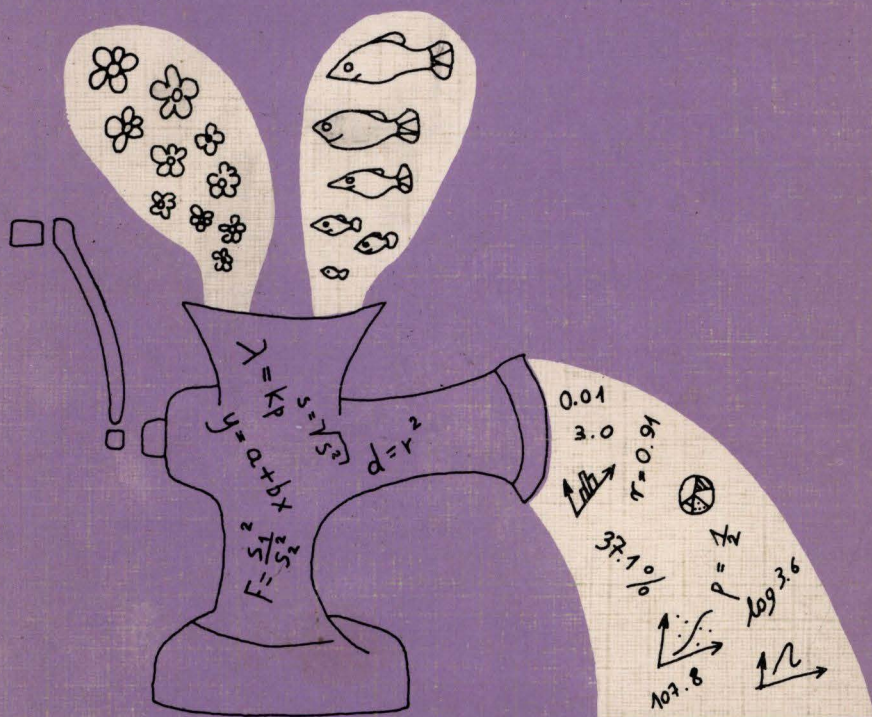


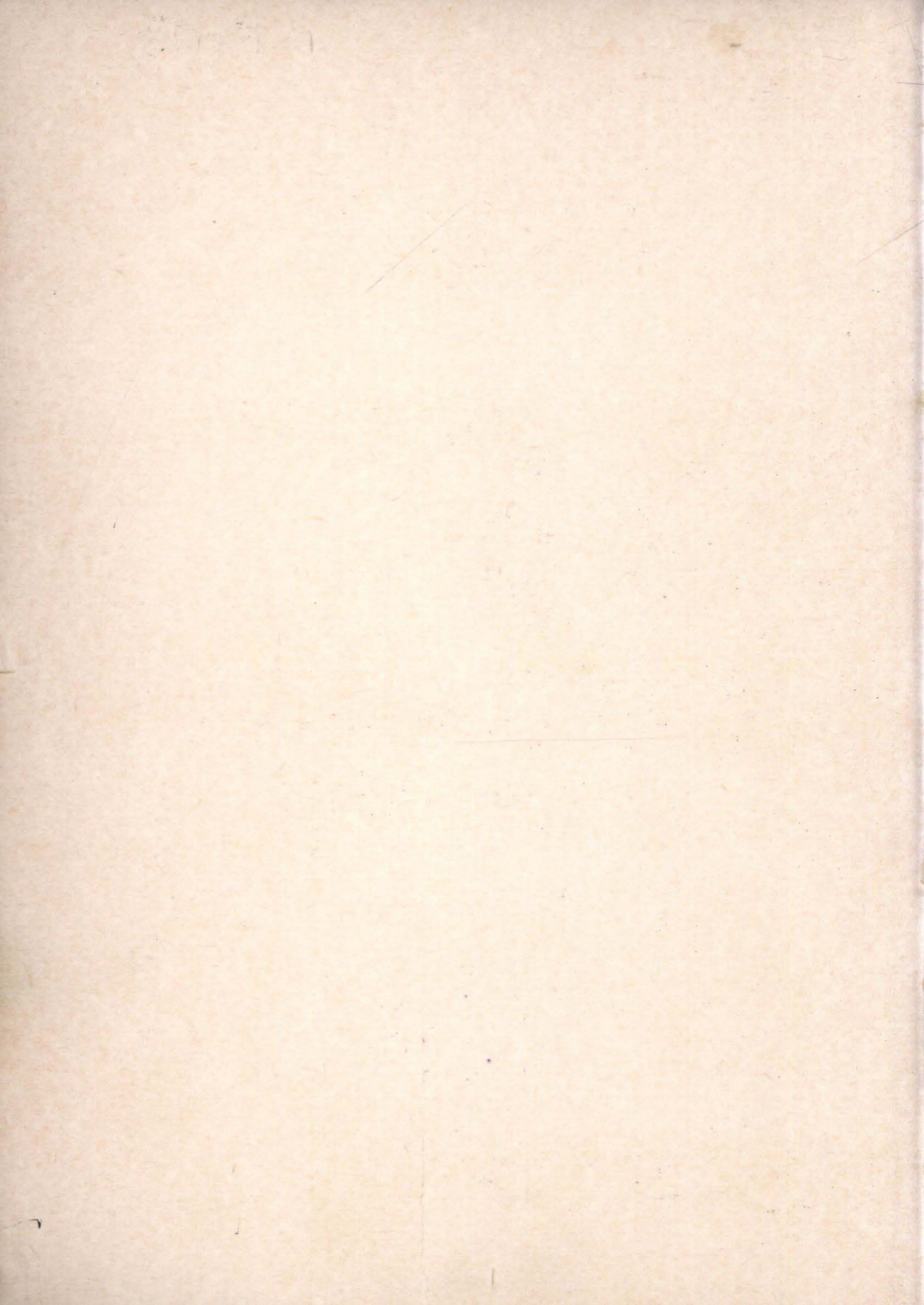
2900

Jerzy Wołek

Vademecum statystyki dla biologów



Instytut Botaniki im. W. Szafera
Polska Akademia Nauk



2900

POLISH BOTANICAL STUDIES
1992 • GUIDEBOOK SERIES • No. 6

Vademecum statystyki dla biologów

Jerzy WOLEK

Biblioteka Instytutu Botaniki



1817004235



KRAKÓW



70438
9.VII.93

POLISH ACADEMY OF SCIENCES

W. SZAFAER INSTITUTE OF BOTANY

1) Biometria Metody Szahyp.

Bj 01 000036287

POLISH BOTANICAL STUDIES
GUIDEBOOK SERIES

1992

No. 6

Guidebook Series is an irregularly issued supplementary series of *Polish Botanical Studies* intending to publish guides to Polish botanical institutions, collections and collectors, botanical literature, and directories as well as botanical field-guides and basic information regarding various branches of botany

Editor-in-Chief : Zbigniew MIREK
Editorial assistant : Jan. J. WÓJCICKI
Editorial Board : Krystyna GRODZIŃSKA
Andrzej JANKUN
Leon STUCHLIK
Jan J. ZAWADZKI

Editorial Office : W. Szafer Institute of Botany, Polish Academy of Sciences
Lubicz 46, PL-31-512 Kraków, Poland

CIP – Biblioteka Narodowa
WOŁEK, Jerzy:
Vademecum statystyki dla biologów / Jerzy Wołek. Kraków:
Polish Academy of Sciences. W. Szafer Institute of Botany,
1992 – (*Polish Botanical Studies. Guidebook Series; No. 6*)

Copyright © W. Szafer Institute of Botany, Polish Academy of Sciences

Published, sold and distributed by W. Szafer Institute of Botany, Polish Academy of Sciences
Lubicz 46, PL-31-512 Kraków, Poland

Issued: 1992
Printed in Poland

ISBN: 83-85444-07-6
ISSN: 0867-0749

Spis treści

1. Wstęp	7
2. Statystyka opisowa	10
Skala pomiaru	10
– skala nominalna	10
– skala porządkowa	10
– skala przedziałowa, skala ilorazowa	12
– przekształcanie skal	12
Populacja	12
Próba	14
– prosta próba losowa	14
Charakterystyki próby i populacji	16
– statystyki, parametry, estymatory	16
Miary statystyczne	17
Miary położenia	18
– kategoria modalna, wartość modalna	18
– obiekt medialny	19
– wartość medialna	20
– średnia arytmetyczna	20
– średnia arytmetyczna ważona	20
– średnia geometryczna	20
– średnia harmoniczna	21
Miary rozproszenia	21
– dyspersja względna klasyfikacji	21
– obiekty kwartyłowe	21
– wariancja	21
– odchylenie standardowe, współczynnik zmienności	22
Miary korelacji	22
– współczynniki siły związku	23
– współczynniki korelacji rang, współczynnik konkordancji	23
– stosunek korelacyjny	23
– współczynnik korelacji liniowej <i>r</i> Pearsona	23
– współczynnik korelacji wielokrotnej, współczynnik korelacji cząstkowej	23

Cechy statystyczne a zmienne losowe	24
Błąd pomiaru	25
Rozkłady empiryczne a teoretyczne	25
Konstruowanie rozkładu empirycznego	26
– szereg statystyczny, szereg rozdzielczy	26
– przedział klasowy	27
Rozkład empiryczny cechy ciągłej	28
– histogram, wielobok liczebności, krzywa liczebności, krzywa liczebności skumulowanych	28
Rozkład empiryczny cechy skokowej	31
– histogram, wielobok liczebności, krzywa liczebności, wielobok liczebności skumulowanych	31
Typy rozkładów liczebności	33
Graficzna prezentacja danych	35
3. Rachunek prawdopodobieństwa	46
ZDARZENIA LOSOWE I PRAWDOPODOBIEŃSTWO	46
Doświadczenie losowe, zdarzenie losowe	46
Zdarzenia elementarne, przestrzeń zdarzeń elementarnych	46
Zdarzenia złożone	47
Zdarzenia pewne, niemożliwe, prawdopodobne	47
Zmienna losowa	47
Prawidłowość statystyczna	48
Prawo wielkich liczb	48
Definicja prawdopodobieństwa	50
ROZKŁADY TEORETYCZNE	53
Zmienna skokowa: rozkład zero–jedynekowy	53
Zmienna skokowa: rozkład dwumianowy	55
Zmienna skokowa: rozkład Poissona	57
Zmienna ciągła: rozkład normalny	59
– gęstość prawdopodobieństwa, prawdopodobieństwo	59
– współczynnik asymetrii, współczynnik kurtozy	61
– standardowy rozkład normalny	62
– reguła trzech odchyłeń standardowych (trzech sigm)	63
– rozkład z próby, błąd standardowy estymatora, centralne twierdzenie graniczne	66
4. Doświadczalnictwo	68
PODSTAWOWE TERMINY	68

80	Observacja, doświadczenie	68
50	Jednostka eksperymentalna	68
20	Zabieg	69
20	Model stały, losowy, mieszany	70
70	Doświadczenie jednoczynnikowe i wieloczynnikowe	70
80	Doświadczenie kontrolne	71
80	Replikacja a powtórzenie	71
0	SCHEMATY DOŚWIADCZALNE	73
0	Metoda zmiennych połączonych	74
70	– bloki kompletnie zrandomizowane	76
80	– kwadrat taciński	78
20	– kwadrat tacińsko-grecki	80
0	Metoda zmiennych niepołączonych	83
0	– układ kompletnej randomizacji	84
0	5. Wnioskowanie statystyczne	85
0	WERYFIKACJA HIPOTEZ STATYSTYCZNYCH	85
0	Etapy analizy statystycznej	86
0	Hipoteza zerowa	87
0	Hipoteza alternatywna	88
0	Testy parametryczne, hipotezy parametryczne	89
0	Testy nieparametryczne, hipotezy nieparametryczne	89
0	Stopnie swobody	90
0	Poziom istotności	90
0	– błąd pierwszego rodzaju	91
0	– błąd drugiego rodzaju	91
0	Moc testu	92
0	ESTYMACJA PARAMETRÓW STATYSTYCZNYCH	92
0	Estymacja punktowa, estymacja przedziałowa	92
0	6. Przegląd zagadnień, które mogą być przedmiotem analizy statystycznej	94
0	Testowanie losowości próby	94
0	Testowanie jednorodności próby	94
0	Testowanie istotności miar położenia	95
0	Testowanie istotności miar rozproszenia	95
0	Testowanie istotności proporcji (frakcji, procentu, wskaźnika struktury) ...	96
0	Testowanie zgodności rozkładów	97
0	Testowanie niezależności cech	97

- Model 1: regresja	98
- Model 2: korelacja	102
Szacowanie przedziałów ufności	105
Testowanie szeregów czasowych	105
7. Wykaz testów stosowanych do statystycznej analizy danych ...	107
A. Testy sprawdzające losowość próby	108
B. Testy sprawdzające jednorodność próby	109
C. Testy istotności miar położenia	109
D. Testy istotności miar rozproszenia	115
E. Testy istotności proporcji (frakcji, procentu, wskaźnika struktury)	117
F. Testy zgodności rozkładów	119
G. Testy niezależności cech	125
H. Szacowanie przedziałów ufności	130
LITERATURA	132
INDEKS	135

1. WSTĘP

W stosunku do podręczników statystyki, pisanych przez profesjonalistów dla odbiorców, którzy nie parają się na co dzień statystyką i nie mają wiele do czynienia z matematyką, można wysunąć kilka zastrzeżeń. Po pierwsze, większość autorów zakłada, że czytelnik zna matematykę na poziomie średnim. Założenie takie jest na ogół nieprawdziwe. Bardzo wielu biologów – jedna z potencjalnych grup odbiorców – ma zaledwie nikłe pojęcie o tym przedmiocie, a i do rzadkości nie należą osoby, których średnia wiedza matematyczna zaczyna się i kończy na obliczaniu procentów. Z tego względu wszelkie dowody matematyczne, wyprowadzanie wzorów, modele, itd., eleganckie być może matematycznie, wywołują u takiego czytelnika jedynie dreszcz przerażenia i wątpienie we własne siły, a w efekcie zniechęcenie i całkowitą obojętność na zagadnienia statystyczne. Oczywiście, trafiają się również tacy czytelnicy, którzy niezrażeni trudnościami merytorycznymi, starają się opanować statystyczne metody opracowywania danych. Im z kolei na przeszkodzie staje niejednokrotnie mało komunikatywny tekst, różnorodna terminologia i symbolika stosowana na określanie tych samych pojęć, często nieprzejrzysty i nielogiczny układ treści, pomijanie jednych zagadnień na korzyść innych, uwypuklanie pewnych, nieistotnych z punktu widzenia nieprzygotowanego matematycznie czytelnika, zagadnień w nadziei, że ułatwią one zrozumienie toku wywodu, itd.

Jest rzeczą zrozumiałą, że każdy podręcznik dotyczący tego samego tematu jest trochę inny, ponieważ jest on odzwierciedleniem selektywnego i subiektywnego podejścia autora do omawianego zagadnienia. Wydaje się jednak, że przez cały czas autor nie powinien tracić z oczu potencjalnych odbiorców i ich możliwości. Wśród nich, biolog jest odbiorcą trochę nietypowym. Z jednej strony nowoczesny rozwój nauk przyrodniczych i ich matematyzacja coraz częściej zmuszają go do sięgania po metody umożliwiające obiektywne opracowywanie danych, z drugiej strony brak odpowiedniego przygotowania utrudnia mu korzystanie z tych metod. Być może już w bliskiej przyszłości ta niekorzystna sytuacja ulegnie zmianie. Tymczasem jednak, istnieje i jeszcze przez

długi czas będzie istniała spora grupa biologów, dla których statystyka matematyczna jest czymś, co wydaje się być nie do pokonania. I do nich przede wszystkim adresowana jest niniejsza praca.

Jak powiada Guilford (1964), są różne sposoby rozumienia tych samych rzeczy. Jeden pojmie zagadnienia statystyki tak, jak to zrozumiałby matematyk, inny opanuje operowanie jej metodami, nie bardzo pojmując dlaczego właściwie wykonuje daną operację w ten a nie inny sposób. Według Guilforda, każdy powinien dążyć do tego, aby na swój sposób włożyć w pojęcia statystyczne tyle sensu, ile potrafi, ponieważ właściwe stosowanie metod statystycznych i statystycznego sposobu myślenia wymaga pewnego, minimalnego poziomu zrozumienia. Właśnie takie minimum, wystarczające, jak sądzę, do opanowania "regul gry" chciałbym zaproponować. Wydaje się, że takie podejście również może spełnić swoje zadanie. W końcu na co dzień obsługujemy różne urządzenia techniczne, korzystając z informacji zawartej w instrukcji, często mając tylko niewielkie pojęcie o współdziałaniu części. Nie chcę twierdzić, że znajomość budowy i działania maszyny, którą kierujemy, jest zbędna. Sądzę tylko, że w pewnych okolicznościach, opanowanie instrukcji obsługi może poprzedzić teorię.

Praca, którą oddaję do rąk czytelnika, jest właśnie taką instrukcją, przeznaczoną dla badacza o minimalnej lub żadnej znajomości statystyki, który chciałby jednak włączyć metody statystyczne do swego warsztatu pracy, a z pewnych względów nie chce tracić całego miesiąca na żmudne studiowanie statystyki. Żeby sprostać temu zadaniu starano się zebrać i przedstawić, w sposób możliwie przystępny, podstawowe pojęcia stosowane w statystyce, sposoby projektowania doświadczeń oraz zapoznać czytelnika z wnioskowaniem statystycznym. Zestawiono również listę testów stosowanych do statystycznego opracowywania danych, klasyfikując je według schematu doświadczalnego, liczby badanych prób oraz w zależności od skali pomiarowej i rodzaju testowanej hipotezy. Przy każdym teście zaznaczono źródło: autora, rok wydania pracy i strony, na których dany test można znaleźć.

W pracy duży nacisk starano się położyć m. in. na sposób projektowania doświadczeń, wydawało się bowiem, że jest to dla biologa sprawą najważniejszą. Może on nie pamiętać, jak oblicza się wariancję, taką czy inną średnią, współczynnik korelacji liniowej itp., ale powinien opanować sztukę zakładania doświadczeń. Myli się bowiem ten, kto są-

dzi, że umiejętność planowania eksperymentów jest mu niepotrzebna, ponieważ w swojej pracy nie posługuje się doświadczeniem, tylko zadawała się obserwacją. Jeżeli swoje obserwacje chce opracować statystycznie, to znajomość planowania eksperymentów jest mu do tego również potrzebna, ponieważ obserwacje te musi najpierw ująć w określony schemat doświadczalny, aby potem móc wybrać właściwy test statystyczny. Obserwacje, to rezultaty eksperymentów, które przeprowadziła sama przyroda. Badacz przychodzi, w tym przypadku, niejako na gotowe. Musi jednak umieć określić, według jakiego schematu doświadczalnego przyroda "eksperymentowała".

