Boris Kožuh

STATYSTYKA DLA PEDAGOGÓW

KRAKOWSKA AKADEMIA im. Andrzeja Frycza Modrzewskiego

2000

Boris Kožuh

STATYSTYKA DLA PEDAGOGÓW

OFICYNA
WYDAWNICZA
AFM

Kraków 2011

Recenzja:
dr hab. Marian Niezgoda, prof. UJ

Projekt okładki
Joanna Sroka

Korekta
Kamila Zimnicka-Warchoł

ISBN 978-83-7571-112-7

Wydawca:
Krakowska Akademia im. Andrzeja Frycza Modrzewskiego
www.ka.edu.pl, Kraków 2011

Sprzedaż prowadzi
Księgarnia u Frycza
Kampus Krakowskiej Akademii im. Andrzeja Frycza Modrzewskiego
ul. Gustawa Herlinga-Grudzińskiego 1
30-705 Kraków, tel./fax: (12) 252 45 93
e-mail: ksiegarnia@kte.pl

Projekt typograficzny, łamanie:
Joanna Sroka

Druk i oprawa:
Krakowskie Towarzystwo Edukacyjne sp. z o.o. – Oficyna Wydawnicza AFM

SPIS TREŚCI

Rozdział I

PODSTAWOWE POJĘCIA STATYSTYKI

1. ZJAWISKA MASOWE

Badania empiryczne w pedagogice opierają się na bezpośrednim, zmysłowym poznawaniu zjawisk i sytuacji edukacyjnych. Badane zjawiska dzielą się na masowe lub pojedyncze. Do zjawisk masowych zalicza się te, które występują wielokrotnie (przynajmniej dwa razy). Zjawiska pojedyncze natomiast to te, które są zdarzeniami niepowtarzalnymi. Czasopismo „Biuletyn Oświatowy" jest zjawiskiem pojedynczym, natomiast czasopismo nauczycielskie należy do zjawisk masowych, ponieważ tylko na terenie Polski publikuje się ich kilkadziesiąt. Nauczycielka Beata Kołowrotkiewicz z miasta R. jest zjawiskiem pojedynczym, natomiast nauczycielka matematyki jest zjawiskiem masowym. Do badań i studiowania zjawisk masowych często używa się metod statystycznych, natomiast nie są one w ogóle wykorzystywane do badań zjawisk pojedynczych.

Pomimo że metody statystyczne umożliwiają badanie zarówno cech jakościowych, jak i ilościowych, to jednak opracowania statystyczne są bardziej charakterystyczne dla cech ilościowych. Określenie zjawisk masowych jest proste. Nie jest ono jednak pomocne w podejmowaniu decyzji o zastosowaniu metod statystycznych. Troje uczniów to już zbiorowość, jednak przy jej badaniu stosowanie metod statystycznych jest niekorzystne i nieużyteczne. Metody statystyczne bowiem stosuje się dopiero dla „bardziej masowych" zjawisk, czyli dla zbiorowości w rodzaju klasy, kółka uczniowskiego, grupy przedszkolnej itp. W przypadku odpowiednio dużej grupy metody statystyczne stają się przydatne.

2. POPULACJA I JEDNOSTKI STATYSTYCZNE

Jednostkami, które tworzą populacje w badaniach pedagogicznych, mogą być uczniowie, wychowankowie, nauczyciele, dyrektorzy, jak również szkoły, przedszkola, internaty, podręczniki, zeszyty, programy nauczania, komputery, meble szkolne, budynki, lekcje, wycieczki, oceny, prace uczniów, pytania itd.
W potocznym języku populacją nazywa się na ogół ludzi. W statystyce populacją określa się zbiorowości ludzi, rzeczy, zdarzeń, idei, itd.
W badaniach empirycznych badacza zawsze interesuje konkretna populacja. Badanie staje się możliwe po jej ścisłym zdefiniowaniu. Polega ono na określeniu cech stałych, które decydują o zakwalifikowaniu danej jednostki do populacji. Wyodrębnia się trzy rodzaje cech stałych (lub kwalifikujących):

1. **rzeczowe**, które określają, kto lub co stanowi jednostkę tej populacji,
2. **czasowe**, które określają czas obserwacji populacji,
3. **przestrzenne**, które określają przestrzeń geograficzną badanej populacji.

Wszystkie jednostki, które posiadają cechy stałe odpowiadające cechom określonym, tworzą określoną zbiorowość lub populację. Oto przykłady odpowiednio określonych populacji:

Tabela 1. Populacje badawcze

Studenci kierunku pedagogika w KSW w Krakowie od 1 października 2006 roku do 31 sierpnia 2008 roku.
Szkoły podstawowe w województwie opolskim w dniu 12 lutego 2008 roku.
Dyrektorzy szkół ponadpodstawowych w Katowicach dnia 6 grudnia 2007 roku.

Najbardziej korzystną dla badań jest sytuacja, w której cecha czasowa nie jest określona w zbyt długim przedziale, takim jak cały rok szkolny, semestr, cykl studiów itp. W przypadkach, gdy cecha czasowa obejmuje cały rok szkolny, należy koniecznie uwzględnić wszystkie jednostki, które w tym czasie z różnych przyczyn trafiły do populacji. Ten wymóg staje się bardzo niepraktyczny, ponieważ wymaga wydłużenia ram czasowych zbierania danych. Z tego powodu populacje określa się w długich przedziałach

jedynie w przypadku gromadzenia danych z już istniejącej dokumentacji, np.: z dokumentacji WOM, kuratorium oświaty, ministerstwa itp. W sytuacjach, w których dane gromadzi się samodzielnie, lepiej określić populacje z konkretną datą.
Badaną zbiorowość w literaturze statystycznej nazywa się populacją statystyczną, a w raportach badawczych populacją badawczą lub krócej – populacją. Inaczej natomiast nazywa się zbiorowości w przypadkach badań na próbach. Całą zbiorowość określa się mianem populacji generalnej, a mniejszą wybraną część populacją próbną lub krócej próbą.

3. ZMIENNE

Jednostki populacji posiadają wiele różnych cech. Niektóre z nich stanowią cechy stałe, które określają populacje i są identyczne dla wszystkich jednostek. Pozostałymi cechami jednostki różnią się między sobą i są to cechy zmienne (krótko nazywa się je zmiennymi). Poniższa tabela prezentuje przykłady często wybieranych do badań jednostek i ich cech.

Tabela 2. Statystyczne jednostki i zmienne

Jednostka	**Badane cechy (zmienne)**
Szkoła	liczba uczniów, liczba klas, liczba zatrudnionych nauczycieli, czas istnienia szkoły, rodzaj i typ szkoły itd.
Uczeń	wzrost, wiek, płeć, narodowość, zainteresowania, oceny, średnia ocen w klasie szóstej, przynależność do szkolnych organizacji itd.
Nauczyciel	staż pracy, wykształcenie, specjalność, płeć, wiek, tryb ukończonych studiów, itd.
Książka	autor, wydawca, cena, liczba stron, format, liczba wydań, rok publikowania itd.
Szkoła lotnicza	programy, które realizuje, liczba instruktorów, miejscowość, w której ma lotnisko szkoleniowe, liczba szybowców, wyposażenie itd.

Badanie jednostek to badanie ich cech, czyli zmiennych. Oznacza to więc, iż zmienne są centralnym pojęciem w statystyce.
Oprócz pojęcia zmienne w metodologii badań pojawia się również termin wskaźniki. Jest to pojęcie bardzo podobne, lecz jednak inne. Z punktu wi-

dzenia badacza wskaźniki i zmienne to nie to samo, natomiast z punktu widzenia statystyka nie ma potrzeby ich odróżniania.

Zmienne jakościowe i zmienne ilościowe

Zmienne można dzielić według rozmaitych kategorii. Najprostszy podział dotyczy formy wyrażania ich wartości.

Zmienne ilościowe wyrażane są w postaci liczb, a **zmienne jakościowe** wyrażane są w postaci słownej. Tabele zaprezentowane poniżej wskazują na przykłady cech jakościowych (tabela 3) i cech ilościowych (tabela 4).

Tabela 3. Cechy jakościowe

Zmienna	**Wartości**
Płeć	kobieta, mężczyzna
Znajomość języków obcych	portugalski, angielski, niemiecki, węgierski, czeski itd.
Kierunek studiów	informatyka, pedagogika, stosunki międzynarodowe, prawo, medycyna, historia, turystyka itd.
Stopień zadowolenia z lekcji	bardzo niski, niski, średni, wysoki, bardzo wysoki
Profil szkoły	ogólnokształcąca, sportowa, artystyczna itd.

Tabela 4. Cechy ilościowe

Zmienna	**Wartości**
Liczba nauczycieli w szkole	26, 28, 32, 37, 38, 43, 57, itd.
Wiek respondentów w latach	10, 16, 20, 31, itd.
Temperatura	8°C, 28°C, 37°C, 39°C, itd.
Pensje w zł	982, 1042, 1320, 1865, itd.
Wzrost w cm	154, 157, 158, 164, 175, itd.
Cena podręcznika w zł	17, 18, 18, 20, 23, 31 itd.

Wartości niektórych zmiennych określa się jednocześnie słownie i wyraża w postaci liczb. Najlepszym przykładem takich zmiennych są oceny szkolne, które mogą być wyrażane w postaci opisowej lub w postaci cyfr. W obliczu tych możliwości powstaje wątpliwość: czy ocena szkolna jest zmienną jakościową czy ilościową? W takich przypadkach należy rozważyć naturę zmiennej. Podstawą podziału jest bowiem natura, a nie (tylko) sposób wyrażania wartości. Analiza tej sytuacji szybko prowadzi do wniosku, iż w przypadku ocen szkolnych istotne jest słowo, a nie liczba. Ocena celująca jest we wszystkich szkolnych systemach oceną najlepszą i najwyższą. Jest ona jednak jednocześnie wyrażana zupełnie różnymi liczbami (na przykład w Polsce cyfrą 6, w Słowenii cyfrą 5, w Czechach cyfrą 1, itd.). W swojej istocie jest to bardziej zmienna jakościowa niż ilościowa.

Skale pomiarowe

Dużo większe znaczenie w statystyce ma kolejny podział według rodzaju informacji, która jest zawarta w rezultatach pomiaru (w danych, w wartościach zmiennych). Według tego kryterium wyodrębnia się cztery rodzaje zmiennych (cztery skale pomiarowe):

1. nominalne,
2. porządkowe,
3. przedziałowe,
4. ilorazowe.

Zmienne nominalne

Zmienne nominalne zawierają informację, według której można jedynie ustalić, czy jednostki różnią się od siebie. Przykładem takiej zmiennej jest płeć. Według niej można ustalić, czy dwoje uczniów jest tej samej, czy różnej płci. W przypadku zmiennych nominalnych zamiast określenia wartości stosuje się te kategorie. Niektóre zmienne nominalne posiadają tylko dwie kategorie, a inne mają ich więcej. Oto przykłady:

Tabela 5. Zmienne nominalne

Zmienna	**Kategorie**
Narodowość	Polak, Serb, Hiszpan, Francuz itd.
Kierunek studiów na uczelni	socjologia, matematyka, historia, informatyka, górnictwo itd.

Płeć	kobieta, mężczyzna
Dostępność źródeł na egzaminie pisemnym ze statystyki	wolno korzystać ze źródeł, nie wolno korzystać ze źródeł
Literatura, z której korzystają nauczyciele w pracy zawodowej	podręczniki, czasopisma, książki, gazety itd.
Czy uczeń ma własny pokój	ma, nie ma

Kategorii zmiennej nominalnej nie da się uporządkować od mniejszych do większych, ponieważ zmienne nominalne nie posiadają wartości wyższych i niższych. Oznacza to jednocześnie, że wszystkie kategorie znajdują się na tym samym poziomie. Zmienne nominalne wyrażają jedynie jakości i nie posiadają podstawy ilościowej. Ustalając, czy pewna zmienna jest nominalna, czy porządkowa, należy uważnie przeanalizować podstawę (ukryte sedno), a nie tylko nazwy jej kategorii. Oto przykład:

W ankiecie uczniowie odpowiadali na pytanie, czy im się podobała wystawa w muzeum. Istniały jedynie dwie możliwe odpowiedzi: TAK i NIE. Zmienna ta nie jest jednak nominalna, ponieważ zadowolenie podlega stopniowaniu. Oznacza to jednocześnie, iż zadowolenia nie można wyrazić sformułowaniem tak lub nie, bowiem posiada ono różne poziomy. Nie można zatem określić, że niektórzy są całkowicie zadowoleni, a inni wcale niezadowoleni. W skali tej bowiem istnieją niższe i wyższe poziomy zadowolenia. W tym przypadku chodzi o źle sformułowane pytanie ankietowe (a raczej źle sformułowane odpowiedzi). Nominalna zmienna powstałaby wówczas, gdyby w ankiecie pytano uczniów na przykład o to, czy byli w muzeum, czy nie. W takim przypadku odpowiedzi TAK i NIE miałyby zupełnie inne znaczenie niż w pytaniu pierwszym.

Podobna sytuacja występuje w wielu innych przypadkach (zainteresowania, motywacja itd.). Nazwa zmiennej lub brzmienie jej odpowiedzi mogą wprowadzić w błąd.

Zmienne porządkowe

Zmienne porządkowe zawierają taką informację, według której można ustalić czy jednostki są jednakowe, czy też nie, a ponadto informację o stopniowaniu wartości. Wartości te można porządkować od najmniejszej do największej. Dla dowolnych dwóch jednostek można ustalić, która z nich jest na skali wyżej, a która niżej. Wartości zwykle nazywa się stopniami. Stopnie na tej skali wyrażają także ilość (posiadają podstawę ilościową). Charakterystyczną zmienną porządkową jest stopień wykształcenia. Zmienne porządkowe najczęściej otrzymuje się w przypadkach, gdy w badaniach użyto ankiety, obserwacji lub wywiadu. Często już forma pytania w ankiecie wskazuje na zmienną porządkową, np. „w jakim stopniu jest Pan/Pani zadowolony/a...” lub „ile Pan/Pani czyta...” itd. Kategorie zmiennej porządkowej też wskazują na podstawę ilościową: dużo, średnio, mało, często, rzadko itd.

O zmiennych porządkowych wiadomo, które stopnie są wyższe i które niższe, jednak nie można w żaden sposób ustalić, jak wielkie są różnice pomiędzy pojedynczymi stopniami (bo brak jednostki pomiaru). Nie można założyć, że przedziały pomiędzy sąsiednimi stopniami są wszędzie jednakowe. Z doświadczeń wiadomo najczęściej, że przedziały nie są jednakowe. Zmienne porządkowe można ilustrować jako schodki z niejednakowymi stopniami.

Stopnie zmiennych porządkowych można wyrazić słowami lub liczbami (np. oceny szkolne). Liczby te stwarzają złudzenie, że przedziały są jednakowe. Wydaje się, że różnica pomiędzy oceną dostateczną (3) i dobrą (4) jest taka sama, jak różnica pomiędzy oceną dobrą (4) i bardzo dobrą (5). Nie jest to jednak prawdą. Rangi takie jak np. kolejność przybycia do mety w biegach również nie odzwierciedlają prawdziwych różnic w wynikach uczniów. Wydaje się, iż wyniki uczniów są równomiernie uporządkowane od pierwszego do ostatniego (że różnice pomiędzy uczniami są jednakowe). Jednak różnice pomiędzy ocenami, a także pomiędzy wynikami w biegu nie są jednakowe. Podobna sytuacja występuje w przypadku osiągnięć uczniów na teście wiadomości – tutaj jednak niejednakowość przedziałów jest bardziej skryta. Wydaje się, że przedział pomiędzy np. 12 i 13 punktów jest taki sam jak przedział pomiędzy 6 i 7 punktów (w obydwu przypadkach przedział wynosi jeden punkt i pozornie punkty te są jednakowe). Nie jest to jednak prawdą, ponieważ trudność zadań testowych nie wzrasta równomiernie od zadania do zadania. Podobieństwo przedziałów w przypadku wyników testu wiadomości jest na ogół większe niż w przypadku niemal wszystkich pozostałych zmiennych porządkowych.

Identyczność przedziałów pomiędzy stopniami określonej zmiennej byłaby pożądaną cechą.

Niektóre zmienne porządkowe często traktuje się jako zmienne przedziałowe (np. wyniki na testach wiadomości w punktach, wyniki na niektórych skalach ocen itd.). Należy jednak pamiętać o tym, jaka jest ich podstawa i być uważnym w interpretacji takich wyników.

Zmienne przedziałowe

Zmienne przedziałowe są zmiennymi ilościowymi. Posiadają taką samą właściwość jak zmienne porządkowe – stopniowanie, jednak przedziały pomiędzy sąsiednimi stopniami są wszędzie **jednakowe**. Które zmienne są przedziałowe? Te, które posiadają precyzyjnie określoną jednostkę pomiaru (temperatura w skali Celsjusza itp.). Zmienne przedziałowe można zilustrować jako schody z jednakowo wysokimi stopniami. Dzięki temu, że przedziały są jednakowe, można ustalić, jaka jest różnica pomiędzy dowolnymi wartościami. Zmienne przedziałowe nie posiadają jednak absolutnego punktu zerowego.

Zmienne ilorazowe

To takie zmienne przedziałowe, które posiadają absolutny punkt zerowy. Ta właściwość umożliwia ustalenie, ile razy jedna wartość jest większa lub mniejsza od drugiej wartości. Zmiennymi ilorazowymi są np. wiek, wzrost, ciężar, liczba uczniów w klasie, czas nauki, liczba przeczytanych książek itd.
Należy jednak zwrócić uwagę, iż ta przewaga zmiennych ilorazowych (w porównaniu ze zmiennymi przedziałowymi) jest niewielka, a tym samym nie umożliwia wyraźnych różnic w opracowaniach statystycznych. Dlatego też często zmienne ilorazowe i przedziałowe opracowuje się w taki sam sposób.

Zmienne ciągłe i nieciągłe

Gdy wartości pewnej zmiennej są całkowite i izolowane, to zmienna ta jest nieciągła. Pomiędzy największą i najmniejszą wartością istnieje ograniczona ilość wartości (punktów). Jeżeli zmienna może posiadać nieograniczoną ilość wartości pomiędzy skrajnymi – największą i najmniejszą wartością, jest to zmienna ciągła. Można to zilustrować, wyobrażając sobie ptaki na drutach telefonicznych. Niektóre gatunki ptaków siedzą na drutach ciasno jeden obok drugiego. Inne ptaki nie tolerują bliskości i siedzą na drutach w odległościach skrzydeł. Ściśnięte jeden przy drugim można porównać do zmiennych ciągłych, ptaki siedzące w odległościach do zmiennych nieciągłych.
Wzrost jest typowym przykładem zmiennej ciągłej. Jeżeli wzrost ucznia na początku roku szkolnego wynosił 152 cm i na końcu 156 cm, oznacza to, że uczeń „przechodził" w ciągu roku wszystkie wartości pomiędzy. Nie istnieje możliwość pominięcia np. wartości 155,37 cm. Przy tym należy pamiętać, że w praktyce wzrost wyraża się w pełnych centymetrach. Jeżeli pomiar

wzrostu u lekarza szkolnego wykonuje się kilka razy w roku, w kartotece tego ucznia zostaną napisane tylko wartości 152, 153, 154, 155 i w końcu 156 cm. Jednak nie zmienia to prawdziwej natury zmiennej: uczeń rósł stopniowo od 152 cm do 156 cm. Nazwa zmiennej ciągłej pochodzi stąd, że wartości są szeregowane nieprzerwanie (w sposób ciągły, pomiędzy wartościami nie ma „skoków").

Liczba uczniów w klasie jest typowym przykładem zmiennej nieciągłej. Jeżeli najmniejsza możliwa liczba uczniów wynosi siedemnaście a największa dwadzieścia siedem, to zmienna może jedynie posiadać wartości 17, 18, 19, 20, 21, 22, 23, 24, 25, 26 i 27. Pomiędzy sąsiednimi wartościami nie ma żadnej wartości.

W praktyce zmienne ciągłe często wyrażone są w postaci wartości nieciągłych. Taka sytuacja występuje w przypadku wzrostu (np. 152, 153, 154 cm), wieku (np. 15, 16, 17 lat), ciężaru (np. 57, 58, 59 kg), temperatury (np. 23°C, 24°C, 25°C, 26°C) itp. Nawet wiedza uczniów jest zmienną ciągłą, pomimo że jest najczęściej wyrażona w postaci punktów z testu lub ocen szkolnych.

Przy opracowaniu danych natomiast postępowanie jest odwrotne. Często zmienne nieciągłe opracowuje się jako ciągłe, ponieważ upraszcza to procedury opracowania.

Zmienne zależne i niezależne

Podział ten powstał według ról, jakie zmienne pełnią we współzależnościach. W najprostszych przypadkach bada się związki dwóch zjawisk (dwóch zmiennych). W każdej parze powiązanych zmiennych jedna odgrywa rolę zmiennej niezależnej (ta, która oddziałuje), a druga zmiennej zależnej (ta, na którą oddziałuje).

Ocena, którą uczeń otrzymał rano w szkole, oddziałuje na rodzinną atmosferę przy obiedzie tego samego dnia, zwłaszcza jeżeli ocena ta jest bardzo dobra lub bardzo zła. W tym przypadku ocena jest zmienną niezależną, a atmosfera zmienną zależną (bo ocena oddziałuje na tę atmosferę).

Role te przydziela się im na podstawie natury zależności pomiędzy nimi oraz zamiaru prowadzonego badania. Ta sama zmienna może być w pewnej parze niezależną, a w innej parze zależną. Nawet w tej samej parze zmienne mogą zmienić swoje role. Często związki zjawisk wychowawczych nie są jednoznaczne i jednostronne. W parze: czas domowego uczenia się i ocena szkolna,

czas traktuje się jako zmienną niezależną, a ocenę szkolną jako zmienną zależną. Wiadomo bowiem, że dłuższy czas uczenia się oznacza na ogół wyższe oceny. Jednak związek ten zachodzi też w przeciwnym kierunku: otrzymana przez ucznia ocena ma wpływ na to, ile uczeń będzie się uczył w przyszłości. Ocena może więc odgrywać role zmiennej niezależnej a czas zależnej. Takie przypadki w obszarze edukacji stanowią częste zjawisko. Podział na zmienne niezależne i zależne jest prawie zawsze tymczasowy.
Istnieje dużo zmiennych, które w badaniach edukacji zawsze posiadają role niezależnych. Płeć jest zmienną, która występuje w badaniach edukacji zawsze jako niezależna (oczywiście, jeżeli w ogóle pojawia się w badaniach). Podobnie jest w przypadku narodowości, wieku itp. Większość jednak prawdziwie edukacyjnych zjawisk ciągle zmienia role niezależnych w zależne i odwrotnie. Oznacza to, iż podział na zmienne niezależne i zależne jest zawsze tymczasowy. Dlatego też nie ma naprawdę potrzeby np. w pracy magisterskiej robić odrębnych list dla zmiennych niezależnych i zmiennych zależnych. Najlepiej zrobić wspólną listę, a w interpretacji każdej tabeli wskazać, która z tych dwóch zmiennych jest niezależną, a która zależną.

Natura zmiennych i stosowanie metod statystycznych

Dlaczego istnieje potrzeba rozróżniania zmiennych? Odpowiedź jest prosta: zastosowanie metod statystycznych zależy od natury zmiennych. W stosowaniu metod przestrzega się dwóch istotnych praw:

1. Im wyżej zmienna znajduje się w podziale (od nominalnych do ilorazowych), tym więcej metod statystycznych można użyć przy opracowywaniu jej danych. Najmniej statystycznych metod można użyć w przypadku zmiennych nominalnych, zaś dużo więcej w przypadku zmiennych porządkowych, jeszcze więcej w przypadku zmiennych przedziałowych i najwięcej w przypadku zmiennych ilorazowych.
2. Wszystkie metody, które można zastosować do opracowania pewnego rodzaju zmiennej, można także użyć dla wszystkich zmiennych, które w omawianym podziale są wyżej położone. W przypadku gdy pewna metoda jest dopuszczalna np. dla zmiennych porządkowych, automatycznie jest dopuszczalna także dla zmiennych przedziałowych i ilorazowych.

4. PARAMETRY STATYSTYCZNE

Parametry to wartości numeryczne charakteryzujące populację (ilościowy wyraz cech populacji). Najczęściej wartość parametru wylicza się z wartości zmiennej dla wszystkich jednostek badanej populacji. Parametry pełnią podobną rolę dla populacji, jak wartości zmiennej dla poszczególnych jednostek. Uczennicę Igę K. charakteryzuje wzrost np. 168,35 cm, natomiast całą klasę (populację) charakteryzuje średnia wzrostu wynosząca np. 169,11 cm. Jednostki porównuje się na podstawie wartości zmiennej, a populacje na podstawie ich parametrów. Oto parametry spotykane najczęściej w badaniach pedagogicznych: procent, średnia arytmetyczna, mediana, odchylenie standardowe, wariancja, współczynniki korelacji.
Wyrazu parametr używa się jeszcze w innym znaczeniu. Jeżeli bada się populacje na podstawie próby (czyli wybranej mniejszej części tej populacji), to wartości wyliczone dla populacji nazywa się parametrami, natomiast te same wartości wyliczone dla próby określa się mianem statystyk.

Rozdział II

PORZĄDKOWANIE DANYCH

Przed opracowaniem statystycznym należy uporządkować dane. Czynność ta ułatwia opracowywanie danych. Od czasu, kiedy pojawiły się komputery, procedury porządkowania danych zmieniły się w sposób zasadniczy. Przed „ręcznym" opracowaniem danych należy je rzeczywiście uporządkować, natomiast do opracowania za pomocą komputera wystarczy je jedynie przygotować. Aby lepiej zrozumieć wyniki opracowania statystycznego, należy zapoznać się z całkowitym postępowaniem: od przygotowania danych do otrzymania końcowych wyników. Ujmując to językiem bardziej beletrystycznym: „oprócz wyników należy znać też drogę do ich otrzymania".

1. PORZĄDKOWANIE DANYCH JAKOŚCIOWYCH

Dane w tym przypadku porządkuje się, robiąc zestawienia tabelaryczne. Do każdej kategorii zmiennej należy policzyć ilość jednostek (liczebność lub "f"). Liczebność zwykle wyrażana jest w postaci liczb absolutnych (np. ile respondentów wybrało pojedyncze odpowiedzi na pytanie w ankiecie) i procentów (ile procent respondentów wybrało daną odpowiedź).
Tak uporządkowane dane nadają się do kolejnych opracowań. Już samo uporządkowanie danych umożliwia pierwszą analizę.

Tabele z jedną zmienną
Pokazują stan każdej zmiennej w badanej populacji. Oto przykład najprostszej tabeli, która ilustruje strukturę zmiennej „tryb studiów" wśród badanej grupy studentów pedagogiki.

Tabela 6. Tryb studiów

	Liczebność	Procent liczebności
Studia zaoczne	34	35,4
Studia dzienne	62	64,6
Razem	96	100,0

W przypadku zmiennych nominalnych posiadających więcej niż dwa stopnie kategorie w tabeli porządkuje się najczęściej według liczebności malejących lub rosnących. Powiększa to przejrzystość wyników i ułatwia zrozumienie interpretacji. Oto przykład takiej tabeli, w której zawarto odpowiedzi respondentów dotyczące oglądalności programów telewizyjnych (rosnąco):

Tabela 7. Odpowiedzi respondentów

	Liczebność	Procent liczebności
Programy publicystyczne	10	4,2
Programy przyrodnicze	21	8,8
Seriale	29	12,2
Programy sportowe	41	17,2
Programy rozrywkowe	67	28,2
Wiadomości	70	29,4
Razem	238	100,0

Jeszcze częściej sporządza się tabele według liczebności malejących:

Tabela 8. Odpowiedzi respondentów

	Liczebność	Procent liczebności
Prawo	29	30,2
Zarządzanie	22	22,9
Socjologia	18	18,8
Pedagogika	17	17,7
Politologia	10	10,4
Razem	96	100,0

W przypadku zmiennych nominalnych nie ma więc potrzeby utrzymywania kolejności np. z kwestionariusza ankiety. W przypadku zmiennych przedziałowych i w ankiecie, i w tabelach zachowuje się naturalną kolejność (np. jak w tabeli nr 9).
Jeżeli zmienna jest porządkową i posiada więcej niż dwa stopnie, można dodać kolumnę liczebności skumulowanych F (lub F%). Te liczebności pokazują, ile jednostek (lub %) jest razem do pewnego stopnia (liczebność tych, którzy są pod tą kategorią wraz z nią). Dla każdej kategorii należy zsumować wszystkie liczebności niższych kategorii. Ma to sens jednak jedynie wtedy, gdy jest to naprawdę niezbędne. Oto przykład takiej tabeli:

Tabela 9. Udział studentów w akcjach charytatywnych

	Liczebność	**Procent liczebności**	**Skumulowany procent liczebności**
Nigdy	100	14,9	14,9
Rzadko	80	11,9	26,8
Czasami	150	22,4	49,2
Często	120	17,9	67,1
Zawsze	220	32,8	100,0
Razem	670	100,0*	

*Suma wszystkich procentów w kolumnie nie wynosi 100%, lecz jedynie 99,9% z powodu zaokrągleń poszczególnych procentów.

Przykład ten pokazuje jednocześnie, jak postępować w rzadkich przypadkach, w których suma poszczególnych procentów nie wynosi 100%. Takiej sytuacji nie można przemilczeć, lecz należy ją czytelnikowi precyzyjnie wyjaśnić. Rozwiązaniem w tej sytuacji jest zaokrąglenie jednego z procentów w kolumnie niezgodnie z zasadami, aby suma wynosiła 100,0% (np. zamiast poprawnie 11,9 należy napisać 12,0). Tę wartość należy oznaczyć gwiazdką i wyjaśnić w przypisie. Innym rozwiązaniem jest pozostawienie końcowego procentu 99,9% i ponowne opisanie w przypisie.
Wyniki w statystyce na ogół zaokrągla się do dwóch miejsc po przecinku. Istnieje jednak pewien wyjątek: **procenty w tabelach** zaokrągla się do jednego miejsca po przecinku. Nie dotyczy to procentów, z których oblicza się kolejne wyniki. Tu znowu obowiązuje prawo dwóch miejsc po przecinku.

W taki sposób przebiega porządkowanie danych dla wszystkich nominalnych i porządkowych zmiennych w badaniu. Te tabele pokazują stan w badanej populacji.

