach analizowanego przykładu Czytelnik dokona zestawienia każdej z tych zmiennych w przedziałowy szereg rozdzielczy, to może dostrzec, że wymagany warunek jest w zasadzie spełniony. Drugim sposobem jest odwołanie się do teorii – z dydaktyki, a dokładniej z teorii pomiaru dydaktycznego wiadomo, że zmienna „wiedza" ma (w populacji) rozkład normalny. Bez dalszych dodatkowych analiz możemy więc zastosować ten współczynnik.

Współczynnik korelacji Pearsona przyjmuje wartości od –1 do 1. Otrzymany w wyniku obliczeń znak (dodatni lub ujemny) informuje bezpośrednio o znaku korelacji. Korelacja dodatnia występuje wówczas, gdy zwiększaniu wartości jednej zmiennej towarzyszy (statystycznie, a nie bezwyjątkowo) również zwiększanie wartości drugiej zmiennej. Korelacja ujemna natomiast oznacza, że wzrostowi wartości jednej zmiennej towarzyszy spadek wartości drugiej. Wartość bezwzględna tego współczynnika informuje natomiast o sile korelacji – im bardziej jego wartość jest oddalona od zera (w stronę liczb –1 lub 1), tym korelacja jest silniejsza.

Wzór do obliczania współczynnika korelacji Pearsona (oznaczanego zwykle *r*) może zostać przedstawiony w następującej postaci:

$$r = \frac{C_{XY}}{S_X \cdot S_Y},$$

gdzie: C_{XY} – kowariancja między zmiennymi *X* i *Y*, natomiast S_X oraz S_Y – odchylenia standardowe odpowiednio zmiennej *X* i *Y*. W literaturze występują najczęściej inne wzory – jednakże zaprezentowany tutaj jest w pełni z nimi zgodny, odwołuje się natomiast do (zakładam, że znanego) pojęcia odchylenia standardowego. Wzór na obliczanie kowariancji:

$$C_{XY} = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i \cdot y_i - n \cdot \overline{X} \cdot \overline{Y}\right),$$

gdzie: $\overline{X}, \overline{Y}$ - średnie odpowiednio zmiennej *X* oraz *Y*.

W zaprezentowanych wzorach użyto oznaczeń *X* i *Y*, dla celów obliczeniowych przyjmijmy więc, że zmienną Z.4 oznaczymy jako *X*, natomiast zmienną Z.5 – jako *Y*.

Korzystając z podanych wcześniej wartości średnich ($\overline{X}$ = 27,7 oraz $\overline{Y}$ = 29,6) i odchyleń standardowych (S_X = 8,874 oraz S_Y = 13,605) obu zmiennych obliczmy najpierw kowariancję:

$$C_{XY} = \frac{1}{19}\left(45 \cdot 37 + 42 \cdot 50 + ... + 11 \cdot 22 - 20 \cdot 27{,}7 \cdot 29{,}6\right) = 46{,}242.$$

Następnie podstawmy dane do wzoru na współczynnik korelacji:

$$r = \frac{46{,}242}{8{,}874 \cdot 13{,}605} = 0{,}383.$$

Obliczanie współczynnika korelacji Pearsona bajecznie łatwo wykonuje się przy użyciu arkusza kalkulacyjnego, nie przywiązujmy więc tutaj nadmiernej wagi do technicznej realizacji obliczeń. Ustaliliśmy w każdym razie, że między wynikami obu testów ma miejsce korelacja dodatnia – mimo istnienia licznych wyjątków (każdy z Czytelników może je łatwo wskazać) „masowa” tendencja w badanej grupie 20 uczniów jest taka, że lepsi z testu humanistycznego są także lepsi z testu matematyczno-przyrodniczego i na odwrót (gorsi z jednego testu są gorsi także z drugiego). Wspomniane liczne wyjątki powodują zaś, że siła tej korelacji nie jest zbyt wysoka. Według rozpowszechnionej w literaturze interpretacji pochodzącej od J.P. Guilforda ma miejsce „wyraźna lecz mała" siła związku (Guilford 1960, s. 171).