Sądzę, że przedstawiony zasób informacji umożliwi czytelnikowi zorientowanie się w całości zagadnienia jakim jest analiza statystyczna i ułatwi mu przyswojenie sobie pewnego logicznego schematu, który będzie mógł wzbogacać o dodatkowe elementy, w miarę dalszego studiowania statystyki. Uważam, że cel jaki sobie postawiłem zostanie osiągnięty również w tym przypadku, gdy badacz, po zapoznaniu się z niniejszą pracą, potrafi, w sposób zrozumiały dla statystyka, przedstawić problem statystyczny, który chciałby rozwiązać z jego pomocą.

Podczas pisania niniejszej pracy korzystano z cytowanej na końcu literatury. Z tych źródeł zaczerpnięto niektóre sformułowania i pomysły przykładów. Oryginalnym wkładem autora było zebranie rozproszonej po różnych podręcznikach informacji – ważnej z punktu widzenia początkującego – w jedną logiczną i zwartą całość. Czy zamiar ten się udał – oceni czytelnik, który w tym miejscu proszony jest o ewentualne uwagi.

Na ostateczny kształt prezentowanej książki znaczny wpływ mieli recenzenci, a mianowicie panowie: prof. dr hab. Jerzy Staszekiewicz (Instytut Botaniki im. W. Szafera PAN), doc. dr hab. Jan Kozłowski (Instytut Zoologii UJ), dr Antoni L. Dawidowicz (Instytut Matematyki UJ) oraz dr Jerzy Kopydłowski (Akademia Medyczna). Ich wnikliwe uwagi krytyczne, jak również sugestie dotyczące treści i układu maszynopisu, pozwoliły mi uniknąć wielu błędów oraz ulepszyć niektóre partie tekstu. Składam im za to w tym miejscu serdeczne podziękowanie. Odpowiedzialność za ewentualne usterki i wady niniejszej książki ponosi wyłącznie autor.

Wyrazy wdzięczności należą się także pani mgr Alinie Sidor za skrupulatne wykonanie rysunków.



2. STATYSTYKA OPISOWA

Skala pomiaru:

– skala nominalna

Wyniki \rightarrow^1 obserwacji lub \rightarrow doświadczeń mogą być przedstawione za pomocą różnych skal pomiaru. *Skala nominalna* polega na tym, że dane grupuje się według arbitralnie przyjętych kategorii (wariantów) danej \rightarrow cechy jakościowej. Na przykład, \rightarrow próbę ludzi można podzielić ze względu na stan zdrowia (\rightarrow cecha) na zdrowych i chorych (kategorie), ze względu na wiek (cecha) na dzieci, młodzież, dorosłych i starców (kategorie) lub ze względu na kolor włosów (cecha) na blondynów, brunetów i rudych (kategorie). Próbę danych liczbowych można podzielić na dane występujące poniżej i powyżej danej wartości cechy lub można wydzielić grupy danych przypadające na określone zakresy badanej \rightarrow cechy ciągłej. W Tab. 1 podano przykład klasyfikacji próby liczącej 125 uczniów ($n = 125$) według 2 cech: koloru włosów i wzrostu. Każda cecha składa się z 3 wariantów.

Tab. 1. Klasyfikacja próby uczniów ($n = 125$) ze względu na dwie cechy: kolor włosów i wzrost (skala nominalna, dane fikcyjne).

Wzrost	Kolor włosów			Razem
	blondyni	bruneci	rudzi	
niski	14	10	1	25
średni	40	35	2	77
wysoki	15	8	0	23
Razem	69	53	3	$n = 125$

– skala porządkowa

Ze *skalą porządkową* mamy do czynienia wówczas, gdy elementy

¹ Strzałka (\rightarrow) kieruje do hasła, które zostanie omówione w innym miejscu (patrz Indeks), a którego treść stanowi istotne uzupełnienie danego tekstu.

próby porządkujemy za pomocą znaku $>$ (większy niż), $=$ (równy) lub znaku $<$ (mniejszy niż). Za pomocą tej skali można, na przykład, uporządkować próbę żołnierzy, ze względu na ich stopnie wojskowe, od najniższego do najwyższego rangą – szeregowy $<$ kapral $<$ sierżant $<$ kapitan $<$..., itd. lub ze względu na ich wiek – od najmłodszego do najstarszego.

Mamy tu więc do czynienia z pomiarem bardziej precyzyjnym niż w przypadku skali nominalnej. Mianowicie, możemy nie tylko pogrupować dane na kategorie, lecz również uporządkować je uwzględniając natężenie badanej cechy, chociaż stopień tego natężenia nie jest określony. Na przykład, kierując się wyglądem (\rightarrow cecha jakościowa), porządkujemy próbę osobników od najmłodszego do najstarszego. W ten sposób wiemy, jakiego typu relacje wiekowe zachodzą między nimi, nie potrafimy jednak wyrazić liczbowo stwierdzonych między nimi różnic wiekowych. Uporządkowanie to możemy również przeprowadzić w oparciu o dane metrykalne (\rightarrow cecha ciągła). W tym przypadku, natężenie badanej cechy można co prawda wyrazić ściśle, w sposób liczbowy, jednakże informacja ta nie jest wykorzystywana w \rightarrow nieparametrycznych testach statystycznych, opracowanych dla skali porządkowej.

Porządkując elementy próby (wyniki) według natężenia badanej cechy, przyporządkowuje się im równocześnie kolejny numer porządkowy, czyli *range*: 1, 2, 3, ..., n . Te numery (rangi) są przedmiotem dalszych operacji statystycznych, zamiast oryginalnych danych liczbowych, których z pewnych względów nie chcemy lub nie możemy wykorzystać.

Uporządkowanie elementów w próbie może być *zupelne* lub *niezupelne* (słabe, rangi związane). Z uporządkowaniem zupełnym mamy do czynienia wtedy, gdy w próbie nie ma elementów identycznych pod względem natężenia danej cechy. W takim przypadku każdy wynik ma swój własny numer porządkowy. W przeciwnym wypadku mamy do czynienia z uporządkowaniem słabym: wszystkie wyniki, identyczne pod względem wartości, dostają taki sam numer porządkowy (*range*). Technika rangowania jest zawsze szczegółowo omówiona tam, gdzie \rightarrow test nieparametryczny opracowany został w oparciu o skalę porządkową.

– skala przedziałowa, skala ilorazowa

W skali przedziałowej (interwałowej) przedstawiane są wyniki, które uzyskujemy poprzez **pomiar** (→ cecha ciągła). Przy pomiarze posługujemy się skalą ciągłą. Pomiar umożliwia więc nie tylko rangowanie (porządkowanie) elementów próby pod względem natężenia danej cechy, lecz także określenie **odległości** (interwału) między tymi elementami. Wadą skali przedziałowej jest umowność jej początku (por. np. skalę temperatur Celsjusza lub Fahrenheita). Jeżeli możliwe jest **nierobitralne** określenie punktu zerowego skali ciągłej, to skalę taką nazywamy **skalą ilorazową**. W skali tej mierzymy liczebność, długość, gęstość, czas, temperaturę w skali Kelvina, głośność, liczbę dni, liczbę ptaków, itp. Za pomocą skali ilorazowej można porównywać pomiary, stwierdzając ile razy dany pomiar jest większy (lub mniejszy) od drugiego. A więc dokładność pomiaru jest wyższa niż w przypadku skali przedziałowej.

– przekształcanie skal

Należy zaznaczyć, że możliwe jest przekształcanie skal pomiarowych silniejszych w słabsze. Na przykład przez uporządkowanie danych przedstawionych w skali przedziałowej według wartości wzrastających lub malejących i przyporządkowanie im kolejnych numerów (rang) otrzymujemy skalę porządkową przedstawienia danych, a przez wyróżnienie na przykład wyników większych i mniejszych od wartości średniej – skalę nominalną.

Takie osłabianie skali pomiaru, jak pokazano, wiąże się jednak z utratą informacji. Istnieją również pewne metody przekształcania skal słabszych w silniejsze, np. skali nominalnej w porządkową a porządkowej w przedziałową. Transformacja taka możliwa jest jednak tylko wówczas, gdy posiadamy pewną informację dodatkową o próbie danych. Więcej informacji na temat transformacji skal można znaleźć u Pocięchy (1986).

Populacja

Wyobraźmy sobie, że jedną z odmian żyta obsiano pole o powierzchni 1 ha. Kiedy zboże dojrzało, postanowiono zbadać, ile średnio ziarniaków zawiera jeden kłos tej odmiany żyta. Badanie to można przeprowadzić na dwa różne sposoby. Po pierwsze, można zebrać **wszystkie**

kłosa żyta, w każdym kłosie policzyć ziarniaki i następnie obliczyć, ile średnio przypada ich na jeden kłos żyta. Sposób ten jednak byłby bardzo pracochłonny. Statystyka oferuje nam łatwiejszy i szybszy sposób realizacji tego zadania. Zamiast badać wszystkie kłosa żyta, wystarczy z całości, jaką stanowi to pole żyta, pobrać tylko pewną liczbę kłosów, opracować ją za pomocą odpowiednich metod statystycznych i, na podstawie tych przebadanych kłosów, wydać sąd o średniej liczbie ziarniaków w kłosie badanej odmiany żyta na tym polu. Takie wydawanie sądów o całości na podstawie analizy jej części jest głównym zadaniem statystyki matematycznej, która zajmuje się metodami wnioskowania statystycznego (statystyką indukcyjną). *Całość*, o której wydajemy sąd, to *populacja* (populacja generalna, populacja ogólna, zbiorowość generalna, zbiorowość statystyczna), natomiast *część* populacji, na podstawie której wnioskujemy o całej populacji, to *próba*.

Populacja jest to więc zbiór wszystkich rozpatrywanych elementów (*jednostek statystycznych*), które zostały wyróżnione ze względu na pewną → cechę lub cechy i które są przedmiotem badania². Populacją może być jedno pole żyta jednej odmiany, wszystkie pola żyta jednej odmiany w całym województwie, wszystkie pola żyta jednej lub różnych odmian w całej Polsce, wszystkie pola żyta na całym świecie. Populację mogą tworzyć wszyscy uczniowie danej szkoły, ale za populację możemy także uznać tylko uczniów klasy czwartej. Decyzja należy do badacza, ponieważ to on właśnie musi zdecydować, jakich informacji oczekuje od swojego obiektu badań: czy interesują go, na przykład, warunki socjalne wszystkich uczniów danej szkoły, czy tylko uczniów klas czwartych w tej szkole.

Mówiąc o populacji uczniów klasy czwartej w danej szkole myślimy zwykle o pewnej cesze wspólnej dla tych uczniów, a będącej przedmiotem naszej statystycznej analizy. Tą cechą może być np. wzrost, kolor włosów, współczynnik inteligencji, odporność na gruźlicę, pochodzenie społeczne, itp. Praktycznie więc, mówiąc o populacji uczniów klasy czwartej, mamy na myśli populację różnych wartości badanej cechy, np. wzrostu.

Zbiorowość jednostek statystycznych może być rozpatrywana ze

² Uwaga: w naukach biologicznych znanych jest około 50 różnych definicji terminu "populacja" (Jonckers 1973), nie należy ich jednak mylić z definicją statystyczną.

względu na wartość jednej cechy (*populacja jednowymiarowa*, szereg jednowymiarowy), dwóch cech (*populacja dwuwymiarowa*) lub wielu cech (*populacja wielowymiarowa*). Na przykład, populacja roślin żyta może być rozważana ze względu na ciężar kłosa (populacja jednowymiarowa), ze względu na ciężar kłosa i liczbę ziarniaków w kłosie (populacja dwuwymiarowa), ze względu na ciężar kłosa, liczbę ziarniaków w kłosie, zawartość skrobi, zawartość białka, zawartość tłuszczu, itp. (populacja wielowymiarowa).

Populacja jest nieskończona (nieskończenie wielka), kiedy składa się z nieskończonej liczby elementów, np. populacja złożona z wyników wszystkich możliwych pomiarów temperatury powietrza, populacja składająca się z wag wszystkich myszy, które żyły, żyją i żyć będą, itp. *Populacja jest skończona*, gdy składa się z pewnej skończonej (określonej) liczby elementów, np. populacja uczniów w danym mieście, populacja roślin żyta na danym polu, itp. Populacja skończona może być traktowana jako nieskończona, gdy liczba składających się na nią elementów jest bardzo duża.

Liczebność populacji oznacza się zwykle literą N .

Próba

– prosta próba losowa

Jak już wspomniano, ze względów praktycznych często nie jesteśmy w stanie przebadać całej populacji, dlatego zadowolamy się jej częścią, czyli *próbą*. Próbę stanowi więc każdy skończony zbiór (a ściślej mówiąc – ciąg) danych liczbowych (wyników), uzyskanych na drodze obserwacji lub pomiarów jednostek statystycznych, pobranych z danej populacji. Aby można było stosować metody statystycznego wnioskowania o populacji na podstawie próby, próba ta musi *reprezentować* populację, z której została pobrana. Znaczy to, że struktura próby (czyli \rightarrow rozkład badanej cechy w próbie) musi możliwie dokładnie odpowiadać strukturze populacji. Aby warunek ten został spełniony, próba musi być pobrana w sposób, który zapewni jej tę reprezentatywność. Jest tylko jeden sposób na to, aby otrzymać *próbę reprezentatywną* – *losowanie*³.

³ Stosując dobór nielosowy (celowy) również można uzyskać próbę reprezentatywną, lecz w przypadku takiej próby wnioskowanie statystyczne nie jest możliwe.

W ten sposób każda jednostka statystyczna populacji ma jednakową szansę (jednakowe \rightarrow prawdopodobieństwo) dostania się do próby. Reasumując, próba jest to liczba obserwacji lub pomiarów wykonanych na wylosowanych elementach populacji. Losowanie elementu populacji nazywamy **doświadczeniem losowym**. W wyniku n -krotnej \rightarrow replikacji doświadczenia losowego otrzymujemy n obserwacji lub pomiarów interesującej nas cechy. Literą n oznaczamy więc liczebność próby. Badanie statystyczne w oparciu o próbę wylosowaną z populacji skończonej przyjęło się nazywać **metodą reprezentacyjną**.

Istnieje wiele **schematów losowania próby**. Ich wyczerpujące omówienie można znaleźć np. u Zasępy (1962), Perkała (1967), Pawłowskiego (1972), Góralskiego (1976), Puchalskiego (1980), Barneta (1982). Tutaj omówimy najprostszy z nich – schemat prostej próby losowej. W oparciu o ten schemat wypracowana została przeważająca liczba metod statystyki matematycznej, będących przedmiotem niniejszego opracowania.

Prosta próba losowa uzyskujemy wówczas, gdy:

- losowanie elementów do próby dokonywane jest z całej populacji;
- każdy z elementów ma jednakowe prawdopodobieństwo dostania się do próby;
- prawdopodobieństwo wylosowania każdego elementu w każdym kolejnym losowaniu jest niezmiennione (stałe).

W celu uzyskania prostej próby losowej stosuje się schemat losowania określanego jako **losowanie niezależne** (zwrotne, ze zwracaniem, z powtórzeniami). Polega ono na tym, że wylosowany element, po zapisaniu wyniku, jest włączany do danej populacji i ponownie bierze udział w losowaniu. W ten sposób ten sam element może być wylosowany kilka razy. Jeżeli wylosowanego elementu nie włącza się z powrotem do populacji, to mamy do czynienia z **losowaniem zależnym** (bezzwrotnym, bez zwracania, bez powtórzeń). W związku z tym, że wylosowany element nie bierze udziału w kolejnych losowaniach, może on być wylosowany tylko jeden raz. W efekcie, prawdopodobieństwo wylosowania poszczególnych elementów do próby zmienia się w czasie trwania losowania.

Jeżeli populacja nie jest liczna, lepiej przeprowadzać losowanie zależne, gwarantuje ono bowiem uzyskanie bardziej reprezentatywnej próby. Jeżeli populacja jest liczna, losowanie zależne i niezależne daje

praktycznie te same wyniki, przy czym różnica między tymi dwoma sposobami losowania dla dużych populacji jest tym mniejsza, im liczniejsza jest badana populacja.

W praktyce, w celu uzyskania prostej próby losowej, zaleca się stosowanie schematu losowania zależnego (bezzwrotnego), ponieważ jest on technicznie prostszy. Musi być tylko spełniony warunek polegający na tym, że każdy pozostający w populacji element w każdym momencie losowania ma to samo prawdopodobieństwo wylosowania. Ponieważ zwykle zajmujemy się populacjami skończonymi ale dużymi, liczącymi wiele tysięcy elementów, a próba stanowi niewielki odsetek populacji generalnej, warunek ten bywa spełniany.

Losowanie elementów populacji do próby stosunkowo łatwo jest przeprowadzić, gdy mamy do czynienia z populacją skończoną, złożoną z niezbyt dużej liczby elementów. W takim przypadku wszystkie elementy tej populacji możemy ponumerować, a następnie numery te (a tym samym – elementy populacji) losować, korzystając na przykład z tablic liczb losowych. Gdy populacja liczy wiele tysięcy elementów lub jest nieskończona, wówczas sposób ten, ze względów praktycznych, jest niewykonalny. W takim razie należy zastosować inny sposób losowania. Przypadek taki szczegółowo omawia np. Perkal (1967).

Niejednokrotnie w podręcznikach statystyki spotykamy się z określeniem *próba mała* – *próba duża*. Nie istnieje kryterium, które pozwoliłoby stwierdzić, przy jakiej liczebności dana próba jest próbą dużą. Zależy to od rozważanych procedur. Na ogół przyjmuje się, że w przypadku szacowania (\rightarrow estymacja) \rightarrow średniej arytmetycznej populacji (\rightarrow parametr) za pomocą średniej arytmetycznej próby (\rightarrow statystyka), próba o liczebności większej niż 30 elementów jest próbą dużą. W przypadku \rightarrow wariancji, \rightarrow mediany czy \rightarrow współczynnika zmienności, dopiero próba o liczebności równej lub większej niż 100 elementów jest próbą dużą.

Charakterystyki próby i populacji

– statystyki, parametry, estymatory

Każdą próbę oraz każdą populację można scharakteryzować za pomocą pewnych charakterystyk liczbowych. W przypadku próby, charaktery-

styki te nazywamy *statystykami*, w przypadku populacji – *parametrami*. Statystyki oznacza się innymi symbolami niż parametry (patrz Tab. 3). Wartość liczbowa parametrów na ogół jest nieznaną. Możemy je → estymować (oceniać, szacować) za pomocą statystyk. Statystyki służące do estymacji wartości liczbowej parametrów nazywamy *estymatorami*.