Tabele krzyżowe

Często oprócz stanu po pojedynczych zmiennych bada się także związki między zmiennymi. Do takich celów należy sporządzić tabele, które zawierają więcej zmiennych. Są to tabele korelacyjne (krzyżowe). Najczęściej taka tabela zawiera dwie zmienne, ponieważ tabela z trzema zmiennymi jest już bardzo skomplikowana i trudna do odczytania (nawet dla specjalistów).

Poniżej zaprezentowano przykład tabeli dwóch zmiennych: stopień wykształcenia i poglądy (dotyczące określonego zjawiska).

Tabela 10. Tabela korelacyjna respondentów według wykształcenia i poglądów

Wykształcenie	**Zgadzam się**	**Niezdecydowany**	**Nie zgadzam się**	**Razem**
Średnie	14	8	30	52
Licencjackie	7	5	9	21
Magisterskie	13	4	6	23
Razem	34	17	45	96

Na podstawie tej tabeli trudno zrozumieć, jaki zachodzi związek pomiędzy zmiennymi. Należy wyliczyć jeszcze procenty we wszystkich komórkach tabeli. Procenty można obliczyć na trzy sposoby, otrzymując trzy odmienne tabele. Zostały one umieszczone poniżej.

W tabeli nr 11 procenty zostały obliczone według wykształcenia, w tabeli nr 12 według udziału, a w tabeli nr 13 z liczebności całej populacji (N=96).

Tabela 11. Procenty według kategorii zmiennej niezależnej

Wykształcenie	Zgadzam się	Niezdecydowany	Nie zgadzam się	Razem
Średnie	14 26,9%	8 15,4%	30 57,7%	52 100,0%
Licencjackie	7 33,3%	5 23,8%	9 42,9%	21 100,0%
Magisterskie	13 56,5%	4 17,4%	6 26,1%	23 100,0%
Razem	34 35,4%	17 17,7%	45 46,9%	96 100,0%

W tej tabeli w wierszach obliczano procenty z sumy na końcu wiersza. W pierwszym wierszu liczebność 14 przedstawia 26,9% od sumy 52 (na prawym krańcu wiersza). Wszystkie procenty w tym wierszu dają razem 100,0% (26,9% + 15,4% + 57,7% = 100,0%). Tak samo oblicza się procenty we wszystkich wierszach.

Tabela 12. Procenty według kategorii zmiennej zależnej

Wykształcenie	Zgadzam się	Niezdecydowany	Nie zgadzam się	Razem
Średnie	14 41,2%	8 47,1%	30 66,7%	52 54,2%
Licencjackie	7 20,6%	5 29,4%	9 20,0%	21 21,9%
Magisterskie	13 38,2%	4 23,5%	6 13,3%	23 24,0%
Razem	34 100,0%	17 100,0%	45 100,0%	96 100,0%*

W tej tabeli w kolumnach obliczano procenty z sumy na końcu kolumny (na dole). W drugiej kolumnie np. liczebność 8 przedstawia 47,1% od sumy 17 (na dnie kolumny). Wszystkie procenty w tej kolumnie dają razem 100,0% (47,1% + 29,4% + 23,5% = 100,0%). Tak samo oblicza się procenty we wszystkich kolumnach. Gwiazdką została oznaczona wartość 100,0% ponieważ suma wszystkich procentów w tej kolumnie (54,2% + 21,9% + 24,0%) nie wynosi 100,0% lecz 100,1%. Stało się to w wyniku zaokrąglenia poszczególnych wartości procentowych.

Tabela 13. Procenty z liczebności całej populacji (N=107)

Wykształcenie	Zgadzam się	Niezdecydowany	Nie zgadzam się	Razem
Średnie	14 14,6%	8 8,3%	30 31,3%	52 54,2%
Licencjackie	7 7,3%	5 5,2%	9 9,4%	21 21,9%
Magisterskie	13 13,5%	4 4,2%	6 6,3%	23 24,0%
Razem	34 35,4%	17 17,7%	45 46,9%	96 100,0%*

W tej tabeli w komórkach obliczano procenty z liczebności całej populacji. W pierwszej komórce (górna lewa) liczebność 14 stanowi 14,6% wszystkich respondentów (N=96). Procenty we wszystkich komórkach tworzą razem 100,0% (14,6% + 8,3% + 31,3% + 7,3% + 5,2% + 9,4% + 13,5% + 4,2% + 6,3% = 100,0%). Procenty w kolumnie „razem" oraz w wierszu „razem" zostały też obliczone z N=96.

Pierwsza tabela pokazuje, jak wykształcenie oddziałuje na poglądy. Z doświadczeń wiadomo, że ludzie z różnym wykształceniem mają odmienne poglądy, o które pytano w ankiecie. Ponieważ obie zmienne są powiązane, należy określić, która z nich jest zmienną niezależną, a która zależną. Można założyć, że w tej parze wykształcenie jest zmienną niezależną, a poglądy zmienną zależną. Dlatego akurat pierwsza tabela jest najbardziej odpowiednia. Przy badaniu związków (zależności) między zmiennymi praktycznie zawsze należy obliczać procenty w przedstawiony sposób.

Druga tabela pokazuje, jak poglądy oddziałują na wykształcenie. Z wyjątkiem bardzo nietypowych i rzadko spotykanych przypadków, taki kierunek wpływu nie ma sensu. Dlatego też nigdy nie korzysta się z takich tabel.

Trzecia tabela nie nadaje się do analizy związku między zmiennymi. Z tego powodu jest zbędna, ponieważ niepotrzebne jest włączenie niezwiązanych zmiennych do wspólnej tabeli.

2. PORZĄDKOWANIE DANYCH ILOŚCIOWYCH

Uporządkowane dane jakościowe nadają się już do interpretacji. Inaczej jest w przypadku danych ilościowych. Porządkuje się je w celach łatwiejszego opracowania. Korzystanie z komputera w opracowaniu danych przyniosło radykalną zmianę: porządkowanie danych ilościowych przed opracowaniem jest już niepotrzebne. Dlatego też zagadnienie dotyczące porządkowania danych ilościowych omawiam w sposób skrócony. Dane ilościowe porządkuje się na dwa sposoby.

Szereg szczegółowy prosty

Wartości zmiennej porządkuje się rosnąco. Sposób ten zostanie zilustrowany jednym przykładem. W pewnym badaniu badano oceny i osiągnięcia testowe uczniów z różnych przedmiotów. Oto nieuporządkowany szereg wyników testu z chemii:

wartości zmiennej *x*	10, 8, 15, 27, 12, 16, 24, 22, 18, 4, 25, 5, 11, 6

Najniższa wartość wynosi 4 punkty, a najwyższa – 27 punktów. Zaczynając od wartości 4, wszystkie wartości są uszeregowane rosnąco.

Tabela 14. Szereg szczegółowy prosty wyników testu z chemii

x	4	5	6	8	10	11	12	15	16	18	22	24	25	27
R	1	2	3	4	5	6	7	8	9	10	11	12	13	14

W szeregu szczegółowym prostym do wartości zmiennej (x) dodaje się liczby porządkowe, czyli rangi (R). Rangi te pokazują kolejność jednostek od początku szeregu. Tak powstałe rangi umożliwiają ustalenie położenia (pozycji) każdej jednostki. Służą też do głębszych analiz, wynikających z położenia jednostek. Jednak powstałe w ten sposób rangi nie mają większego znaczenia w badaniach. Odmienna sytuacja zachodzi względem rang powstałych w inny sposób. W badaniach często analizuje się zjawiska, które bardzo trudno mierzyć ilościowo (lub nie jest to nawet możliwe). Większość zmiennych porządkowych ma taką naturę. Dotyczy to głównie licznych cech ludzkich, takich jak np.: motywacja, pilność, uprzejmość, popularność, komunikatywność itd. Proces edukacyjny umożliwia bardzo skuteczne rozwiązanie problemu pomiaru tych cech. Nauczyciel (wychowawca) jest w stanie stosunkowo rzetelnie uszeregować swoich uczniów według właściwości tego rodzaju. W ten sposób otrzymuje się zmienne porządkowe – rangi. Często jest to jedyny sposób pozwalający na otrzymanie danych potrzebnych m.in.: do porównywania populacji, do obliczania współczynników korelacji itd. Dlatego właśnie rangi mają wielkie znaczenie w badaniach edukacji.

Szereg rozdzielczy z przedziałami klasowymi
Tego sposobu używa się w sytuacjach, gdy liczebność grupy przewyższa liczebność zwykłej szkolnej klasy. Oto szereg rozdzielczy z wynikami studentów z testu:

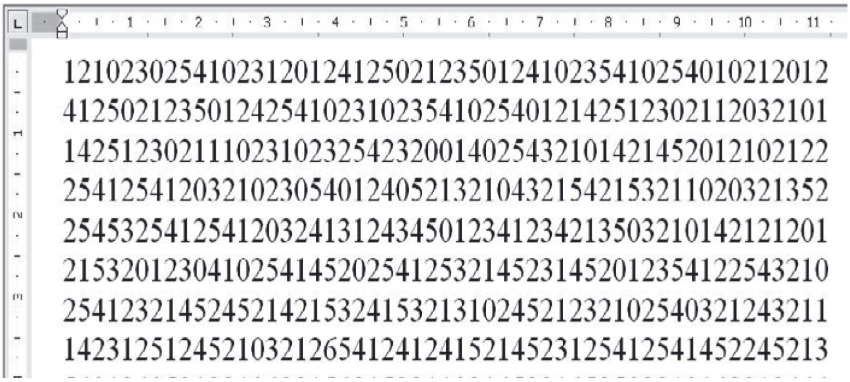
Tabela 15. Wyniki studentów z testu metodologii

Przedziały klasowe	**Liczebność**	**Procent liczebności**	**Liczebność skumulowana**	**Skumulowany procent liczebności**
13–15	4	1,6	4	1,6
16–18	12	4,8	16	6,5
19–21	16	6,5	32	12,9
22–24	21	8,5	53	21,4
25–27	27	10,9	80	32,3
28–30	31	12,5	111	44,8
31–33	33	13,3	144	58,1
34–36	32	12,9	176	71,0
37–39	26	10,5	202	81,5
40–42	23	9,3	225	90,7
43–45	11	4,4	236	95,2
46–48	7	2,8	243	98,0
49–51	5	2,0	248	100,0
	n=248	**100,0**		

3. PRZYGOTOWANIE DANYCH DO OPRACOWANIA KOMPUTEROWEGO

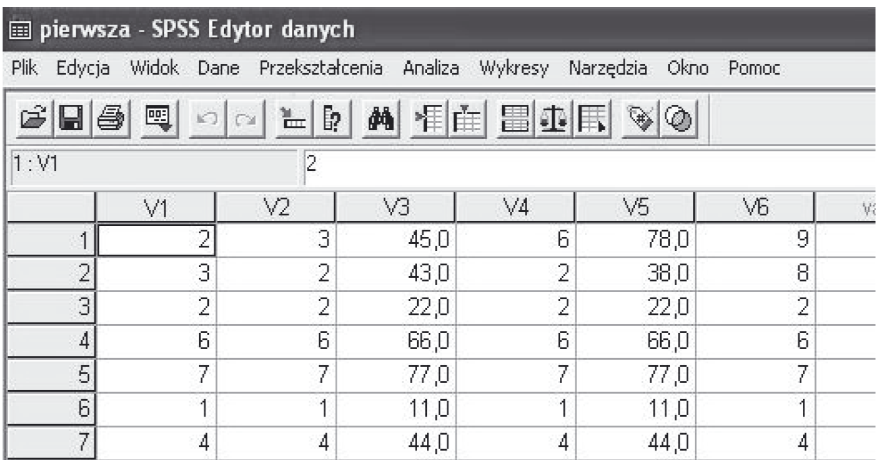
Najczęściej występujące zmienne w empirycznych badaniach edukacji to zmienne nominalne i porządkowe. Rzadziej spotyka się zmienne przedziałowe lub ilorazowe. Zmienne przedziałowe i ilorazowe w tym podrozdziale (w celu łatwiejszego zrozumienia) zostaną omówione wspólnie. Zmienne, które należą do porządkowych, często opracowuje się jak zmienne przedziałowe (np. wyniki uczniów testu). Do najczęściej używanych instrumentów w badaniach pedagogicznych należą kwestionariusze ankiety lub wywiadu, skale postaw, skale szacowania, testy socjometryczne, testy wiadomości i protokoły obserwacji.

W poniższym opracowaniu osoby badane różnymi narzędziami określa się pojęciem „respondent". Na końcu badania dla wszystkich respondentów istnieje taka sama ilość danych (zmiennych) ponieważ dane zbiera się tymi samymi instrumentami dla wszystkich respondentów. To bardzo ułatwia wprowadzanie danych do komputera i ich opracowywanie. Dane dla każdego respondenta zapisuje się w nowym wierszu. W ten sposób powstaje baza danych.



Rysunek 1. Baza danych w programie Word

W bazie danych w wierszach umieszcza się respondentów, natomiast w kolumnach zmienne. Podobna baza powstaje także w programie SPSS.



	V1	V2	V3	V4	V5	V6
1	2	3	45,0	6	78,0	9
2	3	2	43,0	2	38,0	8
3	2	2	22,0	2	22,0	2
4	6	6	66,0	6	66,0	6
5	7	7	77,0	7	77,0	7
6	1	1	11,0	1	11,0	1
7	4	4	44,0	4	44,0	4

Rysunek 2. Baza danych w programie SPSS

Do opracowania takich danych używa się programów statystycznych, takich jak np. SPSS, Excel, Statistica, itp.

Wprowadzanie danych do komputera
Wartości większej części zmiennych z ankiet, skal szacowania, protokołów obserwacji itd. zostają zapisane w postaci jednocyfrowych liczb. Zilustruje to przykład pytania ankietowego zamkniętego z sześcioma proponowanymi odpowiedziami: A, B, C, D, E i F. Do komputera nie wprowadza się liter – zawsze tylko liczby. Odpowiedź A zapisana zostanie jako liczba 1. Dla respondentów, którzy zakreślili odpowiedź B zapisana zostanie liczba 2, dla wszystkich, którzy zakreślili odpowiedź C zostanie zapisana liczba 3 i tak dalej do liczby 6 dla odpowiedzi F. Taka zmienna będzie miała jednocyfrowy zapis i będzie zajmowała w szeregu danych tylko jedno miejsce. W bazie danych będzie to jedna kolumna.
Oto przykład, w którym respondenci odpowiadali na pięć takich pytań. Pierwszy respondent zakreślił przy pytaniu pierwszym odpowiedź C, przy drugim C, przy trzecim D, przy czwartym B i przy piątym A. Odpowiedzi drugiego respondenta to kolejno: B, C, A, F, B; trzeciego: B, B, C, A, C czwartego: D, C, C, E, B, piątego B, B, B, B, B szóstego C, E, A, F, A siódmego A, A, C, D, B itd.
Należy to zapisać w następujący sposób:

1.	pierwszy respondent	33421
2.	drugi respondent	23162
3.	trzeci respondent	22313
4.	czwarty respondent	43352
5.	piąty respondent	22222
6.	szósty respondent	35161
7.	siódmy respondent	11342

itd.

Dla pięciu zmiennych istnieje pięć kolumn, ponieważ wszystkie zapisy są jednocyfrowe. Każda zmienna zajmuje swoją kolumnę. Aby było można opracować takie dane, każda zmienna musi we wszystkich wierszach (czyli u wszystkich respondentów) zachować swoje miejsce. Jeżeli u pierwszego respondenta jest np. na drugim miejscu, to musi być u każdego na drugim miejscu. Kolumna zawsze musi zawierać dane dla tej samej zmiennej.
Nie ma znaczenia fakt, że wszystkie pytania w ankiecie nie są ilościowe (numeryczne). W ten sposób zapisuje się dane także dla wyłącznie jakościowych zmiennych. W przypadku płci należy np. dla kobiet zapisać liczbę 1, dla mężczyzn liczbę 2. Takie postępowanie nie zmienia natury zmiennych. Nie oznacza to, że nie uwzględnia się właściwości zmiennych. Łatwiejsze

jest jedynie wprowadzanie danych do komputera, a także ich późniejsze opracowanie.
Proste rozwiązania są najlepsze i przynoszą najmniej błędów. Dlatego też odpowiedziom na pytania w ankiecie (kategoriom zmiennej) przyporządkowuje się liczby w tej samej kolejności, po której były podane w ankiecie (pierwszej odpowiedzi zawsze 1, drugiej 2, trzeciej 3 itd.). Przy takim postępowaniu możliwość wystąpienia błędów jest najmniejsza.
Nawet w przypadku, gdy w pytaniu ankietowym nr 5 odpowiedzi zostały umieszczone w następujący sposób:

A. zawsze	B. często	C. rzadko	D. nigdy

Natomiast w pytaniu nr 7:

A. nigdy	B. rzadko	C. często	D. zawsze

W obydwu przypadkach odpowiedź A należy zapisać jako 1 (bez względu na to, że to dwie zupełnie przeciwne odpowiedzi). Istnieje wprawdzie możliwość oznaczania jednakowych odpowiedzi zawsze taką samą liczbą (bez względu na fakt, czy były w ankiecie na pierwszym, drugim, trzecim... miejscu), jednak powoduje to komplikacje i zwiększa liczbę błędów. Takie dylematy należy przewidzieć i rozwiązywać wcześniej – przy układaniu pytań ankietowych.
W przypadku, gdy proponowanych odpowiedzi jest więcej niż dziewięć, nie można już zapisać odpowiedzi za pomocą jednej cyfry – konieczny jest zapis dwucyfrowy. Podobnie jest ze zmiennymi ilościowymi. Są to zmienne, takie jak: godziny pracy, rezultaty w testach, wiek, liczba błędów w wypracowaniu itd. W przypadku, gdy najwyższy wynik w teście wynosi np. 105 punktów, należy użyć dla wszystkich odpowiedzi zapisu trzycyfrowego: 004, 025, 067, 100, 105 itd. Dlatego też niektóre zmienne w bazie danych posiadają dwie, trzy lub jeszcze więcej kolumn.
Przy planowaniu liczby kolumn w przypadku pytań zamkniętych na ogół nie ma żadnych komplikacji. Dla każdego pytania ankietowego na podstawie proponowanych odpowiedzi wiadomo, czy wystarczający będzie zapis jednocyfrowy, czy będzie potrzebny dwucyfrowy, trzycyfrowy itd.
W taki sposób powstanie szereg danych dla każdego respondenta. Jeżeli brak jakiejś odpowiedzi (np. respondent nie zakreślił żadnej odpowiedzi

lub odpowiedział, ale nie da się rozpoznać co zakreślił itp.), należy w tym miejscu zapisać zero lub spację. Taki zapis powoduje, że wiersze dla wszystkich respondentów są jednakowej długości.
Narzędzia do zbierania danych nie zawsze są takie proste, jak w omawianych przypadkach. Problemy pojawiają się najczęściej w takich pytaniach ankietowych, w których można zakreślić jednocześnie więcej niż jedną odpowiedź. Niektórzy respondenci zakreślają jedną odpowiedź, niektórzy dwie lub więcej niż dwie, natomiast niektórzy nie zakreślają żadnej. Stąd też dla niektórych respondentów pojawia się więcej danych, dla innych natomiast tylko jedna lub nawet żadna. Jak zatem zapisać odpowiedzi, aby zachować prostotę zapisu i równocześnie umieścić wszystkie dane? Najprostsze rozwiązanie jest następujące: z **jednego** pytania ankietowego zrobić **więcej** zmiennych – tyle, ile jest proponowanych odpowiedzi. Rzecz jasna, iż pytanie w ankiecie było jedno, a ankieta już dawno została przeprowadzona. Podział jednego pytania ankietowego na kilka zmiennych powstaje dopiero przy wprowadzaniu danych do komputera. Wszystkie powstałe w ten sposób zmienne będą jednocyfrowe. W przypadku, gdy respondent zakreślił pewną odpowiedź zapisuje się 1, jeżeli jej nie zakreślił zapisuje się 0. Ilustruje to przykład pytania ankietowego z siedmioma proponowanymi odpowiedziami dla ośmiu respondentów. W przykładzie istniała możliwość zakreślenia dowolnej liczby odpowiedzi.

Tabela 16. Odpowiedzi respondentów

Respondenci							
pierwszy	**drugi**	**trzeci**	**czwarty**	**piąty**	**szósty**	**siódmy**	**ósmy**
A	A	Ⓐ	Ⓐ	A	Ⓐ	Ⓐ	A
Ⓑ	Ⓑ	B	B	Ⓑ	Ⓑ	Ⓑ	B
Ⓒ	C	Ⓒ	Ⓒ	C	Ⓒ	Ⓒ	C
Ⓓ	Ⓓ	Ⓓ	D	Ⓓ	Ⓓ	Ⓓ	D
E	Ⓔ	Ⓔ	Ⓔ	E	E	Ⓔ	E
Ⓕ	F	F	F	Ⓕ	F	Ⓕ	F
G	Ⓖ	G	G	Ⓖ	G	Ⓖ	G

Należy to zapisać następująco:

1. pierwszy respondent 0111010
2. drugi respondent 0101101
3. trzeci respondent 1011100
4. czwarty respondent 1010100

5. piąty respondent 0101011
6. szósty respondent 1111000
7. siódmy respondent 1111111
8. ósmy respondent 0000000

Siódmy respondent zakreślił wszystkie odpowiedzi, więc wpisano siedem razy jeden. Ósmy respondent natomiast nie zakreślił żadnej odpowiedzi, więc wpisano siedem razy zero. W ten sposób powstały wiersze danych jednakowej długości dla wszystkich respondentów. Zjawisko to występuje bardzo często i z tego powodu zostanie dodatkowo wyjaśnione. W ankiecie pytano respondentów: Jakie są motywy podjęcia przez Panią/Pana studiów na kierunku pedagogika? Respondenci mogli zakreślić więcej niż jedną odpowiedź. Oto proponowane odpowiedzi:

A moi rodzice są nauczycielami
B lubię pracę grupową
C będę miał/a długie wakacje
D lubię pracować z dziećmi
E będę miał/a możliwość awansu zawodowego

Zadano tylko **jedno** pytanie, jednak każdy respondent mógł udzielić **kilku** odpowiedzi. W celu wprowadzania i opracowywania danych pytanie to w myślach należy przekształcić w kilka podobnych pytań. Odpowiedzi traktuje się w taki sposób, jak gdyby kolejne pytania brzmiały:

- „Czy Pani/Pan zdecydowała się na studiowanie pedagogiki, ponieważ rodzice są nauczycielami?” TAK NIE
- „Czy Pani/Pan zdecydowała się na studiowanie pedagogiki, ponieważ lubi pracę grupową?” TAK NIE
- „Czy Pani/Pan zdecydowała się na studiowanie pedagogiki z powodu długich wakacji?” TAK NIE
- „Czy Pani/Pan zdecydowała się na studiowanie pedagogiki, ponieważ lubi pracę z dziećmi?” TAK NIE
- „Czy Pani/Pan zdecydowała się na studiowanie pedagogiki z powodu możliwości awansu zawodowego?” TAK NIE

Jeżeli respondent zakreślił A, to tak jak gdyby na pierwsze zasymulowane w myślach pytanie udzielił odpowiedzi TAK. Jeśli go nie zakreślił, to jakby odpowiedział NIE. Tak należy postępować od pierwszej do ostatniej proponowanej odpowiedzi.

Ten rodzaj pytań można traktować jako jedną zmienną tylko w przypadku, jeżeli opracowuje się go samodzielnie (bez poszukiwania związków z innymi zmiennymi). Przy badaniu związków i zależności z innymi pytaniami lub zmiennymi ten rodzaj pytań należy traktować w powyżej opisany sposób (jako więcej zmiennych). Nie można analizować np. różnic według płci dla całego pytania, lecz dla każdej odpowiedzi oddzielnie (jedną po drugiej). Na przykład nie można sporządzić następującej tabeli:

Tabela 17. Odpowiedzi respondentów

	A	B	C	D	E	Razem
Kobiety						
Mężczyźni						
Razem						

Ten sam respondent może się znajdować w kilku okienkach tabeli, np. pewna kobieta mogła udzielić odpowiedzi A, D i E. Wtedy „znajduje" się ona w trzech okienkach pierwszego wiersza i z tego powodu nie można obliczyć procentów. W konsekwencji nie można ustalić (zbadać), jaka jest współzależność płci i motywów wyboru studiów. Można natomiast sporządzić tabelę w następujący sposób:

Tabela 18. Odpowiedzi respondentów

	Odpowiedź A („Czy pani/pan zdecydowała się na studiowanie pedagogiki, ponieważ rodzice są nauczycielami?")		
	Tak	**Nie**	**Razem**
Kobiety	40 58,0%	29 42,0%	69 100,0%
Mężczyźni	12 30,8%	27 69,,2%	39 100,0%
Razem	52 48,1%	56 51,9%	108 100,0%

Tabela 19. Odpowiedzi respondentów

	Odpowiedź B („Czy pani/pan zdecydowała się na studiowanie pedagogiki, ponieważ lubi pracę grupową?")		
	Tak	**Nie**	**Razem**
Kobiety	30 43,5%	39 56,5%	69 100,0%
Mężczyźni	11 28,2%	28 71,8%	39 100,0%
Razem	41 38,0%	67 62,0%	108 100,0%

Powstanie w ten sposób tyle podobnych do siebie tabel, ile istnieje proponowanych odpowiedzi (czyli tyle, ile powstanie zmiennych). Tabele te umożliwiają porównywanie odpowiedzi kobiet i mężczyzn, ale tylko według poszczególnych odpowiedzi, a nie według całego pytania ankietowego.

Problemy opisane powyżej pokazują, iż z takiego pytania ankietowego należy zrobić tyle zmiennych, ile jest proponowanych odpowiedzi. Czynność ta jest skomplikowana, jednak nie istnieje inne rozwiązanie. Jeżeli współzależności nie są ważne, do kwestionariusza można włączyć podobne pytania. Jeżeli jednak współzależności są ważne dla badania, należy postępować w następujący sposób:

- w kwestionariuszu należy ograniczyć możliwość wyboru odpowiedzi do jednej (np. „który z podanych motywów był **najważniejszy** przy wyborze studiów pedagogiki?"),
- do kwestionariusza należy włączyć dwa podobne do siebie pytania (jedno z nich dotyczy wszystkich motywów i istnieje możliwość wyboru kilku odpowiedzi, drugie dotyczy jednego – zwykle najsilniejszego motywu).

Pozostałe, jeszcze bardziej złożone przypadki nie będą omówione, ponieważ książka ta ze względu na ograniczoną objętość koncentruje się na najczęściej spotykanych przypadkach.

W przypadku prostych kwestionariuszy dane wprowadzane są bezpośrednio z kwestionariuszy do komputera, bez wcześniejszego ich przygotowania. W przypadku skomplikowanych kwestionariuszy lepiej wprowadzać dane najpierw z kwestionariusza na kartki (karty kodowe). Każdy respondent posiada swoją kartkę lub przynajmniej swój wiersz. Z wypełnionych kartek dane wprowadza się do komputera. O tym czy dane wprowadza się bezpośrednio, czy pośrednio, decyduje doświadczenie i praktyka w posługiwaniu się komputerem. Przy bezpośrednim wprowadzaniu jest mniej pracy, ale jednocześnie pojawia się więcej błędów.

Rozdział III

WARTOŚCI ŚREDNIE

1. PORÓWNYWANIE POPULACJI

Jednostki (w grupie) według pewnej zmiennej porównuje się w ten sposób, że dokonuje się komparacji ich wartości, osiągniętych w tej zmiennej. Oznacza to, że jeżeli np. na teście wiadomości z matematyki uczniowie Damian i Sebastian osiągnęli 22 i 27 punktów, można wnioskować, iż Sebastian jest lepszy (osiągnął wyższy wynik). W niektórych zmiennych wprawdzie wyższa wartość oznacza gorszy wynik (np. większa liczba błędów w dyktandzie), jednak nie zmienia to sedna porównań osiągnięć jednostek.
Jak postępować, gdy porównuje się całe populacje statystyczne? Dlaczego należy porównywać populacje? Oto odpowiedź – istnieje potrzeba ustalenia:

- Czy test lepiej rozwiązali uczniowie klasy IV A czy klasy IV B?
- Czy pensje nauczycieli w szkołach średnich różną się od pensji nauczycieli w szkołach podstawowych?
- Czy studenci zaoczni studiują dłużej niż studenci dzienni?

itp.

Porównywujac dwie grupy, można porównać wszystkie wyniki z pierwszej grupy z wszystkimi wynikami z drugiej grupy. Jednak od razu można stwierdzić, że ten sposób może okazać się skuteczny tylko i wyłącznie w przypadkach, gdy grupy liczą zaledwie kilka jednostek. Dla trochę większych grup nie jest to możliwie. Wówczas można wybrać z każdej grupy tylko jedną jednostkę do porównywania, pytanie jednak którą? Jeżeli najlepszy uczeń z jednej klasy osiągnął wyższy wynik niż najlepszy uczeń z drugiej klasy, nie oznacza to wcale, że jego klasa jest lepsza. Podobnie jest z wynikami najniższymi. Takie postępowanie nie daje rzetelnej i dobrej komparacji. Najbardziej rzetelne komparacje umożliwiają wartości umieszczone w centrum grupy. Te wartości są typowe i najbardziej liczne. Dlatego też należy porównywać grupy według takich właśnie wartości, które nazywa się miarami tendencji centralnej.

Tendencja centralna to zjawisko polegające na tym, że wartości zmiennych ilościowych na ogół skupiają się w środku (w centrum). Im bliżej centrum, tym wartości bardziej zgęszczone; im dalej od centrum, tym występują rzadziej. Można to zilustrować przykładem: w większej grupie 11-letnich uczniów ich wzrost koncentruje się w przybliżeniu wokół wartości 138 cm. Większość uczniów znajduje się niedaleko tej („przeciętnej") wartości. Uczniów, których wzrost przewyższa 150 cm jest stosunkowo mało, a takich o wzroście ponad 160 cm jest jeszcze mniej. Podobnie jest też w przypadku niższych wartości. Jest to tendencja centralna i z tego powodu w większości przypadków najlepiej porównywać populacje według takich wartości (które znajdują się w centrum).

Naturze zjawisk i zmiennych w obszarze edukacji najbardziej odpowiadają następujące trzy miary tendencji centralnej:

- wartość modalna lub dominanta (Mo),
- mediana (Me),
- średnia arytmetyczna (M).