Kolejnym ważnym krokiem jest ustalenie istotności obliczonego współczynnika. Wzór do obliczania sprawdzianu testowego ma w tym przypadku postać:

$$t = r \cdot \sqrt{\frac{n-2}{1-r^2}}\,.$$

Natomiast wartość krytyczną odczytuje się z tabeli rozkładu Studenta w oparciu o ustalony poziom istotności α oraz liczbę stopni swobody df, obliczaną według wzoru: $df = n - 2$.

Podstawiając dane, otrzymujemy:

$$t = 0{,}383 \cdot \sqrt{\frac{18}{1-0{,}383^2}} = 1{,}759.$$

Dla poziomu istotności α równego 0,05 oraz liczby stopni swobody df równej 18 odczytujemy z tabeli Studenta wartość krytyczną: $t_{\alpha,df}$ = 2,101. Stwierdzamy, że obliczona wartość sprawdzianu testowego jest mniejsza od wartości krytycznej. Oznacza to, że korelacja jest statystycznie nieistotna. Podobnie jak w przypadku zależności rozpatrywanej poprzednio pary zmiennych, stwierdziliśmy istnienie kolejnej zależności, ale tylko w badanej próbie uczniów, nic nas natomiast nie upoważnia do wyciągania wniosku, że takowa zależność ma miejsce w populacji.

Jako ćwiczenie proponuję Czytelnikowi samodzielne (najlepiej przy pomocy komputera) obliczenie współczynników korelacji między tymi samymi zmiennymi, ale dla innych – mniej licznych – zbiorów osób: dla dziewczynek, dla chłopców, dla uczniów wybierających liceum, dla uczniów wybierających inną szkołę. To oczywiście wymaga każdorazowo wstępnej czynności polegającej na „wybraniu" danych dotyczących zmiennych Z.4 i Z.5, ale tylko dla danej grupy osób. W arkuszach kalkulacyjnych nieocenioną usługę oddaje w tym momencie opcja sortowania danych ze względu na wskazaną kolumnę.

Dla kontroli poprawności obliczeń podaję prawidłowe wyniki. Dla grupy dziewczynek: r = 0,736, dla grupy chłopców: r = 0,649, dla grupy uczniów wybierających liceum ogólnokształcące: r = 0,412, dla grupy uczniów wybierających inną szkołę: r = 0,402.

W tym kontekście można pokusić się o próbkę swojego rodzaju analizy trójzmiennowej. Można mianowicie pytać, czy zmienne Z.1 (płeć) oraz Z.2 (typ wybieranej szkoły) mają jakiś wpływ na związek między zmiennymi Z.4 i Z.5. Innymi słowy, czy (każda z osobna) zmienne Z.1 i Z.2 modyfikują ów związek. Analizując podane wyżej wartości łatwo zauważyć, że typ wybieranej szkoły, czyli zmienna Z.2 nie modyfikuje znacząco związku między Z.4 i Z.5, ponieważ wartości współczynników korelacji – dla obu podzbiorów oraz dla całego badanego zbioru uczniów – są dość zbliżone. Inaczej przedstawia się sprawa z wpływem płci na analizowany związek między wynikami testów: współczynniki korelacji obliczone osobno dla dziewczynek (0,736) i osobno dla chłopców (0,649) są wyraźnie wyższe od współczynnika obliczonego dla wszystkich uczniów wspólnie (0,383). Płeć może tu więc zostać uznana za tzw. zmienną modyfikującą albo interweniującą.

Warto przy okazji zauważyć, że korelacja w grupie dziewczynek jest statystycznie istotna i to aż na poziomie istotności 0,01, zaś pozostałe współczynniki korelacji są statystycznie nieistotne – sprawdzenie tego faktu pozostawiam Czytelnikowi.