Parametry są wielkościami stałymi, ale tylko dla krótkich odcinków czasu. Natomiast statystyki zmieniają się dla kolejnych prób pobieranych jednocześnie. Jeżeli na przykład wylosujemy kolejno pięć prób po 10 uczniów z klasy czwartej i dla każdej próby obliczymy ich średni wzrost, to okaże się, że wartość tej statystyki będzie zmieniać się od próby do próby. Dla danej próby wartość statystyki jest znana lub może być wyznaczona. Ponieważ nie wiemy jednak, jak dalece ta próba jest reprezentatywna dla populacji, z której została pobrana, nie wiemy też, jak dokładnie wyznaczona statystyka przybliży poszukiwany, nieznaną parametr.

Losowy wybór próby daje gwarancję, że statystyki próby nie różnią się *istotnie* (tzn. nie różnią się więcej niż to wynika z przypadku) od parametrów populacji, z której została pobrana ta próba. Najczęściej stosowane statystyki, określane mianem *miar statystycznych*, zostaną omówione niżej.

Miary statystyczne

W badaniach statystycznych wykorzystuje się następujące *miary statystyczne*: *miary położenia* i *rozproszenia*, gdy przedmiotem rozważania są populacje jednocechowe (jednowymiarowe) oraz *miary korelacji*, w przypadku populacji dwu- i wielo cechowych. W zależności od stosowanej skali pomiaru, stosuje się odpowiednie miary statystyczne (Tab. 2). Symboliczne oznaczenia niektórych miar przedstawiono w Tab. 3.

Różne techniki obliczania miar statystycznych szczegółowo omawia np. Marszałkiewicz (1980), ale informacje na ten temat można znaleźć w każdym podręczniku statystyki. Z tego powodu, a także ze względu na to, że różne sposoby obliczania statystyk próby integralnie związane są z poszczególnymi testami statystycznymi i, w związku z tym, zawsze są uwzględniane w algorytmie danego testu, metody te nie będą tutaj szczegółowo przedstawiane.

Miary położenia

Miary położenia (miary tendencji centralnej, przeciętne) (por. Tab. 2 i 3) wskazują miejsce, w którym leży wartość najlepiej reprezentująca wszystkie elementy próby (miary klasyczne) lub informują o przeciętnym poziomie wartości lub kategorii rozważanej cechy w badanej próbie (miary pozycyjne).

Tab. 2. Najczęściej stosowane miary statystyczne, charakterystyczne dla różnych skal pomiaru.

Skala	Miary		
	położenia	rozproszczenia	korelacji
nominalna	kategoria modalna	dyspersja względna klasyfikacji	współczynniki siły związku
porządkowa	obiekt medialny	obiekty kwartylowe	współczynniki korelacji rang, współczynnik konkordancji
przedziałowa	klasyczne średnia arytmetyczna, średnia arytmetyczna ważona, pozycyjne wartość modalna, wartość medialna	bezwzględne wariancja, odchylenie standardowe	stosunek korelacyjny, współczynnik korelacji liniowej, współczynnik korelacji cząstkowej i wielorakiej
ilorazowa	klasyczne średnia geometryczna, średnia harmoniczna	względne współczynnik zmienności	—

– kategoria modalna, wartość modalna

Kategoria modalna (kategoria najczęstsza, moda, dominanta) to ta kategoria (wariant) cechy, do której należy najwięcej obserwacji. Natomiast **wartość modalna**⁴ (średnia modalna, wartość najczęstsza, wartość typowa, dominanta, moda) jest tą wartością badanej cechy, której odpowiada maksimum → gęstości prawdopodobieństwa. Kategoria modalna może być wyrażona albo w postaci liczb bezwzględnych albo

⁴ Oprócz wartości modalnej i wartości medialnej (por. Tab. 2) do pozycyjnych miar położenia należą również: *kwartyle* (dzielią próbę danych na cztery części), *kwintyle* (na pięć części), *decyle* (na dziesięć części) oraz *centyle* lub *percentyle* (na sto części).

Tab. 3. Zestawienie niektórych miar statystycznych i ich oznaczeń (wg Góralskiego 1976, nieco zmienione).

Nazwa miary	Oznaczenie	
	dla próby (statystyka)	dla populacji (parametr)
<u>Miary położenia</u>		
kategoria modalna	p_m	ζ_m
wartość modalna	mo	μ_o
wartość medialna	me	μ_E
średnia arytmetyczna	\bar{x}	μ
<u>Miary rozproszenia</u>		
dyspersja względna klasyfikacji	h	χ
wariancja	s^2	σ^2
odchylenie standardowe	s	σ
<u>Miary korelacji</u>		
współczynnik siły związku	r_p	ρ_p
współczynnik korelacji rang Spearmana	r_s	ρ_s
współczynnik korelacji rang Kendalla	$r_k(\tau)$	ρ_k
współczynnik konkordancji	r_w	ρ_w
stosunek korelacyjny	e^2	η^2
współczynnik korelacji liniowej	r	ρ

w postaci liczb względnych (\rightarrow proporcji, frakcji, procentu). Wartość modalna powinna być obliczana wówczas, gdy mamy do czynienia z \rightarrow rozkładem asymetrycznym (patrz Rys. 5), czyli w przypadkach, gdy badana próba danych zawiera bardzo duże lub bardzo małe wartości skrajne oraz wtedy, gdy skrajne \rightarrow przedziały klasowe są niedomknięte i nieznane są dokładne wartości jednostek, które znajdują się w tych przedziałach.

– obiekt medialny

Obiekt medialny to ten obiekt (element) próby, który (ze względu na daną cechę porządkującą) dzieli rozpatrywaną próbę na dwa podzbiory o równej liczebności. Obiektem medialnym jest więc element próby o określonej randze. Wyznaczanie obiektu medialnego dla uporządkowania zupełnego i słabego, dla parzystej i nieparzystej liczby elementów próby szczegółowo omawia Góralski (1976).

– **wartość medialna**

Wartość medialna⁵ (wartość środkowa, mediana), to ta wartość cechy, która dzieli uporządkowany, według wartości rosnących, zbiór danych (próbę) na dwie równe liczebnie części: na elementy mniejsze i większe od mediany. Mediana powinna być obliczana w takich samych przypadkach, jak moda (→ rozkład asymetryczny, Rys. 5).

– **średnia arytmetyczna**

Średnia arytmetyczna to iloraz sumy wyników w próbie podzielony przez liczbę tych wyników.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

W przeciwieństwie do mody i mediany, średniej arytmetycznej **nie można** obliczać dla → rozkładów skrajnie asymetrycznych, → bimodalnych i → wielomodalnych (patrz Rys. 5), ponieważ jest ona wrażliwa na wartości skrajne, duże lub małe.

– **średnia arytmetyczna ważona**

Średnią arytmetyczną ważoną oblicza się wówczas, gdy jest obliczonych kilka średnich arytmetycznych, każda dla grupy (podpróby) danych o innej liczebności. W takiej sytuacji nie wolno obliczać średniej dla całej próby poprzez proste zsumowanie wartości średnich grupowych, lecz należy to zrobić uwzględniając liczebności tych grup.

– **średnia geometryczna**

Średnią geometryczną stosuje się wtedy, gdy w próbie znajdują się wartości skrajne – bardzo duże lub bardzo małe. Stosowana bywa do określania średniego tempa przyrostu, np. masy ciała, wymiarów liniowych, wzrostu populacji.

– **średnia harmoniczna**

Średnia harmoniczna liczona jest wówczas, gdy wartości cechy podane są w postaci odwrotności, np. $1/x_1, 1/x_2, \dots, 1/x_n$, gdzie x oznacza

⁵ Porównaj przypis 4.

wartość cechy. Średnia ta służy do wyliczania przeciętnych wskaźników, np. kształtu, dominacji, zagęszczenia.

Miary rozproszenia

Miary rozproszenia (rozrzutu, dyspersji, zmienności) (por. Tab. 2 i 3) charakteryzują stopień rozproszenia danych wokół miar położenia.

– dyspersja względna klasyfikacji

W przypadku danych wyrażonych za pomocą skal nominalnych (cechy jakościowe) i zgrupowanych w tabeli wielopolowej, w której kolumny stanowią kategorie jednej cechy, a rzędy kategorie drugiej cechy, za miarę rozproszenia przyjmuje się czasami **entropię** rozkładu danych (H). Z powodu braku informacji o rozkładzie miary H , wnioskowanie statystyczne z wykorzystaniem tej miary nie jest możliwe. Z tego powodu Góralski (1976) zaproponował inną miarę rozproszenia, tzw. **dyspersję względną klasyfikacji** (h). Miara ta przyjmuje wartości z przedziału $0 \leq h \leq 1$. Gdy $h = 0$, znaczy to, że wszystkie wyniki próby zgrupowane są w jednym polu tabeli wielopolowej. Gdy $h = 1$, oznacza to, że we wszystkich polach tabeli występuje jednakowa liczba elementów próby.

– obiekty kwartyłowe

Obiekty kwartyłowe (kwartyle), to takie obiekty (elementy) próby, które łącznie z obiektem medialnym dzielą rozpatrywaną próbę na cztery części o równej liczebności. Obiekty kwartyłowe wyznaczone są jako obiekty medialne każdej z dwu części próby, powstałych w efekcie wyznaczenia obiektu medialnego tej próby. Sposób wyznaczania obiektów kwartyłowych omawia szczegółowo Góralski (1976).

– wariancja

Wariancja (s^2) dostarcza informacji na temat rozproszenia danych wokół średniej arytmetycznej. Im rozrzut danych wokół średniej jest większy, tym wartość wariancji jest większa. Aby obliczyć wariancję w próbie, należy kolejne wartości pomiarów, stanowiące daną próbę, odjąć od ich średniej arytmetycznej, każdą z uzyskanych różnic podnieść do kwadratu, następnie zsumować a otrzymaną w ten sposób sumę podzielić przez $n - 1$ jednostek próby (\rightarrow liczba stopni swobody). Otrzy-

many wynik jest właśnie miarą rozproszenia pomiarów wokół średniej arytmetycznej.

$$s^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}$$

– *odchylenie standardowe, współczynnik zmienności*

Odchylenie standardowe (s) (odchylenie średnie) to pierwiastek kwadratowy z wariancji.

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

Odchylenie standardowe (podobnie jak wariancja) zawsze jest liczbą mianowaną, wyrażoną w tych samych jednostkach miary (np. w centymetrach, metrach, kilogramach, tonach, litrach, godzinach, itp.) co dane próby i średnia arytmetyczna w próbie. Odchylenia standardowego obliczonego np. dla cechy wyrażonej w metrach nie można więc bezpośrednio porównać z odchyleniem standardowym obliczonym dla cechy wyrażonej w kilogramach. Odchylenia standardowe nie są również bezpośrednio porównywalne wówczas, gdy obliczone są wprawdzie dla tej samej cechy (np. wyrażonej w jednostkach wagi), ale istnieje znaczna różnica między średnim poziomem cechy w badanych próbach. Na przykład, nie można porównać odchylenia standardowego średniego ciężaru myszy z odchyleniem standardowym średniego ciężaru słonia. Aby to było możliwe, odchylenie standardowe musi być wyrażone w postaci liczby względnej, niemianowanej. Taką względną miarą dyspersji jest tzw. **współczynnik zmienności** (V). Oblicza się go dzieląc odchylenie standardowe w próbie przez średnią arytmetyczną tej próby. Jeżeli uzyskany iloraz pomnożymy przez 100, wartość współczynnika zmienności będzie wyrażona w procentach.

$$V = \frac{s}{\bar{x}} \times 100\%$$

Miary korelacji

Miary korelacji (por. Tab. 2 i 3) określają siłę współzmienności (współzależności) między badanymi cechami, np. siłę współzmienności między długością i szerokością liścia, między wzrostem a masą ciała, między miejscem zamieszkania a stanem zdrowia, itp. Więcej informa-

cji na temat korelacji oraz stosowanych miar podano gdzie indziej (→ Testowanie niezależności cech, str. 97). Tutaj zostaną podane tylko krótkie charakterystyki stosowanych miar.

– *współczynniki siły związku*

Współczynniki siły związku (zwane też współczynnikami asocjacji lub zbieżności) służą do pomiaru korelacji cech jakościowych (skala nominalna), zgrupowanych w postaci tzw. cztero- lub wielopolowej tabeli kontyngencji. Zaliczane są tutaj takie współczynniki, jak np. współczynnik Yule'a, Czuprowa, Bykowskiego, itd.

– *współczynniki korelacji rang, współczynnik konkordancji*

Do oceny zgodności uporządkowania danych, wyrażonych za pomocą dwóch skal porządkowych, służą **współczynniki korelacji rang** (kolejnościowej): Spearmana lub Kendalla. Jeśli dane porządkowane są za pomocą więcej niż dwóch skal porządkowych, wówczas zgodność uporządkowania ocenia się za pomocą **współczynnika konkordancji**. Współczynniki te pozwalają np. stwierdzić, czy ustalając ranking w danej grupie zawodników poszczególni sędziowie różnią się w swoich ocenach (jeżeli tak, to w jakim stopniu), czy też nie.

– *stosunek korelacyjny*

Stosunek korelacyjny stosowany jest wówczas, gdy mamy do czynienia z → korelacją krzywoliniową oraz gdy obie badane cechy są → mierzalne lub przynajmniej → zmienna zależna jest mierzalna.

– *współczynnik korelacji liniowej r Pearsona*

Współczynnik korelacji liniowej r Pearsona stosowany jest wyłącznie dla dwu cech mierzalnych (skala przedziałowa) oraz w przypadku → korelacji liniowej.

– *współczynnik korelacji wielorakiej, współczynnik korelacji cząstkowej*

Współczynnik korelacji wielorakiej i **współczynnik korelacji cząstkowej** służy do mierzenia siły korelacji w przypadku, gdy mamy do czynienia z więcej niż dwoma cechami wyrażonymi za pomocą skali przedziałowej.

Cechy statystyczne a zmienne losowe

Przypuśćmy, że badaną cechą jest liczba ziarniaków w kłosie żyta. Załóżmy, że z populacji roślin żyta wylosowaliśmy próbę o liczebności $n = 10$ kłosów. Jeżeli teraz policzymy ziarniaki w każdym z tych kłosów, to okaże się, że cecha ta nie jest stała. Jej wartość zmienia się od kłosa do kłosa. Taką zmiennością charakteryzuje się w przyrodzie każda badana cecha. Ponieważ nie możemy przewidzieć, jaką wartość każdorazowo będzie przybierać badana cecha, dlatęgo mówimy, że cecha ta ma charakter **zmiennej losowej**.

Pojęcie zmiennej losowej jest jednym z podstawowych pojęć rachunku prawdopodobieństwa (por. str. 47 i 55). Ogólnie biorąc, zmienna losowa to wielkość, która w wyniku doświadczenia lub obserwacji przyjmuje **jedną** i tylko jedną wartość ze zbioru tych wszystkich wartości, jakie ta zmienna może przyjąć, przy czym wartość ta znana jest dopiero **po** przeprowadzeniu doświadczenia lub obserwacji i wcześniej nie można jej przewidzieć w sposób jednoznaczny. Na przykład zmienna "liczba ziarniaków w kłosie" może przyjąć wartość równą 0, 1, 2, ..., k ziarniaków w kłosie. Wartości 0, 1, 2, ..., k stanowią zbiór możliwych wartości tej zmiennej. Wartości, jakie może przyjąć zmienna losowa nazywamy **realizacjami zmiennej losowej**.

Wyróżnia się dwa rodzaje zmiennych losowych: skokową i ciągłą. O **zmiennej losowej skokowej** (dyskretnej, nieciągłej) mówimy wówczas, gdy przybiera ona wartości całkowite: 0, 3, 7, 11, 58, itd. Przykładem zmiennej skokowej może być np. liczba ziarniaków w kłosie żyta, liczba krów w stadzie, liczba gatunków roślin na badanej powierzchni, liczba dzieci w rodzinie, itp. W przypadku zmiennej losowej skokowej dane pochodzą więc z **policzenia**. Jest ich na ogół niewiele i można je wszystkie wymienić.

Z **zmienną losową ciągłą** mamy do czynienia wtedy, gdy w wyniku pomiaru otrzymujemy wartość liczbową należącą do pewnego przedziału skali ciągłej. W skali ciągłej mierzymy np. czas, długość, temperaturę, masę, itp. W przypadku zmiennej losowej ciągłej dane uzyskuje się więc przez **mierzenie**.

Należy pamiętać, że każdy pomiar może być wykonywany z różną **dokładnością**. W efekcie każdy pomiar może dać inny wynik. Fakt ten sprawia, że liczba możliwych wyników (realizacji zmiennej losowej

ciągłej) jest nieskończona i w związku z tym nie można ich wszystkich wyliczyć. Jest to łatwo zrozumieć, jeżeli wyobrazimy sobie, na przykład, że pomiary danej cechy wykonywane są z dokładnością do jednej miliardowej centymetra. W ten sposób na każdym centymetrze może wystąpić miliard realizacji zmiennej losowej ciągłej.

Odpowiednikami zmiennych losowych są w statystyce *cechy statystyczne*. Cecha statystyczna to taka właściwość jednostki statystycznej (elementu populacji), która podlega badaniu statystycznemu. Wyróżnia się cechy statystyczne mierzalne i niemierzalne. *Cecha mierzalna* to taka cecha, która da się wyrazić za pomocą liczby pochodzącej z polizczenia (cecha skokowa) lub z pomiaru (cecha ciągła). Cechy mierzalne zaliczamy do *cech ilościowych*. *Cecha niemierzalna* to taka cecha, której nie da się wyrazić za pomocą liczby, natomiast da się ją wyrazić słownie. Cechą niemierzalną jest np. kolor włosów, płeć, zawód, smak, itp. Cechy niemierzalne zaliczamy do *cech jakościowych*.

Błąd pomiaru

Wykonując pomiary w skali ciągłej nie należy zapominać, że zawsze otrzymujemy wyniki obarczone określonym *błędem*. Powstawanie tych błędów jest nieuniknione. Wyróżnia się *błędy systematyczne*, wynikające z niedokładności przyrządów pomiarowych; *błędy grube*, powstające na przykład w wyniku złego odczytania wskazania przyrządu lub nieprawidłowego zanotowania wyniku, oraz *błędy przypadkowe*. W przeciwieństwie do dwu pierwszych, ta ostatnia grupa błędów jest nie do uniknięcia. Można jedynie próbować zmniejszyć błędy przypadkowe przez wielokrotne powtarzanie pomiarów.

Rozkłady empiryczne a teoretyczne

Podstawowym celem statystyki jest wykrywanie prawidłowości w materiale statystycznym, ponieważ prawidłowość pozwala nie tylko opisać, w sposób zwięzły i jednoznaczny, badaną próbę, ale również pozwala wnioskować o populacji na podstawie tej próby. Wykrywanie prawidłowości w próbie sprowadza się do badania tzw. *rozkładu empirycznego* badanej cechy statystycznej w tej próbie. Ogólnie mówiąc, staramy się stwierdzić, jaki jest obszar zmienności rozważanej cechy w próbie i z jaką częstością występują poszczególne warianty tej cechy.