2. WARTOŚĆ MODALNA

Modalną jest punkt (wartość), gdzie wartości zmiennej są najbardziej zagęszczone. W najprostszym ujęciu jest to wartość, która pojawia się najczęściej. Przykład: jeżeli w klasie większość uczniów posiada ocenę dobrą (4) z określonego przedmiotu, to ocena ta jest modalną (Mo=4,00).

Modalną można ustalić nawet dla zmiennych nominalnych (jednak jest prawdą, że to nie ma prawie żadnego praktycznego znaczenia). Modalna jest najprostszą miarą tendencji centralnej: szybka i prosta do obliczeń, prosta do zrozumienia i interpretacji. Z powodu tego, że dla określenia wartości modalnej nie są potrzebne wszystkie dane w grupie, porównania grup są często nierzetelne. Oprócz tego w grupie może się pojawić więcej wartości modalnych – wówczas porównanie staje się niemożliwe.

3. MEDIANA

Medianą jest wartość (punkt), od której połowa grupy posiada wartości wyższe, a połowa niższe. Szereguje się na przykład uczniów w klasie według wzrostu od najmniejszego do największego. Należy odnaleźć ucznia, który jest w centrum (połowa uczniów w klasie jest wyższa, a połowa niższa). Wzrost tego ucznia jest medianą (np. 161 cm).
Dla określenia mediany potrzebna jest przynajmniej zmienna porządkowa. Dla zmiennych nominalnych nie da się określić mediany, ponieważ nie ma wartości wyższych i niższych. Mediana jest też prostą miarą tendencji centralnej: często można ją określić bez obliczeń, jest ona łatwa do zrozumienia i wystarczająco prosta do interpretacji. Do określania mediany także nie są potrzebne dane dla wszystkich jednostek w grupie. Z tego powodu mediana nie jest bardzo rzetelna. Mimo to często znajduje ona zastosowanie w badaniach edukacji, ponieważ w obszarze tym często bada się zmienne porządkowe.

Obliczanie

W przypadkach małych populacji medianę można łatwo określić bez jakichkolwiek wzorów i obliczeń. Natomiast w przypadkach populacji z większą liczebnością i tak nie używa się ręcznego opracowania. Jedyną potrzebną procedurą jest sortowanie wartości do szeregu prostego. Oto kilka przykładów określania mediany z szeregu prostego.

Przykład 1.

x	4	8	10	12	14	15	19	21	21	24
R	1	2	3	4	5	9	6	7	8	9
	50% wartości					**50% wartości**				
Me= 14,00										

Przykład 2.

x	3	5	9	9	13	14	16	17	19	24
R	1	2	3	4	5	6	7	8	9	10
	50% wartości					**50% wartości**				
mediana jest pośrodku pomiędzy wartościami 13 i 14 $M_e = \frac{13+14}{2}$ $M_e = 13{,}50$										

Przykład 3.

x	**10**	**11**	**15**	**16**	**16**	**22**	**27**	**29**
R	1	2	3	4	5	6	7	8
	50% wartości				**50% wartości**			
mediana jest pośrodku pomiędzy wartościami 16 i 16 $M_e = 16{,}00$								

Przykład 4.

x	**4**	**4**	**5**	**6**	**8**	**15**	**16**	**18**	**19**	**19**
R	1	2	3	4	5	6	7	8	9	10
	50 % wartości					**50 % wartości**				
mediana jest pośrodku pomiędzy wartościami 8 i 15 $M_e = \frac{8+15}{2}$ $M_e = 11{,}50$										

Przykład 5.

x	2	2	3	4	8	11	15	15	18	21	22	22
R	1	2	3	4	5	6	7	8	9	10	11	12
	50 % wartości						**50 % wartości**					
mediana jest pośrodku pomiędzy wartościami 11 i 15 $M_e = \frac{11+15}{2}$ $M_e = 13{,}00$												

4. ŚREDNIA ARYTMETYCZNA

Średnia arytmetyczna jest sumą wszystkich wartości podzieloną przez liczebność grupy. Oblicza się ją według wzoru:

$$M = \frac{\Sigma x}{N}$$

M	średnia arytmetyczna
$\sum x$	suma wszystkich wartości ($x1+ x2+ x3+$ v4$+ x5+$...)
N	liczebność grupy

Sumować można tylko wartości zmiennych przedziałowych. Dla zmiennych nominalnych nie da się zastosować średniej arytmetycznej, a dla zmiennych porządkowych jej stosowanie jest nieuzasadnione (niepoprawne). Fakt, że określa się ją na podstawie wszystkich wyników, powoduje, że umożliwia ona najbardziej rzetelne porównania. Średnia arytmetyczna jest czuła na najmniejsze zmiany w grupie: jeżeli zmienia się wartość nawet jednej jedynej jednostki – od razu zmienia się średnia arytmetyczna. Do obliczania średniej arytmetycznej potrzebne są wszystkie wyniki, dlatego też średnia arytmetyczna wymaga najwięcej pracy. Interpretacja średniej arytmetycznej nie jest już tak prosta, jak w przypadku mody i mediany. Fakt, iż średnia arytmetyczna uwzględnia wszystkie rezultaty w grupie, jest jednocześnie jej słabością: uzależniona jest też od najbardziej ekstremalnych wartości (te mogą być w niektórych przypadkach bardzo nietypowe i dają ułomny obraz grupy).
W badaniach najczęściej stosuje się średnią arytmetyczną. Ma ona zastosowanie w przypadku zmiennych przedziałowych i ilorazowych. W przypadku zmiennych porządkowych stosuje się medianę, natomiast wartość modalną stosuje się rzadko.
Kolejnym problemem, który należy omówić, są sytuacje specjalne, dotyczące okoliczności stosowania miar tendencji centralnej w obszarze edukacji. Tu często pojawiają się zmienne porządkowe, np. oceny szkolne. Dla takich zmiennych stosuje się medianę. Różne okoliczności mogą spowodować, że stosowanie mediany nie umożliwia porównania, a stosowanie średniej arytmetycznej jest nieuzasadnione. Taki przypadek można zilustrować za pomocą danych z dwóch klas szkolnych:

Tabela 20. Oceny w dwóch klasach szkolnych

Klasy	**Oceny**	**Mediana**	**Średnia arytmetyczna**
IV A	3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6	Me = 4,00	M = 4,18
IV B	2, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5	Me = 4,00	M = 3,64

Porównanie median prowadzi do błędnego wniosku, że klasy według ocen są jednakowe. Proste przejrzenie ocen dowodzi, iż nie jest to prawdą. Zastosowana średnia arytmetyczna nie odpowiada naturze ocen szkolnych, ale pokazuje, iż oceny w klasie IV A są wyższe (co jest prawdą). Upieranie się przy stosowaniu całkowicie odpowiednich metod uniemożliwia w tym przypadku dokonanie porównania. Można jednak zastosować średnią arytmetyczną, lecz z zachowaniem wielkiej ostrożności podczas interpretacji. Średnia arytmetyczna umożliwia stwierdzenie, że oceny w klasie IV A są wyższe. Twierdzenie natomiast, iż oceny są wyższe precyzyjne za 0,54, jest nieostrożne i nieprawdziwe, pomimo różnicy średnich arytmetycznych wynoszącej rzeczywiście 0,54.
Można zatem stosować metody w warunkach niezupełnie odpowiednich, jednak należy wiedzieć, co to oznacza i w jaki sposób uwzględnić to w interpretacji.

Obliczanie

Z powodów praktycznych zostanie szczegółowo opisane tylko obliczanie średniej arytmetycznej z danych nie zgrupowanych.
W ankiecie zapytano nauczycieli o staż pracy. Oto odpowiedzi uszeregowane do szeregu prostego:

x 3 3 4 5 6 9 9 10 13 16 20

Do obliczenia średniej arytmetycznej należy zsumować wszystkie wartości, a następnie podzielić tę sumę przez liczebność populacji.

$$M = \frac{\Sigma x}{N} = \frac{3+3+4+5+6+9+9+10+13+16+20}{11} = \frac{98}{11} = 8{,}91$$

Wiadomo, że pokazana procedura służy raczej do wyjaśnienia, niż do praktycznego stosowania. W praktyce nie jest niezbędne tworzenie szeregu. Nie należy również wszystkiego zapisywać w długi (rozbudowany) sposób. Najlepiej przyzwyczaić się do stosowania takich tabel:

x	3	3	4	5	6	9	9	10	13	16	20	Σx= 98

N =11

$$M = \frac{\Sigma x}{N} = \frac{98}{11} = 8,91$$

Mając średnią arytmetyczną można tę populację porównać z którąkolwiek inną populacją (według tej zmiennej). Porównywania te są głównym celem stosowania średnich wartości. Średnia arytmetyczna znajduje zastosowanie także w niektórych innych procedurach, np. przy pomiarze rozproszenia.

Rozdział IV

ROZPROSZENIE

1. POJĘCIE ROZPROSZENIA

Rozproszenie wyników lub zmienność jest zjawiskiem bardzo prostym. Polega ono na tym, że jednostki populacji według wartości pewnej zmiennej różnią się między sobą. Można to zilustrować przykładami zaprezentowanymi w poniższej tabeli.

Tabela 21. Rozproszenie wartości zmiennych

Populacja	**Zmienna**	**Wartości**
Studenci na studiach zaocznych	staż pracy	5, 8, 9, 11, 14, 19
Nauczyciele	płeć	12 kobiet i 6 mężczyzn
Zatrudnieni w WOM	stopień wykształcenia	19 z wykształceniem wyższym, 5 z wykształceniem średnim i 4 z wykształceniem zawodowym
Dyrektorzy szkół podstawowych	wiek	36, 36, 41, 44, 44, 48, 51, 52
Szkoły językowe	liczba nauczycieli	3, 7, 5, 10, 9, 4, 6, 8, 11

We wszystkich przedstawionych przypadkach pojawia się zmienność. Zmienności nie byłoby, gdyby na przykład wszyscy dyrektorzy byli w tym

samym wieku, wszyscy nauczyciele byli tej samej płci, wszyscy zatrudnieni w WOM posiadali ten sam stopień wykształcenia, wszyscy studenci mieli taki sam staż pracy, a szkoły językowe taką samą liczbę nauczycieli.
Z powyższej tabeli wynika, iż w zmiennej „stopień wykształcenia" nie wszystkie jednostki różnią się od siebie. Pomimo to w grupie istnieje zmienność. Wystarczy bowiem, że chociaż jedna jednostka różni się od pozostałych. Podobnie jest z drugą i czwartą zmienną. W pierwszej i piątej zmiennej natomiast wszystkie jednostki się różnią.

2. ŹRÓDŁA ROZPROSZENIA

Dlaczego wszystkie wartości nie są jednakowe? Dlaczego jednostki różnią się od siebie? Dlaczego więc pojawia się zmienność?
W pewnej grupie wszystkie jednostki mogą być jednakowe (bardziej precyzyjnie: wartości pewnej zmiennej mogą być u wszystkich jednostek identyczne). Zmienność w takiej grupie jest zerowa, co oznacza, że jej nie ma. Rodzi się, więc pytanie: dlaczego wszystkie przypadki nie są takie same? Analiza o jakiejkolwiek konkretnej zmiennej szybko prowadzi do poprawnej odpowiedzi. Wszystkie noworodki byłyby tej samej płci, gdyby nic nie miało wpływu na ich płeć. Wszyscy uczniowie otrzymaliy taką samą ocenę z pewnego przedmiotu, jeżeli nic nie miałoby wpływu na te oceny. Przyczyną zmienności są więc wpływy. Oczywistym jest fakt, że na oceny ma wpływ na przykład motywacja (bardziej zmotywowani uczniowie na ogół otrzymują wyższe oceny), warunki do uczenia się w domu (im lepsze warunki, tym lepsze oceny), czas uczenia się i wiele innych czynników. Zmienność pewnej zmiennej jest zawsze skutkiem wpływu innych zmiennych. Zjawisko to otwiera szereg nowych pytań.
Pytanie pierwsze: ile jest czynników, które mają wpływ? Na ogół jest ich ogromna ilość. Pedagogika nie jest w stanie odkryć wszystkich czynników, a tym bardziej ich zbadać. Nawet w najprostszych zjawiskach edukacyjnych nie da się odkryć wszystkich czynników. Zwykle udaje się wyjaśnić tylko najważniejsze (a nawet i te jedynie częściowo!). Nie oznacza to jednak wcale, iż badanie i wyjaśnianie sedna zjawisk nie ma sensu i prowadzi donikąd. Każdy nawet częściowo opisany i wyjaśniony wpływ dostarcza nowej wiedzy pedagogicznej.
Powstaje kolejne pytanie: czy jeżeli w pewnej populacji nie ma zmienności, to oznacza, iż na tę zmienną nic nie ma wpływu? Odpowiedź brzmi: nie! Być może wpływy istnieją, lecz ich efekty zostały zrównoważone (suma

pozytywnych efektów równa się sumie negatywnych efektów). Takie przypadki w praktyce zdarzają się tylko incydentalnie. Oznacza to jednocześnie, że bardzo rzadkie są przypadki, w których pewna populacja posiada zerowe rozproszenie. Zagadnienie to można zilustrować na przykładzie płci: pomimo, że profesja nauczycielska jest wyraźnie sfeminizowana, tylko incydentalnie można znaleźć szkołę, w której nie pracuje nawet jeden nauczyciel. W mniejszych populacjach to zdarza się częściej (im mniejsze populacje, tym częściej). Nauczycielami pierwszej klasy jednej szkoły podstawowej często są same kobiety, jednak wśród nauczycieli pierwszych klas w całym województwie można już znaleźć kilku mężczyzn.
Wyjaśnia to potrzebę badania rozproszenia: głównie po to, aby zbadać wpływy (współzależności między zjawiskami). Jest to powód, dla którego rozproszenie jest centralnym zjawiskiem statystyki. Rozproszenie stanowi fundament większości metod statystycznych.

3. POMIAR ROZPROSZENIA

Rozproszenie to różnice między jednostkami. Wszystkie sposoby pomiaru rozproszenia oparte są na tych różnicach.
Ze względu na fakt, że rozproszenie jest środkiem badania współzależności między zjawiskami, prawie nigdy nie oblicza się go bezpośrednio. Najczęściej pomiar rozproszenia jest częścią innych procedur statystycznych.

Pomiar rozproszenia zmiennych nominalnych
Przy zmiennych nominalnych, których wartości (kategorie) nie podlegają stopniowaniu, można ustalić jedynie to, czy jednostki różnią się od siebie, czy też nie. W takich przypadkach nie można dokonywać pomiaru wielkości różnic, lecz tylko ich istnienia. Pomiar rozproszenia dokonuje się poprzez ustalanie ilości (zliczenie) jednostek, które różnią się od siebie. Sposób takiego pomiaru rozproszenia ilustrują poniższe przykłady. Pierwszym z nich jest najprostszy przypadek rozproszenia według płci.

Przykład 1:

W sali wykładowej znajduje się 50 studentek i nie ma żadnego studenta. Jednostki tej populacji nie różnią się między sobą, czyli rozproszenie jest równe zeru.

Przykład 2:

W sali wykładowej znajduje się jeden student i 49 studentek. Różni się on od pozostałych 49 jednostek płcią, co oznacza, że w tej populacji istnieje określone rozproszenie (niewielkie).

W przykładzie nr 2 tylko jedna jednostka różni się od pozostałych. Oznacza to najmniejsze możliwe rozproszenie. Im więcej różnych jednostek, tym większe rozproszenie. W przypadku płci największe możliwe rozproszenie występuje wtedy, gdy połowa jednostek w grupie jest jednej płci, a połowa – drugiej płci.

W przypadku zmiennych, które posiadają więcej niż dwie kategorie, sytuacja jest nieco odmienna, lecz w istocie taka sama: im więcej różnych jednostek, tym większe rozproszenie. Najprostszą okazuje się sytuacja, w której ilość kategorii jest większa niż ilość jednostek w grupie.

Największe możliwe rozproszenie występuje wtedy, gdy wszystkie jednostki różnią się między sobą. Ilustruje to przykład nr 3.

Przykład 3:

W sali wykładowej znajduje się 19 słuchaczy i każdy z nich jest innej narodowości. Jest to przykład największego rozproszenia.

Najprostszym sposobem dokonywania porównań pomiędzy populacjami według rozproszenia jest porównywanie procentów poszczególnych kategorii. Nie istnieje jednak specjalna miara rozproszenia zmiennych nominalnych.

Pomiar rozproszenia zmiennych porządkowych

Wartości zmiennych porządkowych podlegają stopniowaniu. Stopnie tych zmiennych nie narastają równomiernie (przedziały pomiędzy sąsiadującymi stopniami nie są jednakowe), dlatego też różnice pomiędzy jednostkami są niemożliwe do zmierzenia. Jako przykład można zakładać, że pomiędzy wykształceniem średnim a licencjackim istnieje większa różnica niż pomiędzy wykształceniem licencjackim a magisterskim. Trudno natomiast określić, jak wielkie są te różnice (lub czy w ogóle pierwsza różnica jest większa od drugiej).

Dlatego też na ogół w takich przypadkach stosuje się takie samo rozwiązanie, jak przy zmiennych nominalnych – pomiaru rozproszenia dokonuje się poprzez ustalanie liczby jednostek, które różnią się od siebie.

Zmienne porządkowe posiadają jednak podstawę ilościową, szkoda zatem zupełnie pominąć wielkości różnic między ich poszczególnymi stopniami. Pomysł uwzględnienia wymienionych różnic jest ciekawy, lecz bardzo trudny do realizacji. Rozwiązanie to prowadziłoby do niezwykle skomplikowanych miar rozproszenia, a jednocześnie trudnych w interpretacji. Dlatego też w praktyce stosuje się dwie następujące procedury:

1. zmienne porządkowe traktuje się najczęściej tak samo jak zmienne nominalne – rozproszenie mierzy się poprzez ustalenie, ile różnych wartości pojawia się w grupie,
2. niekiedy zmienne porządkowe traktuje się jak zmienne przedziałowe – wówczas dla obydwu stosuje się identyczną procedurę obliczania rozproszenia (patrz: następny podrozdział).

Identyczne traktowanie zmiennych porządkowych oraz przedziałowych dotyczy tylko procedury obliczenia miary rozproszenia, nigdy natomiast interpretacji otrzymanych rezultatów. Przy interpretacji wyników należy uwzględniać naturę zmiennych porządkowych. Oznacza to tym samym, że w interpretacji rezultatów badacz musi być bardzo ostrożny.

Pomiar rozproszenia zmiennych przedziałowych

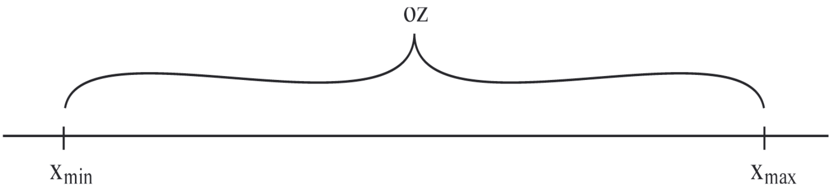
Podrozdział ten traktuje o zmiennych przedziałowych. Jednak wszystkie treści w nim zawarte będą odnosiły się także do zmiennych ilorazowych. W pomiarze rozproszenia zmienne ilorazowe nie są bardziej przydatne niż zmienne przedziałowe. Podstawową właściwością przy pomiarze rozproszenia jest identyczność przedziałów, którą posiadają obydwa rodzaje skal. Dla zmiennych przedziałowych można mierzyć różnice między wartościami (bo istnieje jednostka pomiaru). Dlatego też pomiar rozproszenia zmiennych przedziałowych oparty jest na wielkości różnic, a nie tylko na ich liczebności. Dla zmiennych przedziałowych używa się kilku miar rozproszenia. Oto najbardziej znane:

- obszar zmienności (rozstęp),
- odchylenie ćwiartkowe,
- odchylenie przeciętne,
- wariancja,
- odchylenie standardowe.

Pierwsze dwie miary stosuje się w najprostszych analizach, a więc incydentalnie rzadko. Służą one raczej do rozumienia pozostałych miar. Dlatego miary te można prawie pominąć (wnikliwy czytelnik może ich omówienie znaleźć w innej literaturze). Należy jednocześnie nadmienić, że miary te oblicza się z wartości tylko niektórych (właściwie dwóch) rezultatów z całej populacji. Jest to powód, dla którego miary te są nierzetelne i rzadko stosowane.

Obszar zmienności

Obszar zmienności jest przedziałem między najniższą i najwyższą wartością w populacji, czyli:



Rysunek 3. Obszar zmienności

$$OZ = x_{max} - x_{min}$$

W przypadkach zmiennych nieciągłych należy dodać 1, czyli wzór jest następujący:

$$OZ = x_{max} - x_{min} + 1$$

Odnosi się to do zmiennych, takich jak: liczba punktów w teście, liczba książek w bibliotece szkolnej, wiek wyrażony w latach itd. Bardzo uproszczone wyjaśnienie potrzeby dodawania wartości 1 brzmi tak: przedział od 5 do 6 ma na ogół szerokość jeden. Jeżeli natomiast chodzi o szerokość przedziału od 5 do 6 punktów w teście, to przedział ten zawiera dwa wyniki: 5 punktów i 6 punktów. W takim pojęciu szerokość przedziału wynosi 2, a nie 1 (czyli 6-5+1). Należy jednak podkreślić wyraźnie, że wyjaśnienie to jest zbyt uproszczone. Czytelnik, którego to wyjaśnienie nie zadowala, może skorzystać obszerniejszego podręcznika ze statystyki.

Przykład:
Wyniki uczniów w teście z języka polskiego w pierwszej klasie gimnazjum:

x	7	4	18	4	16	6	15	12	8	10	10	11	10

Wartości należy uporządkować rosnąco:

x	4	4	6	7	8	10	10	10	11	12	15	16	18

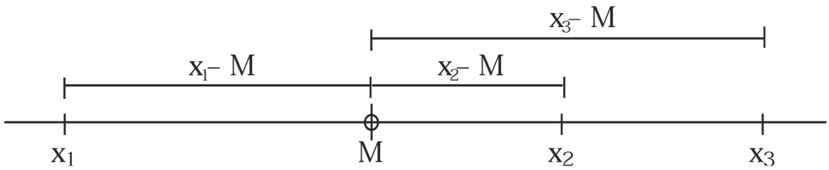
Obszar zmienności wynosi:

$$OZ = x_{max} - x_{min} + 1 = 18 - 4 + 1 = 15,00$$

Trzeciej miary rozproszenia też nie stosuje się w praktyce. Z uwagi na fakt, iż jej znajomość jest niezbędna do zrozumienia bardziej złożonej miary (wariancji), zostanie ona szczegółowo omówiona. Miara ta jest użyteczna w zrozumieniu interpretacji pierwiastka z wariancji (odchylenia standardowego). Wprawdzie zarówno wariancję, jak i odchylenie standardowe **interpretuje** się bardzo rzadko, to jednak zdarza się to w badaniach.

Odchylenie przeciętne

W przypadku, gdy celem jest otrzymanie bardziej rzetelnej i precyzyjnej miary rozproszenia, należy uwzględnić wszystkie wartości w populacji. Uwzględnia się wówczas i mierzy wszystkie różnice. Najprostszym podejściem jest zmierzenie odchylenia każdego rezultatu od średniej arytmetycznej. Są to **odchylenia indywidualne**. Wylicza się z nich średnie odchylenie. Suma wszystkich odchyleń indywidualnych dzieli się przez liczebność grupy. Otrzymany wynik jest odchyleniem przeciętnym. Przy obliczaniu nie uwzględnia się kierunku odchyleń indywidualnych (wartości dodatnich lub ujemnych). Suma wartości ujemnych jest identyczna z sumą wartości dodatnich, dlatego też uwzględnianie znaku plus lub minus prowadziłoby do sumy końcowej równej zero. W konsekwencji takie działanie prowadzi do nietrafnego wniosku, który brzmi: rozproszenie równa się zeru.
Poniżej zaprezentowano graficznie odchylenia indywidualne:



Rysunek 4. Odchylenia indywidualne

Wartość odchylenia przeciętnego interpretuje się, wychodząc z przesłanki: wszystkie dane odchylają się od średniej arytmetycznej – niektóre z nich mniej, inne więcej, a omawiane odchylenie jest średnim odchyleniem wszystkich rezultatów. Odchylenie przeciętne oblicza się według następującego wzoru:

$$OP = \frac{\Sigma|x - M|}{N}$$

OP	odchylenie przeciętne
$|x{-}M|$	odchylenie indywidualne (bezwzględne)
$\Sigma| x{-}M|$	suma wszystkich odchyleń indywidualnych (bezwzględnych)

Obliczanie odchylenia przeciętnego
Pierwszym krokiem jest określenie średniej arytmetycznej. Następnie należy obliczyć odchylenia indywidualne, czyli odchylenie każdej wartości od średniej arytmetycznej. Odchylenia indywidualne należy zsumować, a otrzymaną sumę podzielić przez liczebność populacji. Uzyskany wynik jest odchyleniem przeciętnym.

Przykład:

x	5	3	8	11	6	4	5	∑x=42

$$M = \frac{\Sigma x}{N} = \frac{42}{7} = 6{,}00$$

$$OP = \frac{\Sigma|x - M|}{N} = \frac{|5-6| + |3-6| + |8-6| + |11-6| + |6-6| + |4-6| + |5-6|}{7}$$

$$OP = \frac{14}{7} = 2{,}00$$

Można zaproponować również prosty sposób z użyciem tabeli:

x	5	3	8	11	6	4	5	∑x=42
|*x*-M|	1	3	2	5	0	2	1	∑|x-M|=14

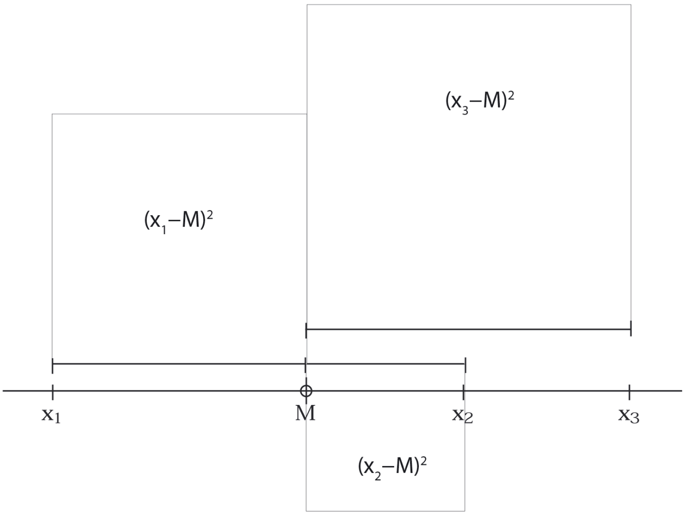
$$M = \frac{\Sigma x}{N} = \frac{42}{7} = 6{,}00$$

$$OP = \frac{\Sigma\left|x - M\right|}{N} = \frac{14}{7} = 2{,}00$$

Odchylenie przeciętne jest odcinkiem, czyli jednowymiarowym wskaźnikiem rozproszenia. Pomimo iż jest obliczone ze wszystkich rezultatów, a tym samym bardzo rzetelne, w praktyce stosuje się je rzadko. Jeżeli celem badań byłoby tylko rozproszenie pojedynczych zmiennych, miara ta byłaby wystarczająca. Ze względu na to, że rozproszenie jest potrzebne do analizy współzależności między dwoma zmiennymi, bardziej przydatny okazuje się wskaźnik dwuwymiarowy.

Wariancja i odchylenie standardowe

Dwuwymiarowy wskaźnik rozproszenia można otrzymać w prosty sposób: wszystkie odchylenia indywidualne należy podnieść do drugiej potęgi (odcinki przekształcić w kwadraty – zob. rysunek nr 4). Z powodu przejrzystości rysunku jeden z kwadratów narysowano pod linią.