6. Badanie zależności dwu zmiennych porządkowych

Jaki wpływ na wyniki testów ma poziom motywacji do nauki? Pytamy tu o ewentualny wpływ zmiennej Z.3 na zmienne: Z.4 oraz Z.5. Rozpatrzmy na początek tylko jedną parę zmiennych: motywacja (Z.3) i wynik testu humanistycznego (Z.4). Każda z tych zmiennych posiada inne możliwości pomiarowe; zmienna Z.3 jest porządkowa, natomiast Z.4 – ilorazowa. W sytuacji, gdy zmienne są na różnych poziomach, należy wybrać procedurę statystyczną właściwą dla zmiennej będącej na niższym poziomie. Odpowiedni jest w tej sytuacji **współczynnik korelacji rangowej Spearmana** (w literaturze opisany jest również inny – niejako konkurencyjny – współczynnik korelacji rangowej Kendalla). Współczynnik korelacji rangowej oblicza się dla pary zmiennych porządkowych (inaczej: rangowych), a więc takich, które pozwalają jedynie porządkować badaną grupę osób ze względu na określone kryterium. Tym samym zmienną Z.4 – ilorazową – będziemy musieli potraktować jak porządkową; abstrahując od faktycznie uzyskanych w teście przez poszczególnych uczniów wartości punktowych, będziemy się interesować jedynie ich kolejnością (rangą) – ma przy tym miejsce pewna utrata informacji.

Współczynnik korelacji rangowej Spearmana, podobnie jak współczynnik Pearsona, może przyjmować wartości od –1 do 1; interpretacja, zarówno znaku, jak i bezwzględnej wartości, jest taka sama dla obu współczynników. Tak samo sprawdza się również istotność. Wzór na obliczanie współczynnika Spearmana (oznaczanego zazwyczaj r_S) ma postać:

$$r_S = 1 - \frac{6 \cdot \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)},$$

gdzie: d_i oznacza różnicę rang i-tej osoby – ustalanych ze względu na każdą zmienną oddzielnie.

Przed przystąpieniem do obliczania samego współczynnika należy więc dokonać wstępnej operacji zwanej rangowaniem obu zmiennych. Wykonajmy to najpierw dla zmiennej Z.4, ponieważ jest to łatwiejsze. Zakładając, że osobom uzyskującym lepszy rezultat w teście przyznamy wyższą rangę (jest to kwestia umowna – rangowanie „w przeciwnym kierunku" da w efekcie taką samą wartość bezwzględną współczynnika korelacji, lecz przeciwny znak – ale interpretacja wyniku pozostanie taka sama), zauważymy, że dla większości osób badanych numer tej osoby jest jednocześnie jej rangą. Jest tak dlatego, że podane w tabeli 2 dane są uporządkowane nierosnąco właśnie ze względu na wartości zmiennej Z.4. Tak więc osobie numer 1 przyznamy rangę 1, osobie numer 2 – rangę 2 itd., inaczej sprawa przedstawia się, gdy dojdziemy do osoby numer 12. Zauważymy, że trzy osoby: o numerach 12, 13 i 14 uzyskały w teście ten sam rezultat: 25 punktów. Zajęły więc one trzy równorzędne lokaty. Jeśli tak, to musimy im przyznać jednakową rangę. Ponieważ ta wspólna lokata może zostać określona jako dzielone miejsce od 12 do 14, nadajemy wszystkim tym trzem osobom rangę będącą średnią tych lokat, a więc 13. Osobom badanym o numerach 15 i 16 nadajemy znowu rangi odpowiadające ich numerom, następnie ponownie napotykamy na równorzędne lokaty – osoby badane o numerach 17 i 18 uzyskały w teście taką samą liczbę punktów. Postępując analogicznie jak poprzednio, przyznajemy im obu taką samą rangę – tym razem 17,5. Z dwiema ostatnimi osobami sprawa jest już jasna (rangi 19 oraz 20). W literaturze owe wspólne lokaty określane są zazwyczaj jako rangi wiązane.