Rozkład empiryczny cechy w próbie możemy następnie porównać z jednym z rozkładów modelowych, tzw. **rozkładów teoretycznych**. Rozkłady teoretyczne to rozkłady zmiennych losowych, wydedukowane matematycznie w oparciu o pewne ogólne hipotezy. Są one przedmiotem badania rachunku prawdopodobieństwa. Zaliczenie lub nie zaliczenie rozkładu empirycznego do jednego ze znanych rozkładów teoretycznych, pozwala nam zdecydować o dalszych etapach analizy statystycznej, dlatego ta faza badania jest tak ważna. Więcej na ten temat powiemy dalej. Obecnie, pokrótce, przedstawiony zostanie sposób konstruowania rozkładu empirycznego dla cechy ciągłej i skokowej.

Wybrane rozkłady teoretyczne zmiennej losowej skokowej i ciągłej zostaną omówione dalej (→ Rozkłady teoretyczne, str. 53).

Konstruowanie rozkładu empirycznego

– szereg statystyczny, szereg rozdzielczy

Wyniki, stanowiące daną próbę, tworzą tzw. **szereg statystyczny**. Szereg ten jest uporządkowany, jeżeli elementy próby zostały uporządkowane za pomocą jakiegoś kryterium, np. relacji $>$ (większy niż), $=$ (równy), $<$ (mniejszy niż), w przeciwnym wypadku mamy do czynienia z szeregiem nieuporządkowanym. W szeregu uporządkowanym wymienione są wszystkie wyniki. Jeżeli dysponujemy bardzo licznym materiałem statystycznym, szereg taki jest nieczytelny – trudno dostrzec w nim jakąś → prawidłowość statystyczną, wyrażającą się np. w określonym rozkładzie liczebności występowania poszczególnych wariantów badanej cechy.

Jak wiemy z doświadczenia, niektóre realizacje zmiennej losowej mogą występować częściej niż inne, np. więcej jest ludzi średniego wzrostu niż osobników bardzo niskich lub bardzo wysokich. Z powodu, o którym wspomniano wyżej, szereg statystyczny przekształcamy najczęściej w tzw. **szereg rozdzielczy** lub, mówiąc inaczej, w **rozkład liczebności** (częstości bezwzględnej, **frekwencji**). Rozkład taki jest rozkładem empirycznym (zaobserwowanym) badanej cechy. Ogólnie biorąc, w przypadku cechy jakościowej i skokowej, szereg rozdzielczy buduje się w oparciu o poszczególne warianty tej cechy. To znaczy, wyliczamy wszystkie warianty cechy (np. biały, żółty, czarny, czerwony).

ny – dla cechy jakościowej lub 0, 1, 2, ..., k – dla cechy skokowej), porządkujemy je, jeśli to możliwe, według natężenia, a następnie notujemy częstość realizacji poszczególnych wariantów tej cechy.

– przedział klasowy

W przypadku cechy ilościowej ciągłej obszar (zakres) zmienności jest zbyt duży, aby postąpić tak, jak ze zmienną skokową. Ponieważ nie można wyliczyć wszystkich wariantów (wartości liczbowych) tej cechy, dlatego szereg rozdzielczy konstruuje się w oparciu o przedziały liczbowe zwane *przedziałami klasowymi* lub *klasami*.

Przedziały klasowe wyznaczane są w ten sposób, aby obejmowały cały obszar zmienności badanej cechy, przy czym liczba klas uzależniona jest od celu badania, charakteru i liczebności próby. Liczba klas nie powinna być ani zbyt mała, ani zbyt duża. Mała liczba klas powoduje, że część szczegółów ginie, natomiast duża zmniejsza efekt uogólnienia. Nie ma jednej reguły, za pomocą której można by, w sposób jednoznaczny, określić liczbę klas w zależności od liczebności próby. Na ogół uważa się, że nie powinno być mniej niż 6 i więcej niż 15 klas. Przy ustalaniu liczby przedziałów klasowych może być pomocna na przykład Tab. 4. Konstruowanie szeregów rozdzielczych szczegółowo omawia Steczkowski (1970) i tam należy szukać więcej informacji na ten temat.

Tab. 4. Maksymalna liczba przedziałów klasowych szeregu rozdzielczego w zależności od liczebności próby (n) (wg Góralskiego 1976).

n	10	20	50	100	200	500	1000
Liczba przedziałów	6	7	8	9	10	11	13

Każdy przedział klasowy określany jest za pomocą *dolnej* i *górnej granicy*. Granice przedziałów mogą się pokrywać (np. 10–20, 20–30, 30–40) lub nie (10–19, 20–29, 30–39). W przypadku gdy granice przedziałów pokrywają się i wyznaczane są z dokładnością do jedności, a pomiary oryginalne wykonywane są również z taką dokładnością, wówczas, zgodnie z obowiązującą zasadą, wynik pomiaru równy górnej granicy przedziału zaliczany jest do przedziału następnego, np. liczba 30 zaliczona zostanie do przedziału 30–40, a nie do przedziału

20–30. Gdy granice przedziałów wyznaczone są z dokładnością do 0.5 jednostki (np. 10.5–20.5, 20.5–30.5, 30.5–40.5), a oryginalne pomiary wykonywane są z dokładnością do jednośc, wówczas liczba 20, której dokładność pomiaru wyznacza zakres 19.5–20.5, zostanie zaliczona do przedziału 10.5–20.5, natomiast liczba 30, której zakresem jest 29.5–30.5, zaliczona zostanie do przedziału 20.5–30.5. Gdy przedziały granic nie pokrywają się, wtedy nie ma kłopotów z przydzielaniem poszczególnych wyników do odpowiednich przedziałów.

Uwaga: jeżeli najmniejsza i/lub największa wartość pomiaru jest nieznaną albo też są to wartości ekstremalne, wówczas przedziały pierwszy i ostatni są *otwarte* (niezamykane), tzn. brak jest dolnej granicy przedziału pierwszego i górnej granicy przedziału ostatniego.

Suma granic przedziału podzielona przez 2 nazywa się *środkiem przedziału klasowego* (środkiem klasy, wariantem klasowym).

Sposób konstruowania rozkładu liczebności dla cechy skokowej i ciągłej zilustrowano niżej.

Rozkład empiryczny cechy ciągłej

– *histogram, wielobok liczebności, krzywa liczebności, krzywa liczebności skumulowanych*

*Przykład 1. Zmierzono wysokość $n = 100$ źdźbeł żyta. Pomiary wykonano z dokładnością do jednośc. Wyniki przedstawiono w postaci szeregu rozdzielczego w Tab. 5 (dane fikcyjne). Jego graficznym obrazem może być *histogram* lub *wielobok liczebności* (diagram) (Rys. 1).*

Jak widać na Rys. 1, linia wieloboku liczebności jest łamana⁶. Łączy ona środki przedziałów klasowych. Przy dostatecznie wąskich przedziałach będziemy mieć wrażenie, że rozkład liczebności zmienia się w sposób ciągły. Możemy wówczas aproksymować (przybliżać) go za pomocą linii krzywej – tzw. *krzywej liczebności* (Rys. 2). Im przedziały klasowe będą węższe, tym aproksymacja za pomocą tej krzywej będzie dokładniejsza. Oczywiście, aby uzyskać taki obraz jak na Rys. 2, próba musi być bardzo liczna, musi liczyć wiele setek elementów.

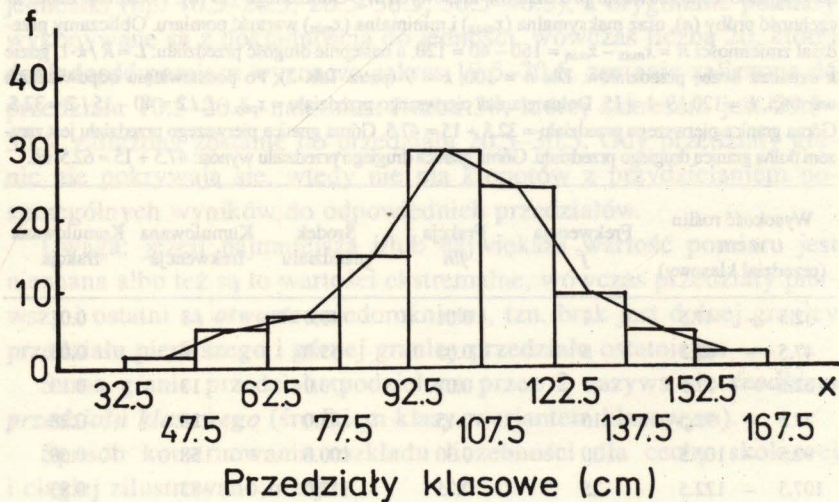
⁶ Jeżeli dane wyrażone są w postaci częstości względnych (frakcji, proporcji), wówczas otrzymujemy, odpowiednio, histogram lub wielobok częstości (frakcji).

Tab. 5. Niekumulowane i kumulowane rozkłady frekwencji i frakcji dla *Przykładu 1*. Poniżej podano sposób konstruowania przedziałów klasowych wg Góralskiego (1976). Potrzebne dane: liczebność próby (n), oraz maksymalna (x_{max}) i minimalna (x_{min}) wartość pomiaru. Obliczamy przedział zmienności $R = x_{max} - x_{min} = 160 - 40 = 120$, a następnie długość przedziału: $L = R / k - 1$, gdzie k oznacza liczbę przedziałów. Dla $n = 100$, $k = 9$ (patrz Tab. 2). Po podstawieniu odpowiednich wartości, $L = 120 / 9 - 1 = 15$. Dolna granica pierwszego przedziału $= x_{min} - L / 2 = 40 - 15 / 2 = 32.5$. Górna granica pierwszego przedziału $= 32.5 + 15 = 47.5$. Górna granica pierwszego przedziału jest zarazem dolną granicą drugiego przedziału. Górna granica drugiego przedziału wynosi: $47.5 + 15 = 62.5$, itd.

Wysokość roślin w cm (przedział klasowy)	Frekwencja f	Frakcja f/n	Środek przedziału	Kumulowana frekwencja	Kumulowana frakcja
32.5 – 47.5	1	0.01	40.0	1	0.01
47.5 – 62.5	5	0.05	55.0	6	0.06
62.5 – 77.5	7	0.07	70.0	13	0.13
77.5 – 92.5	15	0.15	85.0	28	0.28
92.5 – 107.5	30	0.30	100.0	58	0.58
107.5 – 122.5	25	0.25	115.0	83	0.83
122.5 – 137.5	10	0.10	130.0	93	0.93
137.5 – 152.5	5	0.05	145.0	98	0.98
152.5 – 167.5	2	0.02	160.0	100	1.00
Razem	$n = 100$	1.00	–	–	–

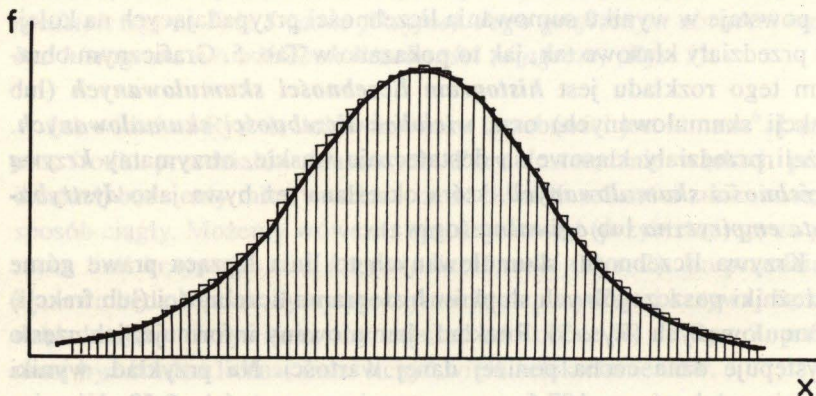
Rozkład liczebności, o którym mowa w *Przykładzie 1*, można również przedstawić w postaci **rozkładu liczebności skumulowanych**, który powstaje w wyniku sumowania liczebności przypadających na kolejne przedziały klasowe tak, jak to pokazano w Tab. 5. Graficznym obrazem tego rozkładu jest **histogram liczebności skumulowanych** (lub frakcji skumulowanych) oraz **wielobok liczebności skumulowanych**. Jeżeli przedziały klasowe są dostatecznie wąskie, otrzymamy **krzywą liczebności skumulowanych**, która określana też bywa jako **dystrybucja empiryczna** lub **ogiwa** (ogiwa).

Krzywa liczebności skumulowanych to linia łącząca prawe górne narożniki poszczególnych słupków histogramu liczebności (lub frakcji) skumulowanych (Rys. 3). Rozkład skumulowany informuje, jak często występuje dana cecha poniżej danej wartości. Na przykład, wyniki mniejsze lub równe 107.5 cm występują z częstością 0.58. Ujmując rzecz w procentach można powiedzieć, że wyników tych jest 58% (Rys. 3).

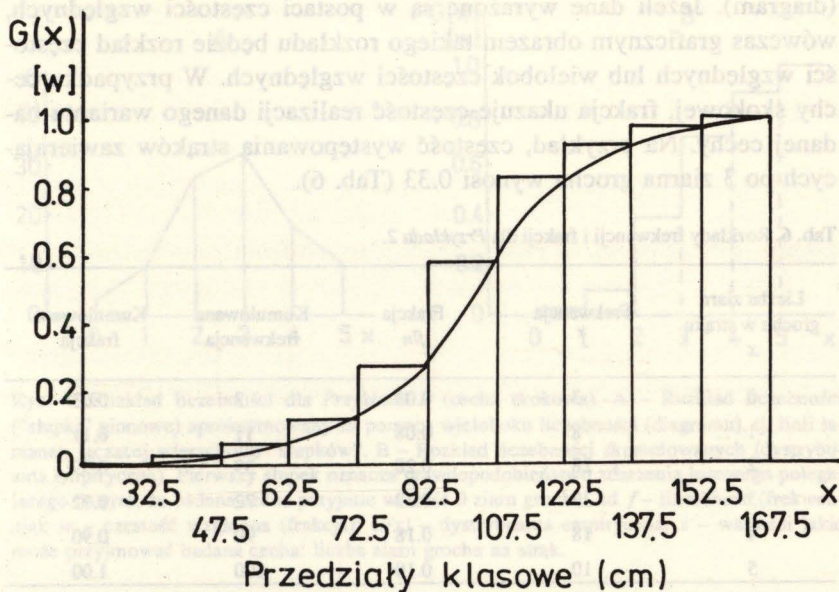


Rys. 1. Histogram liczebności dla *Przykładu 1* (cecha ciągła). Histogram ten aproksymowany jest za pomocą wieloboku liczebności (diagramu), tj. linii łamanej, łączącej środki przedziałów klasowych. f – liczebność (frekwencja).

Uwaga: Granice przedziałów klasowych wyznaczono za pomocą liczb mających o jedną cyfrę znaczącą więcej niż oryginalne pomiary. Sposób wyznaczania granic przedziałów przedstawiono w Tab. 5.



Rys. 2. Histogram liczebności aproksymowany za pomocą linii ciągłej, tzw. krzywej liczebności. f – liczebność (frekwencja).



Rys. 3. Histogram frakcji skumulowanych dla *Przykładu 1* (cecha ciągła). Histogram ten aproksymowany jest za pomocą krzywej liczebności skumulowanych (dystrybuanty empirycznej), tj. linii ciągłej, łączącej prawe górne narożniki poszczególnych słupków histogramu. w – częstość względna (frakcja); $G(x)$ – dystrybuanta empiryczna; x – wartości jakie może przyjmować badana cecha: wysokość roślin (cm).

Rozkład empiryczny cechy skokowej

– *histogram, wielobok liczebności, krzywa liczebności, wielobok liczebności skumulowanych*

Przykład 2. Przebadano $n = 100$ strąków grochu ze względu na liczbę ziarn grochu w strąku. Wyniki zestawiono w Tab. 6 (dane fikcyjne). Ich graficzny obraz przedstawiono na Rys. 4.

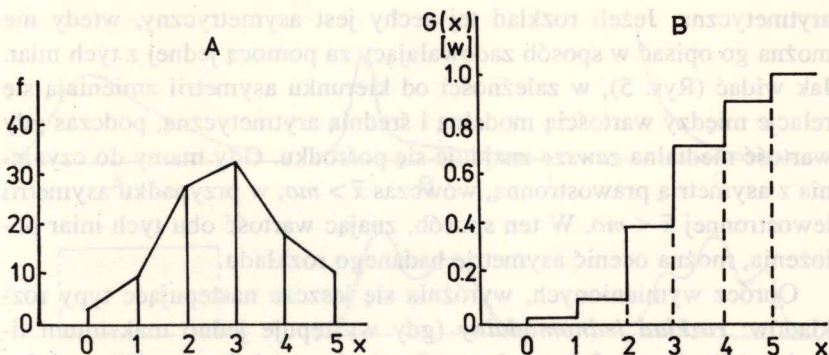
Jak widać w tym przypadku, szereg rozdzielczy zbudowany jest tak, że w każdej z wyróżnionych klas występuje tylko jeden wariant zmiennej X : “liczba ziarn grochu w strąku”, a mianowicie: $x = 0, 1, 2, 3, 4$, lub 5 ziarn grochu. W związku z tym, szerokość słupka *histogramu* zredukowana jest do zera. Dlatego na Rys. 4 zaobserwowane frekwencje przedstawiono w postaci linii pionowych. Jeżeli wierzchołki tych “słupków” połączymy linią łamaną, to otrzymamy *wielobok liczebności*

(diagram). Jeżeli dane wyrażone są w postaci częstości względnych, wówczas graficznym obrazem takiego rozkładu będzie rozkład częstości względnych lub wielobok częstości względnych. W przypadku cechy skokowej, frakcja ukazuje częstość realizacji danego wariantu badanej cechy. Na przykład, częstość występowania strąków zawierających po 3 ziarna grochu wynosi 0.33 (Tab. 6).

Tab. 6. Rozkłady frekwencji i frakcji dla Przykładu 2.