Rysunek 5. Kwadraty odchyleń indywidualnych

Z tych kwadratów należy obliczyć przeciętny kwadrat (patrz drugi wzór na s. 62). Kwadrat ten nazywa się wariancją. Dla wariancji używa się symbolu σ2. Wariancja różni się tylko jedną właściwością od przeciętnego odchylenia. Tzn. zamiast jednego wymiaru wariancja posiada dwa wymiary (nie jest odcinkiem, lecz kwadratem). Cecha ta umożliwia analizę współzależności między dwoma zmiennymi. Dla analiz prostszych jako miarę rozproszenia można przyjąć pierwiastek wariancji.

$$\sigma = \sqrt{\sigma^2}$$

Oznacza to powrót do jednowymiarowego wskaźnika. Jest to bok kwadratu. Miara ta nazywa się odchyleniem standardowym i jest oznaczana symbolem σ. Odchylenie standardowe nie jest równoznaczne z odchyleniem przeciętnym, lecz jest ono interpretowane w taki sam sposób. Błąd powstały w ten sposób jest nieznaczny.
Interpretację odchylenia standardowego (i wariancji) zilustrowano przykładem pensji grupy nauczycieli. W grupie 244 nauczycieli otrzymano M = 1188,00 i σ = 184,00. Średnia arytmetyczna wskazuje, iż pensja wynosi prawie jeden tysiąc dwieście złotych. Odchylenie standardowe pokazuje, że pensje nauczycieli wahają się przeciętnie o 184 zł w dół i w górę (oczywiście niektóre pensje więcej, a niektóre mniej). Wariancja wynosi σ^2=33856 i nie ma możliwości jej sensownej interpretacji. Można interpretować, iż przeciętny kwadrat odchyleń wynosi ponad trzydzieści trzy tysiące, ale co to rzeczywiście oznacza?
Już odchylenie standardowe interpretuje się rzadko, zaś wariancji nie interpretuje się nigdy. Wariancja jest przeznaczona do analizy współzależności. Oto podstawowy wzór do obliczania wariancji:

$$\sigma^2 = \frac{\Sigma\left(x - M\right)^2}{N}$$

σ^2	wariancja
x–**M**	odchylenie indywidualne
$\Sigma(x-\mathbf{M})^2$	suma kwadratów wszystkich odchyleń indywidualnych

Obliczanie wariancji

Do prezentacji obliczeń wykorzystano ten sam przykład, na którym obliczano odchylenie przeciętne.

x	5	3	8	11	6	4	5	$\Sigma x=42$
$(x-M)$	-1	-3	2	5	0	-2	-1	
$(x-M)^2$	1	9	4	25	0	4	1	$\Sigma(x-M)^2=44$

$$M = \frac{\Sigma x}{N} = \frac{42}{7} = 6{,}00$$

$$\sigma^2 = \frac{\Sigma(x-M)^2}{N} = \frac{44}{7} = 6{,}29$$

Oto dłuższy sposób obliczania wariancji:

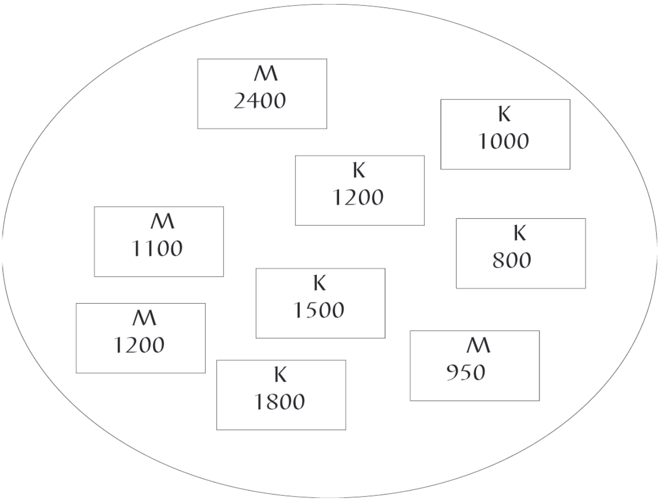
$$\sigma^2 = \frac{(5-6)^2 + (3-6)^2 + (8-6)^2 + (11-6)^2 + (6-6)^2 + (4-6)^2 + (5-6)^2}{7}$$

$$\sigma^2 = \frac{44}{7} = 6{,}29$$

4. ANALIZA ROZPROSZENIA

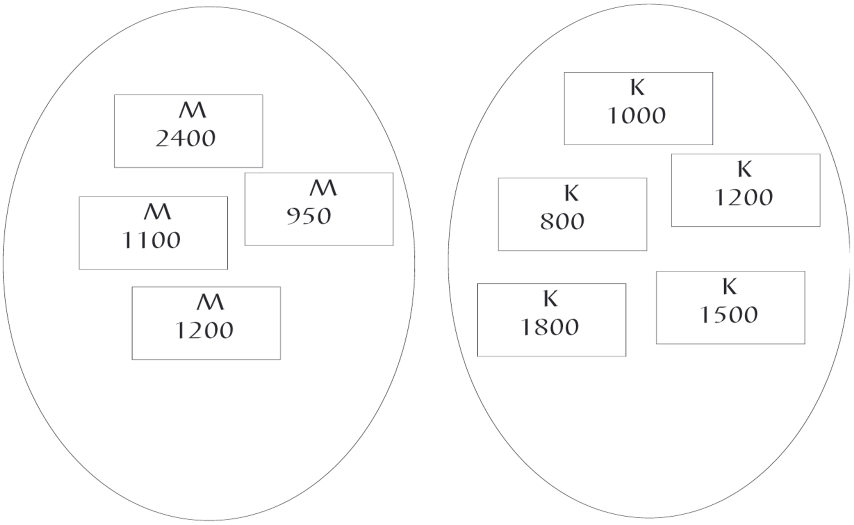
Pomiar zmienności jest zaledwie pierwszym krokiem w analizie. Z obliczonej wariancji prawie nic nie wynika. Do analizy współzależności należy zrobić kolejny krok – jest nim analiza zmienności.

Jako przykład zastosowano populację zatrudnionych, o których znane są dane o pensjach i płeć. Pensje różnią się między sobą. W tej populacji można obliczyć wariancje – wylicza się je z różnic między pensjami. Powstaje pytanie: dlaczego pojawia się zmienność pensji?



Rysunek 6. Pensje i płeć w populacji zatrudnionych

Czy na pensje może mieć wpływ płeć? Czy pensje kobiet różnią się od pensji mężczyzn? Aby odpowiedzieć na to pytanie, należy populację podzielić na dwie grupy: grupę kobiet i grupę mężczyzn.



Rysunek 7. Pensje w grupie mężczyzn i w grupie kobiet

Po podziale można obserwować dwa rodzaje różnic w pensjach: różnice między pensjami w grupie mężczyzn (i podobne różnice między pensjami w grupie kobiet) oraz różnice między grupą kobiet i grupą mężczyzn jako całościami. Pierwsze to różnice wewnątrz grup, a drugie to różnice między grupami (zewnętrzne). Z czego wynikają pierwsze, a z czego drugie?

Na różnice wewnątrz grup płeć nie ma wpływu, ponieważ wszystkie osoby w grupie są takiej samej płci. Żartobliwie można powiedzieć, że na różnice między pensjami mężczyzn płeć nie ma wpływu, bo wszyscy mężczyźni są mężczyznami. Podobnie też można powiedzieć o grupie kobiet. Źródłem zatem tych różnic są wszystkie inne czynniki poza płcią. Odwrotnie jest z różnicami pomiędzy grupami – źródłem tych różnic z pewnością jest płeć!

Z omówionych różnic należy obliczyć wariancje. Z różnic między wszystkimi osobami w całej populacji wylicza się powszechną wariancję (łączną, ogólną).

Z różnic wewnątrz grupy mężczyzn wylicza się wariancję dla tej grupy (wariancja pensji mężczyzn). Z różnic wewnątrz grupy kobiet wylicza się wariancję dla tej grupy (wariancja pensji kobiet).

Oprócz tych trzech wariancji należy wyliczyć jeszcze dwie. Do ich obliczenia potrzebne są nowe wzory (nowa procedura). W swojej statystycznej naturze są to zwyczajne wariancje, pomimo iż niedoświadczony badacz może tego od razu nie zauważyć. Sedno zostaje niezmienne: oblicza się odchylenia, podnosi się je do drugiej potęgi, następnie sumuje i dzieli przez liczebność grupy.

Z różnic między grupą kobiet i grupą mężczyzn oblicza się pierwszą wariancję – jest to **wariancja międzygrupowa**.

Z wariancji w grupie mężczyzn i wariancji w grupie kobiet wylicza się drugą wariancję – jest to **wariancja wewnątrzgrupowa**.

Przyczyną występowania wariancji **międzygrupowej** jest płeć. Ten właśnie wpływ był przedmiotem rozprawy, a dzięki podziałowi na grupy udało się wyjaśnić (określić), jaką część pierwotnych różnic spowodowała płeć. Dlatego można nazwać tę wariancję również **wyjaśnioną** wariancję.

Do wyjaśnienia zostały jeszcze różnice wewnątrz grup. Dlaczego pensje mężczyzn różnią się między sobą i dlaczego pensje kobiet różnią się między sobą? Hipotetycznie można wyliczać źródła tych różnic: stopień wykształcenia, staż pracy, dodatki funkcyjne itd. Takich źródeł może być dużo. W najprostszych przypadkach, które są incydentalne, źródła te są znane, a w bardziej skomplikowanych sytuacjach znane są tylko najważniejsze źródła. Ponieważ nie można wyjaśnić pochodzenie tych różnic w całości, można nazwać tę wariancję jako wariancję **niewyjaśnioną**.

Tabela 22. Wariancja przed podziałem

Wariancja	Różnice	Symbol
całkowita wariancja (powszechna, łączna)	między wszystkimi osobami populacji	σ^2 całkowita

Tabela 23. Wariancje po podziałe

Wariancja w grupie mężczyzn	między wszystkimi mężczyznami	σ^2_1
Wariancja w grupie kobiet	między wszystkimi kobietami	σ^2_2
Międzygrupowa wariancja	między mężczyznami i kobietami	σ^2 WYJAŚNIONA
Wewnątrzgrupowa wariancja	różnice wewnątrz grup	σ^2 NIEWYJAŚNIONA

Analiza wszystkich różnic łatwo prowadzi do wniosku, iż suma wyjaśnionej wariancji i niewyjaśnionej wariancji równa się całkowitej wariancji (w całej populacji):

$$\text{wariancja}_{\text{CAŁKOWITA}} = \text{wariancja}_{\text{WYJAŚNIONA}} + \text{wariancja}_{\text{NIEWYJAŚNIONA}}$$

Wyjaśnioną i niewyjaśnioną wariancję można przedstawić w procentach. Procenty te wskazują, jak silny wpływ na pensje ma płeć, a jak silny jest wpływ innych czynników razem. Dzieląc w ten sposób populację na grupy według dowolnej zmiennej, można ustalić i wyjaśnić wpływ tej zmiennej.

Rozdział V

KORELACJA

1. KORELACJA I WSPÓŁZALEŻNOŚĆ FUNKCYJNA

O zależności pomiędzy zjawiskami mówiono już wcześniej, najwięcej w rozdziale o zmienności. Na ogół pytania o współzależności to najważniejsze pytania każdej naukowej dyscypliny. Dlatego też badanie współzależności między zjawiskami edukacyjnymi jest najważniejszym celem i głównym zadaniem pedagogiki.

Jakie są współzależności między zjawiskami edukacyjnymi? Oto kilka przykładów. Przykłady te pochodzą z praktyki (doświadczeń):

1. Na oceny szkolne ma wpływ to, jak wiele uczeń się uczy. Z prostych obserwacji można stwierdzić, iż ten, który uczy się więcej, otrzymuje lepsze oceny. Oceny jednak nie są uzależnione tylko od uczenia się, a już zupełnie nie tylko od tego, jak wiele czasu uczeń poświęca nauce. Wpływ na oceny mają jeszcze inne czynniki: sposób uczenia się, dotychczasowa wiedza, motywacja, zdolności intelektualne itd. Nie da się jednak zaprzeczyć istnienia związku między czasem uczenia się i ocenami.
2. Wiadomo, że sukcesy szkolne związane są z wykształceniem rodziców: dzieci bardziej wykształconych rodziców osiągają na ogół lepsze wyniki w nauce. Jest prawdą, że istnieją liczne wyjątki, nawet i ekstremalne (bardzo dobrze wykształceni rodzice – bardzo słabi uczniowie i odwrotnie). Pomimo wymienionych wyjątków ta współzależność istnieje i jest na ogół taka, jak wskazano wyżej.
3. Płeć i poglądy są na ogół powiązane: kobiety i mężczyźni różnią się poglądami. Najprawdopodobniej poglądy mężczyzn i kobiet przykładowo

na użyteczność programu Windows nie różnią się między sobą. Różnią ich jednak poglądy dotyczące aborcji, spożywania alkoholu, form spędzania czasu wolnego, obowiązków domowych, socjalnej sytuacji kobiet w społeczności, zawodowych praw kobiet itd. Rzecz jasna, nie można nie zauważyć, iż kobiety również między sobą różnią się poglądami (i mężczyźni też).

Istnienia takich związków nie można nie dostrzegać, ale jednocześnie istnieje przeczucie, że związki te nie są jednoznaczne. Obok wielu przykładów, które potwierdzają te związki, istnieją też liczne, które je obalają stwierdzenia. W celu lepszego wyjaśnienia tych związków należy krótko opisać współzależność innego rodzaju – współzależność funkcyjną.

Współzależność funkcyjna to współzależność dwóch zmiennych. Na zmienną zależną wywiera wpływ jedynie zmienna niezależna i nic innego. Każdej wartości zmiennej niezależnej odpowiada ściśle określona wartość zmiennej zależnej. Współzależność funkcyjną można wyrazić za pomocą tabeli, wykresu i równania.

Jako przykład może posłużyć związek między liczbą kupionych jajek (x) i sumą zakupu w złotych (y). Funkcję tę można wyrazić równaniem:

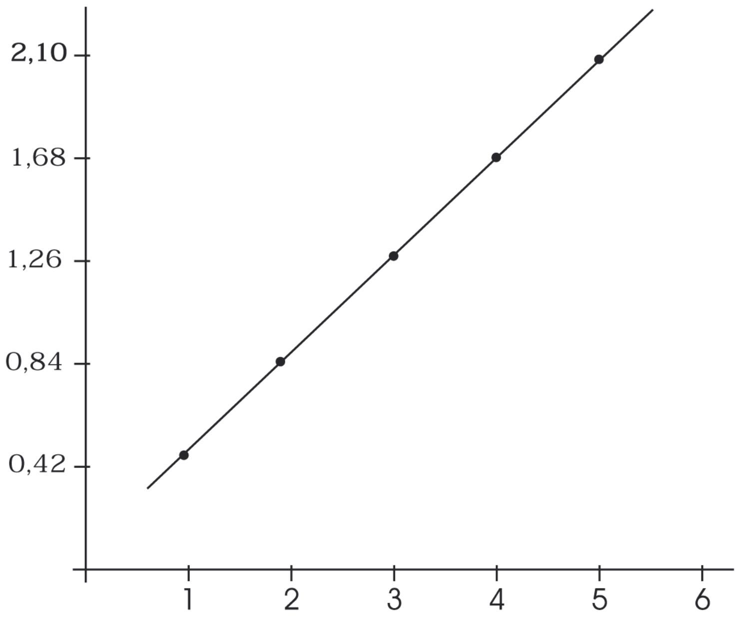
$$y = 0{,}42 \cdot x$$

Oto tabela tej funkcji:

Tabela 34. Związek między liczbą kupionych jajek (x) i sumą zakupu w złotych (y)

x	y
1	0,42
2	0,84
3	1,26
4	1,68
itd.	itd.

Cena jednego jajka wynosi 0,42 zł; dwa kosztują 0,84 zł, trzy 1,26 zł itd. Dla każdej liczby jajek można policzyć końcową sumę. Suma zakupu zależy tylko od liczby jajek (y zależy tylko od x), ponieważ cena jajka jest ściśle określona.



Rysunek 8. Diagram

W przypadku korelacji jest trochę inaczej. Z doświadczeń wiadomo, że czas uczenia się ma wpływ na osiągnięte oceny. Oczywistym jest, iż jedynie na podstawie czasu nie można wnioskować, jaka będzie ocena. Ale dlaczego nie można? Odpowiedź brzmi: na oceny wpływa jeszcze wiele innych czynników, a nie tylko czas uczenia się. Współzależność ta nie jest funkcyjna, ale korelacyjna. Na ogół można stwierdzić, iż dłuższy czas uczenia się powoduje lepsze oceny. Jednak nie jest pewne, iż u wszystkich uczniów dłuższy czas uczenia się daje jednakowe skutki (bo oddziałują i inne czynniki). Współzależność istnieje, ale poszczególne przypadki mogą bardzo odstawać. Im więcej takich przypadków, tym luźniejsza jest współzależność, i odwrotnie – im mniej jest takich przypadków, tym silniejsza współzależność. Dlatego można mówić o sile zależności. Zjawisko to można wytłumaczyć na przykładzie szklanki wody. Funkcję można przedstawić jako pełną szklankę wody, a korelację jako częściowo pełną szklankę. Im jest szklanka pełniejsza, tym silniejsza jest zależność (silniejsza korelacja), im szklanka bardziej pusta, tym luźniejsza jest zależność (słabsza korelacja). W przypadku pełnej szklanki korelacja jest najsilniejsza (najwyraźniejsza) i właśnie już jest funkcją. Taka przenośnia pomaga w zrozumieniu ogólnej miary siły korelacji.

Dotychczasowe uwagi przedstawiały korelację jako zależność między dwoma zmiennymi. Inne czynniki traktowano jako pozostałe, jako coś ubocznego lub nawet zakłócającego. W początkowych przemyśleniach o korelacji jest to koniecznie. Świat to całość, gdzie wszystko jest związane ze wszystkim. Zjawiska zupełnie izolowane nie istnieją. Tak samo jest ze zjawiskami edukacyjnymi. Przy badaniu zależności między zjawiskami pedagogicznymi nie ma możliwości objęcia wszystkich możliwych zależności. Wystarczy przypomnieć, jak wiele czynników ma wpływ na oceny uczniów w szkole. Wiele z nich pedagogika zna i niektóre już zostały wyżej wymienione. Warto posłużyć się przykładem ucznia, który ma chorego brata. Troska o brata może wpływać na jego oceny. Podobna jest sytuacja innego ucznia, który podczas przerwy upadł i boli go kolano. Takich czynników w szkole każdego dnia jest wiele. Można je podzielić na istotne i przypadkowe. Dla każdego badacza ważne są istotne czynniki. O istnieniu przypadkowych czynników wystarczy, że badacz jedynie wie. Jednak istotnych czynników często jest już wiele. Dlatego badania zależności między zjawiskami edukacyjnymi należy rozpocząć od najprostszej sytuacji, którą jest korelacja między dwoma zmiennymi (x i y). Wszystkie inne czynniki (istotne i przypadkowe) są drugoplanowe. Traktuje się je jako czynniki pozostałe. Oddziaływań tych czynników nie bada się szczegółowo. Po zbadaniu i (częściowym) zrozumieniu prostych sytuacji można koncentrować się na oddziaływaniu pozostałych istotnych czynników, czyli na badaniu większej liczby zmiennych **jednocześnie**.

2. PRZYCZYNOWOŚĆ I KORELACJA

W celu lepszego i doskonalszego zrozumienia zależności korelacyjnej należy wyjaśnić jej stosunek do związków przyczynowo-skutkowych. Czy to dwa zupełnie różne zjawiska, czy dwa aspekty tego samego zjawiska? Jak wiele mają wspólnego? Czy silna korelacja oznacza (lub nawet dowodzi) istnienia związku przyczynowo-skutkowego?

By lepiej to zrozumieć, można nawiązać do stylu prezentowania wyników badań w gazetach przed np. trzydziestu laty. Pisano, iż palenie jest przyczyną raku płuc. Może naprawdę tak jest, jednak wtedy w taki sposób interpretowano korelację. Korelacja między paleniem papierosów a rakiem płuc jest rzeczywiście ścisła i pozytywna. Podobnie pisano też, iż pomidory są przyczyną raka, później margaryna itd. Autorzy takich artykułów opierali swoje

twierdzenia na pozytywnej korelacji otrzymanej w praktycznych przypadkach. W przypadku palenia: jeżeli korelacja jest silna i pozytywna, oznacza to rzeczywiście, iż ci, którzy więcej palą, częściej chorują na raka płuc. Jednak nie oznacza to automatycznie, iż palenie jest **przyczyną** raka. Zmienne, między którymi istnieje korelacja, mogą być powiązane przyczynowo-skutkowo lub też nie. Wskaźniki korelacji nie posiadają takiej mocy, by dowodzić lub potwierdzać istnienia związku przyczynowo-skutkowego.
Związki przyczynowo-skutkowe są najważniejszymi związkami między zjawiskami w ogóle. Niestety, korelacja tych związków nie dowodzi!
Można to wyjaśnić znanym przykładem dwóch zmiennych, między którymi korelacja jest bardzo mocna i dodatnia.

Przykład:
Zgromadzono dane dla wszystkich uczniów pewnej szkoły podstawowej o dwóch zmiennych: długość kciuka prawej ręki (cm) i wynik na testu czytania w punktach. Określony wskaźnik korelacji posiada wysoką wartość dodatnią. Oznacza to, że na ogół uczniowie z dłuższym kciukiem lepiej czytają, a uczniowie z krótszym kciukiem czytają gorzej. Byłoby wskazane, gdyby przez moment Czytelnik uwierzył mi, iż korelacja ta jest naprawdę silna i dodatnia (na końcu przykładu będzie oczywiste, iż jest to prawdą).
Jak wytłumaczyć korelację między długością palca a umiejętnością czytania? Nawet bez głębszej analizy merytorycznej można stwierdzić, iż między tymi zmiennymi brak związku przyczynowo-skutkowego. Długość kciuka nie może być przyczyną lepszej umiejętności czytania. A może jest odwrotnie? To jeszcze bardziej śmieszne!
Jest to przypadek, w którym korelacja istnieje, a związku przyczynowo-skutkowego nie ma. Już samo istnienie takiego przypadku dowodzi, iż korelacja **to nie to samo, co związek przyczynowo-skutkowy**.
Dlaczego w tym przypadku występuje korelacja dodatnia i silna? Dlaczego uczniowie z dłuższymi palcami lepiej czytają? Odpowiedź jest oczywista – tkwi w wieku ucznia. Uczniowie np. pierwszej klasy są mali, podobnie jak ich palce. Jednocześnie też gorzej czytają, bo chodzą do szkoły zaledwie niepełny rok.
Uczniowie z klasy drugiej są już więksi (ich palce też) i lepiej czytają, uczniowie trzeciej klasy są jeszcze więksi i czytają jeszcze lepiej, itd.
Uczniowie klas wyższych nie czytają lepiej dlatego, że są **więksi**, ale dlatego, że **chodzą do szkoły dłużej**. Gdyby badaniem objąć uczniów np. tylko szóstej klasy, otrzymany wskaźnik korelacji byłby bardzo niski (bliski

zeru). Często zdarza się, że przyczyną wysokiej korelacji między badanymi zmiennymi jest jakaś trzecia zmienna. Czy to oznacza, iż w badaniach pedagogicznych korelacja nie ma żadnego znaczenia lub nawet sensu? Czy należy odrzucić tę metodę? Pomimo iż korelacja nie oznacza i nie potwierdza związku przyczynowo-skutkowego, jest przydatna. Korelacja musi być jednak dopełniona analizą merytoryczną. Dopiero merytoryczna analiza istoty badanych zjawisk (łącznie z eksperymentami) może pokazać i **potwierdzić** związki przyczynowo-skutkowe. W omówionym przypadku wiedza o istocie i podstawie rozwoju umiejętności czytania ostrzega przed zbyt pochopnym wnioskowaniem w stylu: „korelacja między długością palca a umiejętnościami czytania oznacza, że długość palca jest przyczyną większych umiejętności czytania".

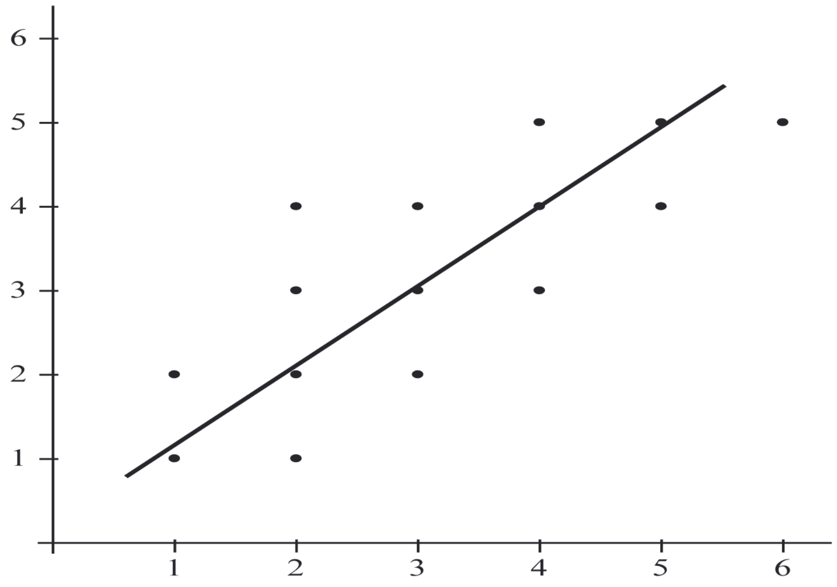
3. DIAGRAM KORELACYJNY

Przy badaniu korelacji można korzystać z diagramów sporządzonych w prostokątnym układzie współrzędnych. Na osiach znajdują się wartości badanych zmiennych: na osi odciętych wartości zmiennej niezależnej i na osi rzędnych wartości zmiennej zależnej. Punkty umieszczone w układzie przedstawiają wartości x i y poszczególnych jednostek badanej populacji. W ten sposób powstaje diagram korelacyjny.
Jako przykład służą wartości zmiennych: średnia ocen w ostatniej klasie szkoły podstawowej (x) i średnia ocen w I klasie gimnazjum (y).

Tabela 23. Oceny uczniów

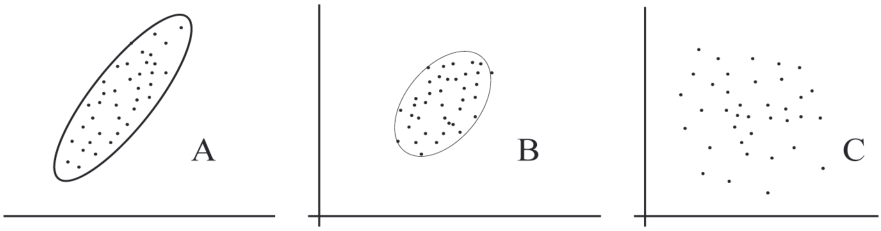
Nr ucznia	1	2	3	4	5	6	7	8	9	10	11	12
x	4,2	3,4	5,2	3,8	3,5	3,2	4,1	5,0	5,1	3,8	3,4	4,4
y	3,5	2,9	4,8	2,8	4,0	3,0	4,0	3,7	4,3	4,0	2,8	4,1

Tabelę sporządzono w sposób „pokazowy": niepotrzebne są liczby porządkowe uczniów. Do tabeli zostały one włączone tylko po to, aby podkreślić, iż każda para ocen odnosi się do określonego ucznia i że nie można tych ocen uporządkować (np. rosnąco do szeregu prostego). Obydwie oceny każdego ucznia muszą zawsze zostać razem.



Rysunek 9. Diagram korelacyjny

Na wykresie umieszczono środkową linię między punktami. Jest to linia regresji. Punkty w diagramie zagęszczają się wokół tej linii. Linia jest tylko wirtualna – w przybliżeniu byłaby taką, jeżeliby nie było wpływu innych czynników. Wpływ *x* na *y* „ciągnie" punkty **do** linii regresji. Pozostałe czynniki „ciągną" punkty **od** linii regresji. Jeżeli „obłok" punktów jest zagęszczony blisko linii regresji, to korelacja między *x* a *y* jest silna (ścisła zależność). Jeżeli punkty znajdują się daleko od linii regresji, korelacja jest słaba (bo wpływ pozostałych czynników jest silny). Kształt „obłoku" punktów w diagramie pokazuje, jak silny jest wpływ *x* na *y*. W przypadku, gdy punkty nie skupiają się w kształt obłoku i są równomiernie rozproszone w diagramie – korelacja między *x* a *y* nie występuje. Takie idealnie równomierne rozproszenie w praktyce jest wielką rzadkością. Dlatego też niemal zawsze otrzymuje się przynajmniej słabą korelację. Jest to kolejny argument przemawiający za wielką ostrożnością w interpretacji wskaźników korelacji.
Warto przyjrzeć się kilku charakterystycznym diagramom. Na wykresie A korelacja jest silna (zagęszczony obłok), na wykresie B słaba (mniejsze zagęszczenie – słabo ukształtowany obłok), a na wykresie C nie ma korelacji (nie ma obłoku).



Rysunek 10. Diagramy korelacyjne

Niepodważalną prawdą jest, że zmienność lub rozproszenie jest podstawą każdej analizy współzależności. W sposób oczywisty pokazuje to też diagram korelacyjny. Nie należy jednak upraszczać związku zmienności i współzależności. Niepoprawne są wnioski takie jak: im mniejsze jest rozproszenie, tym silniejsza jest korelacja lub im większe rozproszenie, tym silniejsza jest korelacja! Do takich wniosków może prowadzić zbyt szybkie i powierzchowne zapoznanie się z diagramem korelacyjnym. Z samego rozproszenia zmiennej zależnej (y) nie można jeszcze niczego wnioskować o korelacji.

Pozornie jesteśmy w ślepej uliczce. Czy w końcu rozproszenie jest podstawą analizy współzależności, czy też nie? Jak jest z tym rozproszeniem, jeżeli ani wielkie, ani też małe rozproszenie nie potwierdza korelacji? Czy stwierdzenia ze strony 63 są prawdziwe, czy nie? Powtórzmy te słowa:

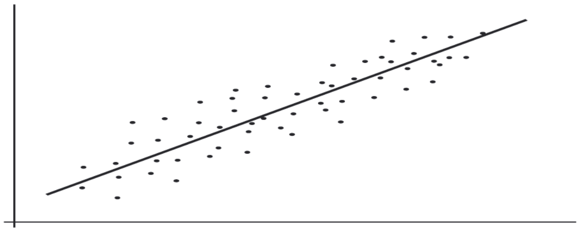
Wpływ x na y „ciągnie" punkty do linii regresji. Pozostałe czynniki „ciągną" punkty od linii regresji. Jeżeli „obłok" punktów jest zagęszczony blisko linii regresji, to korelacja między x a y jest silna (ścisła zależność). Jeżeli punkty znajdują się daleko od linii regresji, korelacja jest słaba (bo wpływ pozostałych czynników jest silny).

Jak teraz wyjść z tej ślepej uliczki? Wyjaśnienie jest na szczęście proste: obłok punktów w diagramie pokazuje rozproszenie dwóch zmiennych a nie jednej! Oprócz tego „zagęszczenie" obłoku odnosi się do rozproszenia wokół linii regresji, a nie do rozproszenia na ogół. Dlatego o korelacji nie należy wnioskować na podstawie tylko jednej zmiennej (np. im mniejsze jest rozproszenie, tym korelacja jest silniejsza lub im większe rozproszenie, tym korelacja jest silniejsza).

Najważniejsze więc nie jest pytaniem jakie jest rozproszenie, ale co je powoduje?! W języku statystycznym powiemy: **nie wystarczy rozproszenia zmierzyć, trzeba je analizować!** A do tego potrzebne są obie zmienne.

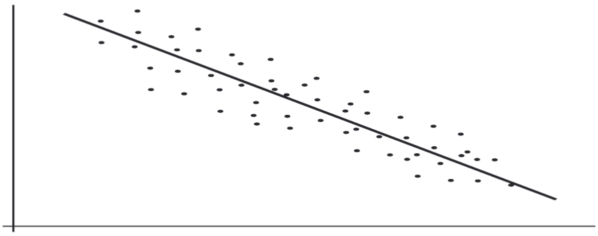
4. KORELACJA DODATNIA I UJEMNA

Dotychczas mówiono o sile korelacji. Korelacja posiada też kierunek oddziaływania. Jeżeli wzrostowi wartości pierwszej zmiennej towarzyszy wzrost wartości drugiej zmiennej, to korelacja jest dodatnia. W diagramie korelacyjnym linia regresji biegnie od lewego dolnego do prawego górnego kąta.