W przypadku zmiennej Z.3 łatwo zauważyć, że będą występować wyłącznie rangi wiązane, bowiem żadna z osób badanych nie została zaliczona do którejś z kategorii „w pojedynkę". Przyznając (co wydaje się rozsądne) wyższe rangi osobom mającym wyższą motywację, zauważamy, że trzem osobom – o numerach 1, 2 i 5 – trzeba przyznać jednakową, najwyższą rangę. Ponieważ tych osób jest 3, więc z punktu widzenia tej zmiennej zajęły one wspólnie lokaty od 1 do 3, więc należy im nadać rangę 2. Kolejno są cztery osoby – o numerach 4, 6, 16 i 17 – które zajęły wspólnie lokaty od 4 do 7, więc średnia ranga, przynależna wszystkim tym czterem osobom, to 5,5 itd. Czytelnik zechce samodzielnie dokończyć tę – z pozoru skomplikowaną – operację. Kolejno dla każdej osoby badanej należy obliczyć różnicę jej rang, każdą z tych różnic należy podnieść do kwadratu, a następnie zsumować. Dla ułatwienia wszystkie te czynności warto jest wykonywać w tabelce – literatura dostarcza tu licznych wzorców.

W warunkach naszego przykładu suma kwadratów różnic rang osiąga wartość 416,5 (można mieć nadzieję, że Czytelnik potrafi dojść do tej liczby samodzielnie). Podstawiając dane do wzoru na współczynnik korelacji, otrzymujemy:

$$r_S = 1 - \frac{6 \cdot 416,5}{20 \cdot \left(20^2 - 1\right)} = 0,687\,.$$

Ustaliliśmy więc, że między motywacją do nauki a wynikiem testu humanistycznego zachodzi wyraźna korelacja dodatnia (interpretacja otrzymanej wartości liczbowej oparta jest – podobnie jak poprzednio – na propozycji J.P. Guilforda). Tak więc lepszej motywacji – mimo pewnej liczby wyjątków – towarzyszą lepsze rezultaty testu (co zresztą nie wydaje się nazbyt odkrywcze).

Badając istotność tej korelacji obliczamy najpierw sprawdzian testowy, którego wartość (obliczenia pozostawiam Czytelnikowi) wynosi 4,011. Dla liczby stopni swobody df równej 18 oraz poziomu istotności α równego 0,05 z tabeli rozkładu Studenta odczytujemy wartość krytyczną: $t_{\alpha,df} = 2,101$. Obliczona wartość jest większa od tabelarycznej, można więc odrzucić hipotezę zerową, inaczej mówiąc: korelacja jest statystycznie istotna. Ale czy koniecznie na poziomie istotności 0,05? Okazuje się, że nawet gdy weźmiemy poziom istotności równy 0,001, wartość krytyczna (oczywiście przy tej samej liczbie stopni swobody) wynosi: $t_{\alpha,df} = 3,992$ i w dalszym ciągu obliczona wartość sprawdzianu testowego jest większa od krytycznej. Konkludując: korelacja jest istotna statystycznie na poziomie istotności 0,001. Istotność zaś oznacza, że – jeśli badana próba gimnazjalistów została pobrana do badań w sposób reprezentatywny dla jakiejś populacji – korelacja jest „ważna” także dla tej populacji, a prawdopodobieństwo tego, że jest inaczej, nie przekracza 0,001, czyli 0,1%.

Podobnie jak w przypadku poprzednio ustalanej korelacji, można zapytać o ewentualny interweniujący wpływ płci. Zachęcając Czytelnika do samodzielnych obliczeń, podam wyniki. Dla dziewczynek wartość współczynnika korelacji rangowej pomiędzy zmiennymi Z.3 i Z.4 wynosi 0,834 i ta korelacja jest istotna na poziomie 0,002, natomiast dla chłopców wartość współczynnika korelacji wynosi 0,542 i jest nieistotna statystycznie. Można się więc pokusić o (ostrożny) wniosek, że płeć nieco modyfikuje związek między motywacją do nauki a wynikami testu humanistycznego.

Badając analogicznie wpływ zmiennej Z.2, stwierdzamy, że wśród uczniów wybierających liceum ogólnokształcące wartość współczynnika korelacji rangowej między zmiennymi Z.3 i Z.4 wynosi 0,762 (i ta korelacja jest statystycznie istotna na poziomie 0,01), natomiast wśród wybierających inną szkołę współczynnik ma wartość 0,560 (korelacja nieistotna). Sprawdzenie tych wyników oraz wnioski pozostawiam Czytelnikowi.