Liczba ziarn grochu w strąku x	Frekwencja f	Frakcja f/n	Kumulowana frekwencja	Kumulowana frakcja
0	3	0.03	3	0.03
1	8	0.08	11	0.11
2	28	0.28	39	0.39
3	33	0.33	72	0.72
4	18	0.18	90	0.90
5	10	0.10	100	1.00
Razem	$n = 100$	1.00	—	—

W rozpatrywanym przypadku, generalizowanie histogramu liczebności za pomocą *krzywej liczebności* nie ma sensu, gdyż nie mogą istnieć strąki posiadające niecałkowitą liczbę ziarn grochu. Załóżmy jednak, że dysponujemy szeregiem statystycznym, w którym da się wyróżnić kilkadziesiąt wariantów badanej cechy. Wówczas szereg rozdzielczy możemy zbudować w ten sposób, że w klasie umieścimy kilka możliwych wariantów tej cechy, np. 0–4, 5–9, 10–14, 15–19, itd. W takim przypadku szerokość słupka histogramu będzie większa niż zero i histogram upodobni się do histogramu dla cechy ciągłej, a wielobok liczebności poprowadzi się, jak poprzednio, przez środki przedziałów (patrz Rys. 1). Jeżeli natomiast będziemy dysponować szeregiem statystycznym, w którym da się wyróżnić wiele tysięcy wariantów badanej cechy, np. więcej niż 100 000, i wydzielimy wiele wąskich przedziałów klasowych, wówczas (w stosunku do liczby wszystkich wariantów) skok o jaki będzie się zmieniać ta cecha od przedziału do przedziału będzie bardzo mały. Taki mały skok możemy przyjąć za przejście ciągłe a dany szereg rozdzielczy można potraktować jak szereg, w którym



Rys. 4. Rozkład liczebności dla Przykładu 2 (cecha skokowa). A – Rozkład liczebności ("słupki" pionowe) aproksymowany za pomocą wieloboku liczebności (diagramu), tj. linii łamanej, łączącej wierzchołki "słupków". B – Rozkład liczebności skumulowanych (dystrybuanta empiryczna). Pierwszy słupek oznacza prawdopodobieństwo zdarzenia losowego polegającego na tym, że badana cecha przyjmie wartość 0 ziarn grochu, itd. f – liczebność (frekwencja); w – częstość względna (frakcja); $G(x)$ – dystrybuanta empiryczna; x – wartości jakie może przyjmować badana cecha: liczba ziarn grochu na strąk.

zmienną jest cecha ciągła i zgeneralizować go za pomocą krzywej liczebności.

Rozkład liczebności, o którym mowa w *Przykładzie 2* można również przedstawić w postaci rozkładu liczebności skumulowanej (Tab. 6). Graficznym obrazem tego rozkładu jest dystrybuanta empiryczna przedstawiona na Rys. 4. W przypadku cech skokowych, wykreślana jest ona jako *wielobok liczebności skumulowanych*.

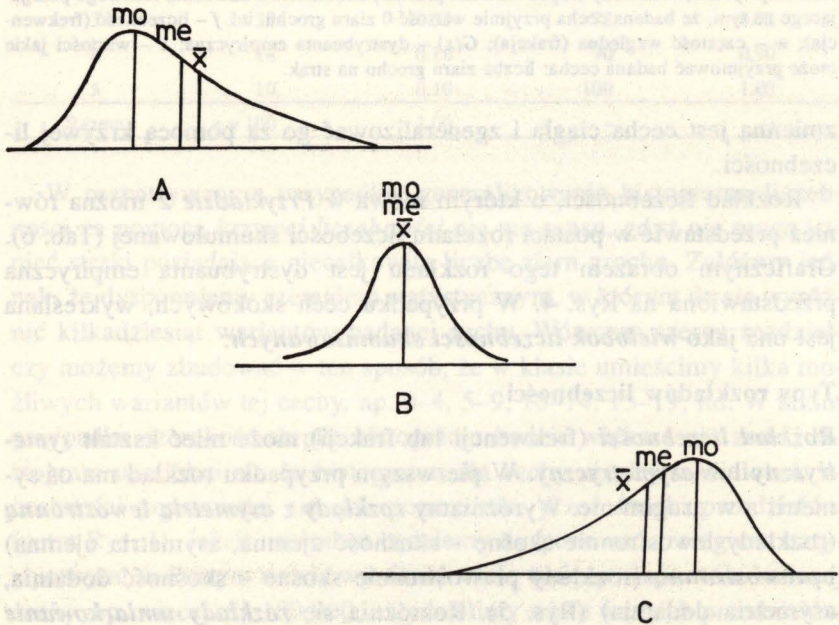
Typy rozkładów liczebności

Rozkład liczebności (frekwencji lub frakcji) może mieć kształt *symetryczny* lub *asymetryczny*. W pierwszym przypadku rozkład ma oś symetrii a w drugim nie. Wyróżniamy *rozkłady z asymetrią lewostronną* (rozkłady lewostronnie skośne – skośność ujemna, asymetria ujemna) i *prawostronną* (rozkłady prawostronnie skośne – skośność dodatnia, asymetria dodatnia) (Rys. 5). Rozróżnia się *rozkłady umiarkowanie i skrajnie asymetryczne* (w kształcie litery J) (Rys. 6).

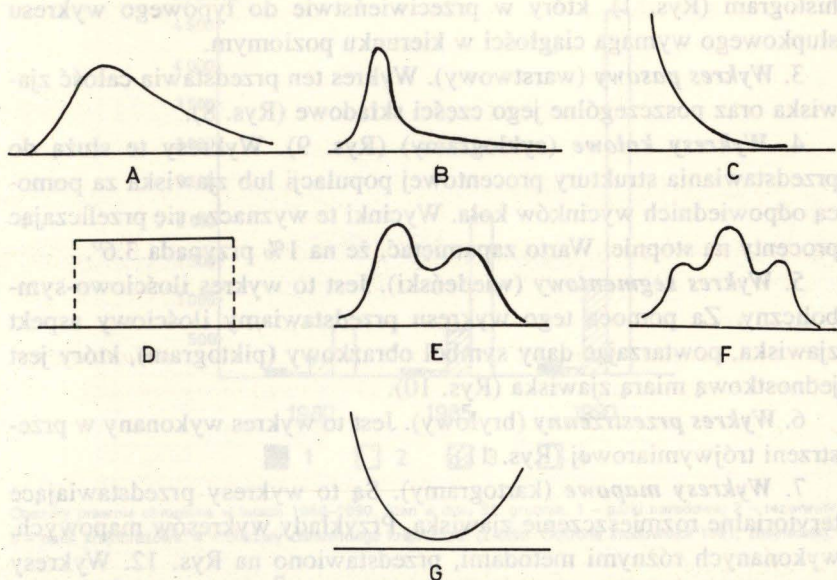
Jeżeli rozkład badanej cechy jest symetryczny, wtedy wartość modalna, wartość medialna i średnia arytmetyczna pokrywają się. W tym wypadku jako jedyną miarę wartości średniej można stosować średnią

arytmetyczną. Jeżeli rozkład tej cechy jest asymetryczny, wtedy nie można go opisać w sposób zadowalający za pomocą jednej z tych miar. Jak widać (Rys. 5), w zależności od kierunku asymetrii zmieniają się relacje między wartością modalną i średnią arytmetyczną, podczas gdy wartość medialna *zawsze* znajduje się pośrodku. Gdy mamy do czynienia z asymetrią prawostronną, wówczas $\bar{x} > mo$, w przypadku asymetrii lewostronnej $\bar{x} < mo$. W ten sposób, znając wartość obu tych miar położenia, można ocenić asymetrię badanego rozkładu.

Oprócz wymienionych, wyróżnia się jeszcze następujące typy rozkładów: *rozkład jednomodalny* (gdy występuje jedno maksimum liczebności), *dwumodalny* (gdy występują dwa maksima), *wielomodalny* (gdy występują więcej niż dwa maksima), *równomierny* (prostokątny, jednostajny – nie posiada maksimum) oraz *antymodalny* (w kształcie



Rys. 5. Położenie wartości modalnej (mo), wartości medialnej (me) i średniej arytmetycznej (\bar{x}) w rozkładach asymetrycznych i w rozkładzie symetrycznym – szczegóły w tekście. A – rozkład prawostronnie asymetryczny (asymetria dodatnia); B – rozkład symetryczny; C – rozkład lewostronnie asymetryczny (asymetria ujemna).



Rys. 6. Typy rozkładów. A – umiarkowanie asymetryczny; B – silnie asymetryczny. C – skrajnie asymetryczny (w kształcie litery J); D – równomierny; A, B, – jednomodalny; E – dwumodalny; F – wielomodalny; G – antymodalny (w kształcie litery U). Rozkład symetryczny pokazano na Rys. 5.

litery U), który zamiast maksimum posiada minimum liczebności (Rys. 6).

Graficzna prezentacja danych

Dane statystyczne można przedstawiać w postaci graficznej za pomocą wykresów. Informacje na ten temat można znaleźć prawie w każdym podręczniku statystyki (patrz np. Zając 1971, Puchalski 1980, Krzysztofiak i Urbanek 1981). Szczegółowo zagadnienie to omawiają Byzow (1951), Osipow (1957), Ziomek (1958) i Krzysztofiak (1971). W tym miejscu wymienimy i zilustrujemy jedynie niektóre rodzaje wykresów.

1. **Wykresy liniowe:** wielobok liczebności, krzywa liczebności, krzywa liczebności skumulowanych (Rys. 1, 2, 3.).
2. **Wykresy słupkowe** (Rys. 7). Odmianą wykresu słupkowego jest

histogram (Rys. 1), który w przeciwieństwie do typowego wykresu słupkowego wymaga ciągłości w kierunku poziomym.

3. **Wykres pasowy** (warstwowy). Wykres ten przedstawia całość zjawiska oraz poszczególne jego części składowe (Rys. 8).

4. **Wykresy kołowe** (cyklogramy) (Rys. 9). Wykresy te służą do przedstawiania struktury procentowej populacji lub zjawiska za pomocą odpowiednich wycinków koła. Wycinki te wyznacza się przeliczając procenty na stopnie. Warto zapamiętać, że na 1% przypada 3.6°.

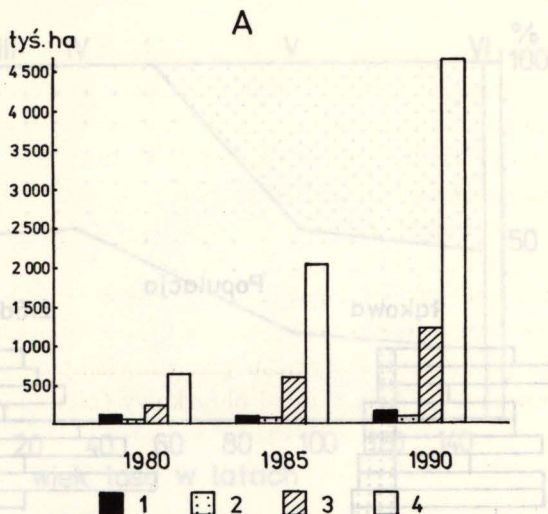
5. **Wykres segmentowy** (wiedeński). Jest to wykres ilościowo-symboliczny. Za pomocą tego wykresu przedstawiamy ilościowy aspekt zjawiska, powtarzając dany symbol obrazkowy (piktogram), który jest jednostkową miarą zjawiska (Rys. 10).

6. **Wykres przestrzenny** (bryłowy). Jest to wykres wykonany w przestrzeni trójwymiarowej (Rys. 11).

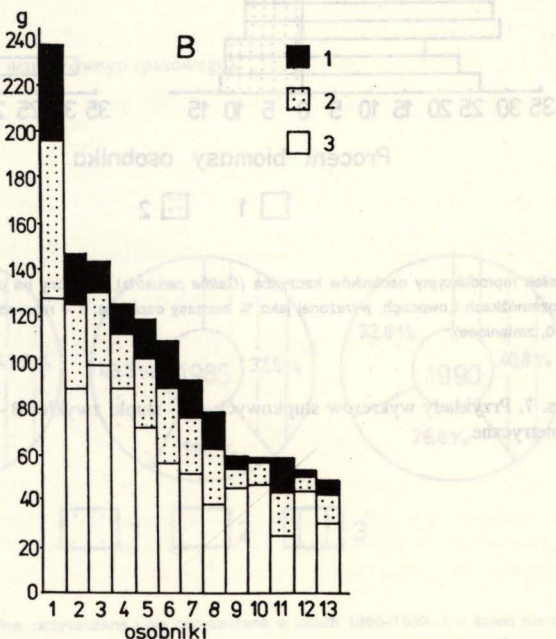
7. **Wykresy mapowe** (kartogramy). Są to wykresy przedstawiające terytorialne rozmieszczenie zjawiska. Przykłady wykresów mapowych, wykonanych różnymi metodami, przedstawiono na Rys. 12. Wykresy mapowe, w odniesieniu do kartografii geobotanicznej, szczegółowo omawia Faliński (1990).

Różne typy wykresów można stosować w postaci "czystej" lub na różne sposoby łączyć ze sobą (patrz np. wykresy mapowe). Nie sposób zaprezentować wszystkich możliwych kombinacji – ich liczba jest praktycznie nieograniczona i zależy tylko od inwencji czytelnika i od celu jaki mu przyświeca. Odpowiednich przykładów prezentacji graficznej danych statystycznych czytelnik powinien szukać w literaturze swojej specjalności.

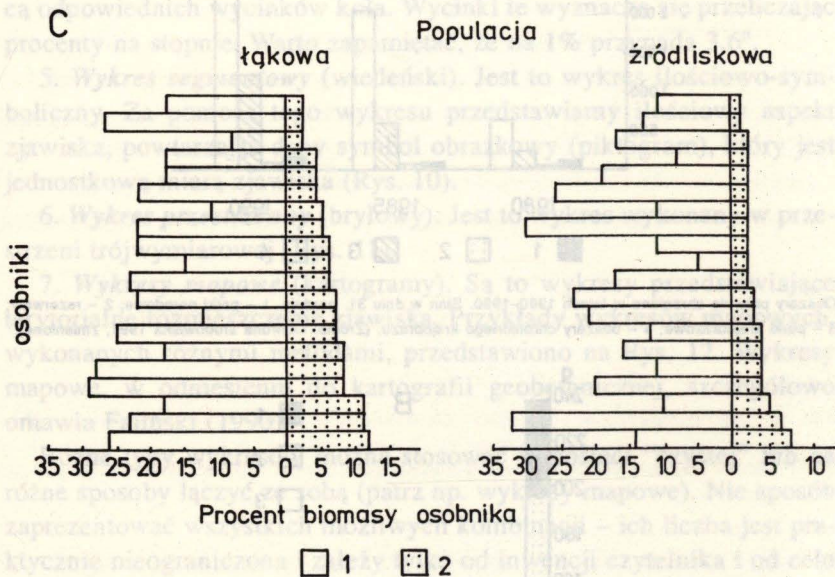
Obecnie dostępne są specjalne komputerowe programy graficzne (np. Fox Graph), które użytkownikowi komputera osobistego mogą ułatwić wybór i wykonanie różnych wersji wykresów liniowych, słupkowych, pasowych, kołowych lub przestrzennych.



Obszary prawnie chronione w latach 1980–1990. Stan w dniu 31. grudnia. 1 – parki narodowe; 2 – rezerwy; 3 – parki krajobrazowe; 4 – obszary chronionego krajobrazu. (Źródło: Ochrona środowiska 1991, zmienione).

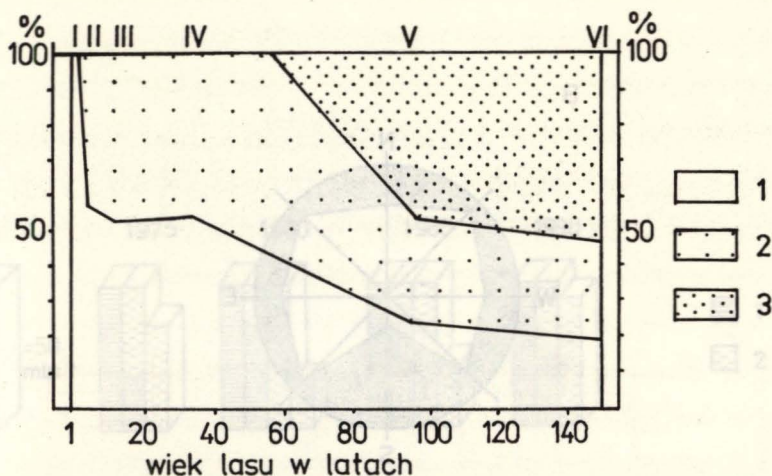


Struktura wielkości (biomasy) osobników krwawnicy pospolitej (*Lythrum salicaria*) w populacji łąkowej. 1 – kwiaty; 2 – pędy; 3 – korzenie. (Źródło: Falińska 1990, zmienione).



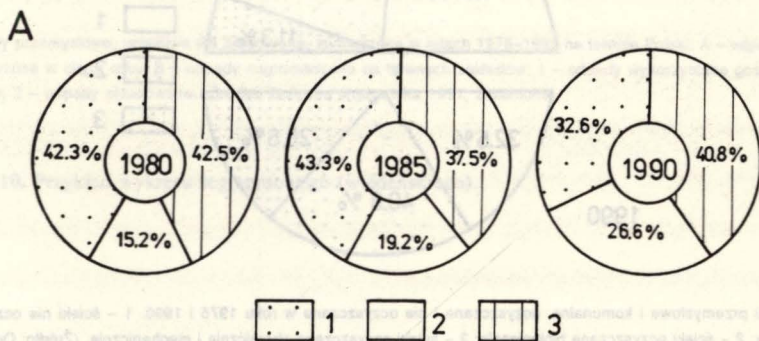
Wysięk reprodukcyjny osobników kaczycica (*Caltha palustris*), oceniony na podstawie biomasy skumulowanej w rozmnóżkach i owocach, wyrażonej jako % biomasy osobnika. 1 – rozmnóżki; 2 – owoce. (Źródło: Falińska 1990, zmienione).

Rys. 7. Przykłady wykresów słupkowych. A – słupki zwykłe; B – słupki złożone; C – słupki symetryczne.

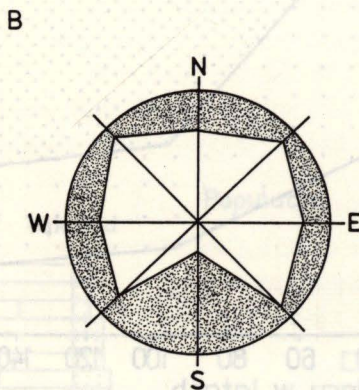


Procentowy udział ptaków o różnym typie gnieźdzenia się w kolejnych stadiach sukcesji wtórnej na zrębie lasu igradowego (*Tilio-Carpinetum*). Stadia sukcesji: I (świeży wyrąb), ..., VI (150-letni igrąd). Typy gnieźdzenia: 1 – gatunki zakładające gniazda na ziemi lub nad nią; 2 – gatunki wijące gniazda na krzewach i drzewach; 3 – dziuplaki. (Źródło: Głowaciński 1975, zmienione).

Rys. 8. Przykład wykresu warstwowego (pasowego).

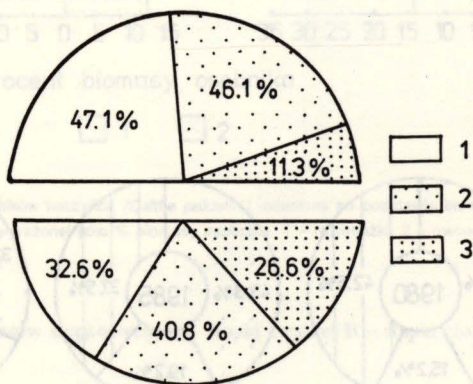


Ścieki przemysłowe i komunalne, oczyszczane i nie oczyszczane w latach 1980–1990. 1 – ścieki nie oczyszczane; 2 – ścieki oczyszczane biologicznie; 3 – ścieki oczyszczane mechanicznie i chemicznie. (Źródło: Ochrona Środowiska 1991, zmienione).