Rysunek 11. Korelacja dodatnia

Jeżeli wzrostowi wartości pierwszej zmiennej towarzyszy spadek wartości drugiej zmiennej, to korelacja jest ujemna. Diagram korelacyjny wygląda zupełnie inaczej:



Rysunek 12. Korelacja ujemna

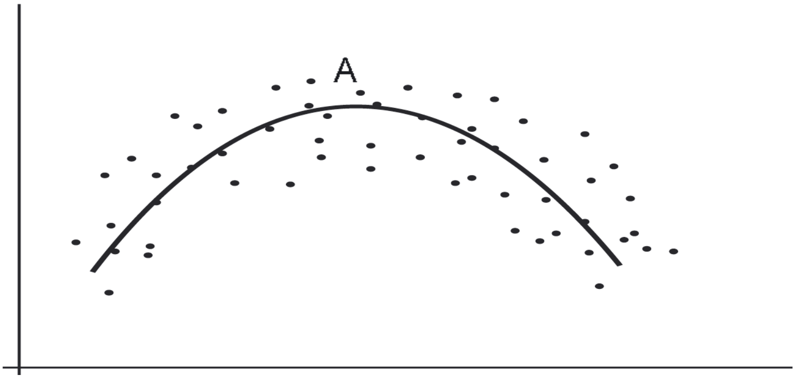
Oto prosty przykład korelacji dodatniej: „im dłuższy czas uczenia się, tym wyższe oceny". Przykładem korelacji ujemnej jest związek między tremą a oceną („im większa trema, tym niższa ocena" i odwrotnie). Obydwa przykłady są uproszczone i służą tylko do łatwiejszego rozumienia kierunku oddziaływania.

Niektóre wskaźniki korelacji informują o kierunku badanej zależności. Znak plus lub minus wskazuje, czy korelacja jest dodatnia, czy też ujemna. Inne wskaźniki korelacji natomiast nie mają znaku plus lub minus, więc nie informują o kierunku zależności. Wówczas konieczny jest przegląd wszystkich rezultatów, aby poprawnie interpretować kierunek korelacji.

kozuh statystyka.indb 65 2011-05-13 12:29:55

5. KORELACJA NIELINIOWA I LINIOWA

Linia regresji może być linią krzywą lub prostą (oczywiście: tylko częścią prostej). W przypadku, gdy linia jest krzywa – występuje korelacja nieliniowa (krzywoliniowa), gdy linia jest prosta – korelacja jest liniowa. W realnych przypadkach linia regresji nigdy nie jest idealną prostą. Dlatego do liniowej korelacji należy zaliczyć przypadki, w których linia regresji chociaż w przybliżeniu jest prostą (jeżeli nie jest wyraźnie krzywa). Do nieliniowej korelacji zaś zalicza się przypadki, gdy linia jest wyraźnie krzywa.



Rysunek 13. Korelacja nieliniowa

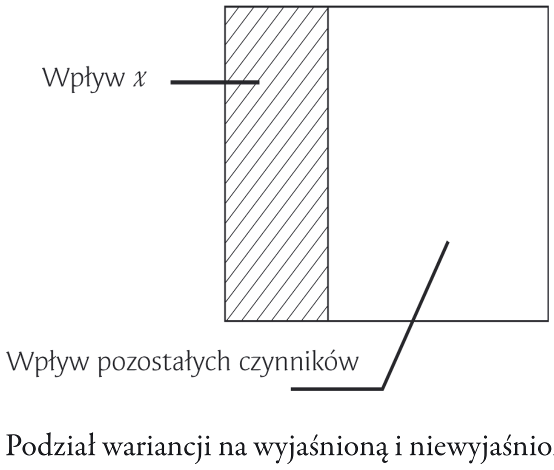
Korelacja nieliniowa jest trudniejsza do interpretacji. Czym charakteryzuje się nieliniowość lub liniowość korelacji (oprócz linii w diagramie)? W przypadku, gdy korelacja jest liniowa, można stwierdzić, iż wartości y wzrastają lub opadają proporcjonalnie (współmiernie) do wzrostu lub spadku wartości x. Kierunek korelacji jest tylko jeden i nie zmienia się.

Przy korelacji nieliniowej istnieją przynajmniej dwie trudności w interpretacji. Pierwsza polega na nieproporcjonalnej przemianie y, podczas gdy x zmienia się równomiernie. Dlatego jest wyraźnie trudniej wyjaśnić zmiany y. Drugim problemem jest to, iż nieliniowa korelacja może być w jednej części dodatnia, a w drugiej ujemna. Proste do zrozumienia jest stwierdzenie: im więcej uczeń się uczy, tym wyższe są jego wyniki. Każdy rozumie też kolejną prawidłowość: im więcej sportowiec trenuje, tym lepsze są jego osiągnięcia. Ale wszystko nie jest tak proste: ostatni przykład może w sposób przejrzysty pokazać trudności w interpretacji korelacji nieliniowej. Osiągnięcia sportowca wzrastają tylko do pewnej granicy. Za tą granicą przedłużanie czasu treningu może spowodować zmniejszanie osiągnięć.

Jest to znane zjawisko przetrenowania (sportowiec zbyt dużo trenował). Diagram takiej korelacji zaprezentowano na rysunku nr 9 (na s. 63). Do punktu A korelacja jest dodatnia, od tego punktu dalej ujemna (więcej treningu przynosi niższe wyniki). Przykład jest wprawdzie nieco uproszczony, bo celowo zaniedbane zostało doświadczenie, iż wzrost wyników ma swoje granice bez względu na trening (czyli: zarówno w przypadku liniowej korelacji wyniki nie wzrastałyby w nieskończoność). Jednak uproszczenie to nie zmienia istoty spostrzeżenia, iż nieliniową korelację interpretuje się o wiele trudniej niż liniową.

6. INDEKS KORELACJI

Sposób pomiaru korelacji został częściowo wyjaśniony w rozdziale o zmienności. Użyto tam wyrazów „wyjaśniona" i „niewyjaśniona" wariancja. Teraz należy do tych wyrazów powrócić. Wpływ zmiennej niezależnej jest wpływem, który znajduje się w centrum uwagi. Rozproszenie z nim związane jest więc wyjaśnione. Wpływem pozostałych czynników badacz jest zainteresowany jedynie ubocznie. Dlatego też rozproszenie powiązane z nimi nazywa się rozproszeniem niewyjaśnionym. Poniższy rysunek ilustruje korelację między zmienną niezależną *x* i zależną *y*.



Rysunek 14. Podział wariancji na wyjaśnioną i niewyjaśnioną

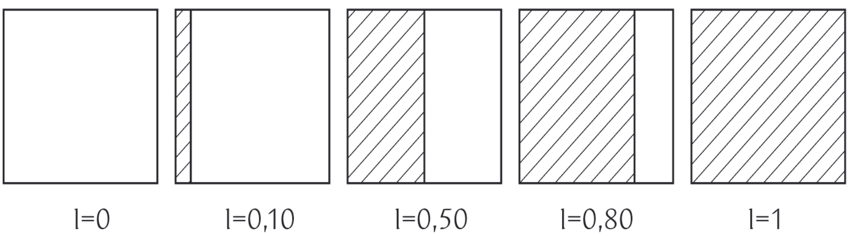
Podział wariancji na wyjaśnioną i niewyjaśnioną jest wyidealizowany. Przesłanką tego podziału jest niezależność *x* od pozostałych czynników. W praktyce zdarza się to jedynie incydentalnie. Takie uproszczenie bardzo

ułatwia zrozumienie zasady pomiaru korelacji. Należy jednak pamiętać, iż procedura ta jest trochę nieścisła. W interpretacji należy uwzględniać różnicę między ideałem i realnością.
Stosunek pomiędzy wariancją wyjaśnioną a wariancją całkowitą wskazuje, z jaką siłą *x* oddziałuje na *y*. Stosunek ten nazywa się indeksem korelacji. Oto wzór obliczania indeksu korelacji:

$$I = \frac{\sigma_w^2}{\sigma_c^2}$$

σ^2_w	wariancja wyjaśniona (międzygrupowa)
σ^2_c	wariancja całkowita

Wartości indeksu wahają się od **0** do **1**. Wartość zero oznacza brak korelacji między *x* i *y* (wyjaśniona wariancja równa się zeru, co oznacza, iż *x* nie oddziałuje na *y*). Wartość 1 oznacza, że korelacja jest najsilniejsza (niewyjaśniona wariancja równa się zeru, co oznacza, iż tylko *x* oddziałuje na *y*). Taka korelacja jest już funkcją.
Należy jeszcze raz podkreślić, iż indeks korelacji nie może przekraczać wartości 1,00! Ta zasada odnosi się do wszystkich miar współzależności.
Oznacza to, że jeżeli w trakcie obliczeń miar współzależności (indeksu korelacji, współczynników korelacji, współczynników zbieżności, itd.) otrzyma się wartość większą niż 1, jest to niewątpliwy znak, iż obliczenia są błędne!
Poniższe wykresy pokazują kilka możliwych przypadków korelacji (obrót kwadratów lewym bokiem w dół ilustruje wymienioną wcześniej szklankę wody):



Rysunek 15. Indeks korelacji

7. WSPÓŁCZYNNIKI KORELACJI

Indeks korelacji służy tylko jako podstawa do zrozumienia pomiaru siły korelacji. W praktyce używa się go incydentalnie rzadko. Zastępują go różne współczynniki korelacji. Współczynników jest więcej – jest to spowodowane rodzajem zmiennych i naturą współzależności.

Współczynnik korelacji liniowej Pearsona
W przypadku, gdy obie zmienne są przedziałowe, a korelacja jest liniowa, można zastosować współczynnik korelacji Pearsona $\mathbf{r}_{xy}$. Współczynnik Pearsona najlepiej odzwierciedla zależność między zmiennymi. Dlatego należy go zastosować wszędzie, gdzie tylko pozwalają na to warunki. W interpretacji pozostałych współczynników korelacji zawsze należy opierać się na nim. Niezwykle istotne jest przy tym ustalenie, czy jest jakiś inny współczynnik porównywalny ze współczynnikiem Pearsona. Współczynniki te można z niewielkim błędem interpretować podobnie. Pozostałe (nieporównywalne) współczynniki interpretuje się samodzielnie i niezależnie. Podstawowy wzór na obliczanie współczynnika Pearsona jest następujący:

$$r_{xy} = \frac{C_{xy}}{\sigma_x \cdot \sigma_y}$$

$\mathbf{C}_{xy}$	kowariancja (miara wspólnego rozproszenia *x* i *y*)
σ_x	odchylenie standardowe zmiennej *x* (pierwiastek wariancji zmiennej *x*)
σ_y	odchylenie standardowe zmiennej *y* (pierwiastek wariancji zmiennej *y*)

Nie należy zbytnio martwić się pytaniem, co to jest kowariancja (wspólne rozproszenie *x* i *y*). Rozumienie tego pojęcia nie jest zupełnie niezbędne. Wnikliwego czytelnika, który chce dowiedzieć się więcej o kowariancji, proszę o poszukanie bardziej szczegółowego podręcznika ze statystyki.

Tabela 24. Wzory dla kowariancji i wariancji

	Podstawowy wzór
Kowariancja	$C_{xy} = \frac{\Sigma(x - M_x) \cdot (y - M_y)}{N}$
Wariancja *x*	$\sigma_x^2 = \frac{\Sigma(x - M_x)^2}{N}$
Wariancja *y*	$\sigma_y^2 = \frac{\Sigma(y - M_y)^2}{N}$

Postępowanie to można skrócić, nie obliczając wszystkiego do końca. Stosowane wzory będą trochę prostsze niż podstawowe.
Do obliczania współczynnika Pearsona skorzystano z tego samego przykładu, który został prezentowany w diagramie.
Należy sporządzić tabelę złożoną z pięciu wierszy. W pierwszym wierszu wpisuje się wartości *x*, w drugim wartości *x* do kwadratu, w trzecim wartości *y*, w kolejnym wartości *y* do kwadratu, a w piątym mnoży się *x* i *y*. Do obliczania współczynnika korelacji potrzebne są sumy wszystkich wierszy.

Tabela 25. Obliczanie współczynnika korelacji Pearsona

Nr ucznia	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
x	4	2	5	2	3	3	4	5	4	3	3	4	$\Sigma x=42$
x^2	16	4	25	4	9	9	16	25	16	9	9	16	$\Sigma x^2=158$
y	3	2	6	3	5	3	6	5	5	4	2	4	$\Sigma y=48$
y^2	9	4	36	9	25	9	36	25	25	16	4	16	$\Sigma y^2=214$
xy	12	4	30	6	15	9	24	25	20	12	6	16	$\Sigma xy=179$

Z sum pierwszych dwóch wierszy oblicza się wartość K_x, z sum trzeciego i czwartego oblicza się wartość K_y, a z sum piątego, pierwszego i trzeciego wartość K_{xy}. Otrzymane trzy wartości podstawia się do wzoru do obliczania współczynnika Pearsona. Nie ma potrzeby precyzyjnego wyjaśniania, co jest K_x, K_y i K_{xy}. Wystarczy jedynie wiedzieć, że są to nie do końca obliczone miary rozproszenia: K_x jest licznikiem z wzoru do obliczania wariancji zmiennej *x*, K_y jest licznikiem z wzoru do obliczania wariancji zmiennej *y* i K_{xy} jest licznikiem z wzoru do obliczania kowariancji.

Oto wszystkie potrzebne wzory:

1	$K_x = \Sigma x^2 - \frac{(\Sigma x)^2}{N}$
2	$K_y = \Sigma y^2 - \frac{(\Sigma y)^2}{N}$
3	$K_{xy} = \Sigma xy - \frac{\Sigma x \cdot \Sigma y}{N}$
4	$r_{xy} = \frac{K_{xy}}{\sqrt{K_x \cdot K_y}}$

Oto kontynuacja przykładu:

1	$K_x = \Sigma x^2 - \frac{(\Sigma x)^2}{N} = 158 - \frac{(42)^2}{12} = 11{,}00$
2	$K_y = \Sigma y^2 - \frac{(\Sigma y)^2}{N} = 214 - \frac{(48)^2}{12} = 22{,}00$
3	$K_{xy} = \Sigma xy - \frac{\Sigma x \cdot \Sigma y}{N} = 179 - \frac{42 \cdot 48}{12} = 11{,}00$
4	$r_{xy} = \frac{K_{xy}}{\sqrt{K_x \cdot K_y}} = \frac{11{,}00}{\sqrt{11{,}00 \cdot 22{,}00}} = 0{,}71$

Wielbiciele krótszych procedur mogą podstawić wszystko w jednym wzorze. Postępowanie to będzie prostsze, ale wzór będzie dłuższy. Wcale nie należy się bać długich wzorów – one tylko wyglądają bardziej skomplikowanie. Jedyne, na co naprawdę trzeba uważać jeszcze bardziej niż w postępowaniu dłuższym, jest dokładne odróżnianie pięciu sum z tabeli (Σx, Σy, Σx^2, Σy^2, Σxy).
Oto wzór do krótkiej procedury:

$$r_{xy} = \frac{\Sigma xy - \frac{\Sigma x \cdot \Sigma y}{N}}{\sqrt{\left(\Sigma x^2 - \frac{(\Sigma x)^2}{N}\right) \cdot \left(\Sigma y^2 - \frac{(\Sigma y)^2}{N}\right)}}$$

Teraz można podstawić wszystkie dane i obliczyć współczynnik korelacji:

$$r_{xy} = \frac{179 - \frac{42 \cdot 48}{12}}{\sqrt{\left(158 - \frac{(42)^2}{12}\right) \cdot \left(214 - \frac{(48)^2}{12}\right)}} = 0{,}71$$

Pojęcie zgodności wyników

Poniżej przytoczono kilka przykładów, w których występuje korelacja. Przykłady te mają pokazać w sposób praktyczny związek pomiędzy *x* i *y*. Związek ten można też opisać słowami: wartości *x* i *y* są zgodne.

Przykład 1.

x	*y*
3	3
5	5
2	2
7	7
8	8

Zgodność wartości *x* i *y* jest niewątpliwie zupełna. Współczynnik korelacji wynosi +1. Większej zgodności nie można osiągnąć, czyli jest to najsilniejsza korelacja. Podobnie silna korelacja może jednak występować w mniej oczywistych przypadkach, co pokazuje kolejny przykład.

Przykład 2.

x	*y*
3	4
5	6
2	3
7	8
8	9

Na podstawie tego nie jest łatwo ocenić, jak silna jest korelacja. Najpierw należy ustalić jej kierunek. Korelacja jest dodatnia, ponieważ większym wartościom x towarzyszą większe wartości y (i mniejszym wartościom x towarzyszą mniejsze wartości y). Uważne przejrzenie się wartościom x i y pokazuje, że i w tym przypadku chodzi o zupełną zgodność! W przykładzie pierwszym stosunek pomiędzy x i y można wyrazić w postaci równania:

$$y = x$$

Stosunek między x i y w przykładzie drugim niewiele się zmienił. Oto równanie:

$$y = x + 1$$

Widać bardzo wyraźnie, że w obydwu przypadkach występuje funkcja, czyli zupełna zgodność wartości. Bez obliczeń wiadomo, że współczynnik korelacji wynosi 1.

Zaprezentowana sytuacja skłania do refleksji i następującego pytania: czy zgodność wartości oznacza, że x i y muszą być identyczne? Odpowiedź brzmi: NIE! Zgodność jest czymś innym niż identyczność. Dodatkowo może to udowodnić następujący przykład:

Przykład 3.

x	y
1145	291,35
2036	518,07
1743	443,51
1007	256,23
756	192,37

Tutaj jeszcze trudniej odkryć, jaka jest zgodność wartości x i y. A zgodność w rzeczywistości jest także i w tym przypadku zupełna. Kolumna x zawiera pensje pięciu zatrudnionych, a kolumna y te same pensje w EURO (wybrano kurs walutowy 1 EURO = 3,93 zł). I w tym przypadku można rozwinąć równanie:

$$y=\frac{x}{3{,}93}$$

We wszystkich trzech przypadkach obliczenia pokazałyby, że współczynnik korelacji wynosi 1. Obliczenia te nie są jednak potrzebne, gdyż z góry wiadomo, iż jest to funkcja, a funkcja jest najsilniejszą możliwą korelacją. Oznacza to, że współczynnik korelacji wynosi 1.
Kolejną konkluzją z tych przykładów jest to, że nie należy wyciągać wniosków o sile korelacji bez obliczeń współczynnika korelacji. Wyjątek stanowi sytuacja, w której posiadamy wielkie doświadczenie w badaniach.
Oto prosty sprawdzian. Proszę, aby czytelnik na podstawie samych wyników ocenił siłę korelacji w tabeli nr 26 (i rozpoznał, czy korelacja jest dodatnia, czy ujemna).

Tabela 26. Wyniki uczniów z dwóch testów

x	*y*
45	11
2	3
17	6
10	7
26	6

Korelacja jest bardzo silna, choć nie wygląda na taką. Współczynnik korelacji wynosi 0,89. To, że korelacja jest dodatnia, można ocenić bez obliczeń. Nie da się jednak ustalić bez obliczeń, że jest tak silna.
Ten ciąg przykładów zakończy przykład z praktyki edukacyjnej. Badano w nim grupę uczniów i ich oceny z matematyki w szkole podstawowej i w pierwszej klasie gimnazjum.

Tabela 27. Oceny uczniów

	Ocena w szóstej klasie szkoły podstawowej	**Ocena w pierwszej klasie gimnazjum**
Paweł	3	2
Jan	3	2
Anna	6	5
Janina	5	4
Damian	4	3

Czy korelacja jest silna? Narzuca się wniosek: ponieważ uczniowie otrzymali w szkoli średniej oceny słabsze, to korelacja nie jest bardzo silna. Inaczej można powiedzieć: jeżeli uczniowie otrzymaliby takie same oceny, jak w szkole podstawowej, to korelacja byłaby nawet zupełna, czyli w omawianym przypadku korelacja nie jest tak silna. Jest to jednak błędny wniosek. Jest prawdą, że oceny uczniów obniżyły się, ale należy zauważyć, że obniżyły się jednakowo u wszystkich uczniów. Z poglądu ich rodziców oceny w szkole średniej wcale nie są zgodne z ocenami wcześniejszymi. Jednak z poglądu liczb samych w sobie, zgodność jest zupełna, czyli ponownie współczynnik korelacji wynosi 1. Większa zgodność nie może istnieć.
Dlaczego należy wyjaśniać tak wiele o pojęciu „zgodności" wyników? Jeżeli badacz nie wie, jakie są wartości x i y w przypadku silnej, średniej, słabej lub innej korelacji, to może polegać tylko na naśladowaniu cudzych interpretacji. Sam nie może merytorycznie (kompetentnie) interpretować wyników. Umiejętność obliczania współczynnika korelacji jest mniej niż połową potrzebnej wiedzy. Ważniejszą umiejetnością jest interpretacja. Obliczeń może bowiem dokonać nawet komputer, ale nie może on interpretować otrzymanych wyników.
Podobne prawa dotyczące zgodności wyników można zauważyć także w przypadkach korelacji ujemnej. Oto przykład zupełniej zgodności w ujemnym kierunku:

Tabela 28. Wyniki uczniów

x	y
2	6
3	5
4	4
5	3
6	2

Tabela przedstawia zgodność wartości x i y, ale w odwrotnym kierunku niż w przykładzie 1. Pomimo że ujemny kierunek zniekształca trochę obraz wyników, nie można nie zauważyć zupełniej zgodności wartości x i y. Ponownie jest to przypadek najsilniejszej korelacji. Współczynnik korelacji wynosi –1.
Tak proste przypadki w praktyce się nie zdarzają. Kolejna tabela przedstawia podobny przypadek do omówionego wyżej, ale mniej oczywisty.

Tabela 29. Wyniki uczniów

x	*y*
4	17
7	13
8	10
11	9
15	6

Tutaj zgodność x i y nie jest zupełna, lecz bardzo duża. Obliczenia pokazują, że współczynnik korelacji wynosi – 0,96.
Nie będziemy kontynuować przykładów, ponieważ z zaprezentowanych wyżej można już wyciągnąć wnioski. Z matematycznego punktu widzenia korelacja pokazuje się jako zgodność wartości x i y. Zgodność tę trudno określić w prosty sposób, ale najważniejsze, aby ją dobrze zrozumieć.

Uwaga!
W niektórych z przytoczonych przykładów obliczano korelacje z ocen szkolnych. Należy nadmienić, że oceny szkolne nie są zmiennymi przedziałowymi i takie postępowanie nie jest całkowicie uzasadnione. Autor tak zrobił tylko z powodu łatwiejszego rozumienia omawianej problematyki.

Interpretacja współczynnika korelacji Pearsona
Wartości współczynnika Pearsona wahają się od –1 do +1. Interpretacja kierunku zależności już została częściowo omówiona wyżej. Przy dodatniej wartości współczynnika kierunek korelacji pokazuje, iż im większy jest x, tym większy jest też y. Im mniejszy jest x tym mniejszy jest także y. W przypadku ujemnej wartości współczynnika korelacji prawidłowość ta jest odwrotna: większym wartościom x odpowiadają mniejsze wartości y, a mniejszym wartościom x odpowiadają większe wartości y.
Należy jednak pamiętać, iż w zależności funkcyjnej powyższe prawo dotyczy wszystkich jednostek, natomiast w korelacji istnieją odchylenia.
Stopień zależności (siłę korelacji) interpretuję się według numerycznej wartości współczynnika korelacji (uwaga: przy interpretacji stopnia zależności znak plus lub minus nie ma żadnego znaczenia!). Wartości współczynnika korelacji do 0,20 interpretuje się jako nieistotną korelację, a praktycznie jako brak korelacji. Można powiedzieć, iż otrzymany współczynnik jest raczej skutkiem przypadku (przypadkowej zgodności danych) niż istnienia

zależności. Wartości współczynnika korelacji od 0,20 do 0,40 interpretuje się jako niewyraźną lub słabą korelację. Współczynniki korelacji od 0,40 do 0,70 interpretuje się jako korelację średnią, od 0,70 do 0,90 jako silną i wyraźną zależność (*x* mocno oddziałuje na *y*) i ponad 0,90 jako bardzo mocną zależność (praktycznie już zależność funkcyjną).
Propozycja ta jest niewystarczająca, dlatego też nie należy stosować jej mechanicznie i bez względu na konkretne okoliczności. W każdym przypadku potrzebna jest analiza merytoryczna. Do interpretacji należy włączyć wcześniejsze wyniki badań, oczekiwania, doświadczenia itd. Potrzebę konkretyzowania interpretacji ilustrują dwa przykłady.

Przykład 1.

Dwóch nauczycieli niezależnie oceniało wypracowania tych samych uczniów. Każdy uczeń otrzymał dwie oceny: jedną od pierwszego nauczyciela i drugą od drugiego. Załóżmy, iż korelacja między ocenami wynosi $r_{xy} = + 0,64$. Wynik ten można zinterpretować następująco: korelacja jest wysoka i nauczyciele są zgodni w swoich ocenach.

Przykład 2.

Dwóch nauczycieli niezależnie oceniało testy wiadomości tych samych uczniów, zawierające głównie zadania zamknięte. Każdy uczeń otrzymał dwie oceny (liczby punktów): jedną od pierwszego nauczyciela i drugą od drugiego. Załóżmy, iż korelacja między ocenami wynosi $r_{xy} = + 0,77$. Wynik ten można zinterpretować następująco: korelacja jest nieoczekiwanie słaba i nauczyciele słabo zgadzają się w swoich ocenach.

Oznacza to, że niższy współczynnik interpretowano jako zależność silną, a wyższy współczynnik jako zależność słabą. W przypadku ocen z wypracowań oczekiwano słabej zgodności, a wartość współczynnika (0,64) okazała się nieoczekiwanie wysoka. W ewaluacji testów oczekiwano całkowitej zgodności punktów (lub przynajmniej prawie zupełnej zgodności), a otrzymany współczynnik zaskoczył swoją niską wartością (0,77).
W tym miejscu warto jeszcze raz uprzedzić o niebezpieczeństwach przyczynowo-skutkowej interpretacji korelacji. Dodatni i wysoki współczynnik korelacji można zinterpretować następująco: uczniowie, którzy lepiej rozwiązali test z języka angielskiego, jednocześnie lepiej rozwiązali też test z matematyki. Jeszcze prościej można powiedzieć: im lepszy uczeń z języka angielskiego, tym lepszy z matematyki.

Niepoprawnie brzmiałaby interpretacja: uczniowie lepiej rozwiązali test z matematyki, **ponieważ** lepiej rozwiązali test z języka angielskiego. Tylko jedno słowo (ponieważ) zupełnie zmienia istotę interpretacji. Wysoka wartość współczynnika korelacji nie oznacza, iż drugi wynik jest skutkiem pierwszego wyniku.

Osiągnięcia uczniów na testach z różnych przedmiotów są zwykle bardzo zgodne, a współczynniki korelacji są dodatnie i wysokie. Dotyczy to praktycznie wszystkich szkolnych przedmiotów. Przy określaniu korelacji wyników testów z dowolnych dwóch przedmiotów niemal zawsze otrzymuje się wysoką dodatnią korelację. Nie znaczy to jednak, iż uczniowie otrzymali dobre wyniki z jednego testu, **ponieważ** osiągnęli dobre rezultaty z drugiego testu. Współczynnik korelacji nie dowodzi ani istnienia, ani też braku związku przyczynowo-skutkowego. Rodzi się więc pytanie, dlaczego korelacja z różnych szkolnych przedmiotów jest dodatnia i wysoka? Można odpowiedzieć w następujący sposób: bardziej motywowani uczniowie osiągają lepsze wyniki z języka polskiego i z matematyki (lub z jakiegoś innego przedmiotu). Podobny wpływ mają zdolności intelektualne, pilność, warunki do uczenia się, itd. Podobnie oddziałują też czynniki ujemne. Uczeń, który opuszczał lekcje (np. z powodu choroby), odczuwa problemy z większości przedmiotów. Stwierdzeń o wysokiej i dodatniej korelacji nie podważa nawet zjawisko, iż u niektórych uczniów osiągnięcia się wyraźnie **nie zgadzają** (obok ocen wybitnych z niektórych przedmiotów, mają oni także oceny bardzo słabe z innych przedmiotów). Takie przypadki powodują, iż korelacja między wynikami z różnych szkolnych przedmiotów nie wynosi 1, lecz „jedynie" np. 0,69.

Należy zwrócić uwagę też na częsty błąd w interpretacji ujemnej korelacji. W badaniu obliczano korelacje wyników testu z matematyki i języka polskiego dla grupy uczniów gimnazjum. Ujemny i wysoki współczynnik korelacji można zinterpretować następująco: uczniowie, którzy lepiej rozwiązali test z języka polskiego, jednocześnie gorzej rozwiązali test z matematyki. Jeszcze prościej można powiedzieć: im lepszy uczeń z języka polskiego, tym gorszy z matematyki. Ale to wcale nie oznacza, że nie ma korelacji między wynikami testu z matematyki i języka polskiego. Zupełnie niepoprawna jest interpretacja: ponieważ uczniowie, którzy lepiej rozwiązali test z języka polskiego, nie rozwiązali także lepiej testu z matematyki, oznacza to, iż nie ma korelacji. Korelacja jest, tylko że jest ujemna. Nieprzyzwyczajenie do takich sytuacji w badaniach często prowadzi do podobnych błędnych wniosków. Brak korelacji byłby w przypadku, gdyby pewnemu wynikowi z matematyki odpowiadał jakikolwiek wynik z języka polskiego (lub odwrotnie). Można to zilustrować prostym przykładem.

Przykład:
Im więcej studenci wydają pieniędzy na rozrywkę, tym mniej mogą wydać na literaturę. Im mniej studenci wydają pieniędzy na rozrywkę, tym więcej mogą wydać na literaturę.
Jaki wniosek można z tego wyciągnąć? Czy te dwie zmienne są powiązane, czy nie? Oczywiście, że są powiązane, czyli istnieje korelacja. Korelacja ta jest jednak ujemna. Błędnym wnioskiem byłoby stwierdzenie, iż te dwie zmienne są niezwiązane, czyli że nie ma korelacji. Przykład ten jest nieco uproszczony, bo studenci nie wydają pieniądzy tylko na rozrywkę i na literaturę. Pokazuje on jednak na konieczność ostrożności przy interpretacji współzależności.