Analogicznie można również obliczać współczynnik korelacji rangowej dla pary zmiennych: Z.3 i Z.5. Mając nadzieję, że Czytelnik samodzielnie wykona obliczenia i analizy dotyczące interpretacji i istotności otrzymanych współczynników, podaję wyniki. Dla wszystkich 20 osób badanych współczynnik korelacji rangowej zmiennych Z.3 i Z.5 wynosi 0,448 (istotny na poziomie 0,05), dla dziewczynek: 0,845 (istotny na poziomie 0,002), dla chłopców: 0,321 (nieistotny), dla wybierających liceum: 0,701 (istotny na poziomie 0,02), oraz dla wybierających inną szkołę: 0 (to rzadki przypadek, by otrzymać dokładnie zerową wartość współczynnika korelacji).

7. Badanie zależności zmiennej ilorazowej (lub interwałowej) i zmiennej nominalnej

Kolejno poddajmy analizie parę zmiennych: Z.1 oraz Z.4. Pytamy więc, jaki jest związek między płcią uczniów a ich wynikami testu humanistycznego, albo inaczej: czy (i ewentualnie w jaki sposób) płeć różnicuje wyniki testu. Płeć jest zmienną nominalną, zaś wynik testu – ilorazową. Badanie zależności można przeprowadzić na dwa – całkowicie odmienne – sposoby.

Pierwszy sposób polega na tym, by „wyrównać" poziom obu zmiennych (w sensie Stevensa) – a więc w tym przypadku zmienną Z.4 potraktować jak nominalną – abstrahując od faktycznych wartości punktowych wyróżnić jedynie kilka rozłącznych kategorii. Liczba tych kategorii jest zależna od liczebności próby, bowiem – jak wiadomo – po zbudowaniu tablicy korelacyjnej nie powinniśmy mieć w niej pól o zbyt małej liczebności. Przy 20 osobach badanych na pewno nie możemy sobie pozwolić na więcej niż dwie kategorie zmiennej Z.4. Najprościej wykonać to definiując kategorie: „wysoki wynik" oraz „niski wynik". Kryterium podziału może stanowić mediana albo średnia. W pierwszym przypadku umowna granica będzie przebiegać w połowie zbioru osób badanych (oczywiście uporządkowanych ze względu na wartości tej zmiennej – w tabeli 2 osoby badane są już tak uporządkowane), czyli między osobami o numerach 10 i 11. W drugim przypadku – ze względu na wartość średniej wynoszącą 27,7 – granica przebiega między osobami o numerach 9 i 10, wyróżniamy więc kategorię powyżej średniej (wysoki wynik) i poniżej średniej (niski wynik). Bez względu na wybrane kryterium podziału w dalszej kolejności należy zbudować (czteropolową) tablicę korelacyjną i przeprowadzić, znany nam już, test niezależności chi-kwadrat. Czytelnik zechce sprawdzić, że w obu wariantach zależność między zmiennymi jest statystycznie nieistotna.

Drugi sposób badania zależności między zmienną liczbową (interwałową lub ilorazową) a nominalną polega na przeprowadzeniu specjalnej procedury statystycznej, zwanej **analizą wariancji** – w tej sytuacji: jednoczynnikową analizą wariancji. Trochę wbrew nazwie istota tej procedury polega na badaniu, w jakim stopniu istotne jest zróżnicowanie średnich zmiennej liczbowej, obliczanych osobno dla każdej z kategorii zmiennej nominalnej. W naszym przypadku trzeba obliczyć średnie wyniki zmiennej Z.4 dla dziewczynek i dla chłopców, a następnie sprawdzić, czy różnica między nimi jest statystycznie istotna.