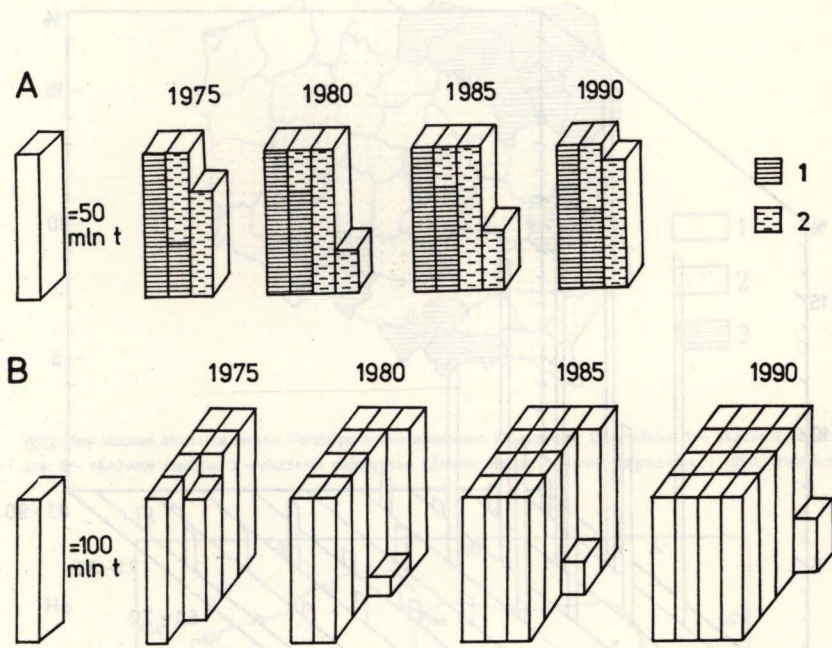
Rozmieszczenie $n = 134$ stanowisk *Campanula rapunculoides* na terenie Pienin w zależności od ekspozycji (zaciemniona część koła). (Źródło: Zarzycki 1976, zmienione).

C



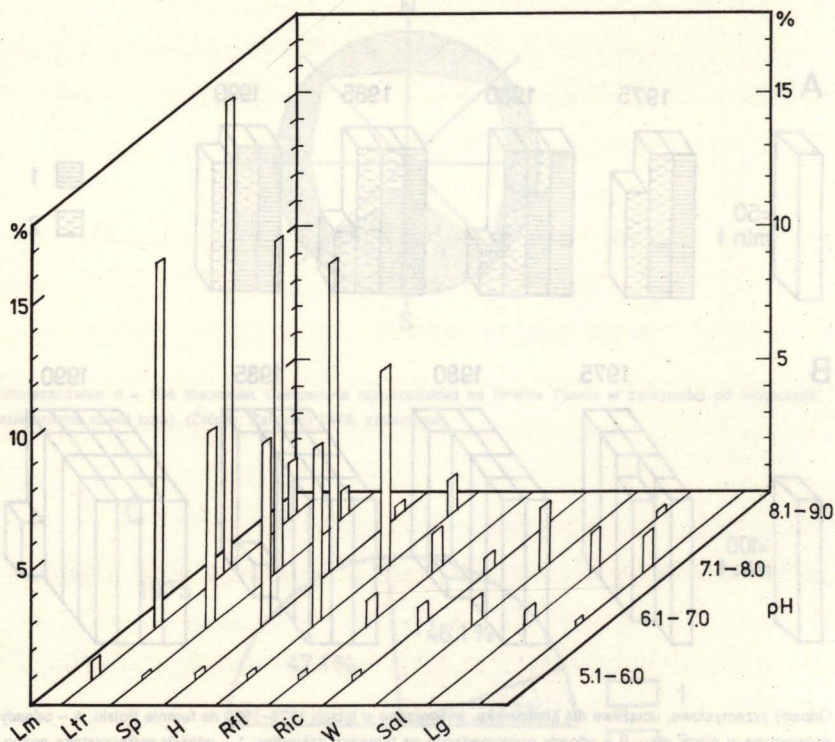
Ścieki przemysłowe i komunalne, oczyszczane i nie oczyszczane w roku 1975 i 1990. 1 – ścieki nie oczyszczane; 2 – ścieki oczyszczane biologicznie; 3 – ścieki oczyszczane chemicznie i mechanicznie. (Źródło: Ochrona środowiska 1991, zmienione).

Rys. 9. Przykłady wykresów kołowych (cyklogramów). A – wycinki kół; B – wykres kołowy z osiami; C – półkoła.



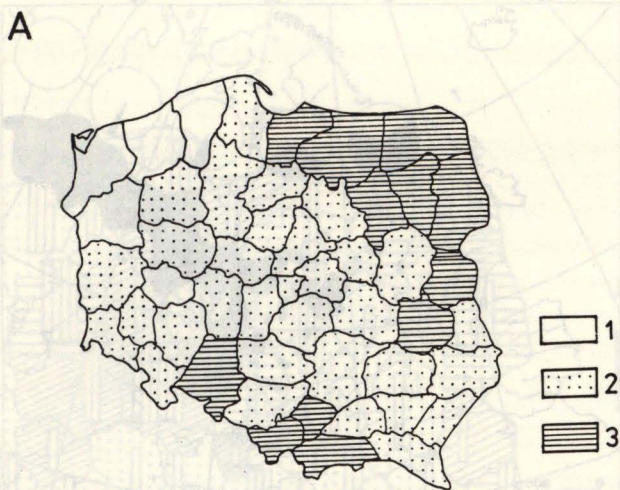
Odpady przemysłowe, uciążliwe dla środowiska, wytworzone w latach 1975–1990 na terenie Polski. A – odpady wytworzone w ciągu roku; B – odpady nagromadzone na terenach zakładów; 1 – odpady wykorzystane gospodarczo; 2 – odpady składowane. (Źródło: Ochrona środowiska 1991, zmienione).

Rys. 10. Przykład wykresu segmentowego (wiedeńskiego).

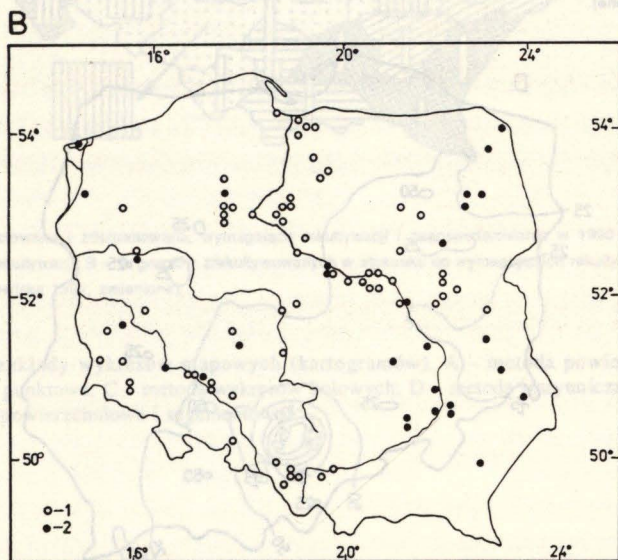


Procentowy udział gatunków roślin wodnych, wolnopływających, w wodach śródlądowych Polski o różnym zakresie pH. Lm - *Lemna minor*, Ltr - *L. trisulca*, Lg - *L. gibba*, Sp - *Spirodela polyrrhiza*, H - *Hydrocharis morsus-ranae*, Rfl - *Riccia fluitans*, Ricc - *Ricciocarpos natans*, W - *Wolffia arrhiza*, Sal - *Salvinia natans* (oryg.).

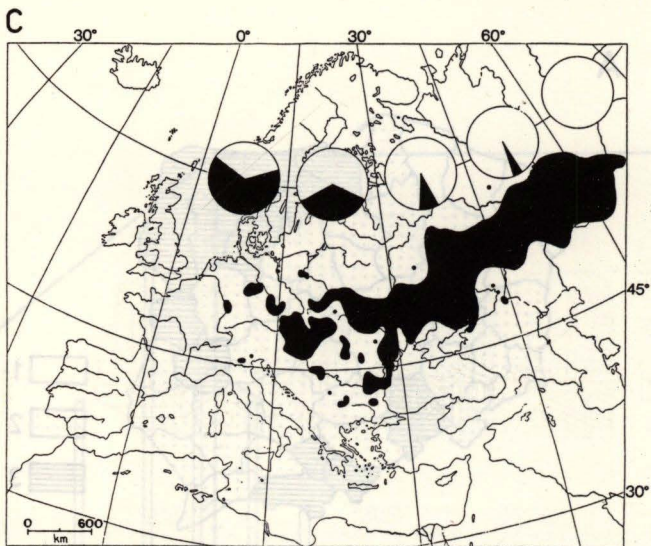
Rys. 11. Przykład wykresu przestrzennego.



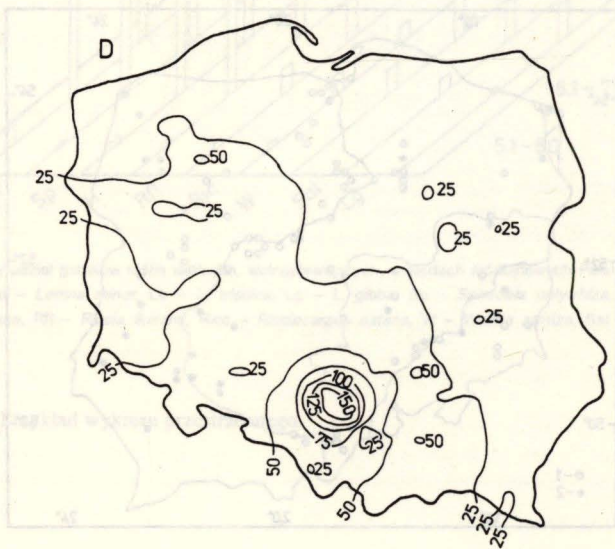
Względny stopień skażenia terenu Polski po awarii elektrowni atomowej w Czarnobylu. 1 – skażenie najmniejsze; 2 – skażenie średnie; 3 – skażenie największe. (Źródło: Raport o stanie, zagrożeniu ... 1990, zmienione).



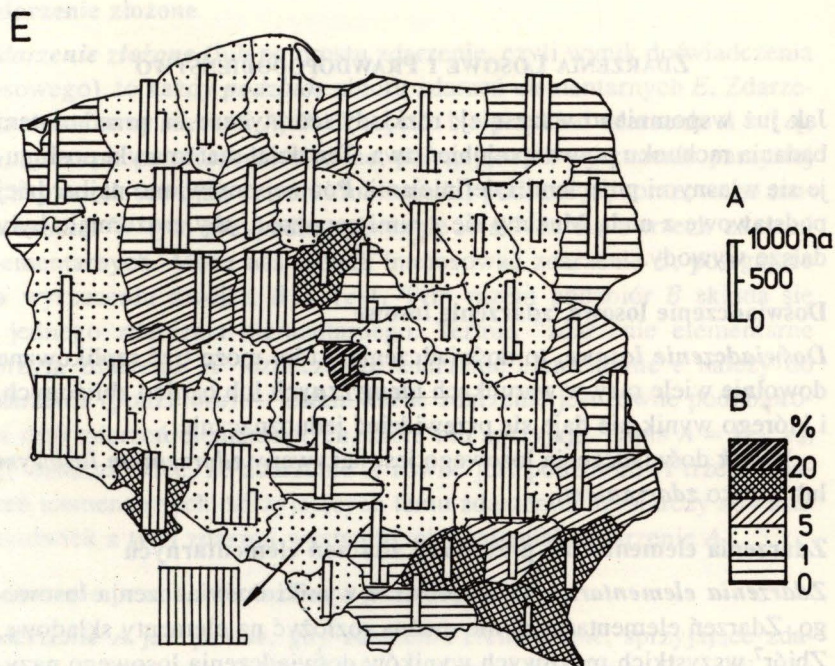
Rozmieszczenie *Riccioarpos natans* (L.) Corda w Polsce. 1 – stanowiska z literatury; 2 – stanowiska na podstawie materiałów zielnikowych. (Źródło: Ochrya, Tomaszewicz 1979, zmienione).



Częstość występowania spontanicznych mieszańców *Prunus fruticosa* x *P. cerasus* w zasięgu *P. fruticosa*. Zaczerniona część koła oznacza % mieszańców w 15° przedziałach długości geograficznej. (Źródło: Wójcicki 1991, zmienione).



Koncentracja ołowiu ($\mu\text{g}\cdot\text{g}^{-1}$) na terenie Polski. Stan w r. 1990. (Źródło: Grodzińska et al., niepubl., zmienione).



Grunty zdewastowane i zdegradowane, wymagające rekultywacji i zagospodarowania w 1990 r. A – grunty wymagające rekultywacji; B – % gruntów zrehabilitowanych w stosunku do wymagających rekultywacji. (Źródło: Ochrona środowiska 1991, zmienione).

Rys. 12. Przykłady wykresów mapowych (kartogramów). A – metoda powierzchniowa; B – metoda punktowa; C – metoda wykresów kołowych; D – metoda izarytmiczna (izolinii); E – metoda powierzchniowa i segmentowa.

3. RACHUNEK PRAWDOPODOBIENSTWA

ZDARZENIA LOSOWE I PRAWDOPODOBIENSTWO

Jak już wspomniano wcześniej, rozkłady teoretyczne są przedmiotem badania rachunku prawdopodobieństwa. Ten dział matematyki posługuje się własnymi pojęciami i definicjami. Poniżej omówiono najbardziej podstawowe z nich. Musimy się z nimi zapoznać, aby zrozumiały były dalsze wywody.

Doświadczenie losowe, zdarzenie losowe

Doświadczenie losowe, to doświadczenie, które może być replikowane dowolnie wiele razy w warunkach *identycznych* lub bardzo zbliżonych, i którego wynik nie daje się przewidzieć jednoznacznie.

Wynik doświadczenia losowego jest nazywany *zdarzeniem losowym* lub krótko *zdarzeniem*.

Zdarzenia elementarne, przestrzeń zdarzeń elementarnych

Zdarzenia elementarne, to najprostsze wyniki doświadczenia losowego. Zdarzeń elementarnych nie można rozłożyć na elementy składowe. Zbiór⁷ wszystkich możliwych wyników doświadczenia losowego nazywa się przestrzenią (zbiorem) elementarnych wyników tego doświadczenia lub *przestrzenią* (zbiorem) *zdarzeń elementarnych*. Zbiór zdarzeń elementarnych oznacza się literą E , a zdarzenie elementarne literą e .

Dla każdego doświadczenia losowego należy oddzielnie określić, co uważamy za zdarzenia elementarne i określić zbiór tych zdarzeń. Dla rzutu monetą $E = \{e_o, e_r\}$, co oznacza, że w pojedynczym doświadczeniu, czyli podczas jednorazowego rzutu monetą, może wypaść albo orzeł (o), albo reszka (r). Dla rzutu kostką do gry $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$.

⁷ Termin "zbiór" jest pojęciem pierwotnym. Każdy zbiór wyznaczany jest przez swoje elementy. Zbiór oznaczamy dużymi literami. Jeśli zbiór A zawiera się w zbiorze B , to zbiór A jest podzbiorem zbioru B .

Oznacza to, że w pojedynczym doświadczeniu (jeden rzut kostką) może wypaść jedyńska, dwójka, ..., lub szóstka.

Zdarzenie złożone

Zdarzenie złożone (lub po prostu zdarzenie, czyli wynik doświadczenia losowego), to każdy **podzbiór** zbioru zdarzeń elementarnych E . Zdarzenia oznaczamy literami A, B, C, \dots , itd. Na przykład zdarzenie $A = \{e_2, e_4, e_6\}$ oznacza zdarzenie złożone, polegające na wyrzuceniu parzystej liczby oczek (dwójki, czwórki lub szóstki) podczas jednorazowego rzutu kostką. W tym przypadku zdarzenie A składa się z trzech zdarzeń elementarnych. Może nas jednak interesować zdarzenie B , polegające na wyrzuceniu dwójki, $B = \{e_2\}$. Tym razem podzbiór B składa się z jednego zdarzenia elementarnego. Termin "zdarzenie elementarne sprzyja zdarzeniu A " oznacza, że zdarzenie elementarne e należy do zdarzenia (podzbioru) A . Zatem każde zdarzenie jest równe podzbiorowi tych zdarzeń elementarnych, które mu sprzyjają. Zapis $A = \{e_2, e_4, e_6\}$ oznacza więc, że zdarzenie A równe jest podzbiorowi trzech zdarzeń elementarnych, sprzyjających temu zdarzeniu. Wystarczy aby którekolwiek z tych zdarzeń wystąpiło, aby zaistniało zdarzenie A .

Zdarzenie pewne, niemożliwe i prawdopodobne

Zdarzenie A jest pewne, gdy zdarzenia elementarne, sprzyjające zdarzeniu A , stanowią jednocześnie przestrzeń wyników E , czyli gdy $A = E$. **Zdarzenie A jest niemożliwe**, gdy nie istnieją zdarzenia elementarne, sprzyjające temu zdarzeniu. Mówimy wówczas, że podzbiór A jest pusty ($A = \emptyset$). Podzbiory przestrzeni wyników różne od podzbiorów charakterystycznych dla zdarzenia pewnego i niemożliwego, to **zdarzenia prawdopodobne** czyli **możliwe**.

Zmienna losowa

Każdemu zdarzeniu elementarnemu skończonego zbioru E można przyporządkować dokładnie jedną liczbę ze zbioru liczb rzeczywistych R . Na przykład, gdy $E = \{e_1, e_2, e_3, e_4, e_5, e_6\}$ (przebieg wyników sześciociennej kostki do gry), wtedy elementowi e_1 można przyporządkować liczbę 1, elementowi e_2 liczbę 2, ..., elementowi e_6 liczbę 6. W przypadku rzutu monetą $E = \{e_o, e_r\}$. Elementowi e_r można przyporządkować liczbę 0, elementowi e_o liczbę 1. Takie przyporządkowanie

elementów jednego zbioru elementom drugiego zbioru nazywamy **odwzorowaniem** lub **funkcją**. W naszym przypadku dokonujemy odwzorowania zdarzeń elementarnych zbioru E w zbiór liczb rzeczywistych R .

Jeżeli każdemu ze zdarzeń elementarnych zbioru E przyporządkowano pewną liczbę rzeczywistą ze zbioru R , to mówimy, że została określona **zmienna losowa**. Mówiąc inaczej, zmienna losowa, to każda funkcja określona na zbiorze E i przybierająca wartości w zbiorze R liczb rzeczywistych. Liczby 1, 2, ..., 6 to wartości zmiennej losowej "liczba oczek kostki sześciostiennej".