Współczynnik korelacji rang
Zjawiska pedagogiczne mają taką naturę, że rzadko wśród nich spotyka się zmienne przedziałowe. Dlatego też stosowanie współczynnika korelacji Pearsona należy do rzadkości. Dotyczy to głównie licznych cech ludzkich, takich jak np.: motywacja, pilność, uprzejmość, popularność, komunikatywność itd. Nauczyciel (wychowawca) jest w stanie stosunkowo rzetelnie uszeregować swoich uczniów według właściwości tego rodzaju. W ten sposób otrzymuje się zmienne porządkowe – rangi. Współczynnik Pearsona nie pasuje do zmiennych porządkowych. Jeżeli nie ma innej możliwości, można go jednak zastosować, zaniedbując fakt, iż zmienne nie należą do przedziałowych. Nie jest to jednak dobre rozwiązanie. Jeżeli natomiast badacz posiada dla każdej jednostki rangi dotyczące dwóch właściwości – można zastosować współczynnik korelacji Spearmana.
Korelacja jest dodatnia w przypadku, gdy wyższym rangom jednostek w jednej zmiennej (np. według pilności) towarzyszą wyższe rangi jednostek w drugiej zmiennej (np. według osiągnięć szkolnych). W przypadku korelacji ujemnej sytuacja ta jest odwrotna.
O sile korelacji decyduje to, jak często się rangi są zgodne (zarówno w kierunku dodatnim, jak i ujemnym). W tabelach nr 30 i 31 zostały zaprezentowane charakterystyczne przypadki korelacji między rangami. Z powodu jasności przekazu przykłady są krótkie (mają małą liczbę jednostek). W praktyce dla tak małych grup w ogóle nie dokonuje się obliczeń współczynników korelacji.

Tabela 30. Korelacje rang

Całkowita dodatnia korelacja		Całkowita ujemna korelacja	
R_1	R_2	R_1	R_2
1	1	1	5
2	2	2	4
3	3	3	3
4	4	4	2
5	5	5	1

Tabela 31. Korelacje rang

Wysoka dodatnia korelacja		Wysoka ujemna korelacja	
R_1	R_2	R_1	R_2
1	2	1	5
2	1	2	4
3	3	3	3
4	5	4	1
5	4	5	2

Obliczanie współczynnika Spearmana

Obliczanie współczynnika Spearmana jest stosunkowo proste. Jeżeli populacja jest wystarczająco liczna, współczynnik Spearmana można porównać ze współczynnikiem Pearsona, a więc jednocześnie podobnie go interpretować. Wzór do obliczania współczynnika Spearmana jest następujący:

$$\rho = 1 - \frac{6 \cdot \Sigma\left(R_1 - R_2\right)^2}{N \cdot \left(N^2 - 1\right)}$$

ρ	współczynnik korelacji Spearmana
$\Sigma(R_1 - R_2)^2$	suma kwadratów różnic między rangami

Oto przykład:

Tabela 32. Obliczanie współczynnika korelacji Spearmana

R_1	R_2	$R_1 - R_2$	$(R_1 - R_2)^2$
1	7	-6	36
2	5	-3	9
3	6	-3	9
4	4	0	0
5	3	2	4
6	1	5	25
7	2	5	25
			$\Sigma(R_1 - R_2)^2 = 108$

$$\rho = 1 - \frac{6 \cdot \Sigma\left(R_1 - R_2\right)^2}{N \cdot \left(N^2 - 1\right)} = 1 - \frac{6 \cdot 108}{7 \cdot \left(49 - 1\right)} = -0,93$$

Korelacja jest wyjątkowo wysoka i ujemna. Zgodność wyników jest praktycznie zupełna. Niższym rangom R_1 odpowiadają wyższe rangi R_2, natomiast wyższym rangom R_1 odpowiadają niższe rangi R_2. W interpretacji należy uwzględnić przynajmniej dwa fakty:

1. Rangi, które określił nauczyciel, są subiektywne. Inny nauczyciel być może uszeregowałby tych samych uczniów inaczej. Wówczas z innych rang otrzymano by inny współczynnik. Wprawdzie rzetelność uszeregowania przez dobrego nauczyciela jest zwykle bardzo wysoka, ale nie idealna!
2. Rangi nie odzwierciedlają właściwości uczniów tak dobrze jak ewentualna przedziałowa zmienna. Różnice między uczniami według rang wydają się jednakowe (a jednak takimi nie są). Najwyraźniej można to zauważyć w skrajnych przypadkach: nawet najbardziej ekstremalna wartość posiada rangę tylko o jeden wyższą lub niższą od sąsiedniej (może nieekstremalnej).

Stosunek korelacyjny

Indeks korelacji jest najprostszą miarą zależności korelacyjnej. Stanowi on podłoże stosunku korelacyjnego. Stosunek korelacyjny stosuje się w dwóch przypadkach. Wtedy, gdy:

1. Zmienne są przedziałowe, ale korelacja nie jest liniową (spełnione są wszystkie pozostałe warunki do stosowania współczynnika Pearsona). Przy stosowaniu współczynnika Pearsona kontroluje się liniowość korelacji tylko w przybliżeniu (wystarczy, że korelacja jest prawie liniowa). Stąd też stosunek korelacyjny stosuje się stosunkowo rzadko – wtedy gdy korelacja jest wyraźnie krzywoliniowa.
2. Zmienna zależna spełnia wszystkie warunki do stosowania współczynnika Pearsona, a zmienna niezależna jest porządkową lub nawet nominalną.

Stosunek korelacyjny według definicji jest pierwiastkiem stosunku pomiędzy wariancją wyjaśnioną i wariancją całkowitą.

$$\eta = \sqrt{\frac{\sigma_w^2}{\sigma_c^2}}$$

σ^2_w	wariancja wyjaśniona
σ^2_c	wariancja całkowita

Stosunek korelacyjny można porównać ze współczynnikiem Pearsona w przypadkach, gdy liczba grup jest wielka lub gdy nieliniowość jest mała. Oznacza to, iż można go interpretować w bardzo podobny sposób. Wariancje nie posiadają znaku plus lub minus, nie ma go też stosunek korelacyjny. Z tego powodu przy interpretacji kierunku korelacji należy uważnie przejrzeć wszystkie rezultaty i ocenić, jaki jest jej kierunek. Zostanie to zilustrowane przykładem.

Tabela 33. Wyniki uczniów na teście (*y*) i ich oceny (*x*).

x	*y*
Dopuszczająca	30 21 22 24 18 20 18 18 19 20 30
Dostateczna	28 24 30 29 25 23 33 30 26 29 34
Dobra	34 29 33 40 26 29 25 20
Bardzo dobra	35 43 38 44 31 48 35 28
Celująca	42 49 44 54 50 42 20

$$\boxed{\eta = 0{,}77}$$

Już z powierzchownego przeglądu łatwo odczytać, iż wyniki testowe uczniów celujących są lepsze niż wyniki testowe uczniów dopuszczających. Wskazuje to na korelację pozytywną. Im wyższa ocena, tym wyższe wyniki testu i odwrotnie – im niższa ocena, tym niższe wyniki testu. Do porównania należy jednak włączyć wszystkie grupy. W przypadkach, gdy różnice między grupami są wyraźne, łatwo rozpoznać, czy korelacja jest pozytywna czy negatywna. Jednak w sytuacjach, gdy różnice nie są wystarczająco wyraźne i nie można rzetelnie rozpoznać kierunku korelacji, należy wyliczyć średnie arytmetyczne wszystkich grup. Wartości średnich arytmetycznych umożliwiają rozpoznanie kierunku korelacji. Oto średnie arytmetyczne dla przypadku opisanego powyżej:

Tabela 34. Średnie arytmetyczne

Ocena	**Średnia arytmetyczna**
Dopuszczająca	21,82
Dostateczna	28,27
Dobra	29,50
Bardzo dobra	37,75
Celująca	43,00

Średnie arytmetyczne wzrastają systematycznie od pierwszej do ostatniej grupy. Oznacza to, że ocenom niższym odpowiadają niższe wyniki testu, ocenom wyższym natomiast wyższe wyniki testu. Korelacja jest więc pozytywna. Należy jednak podkreślić, iż zaprezentowany przykład jest bardzo prosty. Wszystkie średnie arytmetyczne wskazują ten sam kierunek, a dodatkowo różnice pomiędzy nimi są ogromne – nie ma więc wątpliwości co do kierunku współzależności. Nie zawsze jednak badane sytuacje będą tak proste. Dlatego też konieczna jest wielka ostrożność przy interpretacji kierunku korelacji.

Rozdział VI

PODSTAWY METODY REPREZENTATYWNEJ

Rozdział ten zawiera znane już podejścia. W zasadniczo nowy sposób korzysta się tu z metod, które opisywano w poprzednich rozdziałach. Podobnie jak wcześniej w centrum uwagi znajdują się średnie wartości, procenty, wariancja i korelacja.

Często z różnych powodów nie można zgromadzić danych dla wszystkich jednostek badanej populacji. W takich przypadkach należy wybrać mniejszą część populacji. Trzeba jednak pamiętać, iż wszystkie pytania odnoszą się do całej populacji! W takim przypadku całą populację nazywa się **populacją generalną**. Określenie to wyraźnie wskazuje na fakt, iż populacja ta znajduje się w centrum uwagi i ona właśnie podlega badaniu. Wybrana mniejsza część, dla której zostały zgromadzone dane dotyczące badanych zjawisk, nazywa się **próbą**. Na podstawie próby ustala się, jaka jest populacja generalna. Celem badacza ciągle pozostaje zbadanie populacji generalnej, próba natomiast jest tylko środkiem do osiągnięcia tego celu. Próba byłaby niepotrzebna, gdyby dysponowano danymi dla całej populacji. Wówczas też populacji nie określano by mianem populacji generalnej.

Badania wykonywane na próbach (a nie na całej populacji generalnej) prowadzone są według następujących etapów:

- określenie populacji generalnej,
- wybór jednej lub więcej prób z tej populacji generalnej,
- stosowanie metod statystycznych, które umożliwiają generalizację z próby na populację generalną,
- wnioskowanie o populacji generalnej.

Dlaczego dokonuje się badań na próbach?
Populacje, które bada się w obszarze pedagogiki, nie zawsze są łatwo dostępne. Istnieją trzy główne powody, dla których badania pedagogiczne wykonuje się częściej na próbach niż na całych populacjach.

1. Populacje mogą być bardzo wielkie. Uczniów szkół podstawowych w Polsce jest kilka milionów. Rzadko zdarza się jednak, iż badaniami obejmuje się uczniów wszystkich klas szkół podstawowych. Mimo to np. populacja uczniów klas szóstych jest już bardzo liczna – około trzy czwarte miliona. Można przypuszczać, że zbyt duże są też populacje czytelników czasopism młodzieżowych, uczestników szkolnych kół zainteresowań, studentów w dowolnym województwie itd. Z tego powodu, iż populacja jest zbyt duża do zbierania danych wybiera się tylko jej część – próbę.

2. Populacje mogą być też rozproszone geograficznie. Populacja dyrektorów szkół gimnazjalnych w wybranym województwie nie jest już tak wielka, jednak dyrektorzy ci są rozproszeni po całym województwie. Liczba jednostek tej populacji nie stanowi większej przeszkody dla badacza, rozproszenie geograficzne natomiast jest już utrudnieniem. W przypadku ankiety rozpowszechnionej przez pocztę, badaniem można objąć wszystkich dyrektorów. Jeżeli natomiast zamierza się odwiedzić i ankietować dyrektorów bezpośrednio, wymaga to podróżowania po całym województwie. W przypadku gdy populacja jest bardzo rozproszona pod względem geograficznym, do zbierania danych należy wybrać mniejszą jednostkę geograficzną – np. jeden powiat (lub kilka powiatów).

3. Populacje w obszarze edukacji pojawiają się w kilku generacjach. Populację wszystkich studentów, którzy studiowali według określonego programu, tworzą wszystkie generacje studentów w czasie obowiązywania tego programu. Jeżeli zamierza się objąć badaniem całą populację, należałoby powtórzyć część procedur badawczych w każdym roku akademickim, do momentu wygaśnięcia obowiązku realizowania tego programu. Danymi całej populacji dysponuje się dopiero wtedy. Jest to moment, w którym dopiero można rozpocząć opracowywanie zgromadzonych danych, a następnie zakończyć badanie. W przypadku, gdy badaną populację tworzy więcej generacji, jako próbę wybiera się na ogół jedną generację (a często jedynie część jednej generacji).

Podobieństwo próby i populacji

Najważniejszą cechą próby jest jej dobra reprezentatywność. Oznacza to, że próba jest podobna do populacji generalnej, z której została wybrana. Im bardziej jest podobna, tym większa jej reprezentatywność. W próbie znajdują się tylko i wyłącznie jednostki z populacji generalnej, dlatego próba do pewnego stopnia jest zawsze podobna do populacji generalnej. Wskazane jest, aby wnioski dotyczące populacji generalnej były możliwie najbardziej trafne (prawdziwe). Z tego właśnie powodu próba powinna mieć dobrą reprezentatywność. Dobra reprezentatywność nie oznacza podobieństwa według tylko jednej zmiennej. Dobrze, gdy próba jest podobna do populacji generalnej pod względem wszystkich badanych zmiennych, a najlepiej według wszystkich zmiennych, nawet tych, które nie podlegają badaniu. W celu łatwiejszego rozumienia reprezentatywność zostanie omówiona poniżej według jednej tylko zmiennej.

Reprezentatywność próby zależy od trzech czynników:

- rozproszenia zmiennej w populacji generalnej,
- liczebności próby,
- sposóbu wyboru jednostek do próby.

Rozproszenie w populacji generalnej

W przypadku, gdy jednostki populacji generalnej różnią się między sobą, różnią się też jednostki wybrane do próby. Z tego powodu też cała próba będzie różniła się od populacji generalnej. Im bardziej się różni, tym mniejsza jej reprezentatywność. Większe różnice pomiędzy jednostkami oznaczają równocześnie większe różnice między próbą a populacją generalną. W przypadku, gdy różnice pomiędzy jednostkami są małe, próba nie może różnić się od populacji w istotny sposób. Ilustruje to przykład, w którym wszystkie jednostki populacji generalnej są jednakowe. W takim przypadku jednostki w próbie też są jednakowe, a zatem próba nie różni się od populacji generalnej (oczywiście tylko według jednej wybranej zmiennej, np. wzrostu uczniów). Reprezentatywność tej próby byłaby idealna.

Pytania do przemyślenia

- Jaka jest reprezentatywność próby, jeżeli wystepuja w niej niemal same kobiety? Na to pytanie nie można odpowiedzieć jednoznacznie (z powodu braku istotnej informacji). Jeżeli populację generalną stanowią nauczycielki szkół podstawowych, wtedy próba posiada dobrą reprezentatywność według płci. Jeżeli jednak populację generalną stanowią dyrektorzy szkół licealnych, wtedy w próbie jest zbyt dużo kobiet, a zatem reprezentatywność jest słaba.

- Czy w jednej próbie mogą znajdować się jedynie dzieci w wieku od dziesieciu do dwunastu lat? Odpowiedź jest twierdząca w przypadku, gdy populację generalną tworzą uczniowie piątych klas szkoły podstawowej. Problem wyboru prób nie jest jednak tak prosty. Nawet w przypadku, gdy populację generalną tworzą wszyscy uczniowie szkół podstawowych, w próbie mogą znajdować się jedynie uczniowie w wymienionym wieku. W tym przypadku, niestety, reprezentatywność próby będzie bardzo słaba.
- Oznacza to, iż nie można oceniać reprezentatywności próby, nie wiedząc, którą populację ona przedstawia. Próba powinna być podobna do populacji generalnej. Jest to czynnik najważniejszy!

Wielkość próby

Im większa część populacji generalnej została wybrana do próby tym większa jest reprezentatywność próby. Najmniejszą reprezentatywność posiada próba, w której znajduje się tylko jedna jednostka. Natomiast największą reprezentatywność posiada próba, w której znajdują się wszystkie jednostki populacji generalnej. W tym drugim przypadku występuje idealna reprezentatywność, ponieważ próba jest identyczna z populacją generalną. Pod względem reprezentatywności próby takie rozwiązanie jest najlepsze. Jednak w takim przypadku metoda reprezentatywna nie przynosi żadnych korzyści. Dla reprezentatywności przydatne są możliwie największe próby, natomiast z punktu widzenia ekonomii możliwie najmniejsze. Podczas planowania wielkości próby należy uwzględniać obydwa kryteria – reprezentatywność i ekonomiczność. W badaniach pedagogicznych próby wynoszą na ogół zaledwie kilka procent populacji generalnych.

Sposób doboru prób

Problem reprezentatywności jest pozornie łatwy do rozwiązania: do próby należy wybrać te jednostki, które najlepiej odzwierciedlają cechy populacji generalnej. A zatem które? Badacz bowiem nie posiada dokładnej wiedzy o populacji generalnej. Próba byłaby zupełnie niepotrzebna, gdyby badacz tę wiedzę posiadał. Rozważania o wyborze jednostek do próby są więc niemożliwe. W jaki sposób dobrać próbę? W praktyce stosuje się przynajmniej kilka sposobów doboru.

Głównym sposobem doboru próby jest dobór losowy. Z uwagi na to, że dobór losowy jest tak ważny, najczęściej wyodrębnia się:
1. dobór losowy,
2. dobory nielosowe.

Dobór losowy zostanie omówiony szczegółowo poniżej, natomiast z doborów nielosowych zostaną zaprezentowane tylko dwa – te, które najlepiej ukazują istotę różnic pomiędzy poszczególnymi doborami.

Dobór losowy

Dobór, w którym prawdopodobieństwo wyboru dla wszystkich jednostek populacji generalnej jest **takie samo**, nazywa się doborem losowym. Można to osiągnąć jedynie poprzez losowanie, które można prowadzić dwoma sposobami.

Pierwszy sposób polega na tym, że w bębnie losującym znajdują się kartki z nazwami wszystkich jednostek stanowiących populację generalną. Z bębna wyciąga się kartki w liczbie odpowiadającej liczebności próby. Wybrane jednostki stanowią próbę. Procedura ta jest czasochłonna, skomplikowana i trudna do wykonania. Dlatego też w praktyce stosuje się znacznie łatwiejszy sposób, którym jest losowanie pośrednie. Odbywa się ono w następującej kolejności:

- w bębnie umieszcza się kartki z cyframi od 0 do 9,
- z bębna wyciąga się kolejno kartki z cyframi, które zapisuje sie w tabeli,
- po wyciągnięciu i zapisaniu każda kartka wraca do bębna,
- czynność powtarza się do otrzymania obszernej tabeli.

W ten sposób powstaje tabela liczb losowych, którą można wykorzystać podczas wszystkich losowych doborów. Niewielki fragment takiej tabeli został zaprezentowany poniżej (większa tabela znajduje się w aneksie).

Tabela 35. Liczby losowe

16438	26401	35439	92229	69041	29796	74669	22110	96087	24026
91994	00621	09115	37625	18695	12253	75913	67202	11333	61626
62857	82265	92986	47294	07436	70569	67195	42685	03429	11519
32369	89112	48575	91865	06314	77562	50234	34256	77326	78084
78753	05788	24065	04188	30612	76293	95733	99613	13203	16829
35723	47791	13774	67189	64179	78866	37795	34842	27495	27242
35594	65661	75837	08558	86626	33524	65452	53231	38490	64337
86327	06961	84709	41009	62664	61581	64804	57864	47755	11097

Z tabeli po kolei wybiera się liczby, które stanowią jednostki w próbie. Oto przykład takiego postępowania.
Z populacji generalnej liczącej 52 000 jednostek dobiera się próbę (np. studenci UJ w Krakowie). Z tabeli należy wybierać liczby pięciocyfrowe, ponieważ liczebność populacji jest pięciocyfrowa. Liczby można wybierać w dowolnym kierunku, np. od lewej do prawej, od dołu do góry, itd. Zaczynając od lewej strony, jako pierwsza zostanie wybrana jednostka nr 16438, następna 26401, kolejna 35439, 92229 itd. Podczas doboru należy odrzucić liczbę 92229 i wszystkie inne powyżej 52 000, ponieważ takich jednostek w wymienionej populacji nie ma. W tabeli można posuwać się także tylko o jedno miejsce: w tym przypadku wybrane zostaną kolejno jednostki nr 16438, 64382, 43826, 38264, 82640, 26401 itd. Ilustruje to fragment tabeli:

Tabela 36. Dobór jednostek

Wybrane liczby (jednostki)	**Liczby losowe**
Pierwsza	**16438** 26401 35439 92229 69041 29796 74669
Druga	1**6438 2**6401 35439 92229 69041 29796 74669
Trzecia	16**438 26**401 35439 92229 69041 29796 74669
Czwarta	164**38 264**01 35439 92229 69041 29796 74669
Piąta	1643**8 2640**1 35439 92229 69041 29796 74669
Szósta	16438 **26401** 35439 92229 69041 29796 74669
Siódma	16438 2**6401 3**5439 92229 69041 29796 74669
Ósma	16438 26**401 35**439 92229 69041 29796 74669
itd.	16438 264**01 354**39 92229 69041 29796 74669

Czynność tę należy kontynuować do otrzymania próby o wymaganej liczebności. Gotowe tabele liczb losowych zawiera każdy dobry podręcznik ze statystyki.
Największe znaczenie dla badacza mają dwie cechy prób losowych:

1. Jedynie w przypadku próby losowej istnieje możliwość matematycznego określenia relacji pomiędzy populacją generalną a próbą. Relacje te zawierają metody statystyczne.
2. Próby losowe posiadają najlepszą reprezentatywność.

Systematyczny dobór prób

Realizacja losowego doboru próby na ogół nie jest tak prosta. Dlatego też w praktyce sposób ten zastępuje się doborem systematycznym. Do przeprowadzenia takiego doboru potrzebna jest ponumerowana lista jednostek populacji generalnej. Z tej listy wybiera się jednostki w określonym przedziale, np. co piętnasta jednostka lub co czterdziesta jednostka itp. Przedział wyboru otrzymuje się, dzieląc liczebność populacji przez pożądaną liczebność próby (wynik ten zaokrągla się do pierwszej niższej liczby). Postępowanie to ilustrują dwa poniższe przykłady.

Tabela 37. Systematyczny dobór prób

Pierwszy przykład		
Liczebność populacji generalnej (n)	**Liczebność próby (n)**	**Przedział wyboru**
50 000	200	50 000 : 200 = 250
Z populacji generalnej wybiera się co 250. jednostkę		
Drugi przykład		
Liczebność populacji generalnej (n)	**Liczebność próby (n)**	**Przedział wyboru**
11350	100	8540 :100 = 113,5 lub 113
Z populacji generalnej wybiera się co 113. jednostkę		

Należy jednak podkreślić, że doboru nie zaczyna się od pierwszej jednostki na liście. Należy wylosować jednostkę z pierwszego przedziału. Następnie od wybranej jednostki zaczyna się wybór przedziałowy (w pierwszym przykładzie co 250. jednostkę, w drugim co 113.).

Dobór systematyczny dobrze zastępuje dobór losowy, ponieważ pozwala otrzymać próby, których reprezentatywność jest prawie tak samo dobra.

Próby okazjonalne

W badaniach edukacyjnych często występują próby, które nie zostały dobrane według opisanych wyżej sposobów. Oto kilka ich charakterystycznych przykładów.

Tabela 38. Próby okazjonalne

Studentka pedagogiki bada poglądy uczniów szóstej klasy szkoły podstawowej	jako próbę przyjmuje uczniów szóstej klasy jednej tylko szkoły (zwykle tej, do której sama uczęszczała przed laty, ponieważ dobrze zna władze szkoły i dzięki temu jest jej dużo łatwiej uzyskać pozwolenie na zbieranie danych) – jest to próba okazjonalna.
Nauczyciel prowadzi badanie w działaniu na temat domowych obowiązków uczniów	badanie prowadzi na uczniach swojej klasy, traktując ich jako populację, i nie próbuje generalizować rezultatów szerzej.
	uczniów swojej klasy traktuje jako próbę z szerszej populacji – jest to próba okazjonalna.
Pedagog szkolny bada techniki uczenia się uczniów	jako próbę przyjmuje na przykład uczniów jednej klasy w swojej szkole – jest to próba okazjonalna.
Badacz instytutu andragogiki bada wybrane problemy bezrobotnych	jako próbę przyjmuje wszystkich bezrobotnych, zarejestrowanych w urzędzie pracy w określonym mieście – jest to próba okazjonalna.

Jest to dobór oparty na **łatwej dostępności próby.** Powstaje pytanie: czy próba okazjonalna, którą stanowią uczniowie wybranej do badania konkretnej szóstej klasy, posiada jakąkolwiek reprezentatywność w odniesieniu do populacji uczniów szóstych klas wszystkich szkół w Polsce? Odpowiedź jest jednoznaczna: **posiada!** Fakt, iż wszyscy uczniowie stanowiący próbę, są jednocześnie jednostkami całej populacji, gwarantuje próbie określoną reprezentatywność. Dlatego też poprawniej brzmi pytanie: jaka jest reprezentatywność tej próby? Na ogół reprezentatywność prób okazjonalnych jest dużo słabsza niż reprezentatywność prób losowych. W przypadku poważnych badań naukowych próby te są niewystarczające. Kolejnym mankamentem prób okazjonalnych jest brak możliwości rozwinięcia odpowiednich metod statystycznych, umożliwiających generalizację na populacje generalne. W jednym z kolejnych podrozdziałów zostanie omówiony sposób rozwiązania tego problemu.

Próby celowe

Przy doborze próby można też oprzeć się na wiedzy o populacji generalnej (np. wiedzy pochodzącej z poprzednich badań). W tym przypadku badacz celowo wybiera jednostki do próby tak, aby osiągnąć określone cele badania. Częstym przykładem prób celowych jest wybieranie jednostek

odbiegających od normy. Drugim przykładem jest dobieranie jednostek, o których badacz sądzi, że mają większe znaczenie dla badania niż inne osoby, np. wybiera takich nauczycieli, którzy o badanej problematyce wiedzą więcej niż pozostali.
Reprezentatywność takich prób jest bardzo trudna do określenia. Dlatego też rzadko można generalizować wnioski na rzeczywiste populacje. W tych aspektach próby celowe są podobne do prób okazjonalnych. Podstawowa różnica polega na tym, że próby celowe są dobrane pod względem określonych celów badacza, natomiast próbami okazjonalnymi są te, które są najłatwiej dostępne (czyli jednostki nie są w żaden sposób dobierane). Nawet jeżeli badacz stara się dobrać próbę jak najbardziej reprezentatywną, to reprezentatywność ta jest produktem jego arbitralności. Dlatego też próby celowe są często traktowane w podobny sposób jak próby okazjonalne. Późniejsza analiza cech prób może wskazać możliwości wnioskowania na rzeczywiste populacje generalne.

Dobór warstwowy

Zamiast losowania próby z całej populacji (losowanie proste) badacz może wybrać dobór warstwowy. Stosuje się go wtedy, gdy populacja generalna podzielona jest na wyraźne warstwy. Warstwowanie populacji może dotyczyć, np. płci (warstwa kobiet i warstwa mężczyzn), roku studiów studentów (I rok, II rok itd.), miejsca zamieszkania (miasto, wieś) itd. Jeżeli warstwy te są bardziej homogeniczne niż populacja generalna, wtedy badacz może się zdecydować na dobór warstwowy. Próby te dają większą reprezentatywność i w konsekwencji mniejsze błędy. Należy jednak wyraźnie podkreślić, iż korzyści te osiąga się za cenę bardziej złożonej procedury doboru próby.
Z każdej warstwy losuje się odpowiednią liczbę jednostek do próby. Należy sporządzić listę jednostek dla każdej warstwy odrębnie i odrębnie też losować jednostki do próby. W ten sposób więcej częściowych prób tworzy całą próbę. Opracowanie statystyczne też jest podzielone: najpierw opracowuje się próby częściowe odrębnie i w końcowej fazie opracowuje się próbę jako całość.
Próby warstwowe można dobrać w sposób proporcjonalny lub nie. Proporcjonalna próba warstwowa powstaje wtedy, gdy procent różnych warstw w próbie jest taki sam jak procent tych warstw w populacji generalnej. Np. jeżeli procent studentów I roku w całej populacji wynosi 23%, to również w próbie ta warstwa będzie zastąpiona w 23%. Oczywiście, że te modyfikacje można zastosować tylko wtedy, gdy posiada się pewną wiedzę o populacji generalnej (przynajmniej o proporcjach warstw).

Próby warstwowe mogą być dobierane także w sposób systematyczny. Wtedy zamiast losowania z warstw, wybiera się jednostki z każdej warstwy w sposób systematyczny.

Dobór jednostopniowy i wielostopniowy

Podczas doboru jednostopniowego już w pierwszej fazie otrzymuje się jednostki, a tym samym całą próbę. W przypadku jednostopniowego doboru losowego w bębnie znajdują się wszystkie jednostki populacji generalnej (i wybiera się je bezpośrednio). Próbę można wybrać też poprzez dobór wielostopniowy: na pierwszych etapach wybiera się całe podgrupy, a dopiero z nich poszczególne jednostki. Dobór wielostopniowy zostanie zilustrowany przykładami z praktyki.

Tabela 39. Wielostopniowy dobór prób

Populacja generalna	Pierwszy stopień	Drugi stopień	Trzeci stopień
uczniowie czwartej klasy szkół podstawowych w roku 2006/2007	z listy wszystkich szkół podstawowych wybrano co pięćdziesiątą szkołę	w wybranych szkołach wylosowano po dziesięciu uczniów	nie było go
uczniowie, którzy rozpoczęli naukę w pierwszej klasie liceum w roku 2005/2006	z listy wszystkich szkół licealnych wylosowano 75 szkół	co piąty uczeń z generacji 2005/2006 (próba liczyła 1944 uczniów)	nie było go
nauczyciele wczesnej edukacji w szkołach podstawowych w roku 2005/2006	z listy wszystkich powiatów wylosowano 15 powiatów	w tych powiatach wylosowano po dwie szkoły	w wylosowanych szkołach wybrano co czwartego nauczyciela

Korzyści płynące z doboru wielostopniowego zostaną omówione w oparciu o pierwszy przykład z tabeli. W przypadku jednostopniowego doboru losowego badacz powinien posiadać listę wszystkich uczniów, co oznacza ich nawet do kilku milionów. W wyniku losowania np. 1944 uczniów otrzymano by próbę rozproszoną po całym kraju. Oznacza to podróżowanie do bardzo odległych zakątków kraju z powodu czasami tylko jednego ucznia. Na ogół taki obszar doboru próby jest trudny do realizacji. W przypadku

doboru wielostopniowego opisanego wyżej w tabeli badacz odwiedził tylko 75 szkół (podobnie w pozostałych przykładach).
Dobór wielostopniowy umożliwia badaczowi wyraźną oszczędność czasu podczas gromadzenia danych. Pozornie dobór wielostopniowy jest bardziej pracochłonny. Podczas doboru jednostopniowego próbę otrzymuje się już w pierwszej fazie. Dobór wielostopniowy wymaga kilkukrotnych powtórzeń podobnych operacji. Oczywiste jest jednak to, iż selekcja jednostek do próby stanowi zaledwie niewielką część całej procedury badawczej. Zasadnicza część pracy pojawia się wówczas, gdy wybrane do próby jednostki należy objąć badaniem. Jest to ankietowanie, testowanie, wykonanie rozmaitych pomiarów, itp. Przy samym wyborze jednostek dobór wielostopniowy wymaga więcej pracy i czasu (np. wielostopniowy może trwać całą godzinę, a jednostopniowy piętnaście minut). Jednak poszukiwanie wybranych jednostek na terenie (i następnie zbieranie danych) w przypadku doboru wielostopniowego może trwać na przykład jeden dzień, a w przypadku doboru jednostopniowego kilka tygodni lub kilka miesięcy.
Większej ekonomiczności działań badacza towarzyszy niestety spadek reprezentatywności próby. Dobór wielostopniowy oznacza na ogół słabszą reprezentatywność, bowiem im więcej faz, tym reprezentatywność słabsza.
W przypadkach gdy populacja generalna obejmuje cały kraj (lub któreś z dużych województw), istnieje niemal pewność, iż należy skorzystać z doboru wielostopniowego.