Jeśli zmienna nominalna jest dychotomiczna, to jednoczynnikowa analiza wariancji jest tożsama z inną procedurą, którą jest **test istotności różnicy między dwoma średnimi** – w wersji dla prób niezależnych (nieskorelowanych). Jego istota polega na tym, że dla dwu niezależnych prób o liczebnościach n_1 i n_2 obliczamy średnie ($\overline{X}_1$ i $\overline{X}_2$) oraz odchylenia standardowe (S_1 i S_2) zmiennej liczbowej, następnie obliczamy sprawdzian testowy według wzoru:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}.$$

Obliczoną wielkość porównujemy z – odczytaną z tabeli rozkładu Studenta – wartością krytyczną, a następnie podejmujemy decyzję weryfikacyjną według ogólnych zasad. Liczba stopni swobody w tym teście obliczana jest według wzoru: $df = n_1 + n_2 - 2$.

Według danych zawartych w tabeli 2, co Czytelnik zechce sprawdzić, średnia zmiennej Z.4 dla dziewczynek wynosi 30,545, zaś dla chłopców: 24,222. Dziewczynki osiągnęły więc lepszy rezultat niż chłopcy. Aby sprawdzić istotność tego zróżnicowania, potrzebne są jeszcze – jak wynika z powyższego wzoru – kwadraty odchyleń standardowych, czyli wariancje. Ich wartości, to odpowiednio: 60,673 oraz 86,444. Podstawiając te dane do wzoru na sprawdzian testowy, otrzymujemy:

$$t = \frac{30{,}545 - 24{,}222}{\sqrt{\frac{60{,}673}{11} + \frac{86{,}444}{9}}} = 1{,}626.$$

Dla liczby stopni swobody równej 18 ($df = 11 + 9 - 2 = 18$) oraz poziomu istotności równego 0,05 wartość krytyczna wynosi 2,101. Stwierdzamy, że obliczona wartość sprawdzianu testowego jest mniejsza od krytycznej, więc brak jest podstaw do odrzucenia hipotezy zerowej – innymi słowy: różnica między tymi średnimi jest statystycznie nieistotna. Płeć uczniów nie różnicuje więc wyników testu humanistycznego w sposób statystycznie istotny. Dziewczynki osiągnęły w tym teście lepszy rezultat od chłopców – ale dotyczy to z całą pewnością tylko zbadanych 20 uczniów, nie mamy natomiast podstaw, by wypowiadać się w ten sposób o jakiejkolwiek populacji.

W tym przypadku obie metody dały taki sam wynik – zależność między zmiennymi okazała się statystycznie nieistotna. Może się jednak zdarzyć, że każda z metod doprowadzi do innej konkluzji. Którą z nich należy wówczas uznać za bardziej wiarygodną? Odpowiedź jest prosta: drugą, a to dlatego, że w ramach pierwszej metody następuje duża utrata informacji podczas procedury podziału zakresu zmiennej liczbowej na rozłączne kategorie. Operujemy wówczas jedynie jakościowymi kategoriami typu „wysoki wynik" i „niski wynik" – ignorując faktyczne wartości liczbowe uzyskane przez poszczególne osoby badane. Nie zmienia tego faktu również sytuacja dużej liczebnie próby, która mogłaby pozwolić na wyróżnienie większej niż dwie liczby bardziej szczegółowych kategorii. Pierwszą metodę stosujemy niejako w ostateczności, gdy analiza wariancji (lub test istotności różnicy między średnimi) jest poza zasięgiem naszych możliwości.

Można wspomnieć o jeszcze jednym narzędziu, które może być używane do badania związku między zmienną liczbową a nominalną – jest nim tzw. współczynnik korelacji punktowej dwuseryjnej. Do pewnego stopnia można go traktować jako modyfikację omówionego testu istotności różnicy między średnimi. Współczynnik ten nie jest tu jednak omawiany, ponieważ nie jest typowym współczynnikiem korelacji – nie osiąga, nawet czysto teoretycznie, wartości 1 (ani –1), stąd interpretacja ewentualnie obliczonej jego wartości jest utrudniona. W ostatnich latach używa się go zresztą do specyficznych celów, głównie na etapie wewnętrznej statystycznej analizy podczas konstruowania testów lub skal (patrz np. Ferguson, Takane 1997, s. 481–484).