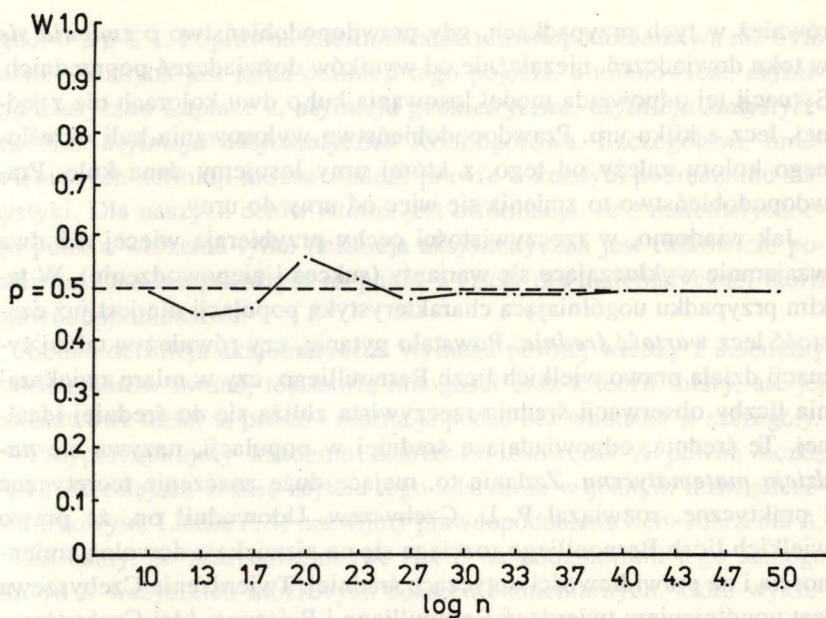
Prawidłowość statystyczna

W pojedynczym doświadczeniu losowym dane zdarzenie, np. zdarzenie A , może zajść lub może nie zajść. Jeśli dane doświadczenie zreplikowano n razy, a zdarzenie A zaszło przy tym m razy, to stosunek m/n nazywamy **częstością względną** zdarzenia A . Jeżeli, dokładnie w tych samych warunkach, doświadczenie, o którym mowa, przeprowadzilibyśmy wiele razy, przy czym za każdym razem zwiększano by liczbę replikacji tego doświadczenia, wówczas zauważylibyśmy zjawisko zwane **prawidłowością statystyczną**, polegające na tym, że przy coraz większej liczbie replikacji (n), częstość zdarzenia coraz bardziej zbliża się do pewnej stałej liczby. Ta liczba, wokół której oscyluje częstość danego zdarzenia, jest **idealną** częstością tego zdarzenia. Oznaczamy ją literą p i nazywamy **prawdopodobieństwem**.

Jak powiedziano wcześniej, prawidłowość statystyczna uwidacznia się wtedy, gdy istnieje duża liczba obserwacji, czyli – jak mówimy – w **zjawiskach** (procesach) **masowych**. Kształtowanie się prawidłowości statystycznej można prześledzić, wykonując na przykład wielokrotne rzuty monetą symetryczną. Jak widać na Rys. 13, ze wzrostem liczby rzutów n , częstość wyrzucenia orła coraz bardziej zbliża się do częstości idealnej p równej 0.5. Taką samą prawidłowość można zaobserwować rzucając symetryczną sześciostenną kostką do gry. Tym razem jednak, częstość wystąpienia dowolnej liczby oczek (1, 2, ..., lub 6) będzie zbliżać się do $1/6$.

Prawo wielkich liczb

Naukowe wyjaśnienie prawidłowości statystycznej, podane po raz pierwszy przez J. Bernoulliego, znane jest od roku 1713. Sformułowane



Rys. 13. Wyniki 10, 20, 50, ..., n rzutów monetą symetryczną (opracowano wg Żuka 1989, Tab. 1.2.1.). w – empiryczna częstość względna (frakcja) wyrzucenia “orka”; p – idealna częstość względna (prawdopodobieństwo) wyrzucenia “orka” w jednym rzucie; n – liczba rzutów.

wówczas przez niego twierdzenie znane jest obecnie jako **prawo wielkich liczb Bernoulliego**. Od czasów Bernoulliego udowodniono kilka dodatkowych twierdzeń, które wraz z twierdzeniem Bernoulliego występują zwykle pod wspólną nazwą **prawa wielkich liczb**. Warto może w tym miejscu powiedzieć kilka słów na ten temat.

Twierdzenie Bernoulliego odnosiło się do bardzo prostego modelu: możliwe były tylko dwa wykluczające się wyniki (sukces i niepowodzenie), a prawdopodobieństwo sukcesu p w pojedynczym doświadczeniu było **stałe**. Modelowi temu odpowiada losowanie z urny kul o dwu kolorach, przy czym frakcje kul w urnie są stałe. Poza tym, twierdzenie Bernoulliego dotyczyło **wyłącznie** prawidłowości związanej z **częstością zdarzenia**. Dalszy rozwój idei Bernoulliego polegał na komplikowaniu jego modelu i rozciąganiu go na inne charakterystyki.

S. D. Poisson był pierwszym badaczem, który uogólnił twierdzenie Bernoulliego. Udowodnił on, że twierdzenie to zachowuje słuszność

również w tych przypadkach, gdy prawdopodobieństwo p *zmienia się* w toku dowiadczeń, niezależnie od wyników doświadczeń poprzednich. Sytuacji tej odpowiada model losowania kul o dwu kolorach nie z jednej, lecz z kilku urn. Prawdopodobieństwo wylosowania kuli określonego koloru zależy od tego, z której urny losujemy daną kulę. Prawdopodobieństwo to zmienia się więc od urny do urny.

Jak wiadomo, w rzeczywistości cechy przybierają więcej niż dwa wzajemnie wykluczające się warianty (sukces i niepowodzenie). W takim przypadku uogólniającą charakterystyką populacji nie jest już częstość lecz *wartość średnia*. Powstało pytanie, czy również w takiej sytuacji działa prawo wielkich liczb Bernoulliego, czy w miarę zwiększania liczby obserwacji średnia rzeczywista zbliża się do średniej idealnej. Tę średnią, odpowiadającą średniej w populacji, nazywa się *nadzieją matematyczną*. Zadanie to, mające duże znaczenie teoretyczne i praktyczne, rozwiązał P. L. Czebyszew. Udowodnił on, że prawo wielkich liczb Bernoulliego rozciąga się na zjawiska z dowolną zmiennością i na prawidłowości dotyczące średniej. Twierdzenie Czebyszewa jest uogólnieniem twierdzeń Bernoulliego i Poissona. Idei Czebyszewa odpowiada model urny, w której znajdują się kule różnego koloru.

Prawidłowość statystyczna przejawia się również w charakterze *rozkładu częstości*. Tym zagadnieniem zajmowało się wielu badaczy. Spośród kilku, bardzo ważne twierdzenie (tzw. *centralne twierdzenie graniczne*) udowodnił A. M. Lapunow. Ogólnie biorąc mówi ono, że *bez względu* na kształt rozkładu wyjściowego, \rightarrow rozkład z próby dąży do teoretycznego (idealnego) rozkładu symetrycznego, tzw. \rightarrow *rozkładu normalnego*, pod warunkiem, że liczebność próby n będzie dostatecznie duża. Wiecej na ten temat będzie powiedziane dalej. Natomiast w tym miejscu jeszcze raz podkreślmy, że w procesie masowym charakter zbliżenia empirycznego rozkładu częstości do ich rozkładu teoretycznego jest taki sam, jak w przypadku zbliżenia częstości do prawdopodobieństwa i średniej rzeczywistej do nadziei matematycznej.

Więcej szczegółów na temat *prawa wielkich liczb* można znaleźć w przystępnie napisanej książce Paschawera (1967).

Definicja prawdopodobieństwa

Prawdopodobieństwo zdarzenia, podobnie jak i częstość względna, jest dowolną liczbą dodatnią z przedziału $(0,1)$, co zapisujemy następu-

jąco: $0 \leq p \leq 1$. Poprawne zdefiniowanie prawdopodobieństwa nie było łatwe. Znanych jest kilka definicji tego pojęcia, a mianowicie: *definicja klasyczna* Laplace'a, *definicja geometryczna*, *definicja statystyczna* oraz *definicja aksjomatyczna* Kołmogorowa. Szczegółowe omówienie tych definicji można znaleźć prawie w każdym podręczniku statystyki. Dla naszych celów istotna jest informacja, że z matematycznego punktu widzenia tylko definicja aksjomatyczna jest całkowicie poprawna. Ona też posłużyła za punkt wyjścia dla matematycznej teorii prawdopodobieństwa.

Ścisła definicja aksjomatyczna wymaga pewnej wiedzy z dziedziny teorii zbiorów zwanej też teorią mnogości oraz z teorii miary, ale jej podstawowe treści są proste i można je podać bez wnikania w szczegóły.

Przyporządkujemy każdemu zdarzeniu losowemu A pewną liczbę $P(A)$, określającą szansę zajścia tego zdarzenia w jednym doświadczeniu losowym. Liczbę $P(A)$ nazwijmy prawdopodobieństwem zdarzenia A .

Żałóżmy, że zdarzenia losowe A i B są podzbiórami tego samego zbioru E wszystkich możliwych zdarzeń elementarnych, które wykluczają się wzajemnie, tzn. że dane zdarzenie może być jednym i tylko jednym zdarzeniem elementarnym i żadnym innym. Jeżeli teraz, każdemu zdarzeniu losowemu A przyporządkowano liczbę rzeczywistą $P(A)$, zwaną prawdopodobieństwem zdarzenia A , w taki sposób, aby spełnione były następujące warunki (aksjomaty):

I. prawdopodobieństwo wystąpienia zdarzenia A należy do przedziału $(0,1)$

$$0 \leq P(A) \leq 1$$

II. prawdopodobieństwo zdarzenia pewnego jest równe 1

$$P(E) = 1$$

III. Prawdopodobieństwo sumy (alternatywy) dwóch wykluczających się zdarzeń jest równe sumie prawdopodobieństw tych zdarzeń

$$P(A \text{ lub } B) = P(A) + P(B)$$

to określoną w ten sposób funkcję P nazywamy **prawdopodobieństwem**. Powyższy system warunków, zaproponowany przez A. Kołmogorowa w 1933 r., stanowi podstawę rachunku prawdopodobieństwa. Definicja aksjomatyczna jest uogólnieniem wyżej wspomnianych defi-

nicji. Nie zakłada ona niczego o zdarzeniach ani o sposobie przyporządkowywania tym zdarzeniom określonych prawdopodobieństw, co pozwala **połączyć teorię z praktyką**.

Założmy na przykład, że wykonujemy rzut monetą. Zdarzeniami elementarnymi są: wyrzucenie orła (e_o), wyrzucenie reszki (e_r). Zbiorem zdarzeń elementarnych jest zbiór $E = \{e_o, e_r\}$. Możliwe są cztery zdarzenia losowe: zdarzenie pewne polegające na wyrzuceniu albo orła albo reszki, $A = E$; zdarzenie niemożliwe polegające na tym, że nie zostanie wyrzucony ani orzeł ani reszka, $B = \emptyset$; zdarzenie polegające na wyrzuceniu orła, $C = \{e_o\}$; zdarzenie polegające na wyrzuceniu reszki, $D = \{e_r\}$. Na tych zdarzeniach możemy określić funkcję P na wiele sposobów, pod jednym wszakże warunkiem – aby funkcja ta spełniała wyliczone powyżej aksjomaty. Na przykład, dla monety symetrycznej funkcja ta będzie wynosiła: $P(A) = 1$, $P(B) = 0$, $P(C) = 1/2$, $P(D) = 1/2$. Dla monety niesymetrycznej, funkcja ta może wynosić np.: $P(A) = 1$, $P(B) = 0$, $P(C) = 1/3$, $P(D) = 2/3$. Jak widać, z teoretycznego punktu widzenia jest wszystko jedno, czy przyjmiemy, że prawdopodobieństwo wyrzucenia orła jest równe $1/2$, $1/3$, czy też jakiejś innej wartości, byle prawdopodobieństwa pozostałych zdarzeń określone były zgodnie z powyższymi aksjomatami. Jak z tego wynika, dysponujemy w tym względzie pewną dowolnością.

Prawdopodobieństwo zdarzenia możemy wyznaczać albo a priori, lub eksperymentalnie. W pierwszym przypadku kształt przyrządu losującego (np. kostki do gry, monety, itp.) oraz charakterystyczne dla tego kształtu symetrie pozwalają ocenić wartość prawdopodobieństwa określonych zdarzeń bez eksperymentowania. Na przykład, w przypadku monety symetrycznej możemy wnioskować, że prawdopodobieństwo wyrzucenia orła jest takie samo jak prawdopodobieństwo wyrzucenia reszki i wynosi $1/2$. Gdy prawdopodobieństwa nie można ustalić a priori (np. gdy moneta jest asymetryczna i nic nie wiemy bliższego o tej asymetrii), wówczas przeprowadzamy doświadczenie losowe i przyjmujemy, że prawdopodobieństwo zdarzenia równa się zaobserwowanej częstości względnej tego zdarzenia, a ściślej mówiąc jest równe p . Mamy więc w tym przypadku, jak widać, do czynienia z częstością (statystyczną) interpretacją prawdopodobieństwa.

ROZKŁADY TEORETYCZNE

Znana jest wielka liczba teoretycznych rozkładów prawdopodobieństwa zmiennej losowej, ale nas najbardziej będą interesowały następujące typy rozkładów: *rozkład zero-jedynkowy*, *rozkład dwumianowy* (binomialny, Bernoulliego) i *rozkład Poissona* zmiennej skokowej oraz *rozkład normalny* (Gaussa, Gaussa-Laplace'a) zmiennej ciągłej. Rozkłady te szczegółowo zostały omówione m. in. przez Perkała (1963). W tym miejscu zostaną one tylko pokrótce scharakteryzowane. Rozważania nasze zaczniemy od zmiennej skokowej.

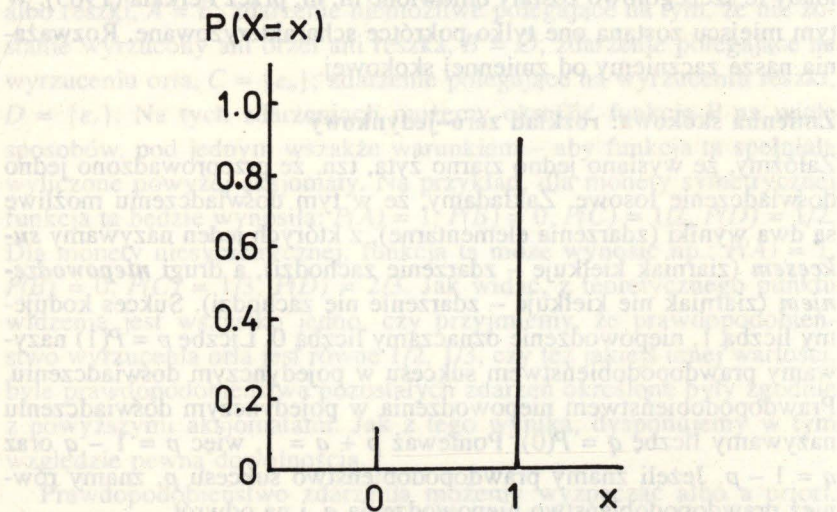
Zmienna skokowa: rozkład zero-jedynkowy

Założmy, że wysiano jedno ziarno żyta, tzn. że przeprowadzono jedno doświadczenie losowe. Zakładamy, że w tym doświadczeniu możliwe są dwa wyniki (zdarzenia elementarne), z których jeden nazywamy *sukcesem* (ziarniak kiełkuje – zdarzenie zachodzi), a drugi *niepowodzeniem* (ziarniak nie kiełkuje – zdarzenie nie zachodzi). Sukces kodujemy liczbą 1, niepowodzenie oznaczamy liczbą 0. Liczbę $p = P(1)$ nazywamy prawdopodobieństwem sukcesu w pojedynczym doświadczeniu. Prawdopodobieństwem niepowodzenia w pojedynczym doświadczeniu nazywamy liczbę $q = P(0)$. Ponieważ $p + q = 1$, więc $p = 1 - q$ oraz $q = 1 - p$. Jeżeli znamy prawdopodobieństwo sukcesu p , znamy również prawdopodobieństwo niepowodzenia q , i na odwrót.

Jak już wspomniano wcześniej, prawdopodobieństwo danego zdarzenia można *szacować* tylko na podstawie pewnej liczby doświadczeń losowych. Na podstawie pojedynczego doświadczenia zrobić tego nie można. Można jedynie stwierdzić, że pewne zdarzenie zaszło lub nie. Dopiero gdy dane doświadczenie zreplikujemy wielokrotnie w tych samych warunkach, możemy oszacować prawdopodobieństwo interesującego nas zdarzenia. Założmy, że chcemy oszacować prawdopodobieństwo sukcesu w doświadczeniu z żytem. Możemy to zrobić na dwa sposoby: albo *kolejno* wysiewać poszczególne ziarna żyta i za każdym razem notować wynik albo wysiać *równocześnie* wiele nasion i po skielkowaniu zapisać wynik. Rozpatrzmy sposób pierwszy.

Założmy, że po kolei wysiewano 10 ziarn żyta, tzn. że przeprowadzono 10 *nierównoczesnych* doświadczeń. Mówiąc inaczej, $n = 10$ razy zreplikowano (\rightarrow replikacja) $k = 1$ doświadczenie losowe. Założmy da-

lej, że na $n = 10$ poddanych kiełkowaniu ziarniaków skiełkowało 9 ziarniaków. Na tej podstawie stwierdzamy, że prawdopodobieństwo wykiełkowania ziarniaka w pojedynczym doświadczeniu wynosi $p = 9/10 = 0.9$. Prawdopodobieństwo, że w pojedynczym doświadczeniu ziarno żyta nie wykiełkuje równe jest $q = 1 - p$, czyli $q = 0.1$. Uzyskane rezultaty przedstawiono na Rys. 14.



Rys. 14. Rozkład zero-jedynkowy: prawdopodobieństwo kiełkowania ziarn żyta (dane fikcyjne) dla $n = 10$, $k = 1$. 0 – oznacza, że ziarno nie wykiełkowało (niepowodzenie), 1 – oznacza, że ziarno wykiełkowało (sukces). $P(X=x)$ – prawdopodobieństwo zdarzenia losowego polegającego na tym, że zmienna losowa X przyjmie wartość $x = 0, 1$.