Duże i małe próby

W badaniach pedagogicznych próby zawierają na ogół od kilkudziesięciu do kilkuset jednostek. Próby z kilkoma tysiącami jednostek należą do rzadkości. Jeżeli jednak badaniu podlegają na przykład jedynie oceny uczniów na progach edukacyjnych (przejście ze szkoły podstawowej do gimnazjum, z gimnazjum do szkoły ponadgimnazjalnej itd.) nie ma trudności z opracowaniem próby, która zawiera kilka tysięcy jednostek. Dane pobiera się z dokumentacji szkół i nie ma większej różnicy, czy próba obejmuje kilkudziesięciu uczniów, czy też kilkuset. Na ogół badania nie obejmują tak prostych danych empirycznych, dlatego też wielkie próby nie należą do częstych zjawisk. Próby najczęściej nie obejmują więcej niż kilkaset jednostek.
Z małymi próbami należy postepować ostrożniej. W zasadzie zarówno w przypadku mniejszych, jak i większych prób należałoby stosować te same metody statystyczne. Praktyka pokazuje jednak, iż przy większych próbach można stosować metody uproszczone (z nieistotnym błędem). Dlatego też

wszystkie próby można podzielić na dwie kategorie: duże i małe. Między nimi nie ma wyraźnej granicy. W niektórych procedurach granica ta wynosi 30 jednostek, w innych 100 (a nawet i więcej).
W przypadku małych prób nie wolno stosować metod uproszczonych. Oznacza to, że jeżeli można stosować określoną metodę w małych próbach, to bez wątpliwości można jej użyć w przypadku dużych prób – duże próby są przynajmniej tak samo dobre, jak próby małe. Zasada ta nie działa w przeciwnym kierunku, ponieważ nie zawsze to, co poprawne w przypadku dużych prób, jest też poprawne w odniesieniu do małych prób.

Prosta próba losowa
Próba ta jest podstawą wszystkich procedur metody reprezentatywnej. Wszystkie metody statystyczne opisane poniżej dotyczą tylko i wyłącznie tej próby. Próba ta jest dobierana **losowo i jednostopniowo**. Do tej kategorii nie należą wszystkie badane próby. Metody statystyczne znajdują zastosowanie także w przypadku prób, które dobrze zastępują próby losowe (np. próby dobrane w sposób systematyczny). Reprezentatywność osiągana doborem systematycznym jest niemal identyczna, jak w przypadku doboru losowego. Błędy powstałe przy stosowaniu tych samych procedur statystycznych są nieznaczne, a tym samym nieistotne. Nie dotyczy to jednak prób okazjonalnych.

Stosowanie prób okazjonalnych
Jeśli podczas badania korzystano z prób okazjonalnych, to uogólnienia wydają się niemożliwe, a sytuacja nierozwiązywalna. Nie istnieją bowiem metody statystyczne, które umożliwiają generalizowanie. Istnieje jednak pewne rozwiązanie.
W sytuacjach tych należy wyimaginować hipotetyczną populację generalną, z której próba okazjonalna mogłaby być dobrana losowo. Przypadek ten jest jednak wyłącznie hipotetyczny, ponieważ próba faktycznie nie została dobrana losowo. Sytuacja ta pozwala na generalizowanie z tej próby na populację hipotetyczną, tak samo jak w przypadku próby losowej i realnej populacji generalnej.
Jakie korzyści osiąga badacz, generalizując wnioski na populację hipotetyczną? Powstaje zabawne pytanie: czy sytuację tę można porównać do wyimaginowanego urlopu pod wyimaginowaną palmą na wyimaginowanej tropikalnej wyspie? W całej sytuacji rzeczywisty jest jedynie pytający, a wszystko inne jest hipotetyczne. Zdaje się, że w przypadku populacji hi-

potetycznej powstaje taka sama sytuacja, ponieważ tylko próba okazjonalna jest rzeczywista, wszystko pozostałe jest hipotetyczne.

Poszukując odpowiedzi na to pytanie, należy rozważyć, jakie cechy posiada ta populacja. Populacja jest hipotetyczna, a próba z niej wybrana – jest próbą losową. Próba jest zatem podobna do tej populacji – lub jeszcze bardziej poprawnie: hipotetyczna populacja generalna jest podobna do próby. Wnioski powstałe w badaniach nad taką próbą odnoszą się do wszystkich podobnych populacji. Dlatego też każdy, kto zamierza korzystać z rezultatów badawczych, może sam ocenić, w jakim stopniu jego populacja jest zbliżona do próby z tego badania. Jeżeli jest ona podobna, można korzystać z tych rezultatów, jeżeli nie – nie należy tego robić.

Próby okazjonalne wraz z omówioną generalizacją stosuje się w mniej istotnych badaniach, w badaniach pilotażowych (jako wstęp do późniejszych poważnych badań) itd. Próby te najczęściej stosują studenci w swoich pierwszych badaniach: w pracach seminaryjnych, licencjackich oraz magisterskich. Nie ma potrzeby, aby wymagać od studentów doboru prób losowych, ponieważ nie zakłada się, że ich badania mogą mieć istotny wpływ na szkolną rzeczywistość.

Warto zwrócić uwagę na kolejną różnicę pomiędzy próbami losowymi i okazjonalnymi. W przypadku, gdy próba jest losowa i wnioskowanie dotyczy realnej populacji generalnej, należy możliwie najdokładniej opisać populację generalną. W opis należy włączyć wszystko, co może mieć wpływ na rozumienie ostatecznych rezultatów. Próbę opisuje się wówczas jedynie krótkim zdaniem, które zawiera stwierdzenie, że próba została dobrana losowo z omówionej populacji. Zupełnie inaczej postępuje się w sytuacjach z próbą okazjonalną. Należy wówczas dokładnie opisać dobraną próbę oraz omówić dodatkowo populację hipotetyczną. Oto przykład takiego dodatkowego opisu: „rezultaty generalizuje się na hipotetyczną populację nauczycieli, podobnych do nauczycieli z badanej próby".

Oznacza to, że każdą próbę okazjonalną można traktować jako losową, jeżeli generalizacje dotyczą hipotetycznej populacji generalnej.

Znaki dla parametrów

W tej części statystyki ważna jest zasada, która zakłada, iż parametry populacji generalnej oznaczane są dużymi literami, natomiast parametry próby małymi literami. Niestety zasadę tę czasami trudno stosować, ponieważ niektóre znaki już nie były „wolne" w momencie powstania tej zasady. Oto lista znaków najważniejszych parametrów:

Tabela 40. Znaki dla parametrów

Parametr	**Populacja generalna**	**Próba**
Liczebność	N	n
Procent	P %	p %
Średnia arytmetyczna	M	$\bar{x}$
Odchylenie standardowe	σ	s
Współczynnik korelacji Pearsona	r	r
Współczynnik korelacji Spearmana	ρ	ρ

Gdy z tekstu nie wynika jednoznacznie, czy współczynnik dotyczy populacji, czy też próby, należy to precyzyjnie opisać.

Wnioskowanie o populacji generalnej

Wszystkie pytania badawcze odnoszą się na populację generalną. Pytania ogólne często rozdziela się na więcej pytań szczegółowych (już właśnie statystycznych).

Tabela 41. Uszczegółowienie pytań badawczych

Ogólne pytania badawcze	**Szczegółowe pytania badawcze**
Jaka jest populacja generalna?	Jaki jest procent badanego zjawiska w populacji generalnej?
	Jaka jest średnia arytmetyczna w populacji generalnej?
	Jaka jest wariancja w populacji generalnej?
	Jaki jest współczynnik korelacji w populacji generalnej?
Czy populacje generalne różnią się między sobą?	Czy procenty w populacjach generalnych różnią się między sobą?
	Czy średnie arytmetyczne populacji generalnych różnią się między sobą?
	Czy wariancje populacji generalnych różnią się między sobą?
	Czy współczynniki korelacji w populacjach generalnych różnią się między sobą?

Pierwszym celem badacza jest ustalenie wartości parametru w populacji generalnej. Jeżeli badacz dysponuje danymi dla całej populacji, to parametr **się wylicza**. Dane te nie są jednak dostępne (dostępne są tylko dane w próbie), dlatego też należy oszacować parametr populacji generalnej na podstawie próby. Procedurę tę nazywa się **estymacją parametryczną**.
Drugim celem jest odpowiedź na pytanie: czy populacje generalne różnią się między sobą. W najprostszym przypadku pytanie to odnosi się do dwóch populacji (np. czy średnia stypendiów studentów pedagogiki różni się od średniej stypendiów studentów ekonomii ?). Jako punkt wyjścia przyjmuje się hipotezę, iż populacje generalne nie różnią się od siebie. W takim przypadku potrzebne są dwie próby. Z różnic tych prób ustala się, czy różnią się też populacje generalne. Wnioskowanie to nazywa się **weryfikacją hipotez statystycznych**.

Rozdział VII

ESTYMACJA PARAMETRYCZNA

Celem oszacowania (estymacji) jest ustalenie wartości wybranego parametru w populacji generalnej. Najczęściej występują cztery parametry: procent, średnia arytmetyczna, wariancja i współczynnik korelacji.
W przypadku, gdy badacz posiada dane całej populacji generalnej, wartość parametru z tych danych **się wylicza**. Gdy brak danych dla całej populacji – parametru nie można wyliczyć. Na podstawie danych z próby można jedynie **oszacować** parametr populacji generalnej. Istnieją dwa sposoby oszacowywania parametrów: oszacowanie punktowe oraz oszacowanie przedziałowe.

1. ESTYMACJA PUNKTOWA

W estymacji punktowej wylicza się parametr próby (statystyka) i przyjmuje (twierdzi), iż parametr populacji generalnej jest taki sam. W tabeli nr 41 na kolejnej stronie przedstawiono kilka przykładów estymacji punktowej.
W sytuacji, gdyby próba obejmowała całą populację generalną (co jest, rzecz jasna, zupełnie pozbawione sensu), estymacja punktowa byłaby zupełnie trafna. Ze względu na to, że w próbie jest tylko część tej populacji, estymacja punktowa nie daje trafnej wartości. Prawdopodobieństwo, że otrzymana wartość jest trafna (czyli, że parametr populacji generalnej jest rzeczywiście taki, jak oszacowano), jest zerowe. Prawdą jest, iż estymacja ta jest bardzo prosta, jednak z powodu wymienionej nietrafności praktycznie zupełnie nieużyteczna. Można ją zastosować jedynie jako pomoc w trakcie lepszej estymacji (bardziej trafnej).

Tabela 42. Estymacja punktowa

Dane próby	Statystyka (parametr próby)	Parametr (parametr populacji generalnej)
W próbie n=142 ocenę dobrą posiada 54 uczniów	procent uczniów z oceną bardzo dobrą p = 38,03%	P = 38,03%
Pensje nauczycieli	średnia arytmetyczna $\overline{x}$= 1325,23	M = 1325,23
Staż pracy	wariancja s^2 = 187,37	σ^2 = 187,37
Staż pracy i pensje	współczynnik korelacji r = -0,68	r = -0,68

Jest to dobry moment na wprowadzenie pojęcia ryzyka. Ryzyko jest stałym towarzyszem procedur związanych z próbami. Jeżeli parametr populacji generalnej szacuje się za pomocą estymacji punktowej, ryzyko wynosi 100% – już z góry wiadomo, że parametr w populacji generalnej i otrzymana wartość (ocena) nie są identyczne. Istnieje jedynie teoretyczna możliwość, że w próbie i w populacji generalnej parametr jest identyczny.
Parametr w próbie nie może jednak być dowolny. Jeżeli próbę charakteryzuje dobra reprezentatywność (czyli próba jest podobna do populacji generalnej), wtedy parametr próby jest bliski parametrowi populacji generalnej. Im bardziej próba jest podobna do populacji generalnej, tym bardziej też parametr próby jest „podobny" do parametru populacji generalnej (czyli różnica pomiędzy nimi jest mniejsza). Ale parametry te nigdy nie są identyczne. Z powodu tak wielkiego ryzyka estymacji punktowej (100%), w praktyce stosuje się wyłącznie estymację przedziałową.

2. ESTYMACJA PRZEDZIAŁOWA

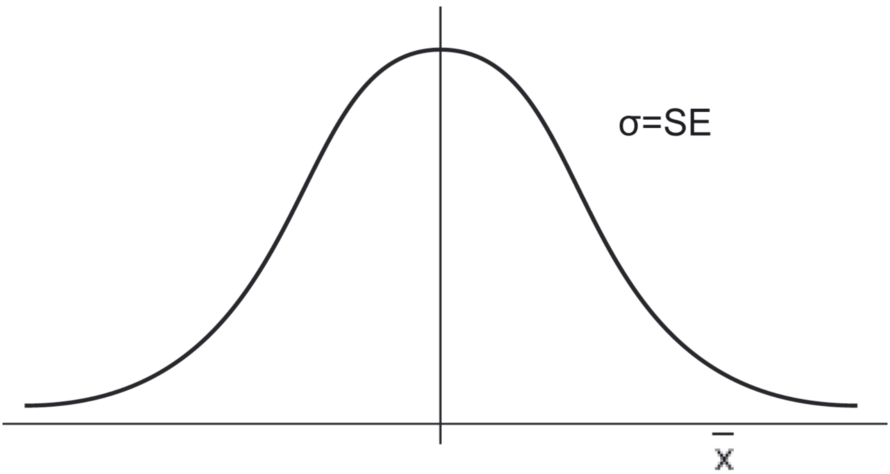
Pojęcie estymacji przedziałowej zostanie przedstawione na tych samych przykładach. Dla podkreślenia istoty tej estymacji użyto skrajnych przykładów (uwaga: nie są to realne przykłady estymacji!).

Tabela 43. Pojęcie estymacji przedziałowej

Dane próby	Parametr próby	Parametr populacji generalnej
W próbie n=142 ocenę dobrą posiada 54 uczniów	procent uczniów z oceną bardzo dobrą p = 38,03%	od 0% do 100%
Pensje nauczycieli	średnia arytmetyczna $\overline{x}$ = 1325,23	od 0 do ∞
Staż pracy	wariancja s^2 = 187,37	od 0 do ∞
Staż pracy i pensje	współczynnik korelacji r = -0,68	od –1 do +1

Czego dowodzi pierwszy przykład? Procent uczniów z oceną bardzo dobrą nie może być mniejszy od 0% i większy od 100%, dlatego też jest oczywiste, iż estymacja (oszacowanie) ta jest zupełnie trafna. Estymacja ta nie jest ryzykowna (nie ma ryzyka, że wartość parametru populacji generalnej różni się od otrzymanej wartości). Taka sama sytuacja jest w pozostałych przypadkach. Estymacja punktowa daje ocenę w postaci jednego punktu, natomiast estymacja przedziałowa wyznacza dolną i górną granicę: zakłada się, że parametr populacji generalnej znajduje się pomiędzy tymi granicami.
Oczywiście oszacowania z ostatniej tabeli są nieużyteczne i pozbawione sensu, pomimo iż są trafne! Wymienione cztery oceny można otrzymać bez żadnej próby i bez wykorzystania statystyki. Tak uproszczonej procedury nie stosuje się w praktyce. Przedział powinien być węższy, co jednocześnie oznacza, że będzie mniejsza pewność, iż jest trafny! Pojawia się ryzyko, że parametr populacji generalnej nie znajduje się w tym przedziale, lecz poza jego granicami. Im szerszy przedział, tym mniejsze ryzyko, im węższy przedział – tym ryzyko większe. Węższy przedział jest bardziej precyzyjny (skrajny przypadek stanowi ocena punktowa), natomiast szerszy przedział przynosi mniejsze ryzyko (skrajnym przypadkiem są nieskończenie szerokie przedziały z ostatniej tabeli). Kompromisem między precyzją i rzetelnością (małym ryzykiem) jest prawo mówiące o tym, iż ryzyko nie powinno przekraczać 5%. Tej zasady przestrzega się bez wyjątków!
Aby wyliczyć granice tego przedziału, należy zapoznać się z prawami dotyczącymi stosunku pomiędzy parametrami prób i populacji generalnych. Podstawą tych obliczeń jest rozkład parametru ze wszystkich prób.
W dalszej części zostanie omówiona jedynie estymacja średniej arytmetycznej, ponieważ zasady estymacji pozostałych parametrów są takie same. Jako wyja-

śnienie posłuży rysunek rozkładu średnich arytmetycznych z prób. W przypadku, gdy próby są wystarczająco duże (ich liczebność przekracza 30), rozkład jest w przybliżeniu rozkładem normalnym, czyli zgodnym z krzywą Gaussa. Odchylenie standardowe takiego rozkładu nazywa się błędem standardowym i oznacza się go symbolem SE (z języka angielskiego: *standard error*).



Przedział, w którym znajduje się średnia arytmetyczna populacji generalnej oblicza się w następujący sposób:

$$\bar{x} \pm z \cdot SE$$

$\bar{x}$	średnia arytmetyczna w próbie
z	wartość z tabeli krzywej normalnej
SE	błąd standardowy

Średnia arytmetyczna populacji generalnej znajduje się w tych granicach (z ryzykiem pięciu procent).

Powyższa procedura statystyczna umożliwia oszacowanie średniej arytmetycznej populacji generalnej na podstawie danych z próby. Procedura ta wymaga tylko jednej próby i jej danych. Jest to, niestety, ocena nieco ryzykowna, jednak nie istnieją metody reprezentatywne bez ryzyka.
Podczas stosowania takiej estymacji w praktyce należy mieć świadomość, iż powstałe wnioski są dla około pięciu procent przykładów nieprawdziwe.

Najgorsze nie jest to, że ocena może być nieprawdziwa (takich przypadków jest niewiele), lecz to, iż **nigdy** nie wiadomo, kiedy ocena jest prawdziwa, a kiedy nie. Istnieje tylko jeden sposób, który pozwala to ustalić: należałoby zbadać całą populację! Nie jest to możliwe do wykonania.
Nie ma konieczności stosowania zawsze pięcioprocentowego ryzyka. Może być ono mniejsze, czyli można wybrać jakiekolwiek ryzyko nieprzekraczające 5%. W takich sytuacjach wybiera się na ogół ryzyko wynoszące 1% lub 0,1%, ponieważ jako ludzie jesteśmy skłonni wybierać okrągłe liczby.
W jakich przypadkach stosuje się ryzyko 1% lub 0,1%? Ma to zastosowanie w przypadkach, gdy następstwem wyników badawczych mogą być bardzo kosztowne zmiany w szkolnej praktyce. Mniejsze ryzyko chroni przed wprowadzaniem niewystarczająco dobrze sprawdzonych nowości. Podobnie byłoby w przypadkach, gdyby przeprowadzone badania dowodziły czegoś innego niż dotychczasowe teorie. Większa ostrożność (mniejsze ryzyko) chroni nas wówczas przed zbyt pochopnymi wnioskami.
Bardzo podobne sytuacje zdarzają się w codziennym życiu. Na przykład mniejszą sumę pieniędzy pożycza się znajomemu, nawet w przypadku większego ryzyka, że pieniędzy nie odda. Jednak przy dużej sumie postępuje się bardziej ostrożnie. Podobnie jest z ubezpieczeniem domu. Ubezpiecza się go przed pożarem, natomiast np. koszuli już nie, ponieważ jej strata nie jest tak dotkliwa.
Oszacowanie średniej arytmetycznej zostanie zilustrowane przykładem w poniższych tabelach.

Tabela 44. Wyniki z testu wiadomości w próbie 209 uczniów gimnazjum

Wyniki w punktach	**f**
26–30	4
31–35	11
36–40	14
41–45	16
46–50	30
51–55	14
56–60	17
61–65	10
66–70	4
N	120

Należy ustalić średnią arytmetyczną populacji generalnej (z ryzykiem 5%).

Tabela 45. Obliczanie przedziału ufności

	Obliczanie	Wynik
Liczebność próby		120
Średnia arytmetyczna próby		48,04
Odchylenie standardowe próby		9,99
Błąd standardowy		0,91
Z-wartość		1,96
Przedział ufności	48,04 – 1,96·0,91 48,04 + 1,96·0,91	46,26 49,82
Twierdzenie o populacji generalnej	**46,26 ≤ M ≤ 49,82** Średnia arytmetyczna populacji generalnej znajduje się pomiędzy 46,26 i 49,82 (oczywiście z ryzykiem 5%).	

Rozdział VIII

WERYFIKACJA HIPOTEZ STATYSTYCZNYCH

1. WERYFIKACJA HIPOTEZ ZEROWYCH

Badacze często stawiają pytanie, czy populacje generalne różnią się od siebie. Różnice w tym przypadku odnoszą się na ich parametry. Czy różnią się, zatem średnie arytmetyczne populacji generalnych? Czy różnią się ich procenty? Czy różnią się ich współczynniki korelacji?

Pytania te dotyczą najczęściej różnic pomiędzy dwiema populacjami. Porównywanie dwóch populacji jest łatwiejsze niż porównywanie większej liczby populacji. Gdy w badaniu porównuje się kilka populacji, czynność tę można rozdzielić na kolejne porównania po dwie próby.

Poniżej zostanie omówiona procedura weryfikacji hipotez odnoszących się do średnich arytmetycznych. Zasady weryfikacji w przypadku pozostałych parametrów są takie same.

Posiadając dane całych populacji generalnych można wyliczyć ich średnie arytmetyczne, a następnie je porównać. Nawet najmniejsza różnica pomiędzy nimi oznacza, że populacje różnią się od siebie. Nie ma potrzeby dodatkowego udowadniania istniejącej różnicy.

W przypadku posiadania danych jedynie dla prób z tych populacji, procedura nie jest już tak prosta. Średnie arytmetyczne prób nie są wystarczające. Nawet jeżeli różnią się one między sobą, nie dowodzi to jeszcze wcale różnicy pomiędzy populacjami. Potwierdza to proste przemyślenie. Gdyby obie próby zostały wybrane z jednej populacji, ich średnie arytmetyczne byłyby różne. Różnice te są jeszcze większe w przypadku wyboru prób z dwóch różnych populacji. Stąd też istnieje potrzeba testu, w którym po-

szukuje się odpowiedzi na następujące pytanie: czy średnie arytmetyczne prób różnią się między sobą na tyle wystarczająco, by stanowiło to dowód różnic populacji?
Ta procedura nazywa się weryfikacją hipotez zerowych.
Średnie arytmetyczne prób niemal zawsze różnią się między sobą. Należy zatem ustalić, czy średnie arytmetyczne populacji generalnych również się różnią. W związku z tym przyjmuje się następującą hipotezę:

Średnie arytmetyczne populacji generalnych nie różnią się między sobą.

Jest to hipoteza zerowa. Oto jej matematyczny zapis:

$$M_1 - M_2 = 0 \text{ lub } M_1 = M_2$$

Na tym etapie ciągle jeszcze nie wiadomo, czy hipoteza zerowa jest poprawna. Dopiero test pozwala rozpoznać, jakie jest prawdopodobieństwo, że hipoteza zerowa jest poprawna, a jakie jest prawdopodobieństwo jej niepoprawności. Należy zatem przyjąć dodatkowo hipotezę przeciwną, która brzmi:

Średnie arytmetyczne populacji generalnych różnią się między sobą.

Z pierwszej populacji generalnej wybrano próbę z n_1, a z drugiej próbę z n_2. Podstawą testu zostanie różnica średnich arytmetycznych obydwu prób.
Jeżeli różnica ta jest **wystarczająco duża** (w języku statystyki – istotna), odrzuca się hipotezę zerową. Odrzucenie hipotezy zerowej dowodzi poprawności hipotezy odwrotnej, a tym samym różnic w populacjach generalnych. Istnieje jednak możliwość, że hipoteza zerowa jest poprawna. Wynika to z ryzyka testu, które nie może przekraczać 5%. Przy odrzuceniu hipotezy zerowej istnieje 5% możliwości, że populacje generalne nie różnią się.
Jeżeli różnica średnich arytmetycznych z prób **nie przekracza** pewnej granicy, hipotezy zerowej nie odrzuca się. Jakiej informacji dostarcza taki wynik o populacjach generalnych? Właściwie żadnej! W przypadku, gdy hipoteza zerowa nie została odrzucona, nie ma dowodu na różnice między populacjami.
Czy brak dowodu na różnice między populacjami oznacza jednocześnie, że są one identyczne? Pozornie są identyczne, ponieważ wydaje się, iż brak dowodu na ich różnice oznacza jednocześnie, że są takie same. Wniosek

ten jest jednak pochopny. W przypadku, gdy hipoteza zerowa nie została odrzucona, o populacjach generalnych **brak jakichkolwiek stwierdzeń**. Przedwczesność wymienionego wcześniej wniosku ilustrują zgromadzone niżej przykłady, do których wybrano charakterystyczne, a jednocześnie życiowe zdarzenia. Należy przy tym pamiętać, że zaprezentowane przykłady nie stanowią dowodów, lecz są jedynie ilustracją.

Tabela 46. Przykłady wniosków

	Niewłaściwy wniosek	**Poprawny wniosek**
Jana widziano na rynku.		Jest to dowód, że był na rynku.
Marka nie widziano na rynku.	Jest to dowód, że nie był na rynku.	Nie wiadomo, czy był na rynku, czy nie.
Oskarżonego złapano podczas rabunku.		Jest to dowód, że dokonał rabunku.
Oskarżonego nie złapano podczas rabunku.	Oznacza to, że oskarżony jest niewinny.	Nie wiadomo, czy oskarżony dokonał rabunku, czy nie.

W statystyce używa się wielu podobnych testów, np.: t-test, F-test, χ^2-test itd.

2. TEST CHI-KWADRAT

Zjawiska pedagogiczne bywają rzadko opisywane zmiennymi ilościowymi. Dlatego też weryfikacja hipotez o różnicach średnich arytmetycznych nie jest wystarczająca. Testowanie hipotez o średnich arytmetycznych, o odchyleniach standardowych oraz o współczynnikach korelacji Pearsona dotyczy tylko zmiennych ilościowych. Jedynie test hipotez o procentach stosuje się w przypadku zmiennych jakościowych. Nawet ten test nie rozwiązuje wszystkich sytuacji badawczych. Bardzo często w centrum zainteresowania znajduje się współzależność zmiennych.

W przypadku zmiennych ilościowych badacz może zastosować różne współczynniki korelacji. Do ustalania współzależności między zmiennymi jakościowymi metody rozpatrywane powyżej nie są wystarczające. Test, który zostanie omówiony poniżej, należy do najbardziej uniwersalnych metod statystycznych. Pozwala on bowiem odnaleźć odpowiedź na wiele pytań w badaniach obszaru edukacyjnego. Test ten znajduje najczęściej zastosowanie jako test niezależności oraz test zgodności.

Test niezależności

Opisywana procedura służy badaniu związków pomiędzy zmiennymi jakościowymi. Dane zestawione w tabelach, dostarczają wielu informacji o współzależności pomiędzy zmiennymi. Gdyby celem badań była próba, wówczas wystarczającą okazałaby się tabela z procentami. Celem badań jest jednak populacja generalna, a zatem z danych prób nie można o niej wnioskować niczego rzetelnego.

Pierwszym etapem jest postawienie następującego pytania: czy dwie badane zmienne z populacji generalnej są z sobą powiązane? Przyjmuje się następującą hipotezę niezależności w populacji generalnej:

Zmienne w populacji generalnej są niezależne.

Jest to jedna z postaci hipotezy zerowej, ponieważ zakłada, że nie ma zależności w populacji generalnej. Przeciwna byłaby hipoteza mówiąca o tym, że zmienne w populacji generalnej są zależne. Hipotezy tej nie trzeba zakładać, ponieważ niemal zawsze jest ona identyczna z główną hipotezą badawczą. Hipotezy zerowej nie traktuje się poważnie, ponieważ służy ona tylko do prowadzenia procedury statystycznej. Hipotezę zerową weryfikuje się za pomocą testu chi kwadrat. Podczas testowania zostaje ona odrzucona lub nie.

Dane o dwóch zmiennych w próbie są umieszczone w tabeli. Jako przykład wybrano prosty i często spotykany przypadek z obszaru edukacji, którym jest współzależność pomiędzy płcią a postawami. Prostota przykładu zawiera się w tym, że zmienna niezależna posiada dwie kategorie (kobiety/mężczyźni), zmienna zależna trzy kategorie. Przykładem bardziej prostym od przedstawionego jest już tylko tabela 2x2.

Tabela 47. Odpowiedzi w próbie badawczej (płeć a postawy)

	Zgadzam się	**Niezdecydowani**	**Nie zgadzam się**	**Razem**
Mężczyźni	21	7	8	36
Kobiety	13	10	37	60
Razem	34	17	45	96

Z tabeli jasno wynka, że mężczyźni bardziej zgadzają się z postawą zaprezentowaną w ankiecie. Dane te odnoszą się jednak jedynie do próby. Jaki zatem jest rozkład odpowiedzi w populacji generalnej?
Liczebności w tabeli określa się jako „otrzymane", ponieważ zostały zgromadzone w wyniku procedury badawczej. Oznacza się je symbolem f_E. Liczebności te odzwierciedlają stan rzeczywisty badanego zjawiska. W celu łatwiejszego rozumienia sedna przeprowadzanej procedury, liczebności zostały przekształcone w procenty. Pragnę jednak wyraźnie podkreślić, że do obliczania wartości chi-kwadrat operacja przekształcania nie jest konieczna, a nawet zupełnie zbędna. Oto tabela z przekształconymi liczebnościami.