Czytelnik zechce sprawdzić – przy użyciu testu istotności różnicy między średnimi – jakie są związki między parami zmiennych: Z.2 i Z.4, Z.1 i Z.5 oraz Z.2 i Z.5. Oto – dla kontroli – skrótowe wyniki. W przypadku pary zmiennych Z.2 i Z.4 wartość sprawdzianu testowego wynosi 2,353, co w konsekwencji daje istotność różnicy między średnimi na poziomie 0,05; wyższą średnią z testu humanistycznego uzyskała grupa uczniów wybierających jako przyszłą szkołę liceum. Dla pary Z.1 i Z.5 sprawdzian testowy ma wartość –2,377 (czyli jego wartość bezwzględna wynosi 2,377), co również oznacza istotność różnicy między średnimi na poziomie 0,05; w tym przypadku wyższą średnią z testu matematyczno-przyrodniczego uzyskali chłopcy. Dla pary zmiennych: Z.2 i Z.5 zależność jest statystycznie nieistotna.

8. Uwagi końcowe

Zaprezentowane przykłady analiz nie wyczerpują oczywiście wszystkich możliwości badania zależności między zmiennymi. Przeciwnie – należy stwierdzić, że jest to jedynie zarys wprowadzenia w tę problematykę. W polu widzenia są jeszcze przecież: analiza regresji, także wielokrotnej, analiza wariancji, także wieloczynnikowa, tzw. analiza czynnikowa i inne procedury.

Powyższy tekst miał w zamierzeniu wskazać różnorodne możliwości analiz między zmiennymi – już przy użyciu „elementarza” statystyki dwuzmiennowej. Nacisk nie został tu położony na obliczeniowy aspekt statystycznych dociekań, lecz na ich stronę „interpretacyjną”. Jednocześnie podkreślone zostały dwa istotne aspekty sprawy. Po pierwsze, dla prawidłowego wyboru właściwej procedury statystycznej ważne jest poprawne umiejscowienie analizowanych zmiennych na jednej ze skal pomiarowych: nominalna, porządkowa czy też interwałowa lub ilorazowa. Po drugie, dokonując ustalania, czy i jaka zachodzi korelacja (zbieżność, zależność) między danymi otrzymanymi z przebadanej próby, należy starać się „wyjść poza próbę” i dążyć do ustalenia, co otrzymane wyniki oznaczają dla populacji – jakkolwiek została ona w danym badaniu zdefiniowana.

Autor wyraża nadzieję, że lektura tego tekstu może się przyczynić do podniesienia na wyższy poziom analizy danych w pracach licencjackich i magisterskich przynajmniej niektórych Czytelników.

9. Literatura

Babbie E. (2003). *Badania społeczne w praktyce*. Warszawa: PWN.

Brzeziński J. (2003). *Metodologia badań psychologicznych*. Warszawa: PWN.

Ferguson G.A., Takane Y. (1997). *Analiza statystyczna w psychologii i pedagogice*. Warszawa: PWN.

Frankfort-Nachmias C., Nachmias D. (2001). *Metody badawcze w naukach społecznych*. Poznań: Zysk i S-ka.

Góralski A. (1987). *Metody opisu i wnioskowania statystycznego w psychologii i pedagogice*. Warszawa: PWN.

Guilford J.P. (1960). *Podstawowe metody statystyczne w psychologii i pedagogice*. Warszawa: PWN.

Jóźwiak J., Podgórski J. (1997). *Statystyka od podstaw*. Warszawa: Polskie Wydawnictwo Ekonomiczne.

Krajewska A. (2001). *Statystyka dla pedagogów: wybrane zagadnienia*. Białystok: Trans Humana.

Łaniec J.D. (1999). *Elementy statystyki dla pedagogów*. Olsztyn: Wydawnictwo Uniwersytetu Warmińsko-Mazurskiego.

Nowaczyk C. (1995). *Podstawy metod statystycznych dla pedagogów*. Bolków: Avis.