Tabela 48. Liczebności procentowe

	Zgadzam się	Niezdecydowani	Nie zgadzam się	Razem
Mężczyźni	58,3%	19,4%	22,2%	100,0%
Kobiety	21,7%	16,7%	61,7%	100,0%
Razem	35,4%	17,7%	46,9%	100,0%

Do przeprowadzenia testu potrzebne są dodatkowo liczebności, których oczekiwałby badacz w przypadku, gdyby hipoteza niezależności była uzasadniona (ważna). Liczebności te określa się mianem „liczebności oczekiwane", które oznacza się symbolem f_T. W jaki sposób otrzymuje się te liczebności? W przypadku, gdy postawy nie zależą od płci, odpowiedzi kobiet i mężczyzn na ogół nie różnią się od siebie. Należy przyjrzeć się liczebności na przykład w lewym dolnym okienku tabeli powyżej. Zawiera ono procent tych w całej badanej próbie, którzy udzieli odpowiedzi „zgadzam się", czyli 35,4%. Jeżeli odpowiedzi kobiet i mężczyzn nie różnią się od siebie, procent osób, które udzieliły odpowiedzi „zgadzam się" wśród kobiet i mężczyzn musi być identyczny, a więc 35,4%. Respondentów niezdecydowanych wśród kobiet i mężczyzn musi być 17,7%, natomiast tych, którzy mają inne zdanie – 46,9%. Oto tabela oczekiwanych liczebności.

Tabela 49. Liczebności oczekiwane

	Zgadzam się	Niezdecydowani	Nie zgadzam się	Razem
Mężczyźni	35,4%	17,7%	46,9%	100,0%
Kobiety	35,4%	17,7%	46,9%	100,0%
Razem	35,4%	17,7%	46,9%	100,0%

Liczebności oczekiwane i otrzymane są niezgodne. Oznacza to, że rzeczywistość różni się od założonej hipotezy. Dlaczego liczebności oczekiwane i otrzymane są różne? Pierwszym powodem może być losowy dobór jednostek do próby. Różnice powstałe w ten sposób można przewidzieć matematycznie. Drugą przyczyną różnic może być fakt, że stan populacji generalnej jest odmienny od założeń hipotezy niezależności. Tym samym: kobiety i mężczyźni w próbie różnią się postawami z powodu losowania **oraz** różnic w populacji generalnej. Zasada ta stanowi sedno testu chi kwadrat. Jeżeli różnice pomiędzy liczebnościami oczekiwanymi i otrzymanymi są małe, na tej podstawie wnioskuje się następująco: najprawdopodobniej powodem różnic jest dobór losowy. W przypadku, gdy różnice pomiędzy liczebnościami oczekiwanymi i otrzymanymi są duże, można wnioskować, że ich powodem niemal na pewno są różnice w populacji generalnej. Rozbieżność, która przekracza z góry określoną granicę, wymaga odrzucenia hipotezy niezależności. Odrzucenie hipotezy niezależności traktuje się jako dowód zależności postaw od płci w populacji generalnej.
Głównym etapem testu chi-kwadrat jest pomiar rozbieżności pomiędzy liczebnościami oczekiwanymi i otrzymanymi. Miarą tej rozbieżności jest wartość chi-kwadrat (χ^2).

Test zgodności

W badaniu hipotetycznym respondenci odpowiadali na określone pytanie. Powstaje pytanie, czy którakolwiek z proponowanych odpowiedzi dominuje w populacji generalnej. Dane w próbie zawierają informację, że częstotliwości poszczególnych odpowiedzi są różne. Niektóre odpowiedzi respondenci wybierają częściej, inne rzadziej. Taki wybór zdarza się niemal zawsze, jedynie incydentalnie częstotliwości wszystkich odpowiedzi są identyczne. Jeśli celem badań byłaby próba, już dane o częstotliwościach stanowią odpowiedź na powyższe pytanie. Jeżeli częstotliwości są jednakowe, oznacza to, że żadna z odpowiedzi nie dominuje. Jeżeli natomiast częstotliwości różnią się, oznacza to, że niektóre odpowiedzi respondenci wybierają częściej niż inne. Stwierdzeń tych nie trzeba dodatkowo udowadniać sposób. Ilustrują to dwa przykłady.

Przykład 1.

Uczniów ostatniej klasy liceum pytano, jaki kraj wybraliby na miejsce swojego wakacyjnego pobytu. Proponowano trzy odpowiedzi. Oto liczebności otrzymanych odpowiedzi w badanej próbie:
Francja 31
Czechy 13
Niemcy 10
Oznacza to, że najwięcej respondentów pragnie spędzić wakacje we Francji. Fakt ten jest niepodważalny, dlatego też nie ma potrzeby udowadniania go w dodatkowy sposób.

Przykład 2.

Gdyby odpowiedzi uczniów były następujące:
Francja 18
Czechy 18
Niemcy 18
oznaczałoby to niepodważalnie, że odpowiedzi w próbie są rozłożone równomiernie. Takiego rozkładu wyników również nie trzeba dodatkowo udowadniać.

Celem badania nie jest wybrana próba, lecz populacja generalna. Liczebności w próbie same w sobie nie stanowią odpowiedzi na pytanie dotyczące populacji generalnej. Do wnioskowania o populacji generalnej jest potrzebny test chi-kwadrat.
Pierwszym etapem jest postawienie hipotezy zerowej, która zakłada identyczność liczebności wszystkich kategorii zmiennej w populacji generalnej. Hipoteza brzmi następująco:

Liczebności wszystkich kategorii zmiennej w populacji generalnej są identyczne.

W zaprezentowanym powyżej przykładzie wszystkie trzy odpowiedzi (Francja, Czechy, Niemcy) byłyby wybierane z taką samą popularnością. Gdyby odpowiedzi w populacji generalnej kształtowałyby się zgodnie z hipotezą zerową, wówczas wszystkie liczebności w tabeli byłyby identyczne. Weryfikacja hipotezy zerowej odbywa się poprzez niezgodność między liczebnościami oczekiwanymi a otrzymanymi.
Kolejnym etapem testu chi-kwadrat jest pomiar tej rozbieżności. Miarą rozbieżności jest wartość chi-kwadrat. Jeżeli wyliczona wartość przekracza z góry określoną granicę, wówczas odrzuca się hipotezę zerową. Odrzuce-

nie hipotezy zerowej traktuje się jako dowód na to, że odpowiedzi w populacji generalnej nie były wybierane jednakowo często.
W omawianym przypadku oznaczałoby to, że wszystkie trzy odpowiedzi (Francja, Czechy, Niemcy) w populacji generalnej nie cieszą się taką samą.

3. WSPÓŁCZYNNIKI ZBIEŻNOŚCI

Weryfikacja hipotezy niezależności prowadzona na podstawie testu chi-kwadrat potwierdza jedynie, iż zmienne w populacji generalnej są współzależne. Ciągle nieznaną pozostaje siła tej współzależności. Do jej ustalenia konieczna jest miara współzależności, podobna do współczynników korelacji. Miarę tę można otrzymać z wartości chi-kwadrat. Wartość ta wyraża stopień niezgodności pomiędzy otrzymanymi a oczekiwanymi liczebnościami. Im większa niezgodność, tym większa siła współzależności (oczywiście w próbie). Miarami współzależności dla zmiennych jakościowych są współczynniki zbieżności.
Pomiędzy współczynnikami zbieżności i testem chi-kwadrat dostrzega się inną, jeszcze istotniejszą różnicę. Test chi-kwadrat odnosi się do współzależności zmiennych w **populacji generalnej**. Pomimo że test chi-kwadrat został wyliczony z danych próby, odpowiada jedynie na pytania dotyczące populacji generalnej (podobnie jak wszystkie testy hipotez). W badaniach na całych populacjach generalnych (bez próby) korzystanie z testu chi-kwadrat jest pozbawione sensu. Nawet w takich badaniach wylicza się jednak wartość chi-kwadrat, a z niej następnie współczynnik zbieżności. Współczynniki zbieżności odnoszą się zawsze do grupy umieszczonej w tabeli (bez względu na to, czy dane dotyczą tylko próby, czy całej populacji generalnej). W przypadkach, gdy badaniu podlega próba – współczynnik zbieżności dotyczy próby, w badaniu prowadzonym na całej populacji – współczynnik zbieżności dotyczy populacji.
Oznacza to, że współczynniki zbieżności można stosować we wszystkich badaniach, natomiast test chi-kwadrat jedynie w badaniach na próbach. Uwaga: należy odróżniać test chi-kwadrat od wyliczania wartości chi-kwadrat (χ^2). Wartość chi-kwadrat można wyliczać w każdej sytuacji, natomiast całkowity (kompletny) test chi-kwadrat stosuje się jedynie w przypadku wnioskowania na populację generalną.

Współczynnik zbieżności Pearsona

Spośród wielu współczynników zbieżności najważniejszy jest współczynnik zbieżności Pearsona.

Współczynnik Pearsona oblicza się według następującego wzoru:

$$C = \sqrt{\frac{\chi^2}{n + \chi^2}}$$

Wartości tego współczynnika wahają się pomiędzy **0** i **1**. Interpretuje się je podobnie jak współczynniki korelacji. W przypadku zmiennych jakościowych na ogół otrzymuje się niższe stopnie współzależności niż w przypadku zmiennych ilościowych. Dlatego też niższe wartości tych współczynników interpretuje się jako znak solidnej współzależności. Drugą różnicą w interpretacji tych współczynników jest brak posiadania znaku plus lub minus. W przypadku zmiennych nominalnych nie tworzy to żadnych problemów, ponieważ brak tam dodatniego i ujemnego kierunku. W zmiennych porządkowych obydwa te kierunki istnieją, pomimo że współczynnik zbieżności ich nie wskazuje. Dlatego też w takich sytuacjach istnieje konieczność szczegółowej analizy liczebności w tabeli. Analiza ta umożliwia ustalenie kierunku współzależności. Do interpretacji kierunku współzależności potrzebna jest tabela z wartościami procentowymi.

Tabela 50. Liczebności procentowe w próbie badawczej

	Zgadzam się	**Niezdecydowani**	**Nie zgadzam się**	**Razem**
Mężczyźni	58,3%	19,4%	22,2%	100,0%
Kobiety	21,7%	16,7%	61,7%	100,0%
Razem	35,4%	17,7%	46,9%	100,0%

Wnioskowanie w przypadku niewielkich tabel jest proste, ponieważ procenty w bardzo czytelny i przejrzysty sposób wskazują na kierunek współzależności. Wyniki zaprezentowane w tabeli wskazują, że mężczyźni dużo częściej niż kobiety zgadzają się z badaną postawą.

Duże tabele wymuszają na badaczu dokonywanie porównań całych kolumn z tabeli (kolumny po kolumnie). W przypadku tabeli 50 jako pierwsze zostałyby porównane procenty z pierwszej kolumny: 58,3% i 21,7%. Więcej niż połowa mężczyzn wybrała odpowiedź „zgadzam się", natomiast

wśród kobiet takiego wyboru dokonała zaledwie jedna piąta. Porównania te kontynuuje się dla wszystkich kolumn, tworząc tym samym całkowity obraz kierunku współzależności.
Najważniejszym problemem przy interpretacji współczynnika zbieżności Pearsona jest jego uzależnienie od wielkości tabeli. Nawet w przypadku całkowitych związków pomiędzy zmiennymi, współczynnik zbieżności nie może posiadać wartości 1. Wartość tę może osiągnąć dopiero wtedy, gdy liczba kategorii obydwu zmiennych jest nieskończenie wielka, co oznacza tym samym nieskończenie wielką tabelę.
Można to zilustrować przykładem tabeli zawierającej te same zmienne, które występowały w poprzednich tabelach. Przy założeniu, że wszystkie kobiety wybrały odpowiedź C, a wszyscy mężczyźni odpowiedź A, otrzymany wynik jest bardzo frapujący. Związek pomiędzy zmiennymi jest całkowity – odpowiedzi są uzależnione jedynie od płci. Przypadek ten ilustruje kolejna tabela.

Tabela 51. Odpowiedzi respondentów

	Zgadzam się	**Niezdecydowani**	**Nie zgadzam się**	**Razem**
Mężczyźni	36	0	0	36
Kobiety	0	0	60	60
Razem	36	0	60	96

Współczynnik zbieżności wyliczony z tabeli wynosi zaledwie C = 0,71. Brak wiedzy o omówionej wyżej właściwości współczynnika zbieżności Pearsona może spowodować jego niewłaściwą interpretację. Wartość 0,71 wskazuje na ogromną, choć niecałkowitą współzależność, w rzeczywistości jednak współzależność ta jest całkowita (zupełna).
Przy porównywaniu współczynników z tabel różnej wielkości w interpretacji występują jeszcze większe problemy. Komparacje te są jednak najczęściej nieodzowne. Przeprowadzanym badaniom towarzyszą rozmaite liczby kategorii zmiennych: płeć ma dwie kategorie, postawy mogą posiadać np. trzy kategorie, wykształcenie np. cztery kategorie itd. W interpretacji często występuje potrzeba porównywania otrzymanych współczynników. Podobna sytuacja towarzyszy porównywaniu wyników badań własnych z wynikami wcześniejszych badań. W rozmaitych badaniach obszaru edukacji dokonuje się pomiaru tych samych lub podobnych zjawisk różnymi narzędziami. Z tego powodu zdarzają się tabele z różną liczbą kategorii.

Współczynnik zbieżności Cramera

Uniknięcie wymienionych wyżej kłopotów interpretacyjnych umożliwia zastosowanie współczynnika zbieżności Cramera. Nie można go jednak porównywać ze współczynnikiem Pearsona. Dlatego też w badaniach stosuje się albo współczynnik Pearsona, albo Cramera, natomiast nigdy obu równocześnie. Współczynnik Cramera oblicza się według następującego wzoru:

$$V = \sqrt{\frac{\chi^2}{n \cdot (m-1)}}$$

m	liczba kategorii tej zmiennej w tabeli, która posiada mniej kategorii

Sposób obliczania współczynnika Cramera powoduje jego niezależność od wielkości tabel, a tym samym stwarza możliwość porównywania współczynników z tabel różnej wielkości. Pomimo że współczynniki Cramera mogą osiągnąć wartość 1, mają na ogół wartości mniejsze niż wartości współczynnika Pearsona. Z tego powodu współczynników tych nie można porównywać. Decyzja o wyborze współczynnika Cramera wyklucza możliwość jednoczesnego zastosowania współczynnika Pearsona.

LITERATURA

Borenstein M i Cohen J., *Statistical power analysis: A computer program*, Hillsdale 1988.

Claus G., Ebner H., *Podstawy statystyki dla psychologów, pedagogów i socjologów*, Warszawa 1972.

Cohen J., *Statistical power analysis for the behavioral sciences*, Hillsdale 1988.

Ferguson G. A., Takane Z., *Analiza statystyczna w psychologii i pedagogice*, Warszawa 1999.

Góralski A., *Podstawowe metody statystyczne w psychologii i pedagogice*, Warszawa 1994.

Guilford, J. P., *Podstawowe metody statystyczne w psychologii i pedagogice*, Warszawa 1960.

Kozlov, V. S. in Erlih, J. M., *Obščaja teorija statistiki*, Moskva 1975.

Kožuh B., *Statystyka*, Częstochowa 2005.

Kožuh B., *Analiza danych w badaniach edukacji*, Kraków 2006.

Krzysztofiak, M., Urbanek, D., *Metody statystyczne*, Warszawa 1977.

Major, M., Niezgoda, J., *Elementy statystyki, część I. Statystyka opisowa*, Kraków 2003.

Malarska A., Mikulska H., *Statystyka stosowana nie tylko przez psychologów i pedagogów*, Łódź 2000.

Nowaczyk, Cz., *Podstawy metod statystycznych dla pedagogów*, Warszawa-Poznań 1985.

Sobczyk M., *Statystyka*, Warszawa 2002.

Winer, B. J., *Statistical Principles in Experimental Design*, London 1971.

Zaborowski Z., *Wstęp do metodologii badań pedagogicznych*, Wrocław 1973.

Zaczyński W. P., *Statystyka w pracy badawczej nauczyciela*, Warszawa 1997.

ANEKS

Tabela A. Rozkład krzywej normalnej (Gaussa)

z	%	z	%	z	%	z	%	z	%
0,00	0,00	0,40	15,54	0,80	28,81	1,20	38,49	1,60	44,52
0,01	0,40	0,41	15,91	0,81	29,10	1,21	38,69	1,61	44,63
0,02	0,80	0,42	16,28	0,82	29,39	1,22	38,83	1,62	44,74
0,03	1,20	0,43	16,64	0,83	29,67	1,23	39,07	1,63	44,84
0,04	1,60	0,44	17,00	0,84	29,95	1,24	39,25	1,64	44,95
0,05	1,99	0,45	17,36	0,85	30,23	1,25	39,44	1,65	45,05
0,06	2,39	0,46	17,72	0,86	30,51	1,26	39,62	1,66	45,15
0,07	2,79	0,47	18,08	0,87	30,78	1,27	39,80	1,67	45,25
0,08	3,19	0,48	18,44	0,88	31,06	1,28	39,97	1,68	45,35
0,09	3,59	0,49	18,79	0,89	31,33	1,29	40,15	1,69	45,45
0,10	3,98	0,50	19,15	0,90	31,59	1,30	40,32	1,70	45,54
0,11	4,38	0,51	19,50	0,91	31,86	1,31	40,49	1,71	45,64
0,12	4,78	0,52	19,85	0,92	32,12	1,32	40,66	1,72	45,73
0,13	5,17	0,53	20,19	0,93	32,38	1,33	40,82	1,73	45,82
0,14	5,57	0,54	20,54	0,94	32,64	1,34	40,99	1,74	45,91
0,15	5,96	0,55	20,88	0,95	32,89	1,35	41,15	1,75	45,99
0,16	6,36	0,56	21,23	0,96	33,15	1,36	41,31	1,76	46,08
0,17	6,75	0,57	21,57	0,97	33,40	1,37	41,47	1,77	46,16
0,18	7,14	0,58	21,90	0,98	33,65	1,38	41,62	1,78	46,25
0,19	7,53	0,59	22,24	0,99	33,89	1,39	41,77	1,79	46,33
0,20	7,93	0,60	22,57	1,00	34,13	1,40	41,92	1,80	46,41
0,21	8,32	0,61	22,91	1,01	34,39	1,41	42,07	1,81	46,49
0,22	8,71	0,62	23,24	1,02	34,61	1,42	42,22	1,82	46,56
0,23	9,10	0,63	23,57	1,03	34,85	1,43	42,36	1,83	46,64
0,24	9,48	0,64	23,89	1,04	35,08	1,44	42,51	1,84	46,71
0,25	9,87	0,65	24,22	1,05	35,31	1,45	42,65	1,85	46,78
0,26	10,26	0,66	24,54	1,06	35,54	1,46	42,79	1,86	46,86
0,27	10,64	0,67	24,86	1,07	35,77	1,47	42,92	1,87	46,93
0,28	11,03	0,68	25,17	1,08	35,99	1,48	43,06	1,88	46,99

z	%	z	%	z	%	z	%	z	%
0,29	11,41	0,69	25,49	1,09	36,21	1,49	43,19	1,89	47,06
0,30	11,79	0,70	25,80	1,10	36,43	1,50	43,32	1,90	47,13
0,31	12,17	0,71	26,11	1,11	36,65	1,51	43,45	1,91	47,19
0,32	12,55	0,72	26,42	1,12	36,86	1,52	43,57	1,92	47,26
0,33	12,93	0,73	26,73	1,13	37,08	1,53	43,70	1,93	47,32
0,34	13,31	0,74	27,04	1,14	37,29	1,54	43,82	1,94	47,38
0,35	13,68	0,75	27,34	1,15	37,49	1,55	43,94	1,95	47,44
0,36	14,06	0,76	27,64	1,16	37,70	1,56	44,06	1,96	47,50
0,37	14,43	0,77	27,94	1,17	37,90	1,57	44,18	1,97	47,56
0,38	14,80	0,78	28,23	1,18	38,10	1,58	44,29	1,98	47,61
0,39	15,17	0,79	28,52	1,19	38,30	1,59	44,41	1,99	47,67

z	%	z	%	z	%	z	%
2,00	47,72	2,40	49,18	2,80	49,74	3,20	49,93
2,01	47,78	2,41	49,20	2,81	49,75	3,21	49,93
2,02	47,83	2,42	49,22	2,82	49,76	3,22	49,94
2,03	47,88	2,43	49,25	2,83	49,77	3,23	49,94
2,04	47,93	2,44	49,27	2,84	49,77	3,24	49,94
2,05	47,98	2,45	49,29	2,85	49,78	3,25	49,94
2,06	48,03	2,46	49,31	2,86	49,79	3,26	49,94
2,07	48,08	2,47	49,32	2,87	49,79	3,27	49,95
2,08	48,12	2,48	49,34	2,88	49,80	3,28	49,95
2,09	48,17	2,49	49,36	2,89	49,81	3,29	49,95
2,10	48,21	2,50	49,38	2,90	49,81	3,30	49,95
2,11	48,26	2,51	49,40	2,91	49,82	3,31	49,95
2,12	48,30	2,52	49,41	2,92	49,82	3,32	49,95
2,13	48,34	2,53	49,43	2,93	49,83	3,33	49,96
2,14	48,38	2,54	49,45	2,94	49,84	3,34	49,96
2,15	48,42	2,55	49,46	2,95	49,84	3,35	49,96
2,16	48,46	2,56	49,48	2,96	49,85	3,36	49,96
2,17	48,50	2,57	49,49	2,97	49,85	3,37	49,96
2,18	48,54	2,58	49,51	2,98	49,86	3,38	49,96
2,19	48,57	2,59	49,52	2,99	49,86	3,39	49,97
2,20	48,61	2,60	49,53	3,00	49,87	3,40	49,97
2,21	48,64	2,61	49,55	3,01	49,87	3,41	49,97
2,22	48,68	2,62	49,56	3,02	49,87	3,42	49,97
2,23	48,71	2,63	49,57	3,03	49,88	3,43	49,97
2,24	48,75	2,64	49,59	3,04	49,88	3,44	49,97
2,25	48,78	2,65	49,60	3,05	49,89	3,45	49,97
2,26	48,81	2,66	49,61	3,06	49,89	3,46	49,97

kozuh statystyka.indb 122 2011-05-13 12:30:01

z	%	z	%	z	%	z	%
2,27	48,84	2,67	49,62	3,07	49,89	3,47	49,97
2,28	48,87	2,68	49,63	3,08	49,90	3,48	49,97
2,29	48,90	2,69	49,64	3,09	49,90	3,49	49,98
2,30	48,93	2,70	49,65	3,10	49,90	3,50	49,98
2,31	48,96	2,71	49,66	3,11	49,91	3,60	49,98
2,32	48,98	2,72	49,67	3,12	49,91	3,70	49,99
2,33	49,01	2,73	49,68	3,13	49,91	3,80	49,99
2,34	49,04	2,74	49,69	3,14	49,92	3,90	49,99
2,35	49,06	2,75	49,70	3,15	49,92	4,00	49,997
2,36	49,09	2,76	49,71	3,16	49,92	4,50	49,9998
2,37	49,11	2,77	49,72	3,17	49,92	5,00	
2,38	49,13	2,78	49,73	3,18	49,93		
2,39	49,16	2,79	49,74	3,19	49,93		

Tabela B. Rozkład chi-kwadrat (χ^2)

df	p=0,05	p=0,01	p=0,001
1	3,84	6,635	10,83
2	5,99	9,21	13,82
3	7,815	11,34	16,27
4	9,488	13,28	18,47
5	11,07	15,09	20,52
6	12,59	16,81	22,46
7	14,07	18,48	24,32
8	15,51	20,09	26,12
9	16,92	21,67	27,88
10	18,31	23,21	29,59
11	19,68	24,72	31,26
12	21,03	26,22	32,91
13	22,36	27,69	34,53
14	23,68	29,14	36,12
15	25,00	30,58	37,70
16	26,30	32,00	39,25
17	27,59	33,41	40,79
18	28,87	34,80	42,31
19	30,14	36,19	43,82
20	31,41	37,57	45,32
21	32,67	38,93	46,80
22	33,92	40,29	48,27
23	35,17	41,64	49,73
24	36,42	42,98	51,18
25	37,65	44,31	52,62
26	38,88	45,64	54,05
27	40,11	46,96	55,48
28	41,34	48,28	56,89
29	42,56	49,59	58,30
30	43,77	50,89	59,70
35	49,80	57,34	66,62
40	55,76	63,69	73,40
45	61,66	69,96	80,08
50	67,50	76,15	86,66

Tabela C. Rozkład Studenta (t)

2p p	0,10 0,05	0,05 0,025	0,02 0,01	0,01 0,005	0,001 0,0005
1	6,314	12,706	31,821	63,657	636,619
2	2,920	4,303	6,965	9,925	31,598
3	2,353	3,182	4,541	5,841	12,941
4	2,132	2,776	3,747	4,604	8,610
5	2,015	2,571	3,365	4,032	6,859
6	1,943	2,447	3,143	3,707	5,959
7	1,895	2,365	2,998	3,499	5,405
8	1,860	2,306	2,896	3,355	5,041
9	1,833	2,262	2,821	3,250	4,781
10	1,812	2,228	2,764	3,169	4,587
11	1,796	2,201	2,718	3,106	4,437
12	1,782	2,179	2,681	3,055	4,318
13	1,771	2,160	2,650	3,012	4,221
14	1,761	2,145	2,624	2,977	4,140
15	1,753	2,131	2,602	2,947	4,073
16	1,746	2,120	2,583	2,921	4,015
17	1,740	2,110	2,567	2,898	3,965
18	1,734	2,101	2,552	2,878	3,922
19	1,729	2,093	2,539	2,861	3,883
20	1,725	2,086	2,528	2,845	3,850
21	1,721	2,080	2,518	2,831	3,819
22	1,717	2,074	2,508	2,819	3,792
23	1,714	2,069	2,500	2,807	3,767
24	1,711	2,064	2,492	2,797	3,745
25	1,708	2,060	2,485	2,787	3,725
26	1,706	2,056	2,479	2,779	3,707
27	1,703	2,052	2,473	2,771	3,690
28	1,701	2,048	2,467	2,763	3,674
29	1,699	2,045	2,462	2,756	3,659
30	1,697	2,042	2,457	2,750	3,646
40	1,68	2,02	2,42	2,71	3,55
50	1,68	2,01	2,40	2,68	3,50
60	1,67	2,00	2,39	2,66	3,46
70	1,67	2,00	2,38	2,65	3,44
80	1,66	1,99	2,38	2,64	3,42
90	1,66	1,99	2,37	2,63	3,40
100	1,66	1,98	2,36	2,63	3,39
300	1,65	1,97	2,34	2,59	3,32
500	1,65	1,96	2,33	2,59	3,31
1000	1,65	1,96	2,33	2,58	3,30
∞	1,645	1,96	2,33	2,58	3,29

∞

Tabela D. Liczby losowe

13774	77562	50234	11097	62857	62664	91994	65452	35723	06314	96087
41009	03535	34593	45170	48858	49835	55466	80349	67391	20558	62668
97689	44203	06830	59868	36326	45375	29487	98595	52964	11993	38422
72534	60334	07950	39835	61064	44547	76729	33277	49579	29284	53561
58779	41268	45552	10725	25918	39936	61593	78051	84673	74682	35567
67760	00549	07026	49173	18205	52686	35724	50609	54710	38438	76621
17204	90719	19482	67176	46776	84718	87547	78500	14619	55543	59702
62032	88185	09183	27701	46738	16351	89102	16574	10712	28924	99170
65299	00202	26803	60920	34401	44303	71954	78004	19142	47975	60783
31453	81651	29587	19778	02609	09317	83032	74027	31564	92495	21745
94856	57020	31208	63774	97504	84370	52333	73659	03468	64631	56056
87582	24888	47099	64090	42111	86586	75225	32150	82119	02385	17015
32152	44047	83026	63767	91256	67199	51349	31441	33674	08363	61429
94173	93871	06195	75910	58037	60368	06190	93131	04742	09177	30821
03215	26570	17988	74651	96423	43604	06600	79239	69587	21452	60178
27207	62217	58467	42564	89064	07145	55207	81968	83451	14733	78095
33828	98117	03938	68478	11624	45901	53945	75115	12133	57298	12989
80068	94545	03855	54389	61175	72022	28338	53618	95180	31199	60119
39101	14402	84960	76482	81823	53960	07910	29182	57082	71799	27315
44978	81352	00603	33510	69660	67934	95639	46584	99173	57310	83794
54539	12829	85685	65426	32932	94233	09095	65842	26698	60650	88432
26749	83962	89037	21569	10308	61823	50943	19261	91088	16723	51106
42309	54961	06087	23158	11218	09196	23256	88923	28386	61660	85345
25145	07142	39101	24032	95180	26426	88432	40815	73758	69835	25024
39854	38407	53245	48790	42564	32692	68206	87868	76125	40658	81553
39096	61423	00376	70678	23487	65441	09672	20301	89730	94576	45531
87302	43285	66525	23194	77620	35653	93823	87246	89152	92285	09321
53963	04439	57359	23378	02519	17911	12696	71359	63870	40041	68304
18799	46464	21472	94364	21369	79773	58130	54918	48906	29357	59706
67355	92549	61985	71267	11562	53208	52975	65543	21096	57569	87142
60457	20741	20882	93468	59323	21396	44531	45832	37560	08626	06155
44394	69673	94474	24634	73287	06840	41560	54649	18292	51265	20221
77572	69532	24632	35636	87694	40697	35308	62773	30312	24214	96307
12830	09547	34243	46987	64328	41154	39284	73905	87471	12495	69920
19512	23547	16210	48732	13250	25001	25470	91168	53451	00368	58809
27204	28338	76527	23598	00651	23857	75749	32974	65945	35841	76963
26001	24589	79450	18079	62324	08068	40656	92063	39962	89600	13871
79251	67518	49691	71990	12986	85068	94542	38405	65984	25467	52964
11993	38422	72534	60334	07950	39835	87694	40697	35308	46464	21472

INDEKS RZECZOWY