Generalnie rzecz biorąc, gdy n razy przeprowadzamy pojedyncze doświadczenie ($k = 1$), przy czym zbiór wszystkich możliwych wyników takiego doświadczenia jest dwuelementowy, np. gdy wykonujemy n rzutów pojedynczą monetą ($k = 1$), gdy n razy sprawdzamy czy $k = 1$ produkt jest dobry czy wadliwy, gdy n razy sprawdzamy czy $k = 1$ jaja wykluje się kogut czy kura, gdy n razy sprawdzamy czy $k = 1$ nasienie wykiełkuje czy też nie, itd., wówczas otrzymujemy tzw. **dwupunktowy** lub **zero-jedynkowy rozkład prawdopodobieństw**. Rozkład ten jest szczególnym przypadkiem \rightarrow rozkładu dwumianowego, gdy $k = 1$. **Rozkład prawdopodobieństw** jest to więc zbiór prawdopodo-

bieństw, odpowiadających wszystkim wariantom zmiennej losowej. Jest to zarazem rozkład cechy w populacji.

Wprowadzenie pojęcia "rozkład prawdopodobieństwa" pozwala zdefiniować zmienną losową w następujący sposób: **zmienna losowa** to taka zmienna, która przybiera każdy ze swoich wariantów z określonym prawdopodobieństwem, lub – jest to zmienna o znanym rozkładzie prawdopodobieństwa.

Zmienna skokowa: rozkład dwumianowy

Przyjmijmy obecnie, że zastosowano drugi schemat doświadczalny, tzn. że **równocześnie** wysiano $k = 10$ ziarn żyta, czyli, mówiąc inaczej, przeprowadzono jedną serię ($n = 1$) 10 równoczesnych doświadczeń losowych ($k = 10$). Również w tym przypadku może się zdarzyć, że na 10 wysianych ziarniaków, wykiełkuje dokładnie 9 ziarn żyta. Zauważmy jednak, że teoretycznie rzecz biorąc, gdy **równocześnie** wysiewamy $k = 10$ ziarniaków, możemy oczekiwać **różnych** wyników. Możemy się spodziewać, mianowicie, że w rozważanym przypadku wykiełkuje 0, 1, 2, ..., lub 10 ziarniaków, tzn. że wystąpi 0, 1, 2, ..., lub 10 sukcesów – wynik tej serii równocześnie przeprowadzonych doświadczeń nie może być przewidziany w sposób jednoznaczny.

Jak widzimy, w tym przypadku przestrzeń wyników zawiera więcej niż dwa elementy (zdarzenia elementarne). Wszystkie możliwe wyniki danej serii (ciągu) doświadczeń możemy zapisać w następujący sposób: $x = 0, 1, 2, \dots, k$, gdzie x oznacza możliwą liczbę sukcesów, a k oznacza maksymalną liczbę sukcesów, którą można uzyskać dla danej serii doświadczeń. Mówiąc inaczej, k oznacza liczbę realizowanych równocześnie doświadczeń. Czasami k bywa też określane jako **długość serii**. W naszym przykładzie $k = 10$, ponieważ jednocześnie wysiano 10 ziarniaków. Tyle też ziarn żyta, teoretycznie biorąc, może skiełkować.

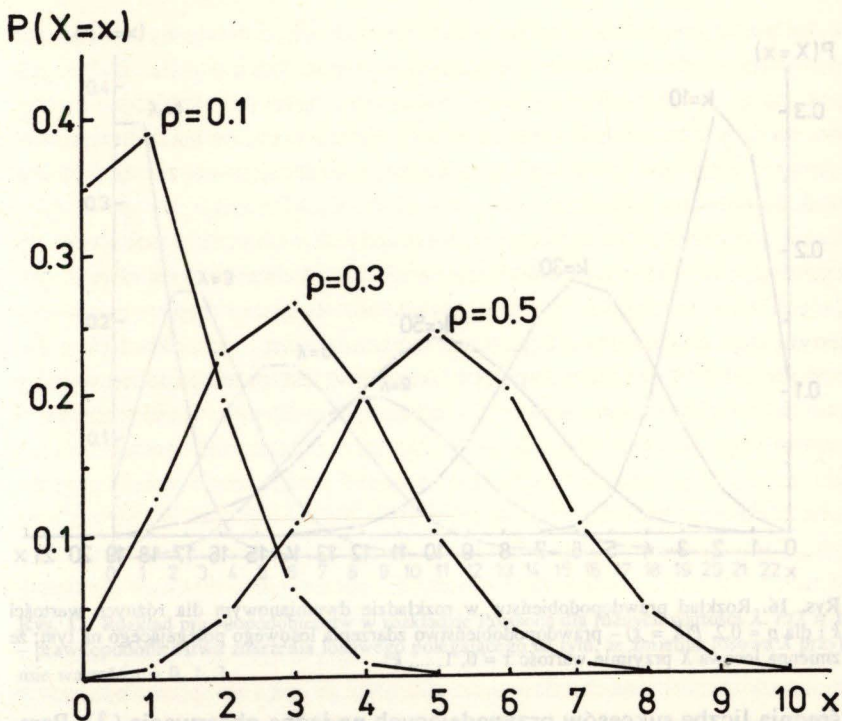
Prawdopodobieństwa zajścia lub nie zajścia konkretnego zdarzenia losowego nie można określić realizując jedną serię doświadczeń. Jeżeli przeprowadzamy tylko **jedną** serię doświadczeń ($n = 1$), to otrzymujemy tylko **jeden** wynik, który mówi nam tylko tyle, że dane zdarzenie zaszło – tzn. znamy liczbę sukcesów właściwych dla tego wyniku. Tak było w naszym przykładzie z wysianymi równocześnie 10 ziarnami żyta ($k = 10$). Aby móc oszacować prawdopodobieństwo realizacji różnych wyników, potrzeba większej liczby obserwacji, dlatego daną serię

k doświadczeń trzeba zreplikować n razy. Na podstawie uzyskanych n wyników możemy obliczyć częstość występowania $x = 0, 1, 2, \dots, k$ sukcesów, a następnie możemy oszacować prawdopodobieństwa uzyskania tych sukcesów. Przypisując każdemu wariantowi zmiennej losowej (x), określone prawdopodobieństwo, otrzymujemy rozkład prawdopodobieństw tej zmiennej.

Gdy wyniki poszczególnych doświadczeń są zdarzeniami *niezależnymi*, tzn. wynik jednego doświadczenia nie ma wpływu na wynik drugiego doświadczenia, czyli gdy prawdopodobieństwo sukcesu p w pojedynczym doświadczeniu jest zawsze takie samo, to rozkład prawdopodobieństw w takim przypadku określany jest przez kolejne wyrazy rozwinięcia dwumianu $(p + q)^k$. Mówimy wówczas o *rozkładzie dwumianowym*. Rozkład dwumianowy oparty jest więc na następujących założeniach: 1. przeprowadza się k niezależnych, równoczesnych doświadczeń; 2. przestrzeń wyników pojedynczego doświadczenia jest dwuelementowa (0,1); 3. prawdopodobieństwo zaistnienia danego zdarzenia (sukces) w pojedynczym doświadczeniu jest stałe i równe p ; 4. znana jest długość serii k .

Kształt rozkładu dwumianowego zależy od dwóch parametrów: k oraz p . Im bardziej p różni się od 0.5, oraz im mniejsze jest k , tym bardziej skośny jest kształt tego rozkładu (Rys.15 i 16).

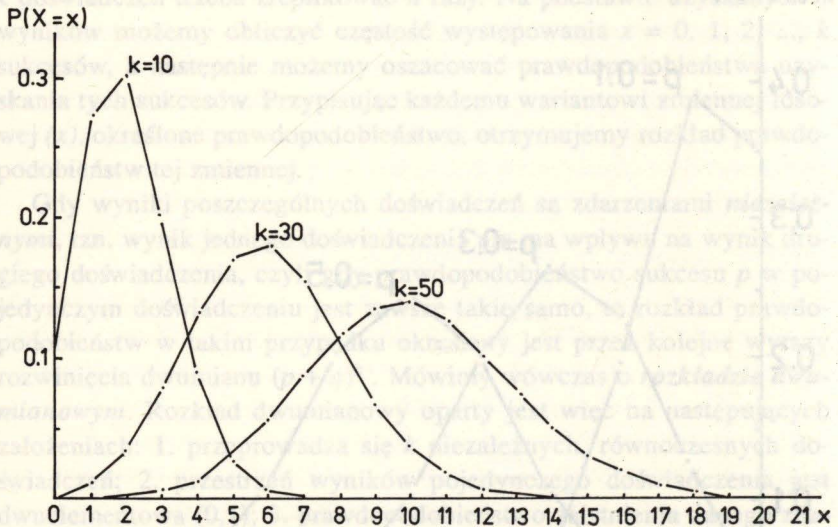
Odwołajmy się jeszcze raz do naszego przykładu z kiełkowaniem żyta. Załóżmy, że do doniczki z ziemią wysiano 10 ziarn żyta, i że tę serię doświadczeń zreplikowano (\rightarrow replikacja) 100 razy ($n = 100$ doniczek po $k = 10$ ziarn żyta każda). Przyjmijmy również, że prawdopodobieństwo pojedynczego sukcesu wynosi $p = 0.9$. Prawdopodobieństwo to oszacowano na podstawie wcześniej przeprowadzonego doświadczenia. Za pomocą rozkładu dwumianowego możemy obliczyć, jakie jest teoretyczne (oczekiwane) prawdopodobieństwo wykiełkowania 0, 1, ..., lub 10 ziarniaków w pojedynczej serii doświadczeń, a następnie możemy wyliczyć oczekiwaną liczbę doniczek, w których wykiełkuje 0, 1, ..., lub 10 ziarniaków żyta. Oczywiście, rozkład dwumianowy możemy wykorzystać do tego typu obliczeń tylko wtedy, gdy spełnione są wymienione wyżej warunki, tzn. gdy mamy do czynienia ze zmienną dwumianową.



Rys. 15. Rozkład prawdopodobieństw w rozkładzie dwumianowym dla różnych wartości p i dla $k = 10$. Uwaga: im bardziej $p > 0.5$, tym silniej w rozkładzie dwumianowym zaznacza się asymetria lewostronna. Przypadek ten nie został zilustrowany. $P(X = x)$ – prawdopodobieństwo zdarzenia losowego polegającego na tym, że zmienna losowa X przyjmie wartość $x = 0, 1, \dots, 10$.

Zmienna skokowa: rozkład Poissona

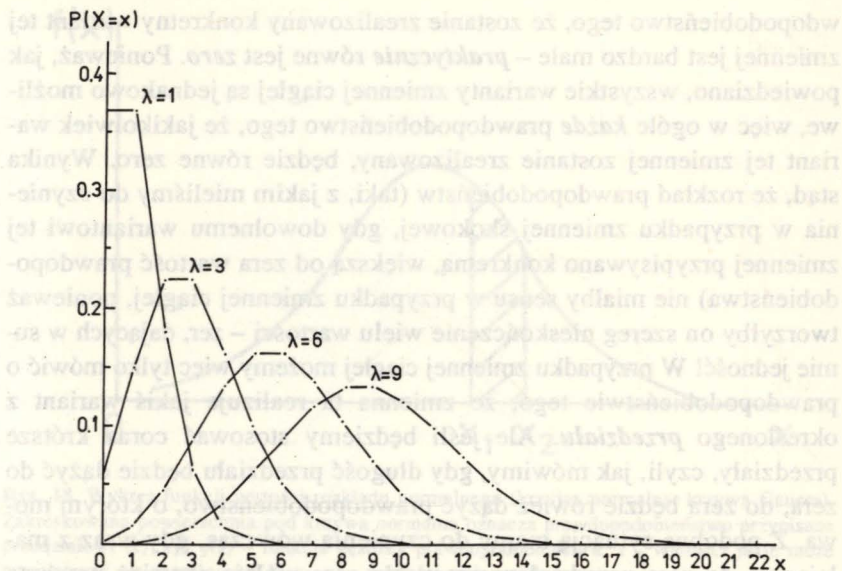
Gdy długość serii k jest większa niż 20, a prawdopodobieństwo sukcesu p jest bardzo małe, rozkład dwumianowy aproksymuje się za pomocą **rozkładu Poissona**. Rozkład ten w przybliżeniu można traktować jako szczególny przypadek rozkładu dwumianowego, w którym k jest duże, praktycznie **nieznane**, a p jest mniejsze niż 0.1. Mówiąc inaczej, rozkład Poissona jest rozkładem granicznym, do którego zmierza rozkład dwumianowy wówczas, gdy k zmierza do nieskończoności, a p zmierza do zera. Aby obliczyć rozkład dwumianowy trzeba znać dwa parametry: k i p . Aby obliczyć rozkład Poissona wystarczy znać tylko



Rys. 16. Rozkład prawdopodobieństw w rozkładzie dwumianowym dla różnych wartości k i dla $p = 0.2$. $P(X = x)$ – prawdopodobieństwo zdarzenia losowego polegającego na tym, że zmienna losowa X przyjmie wartość $x = 0, 1, \dots, k$.

średnią liczbę sukcesów przypadających na jedną obserwację (λ). Parametr λ związany jest z parametrami rozkładu dwumianowego k i p następującą zależnością: $\lambda = kp$. Im mniejsza jest wartość λ , tym bardziej prawostronnie skośny jest rozkład Poissona (Rys. 17).

Rozkład Poissona stosunkowo najczęściej napotykamy wówczas, gdy np. liczymy organizmy występujące na losowych powierzchniach badawczych, w czasie lub w przestrzeni. W takich przypadkach k jest nieokreślone, ponieważ nie wiemy, jaka maksymalna liczba organizmów może być zaobserwowana na tych powierzchniach – nieznaną jest więc maksymalna liczba sukcesów w serii, tak jak to miało miejsce w przypadku zmiennej dwumianowej. Rozkład ten może posłużyć na przykład do sprawdzania, czy rozmieszczenie roślin określonego gatunku, na n powierzchniach ma charakter losowy, czy też nie. W celu sprawdzenia hipotezy o losowym rozmieszczeniu roślin, zliczamy wszystkie powierzchnie, na których stwierdzono 0, 1, 2, ... osobników danego gatunku (długość serii k jest nieznaną), a następnie otrzymany rozkład empiryczny porównujemy z rozkładem oczekiwanym, obliczo-



Rys. 17. Rozkład prawdopodobieństw w rozkładzie Poissona dla różnych wartości λ . $P(X=x)$ – prawdopodobieństwo zdarzenia losowego polegającego na tym, że zmienna losowa X przyjmie wartość $x = 0, 1, 2, \dots$.

nym za pomocą rozkładu Poissona. Statystyczna zgodność obu rozkładów sugeruje, że, na badanych powierzchniach, rośliny te rozmieszczone są losowo.

Zmienna ciągła: rozkład normalny

Gdy długość serii k dąży do nieskończoności, wówczas, niezależnie od tego czy $p = q$, czy $p \neq q$, rozkład dwumianowy dąży do drugiego rozkładu granicznego – **rozkładu normalnego**. Rozkład ten jest dobrym przybliżeniem rozkładu dwumianowego, nawet dla względnie małych wartości k , przy założeniu jednak, że p i q nie różnią się znacznie od siebie.

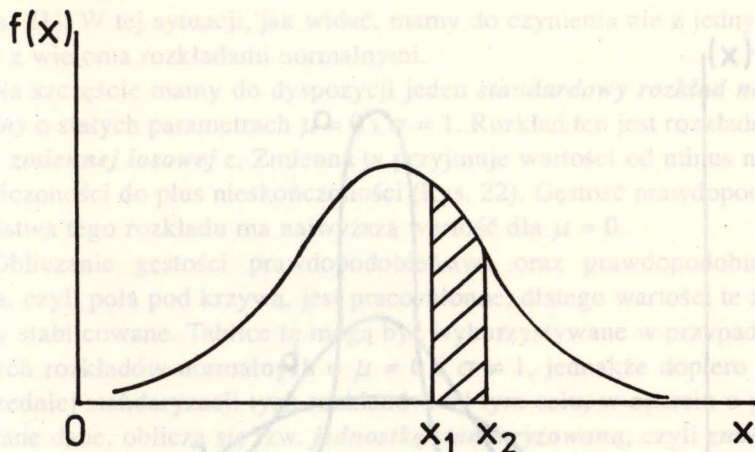
– gęstość prawdopodobieństwa, prawdopodobieństwo

Rozkład normalny jest rozkładem symetrycznym **zmiennnej losowej ciągłej**. Jak pamiętamy, w przypadku zmiennnej ciągłej, liczba wszystkich jednakowo możliwych i wzajemnie wykluczających się wariantów tej zmiennnej (zdarzeń elementarnych) jest nieskończona. A więc pra-

wpodobieństwo tego, że zostanie zrealizowany konkretny wariant tej zmiennej jest bardzo małe – **praktycznie równe jest zero**. Ponieważ, jak powiedziano, wszystkie warianty zmiennej ciągłej są jednakowo możliwe, więc w ogóle **każde** prawdopodobieństwo tego, że jakiegokolwiek wariant tej zmiennej zostanie zrealizowany, będzie równe zero. Wynika stąd, że rozkład prawdopodobieństw (taki, z jakim mieliśmy do czynienia w przypadku zmiennej skokowej, gdy dowolnemu wariantowi tej zmiennej przypisywano konkretną, większą od zera wartość prawdopodobieństwa) nie miałby sensu w przypadku zmiennej ciągłej, ponieważ tworzyłby on szereg nieskończenie wielu wartości – zer, dających w sumie jedność! W przypadku zmiennej ciągłej możemy więc tylko mówić o prawdopodobieństwie tego, że zmienna ta realizuje jakiś wariant z określonego **przedziału**. Ale jeśli będziemy stosować coraz krótsze przedziały, czyli, jak mówimy, gdy długość przedziału będzie dążyć do zera, do zera będzie również dążyć prawdopodobieństwo, o którym mowa. Z podobną sytuacją mamy do czynienia wówczas, gdy wraz z malejącym do zera przedziałem czasu, do zera maleje również przebyta droga. Ale, jak pamiętamy z fizyki, stosunek drogi do czasu pozostaje wielkością skończoną – chwilową prędkością. Okazuje się, że również stosunek prawdopodobieństwa do długości przedziału pozostaje liczbą skończoną. **Liczbę** tę nazywamy **gęstością prawdopodobieństwa** lub **funkcją gęstości prawdopodobieństwa** (funkcją Laplace'a). Wykres tej funkcji nazywamy **krzywą normalną** lub **krzywą Gaussa** (Rys. 18).

Funkcję gęstości prawdopodobieństwa można interpretować w ten sposób, że jest to średnia ilość prawdopodobieństwa, przypadająca na jednostkę długości przedziału, gdy długość tego przedziału dąży do zera. Gęstość prawdopodobieństwa odpowiada rozkładowi prawdopodobieństw **zmiennej losowej skokowej** z tym, że rozkład prawdopodobieństw tej zmiennej przypisywał każdemu wariantowi tej zmiennej ($x = 1, 2, \dots, k$) pewne prawdopodobieństwo p , natomiast rozkład gęstości prawdopodobieństwa **zmiennej losowej ciągłej** każdemu wariantowi tej zmiennej przypisuje pewną **gęstość** $f(x)$, a **prawdopodobieństwo** p przypisuje przedziałowi (x_1, x_2) . Prawdopodobieństwo to jest równe **polu** pod krzywą normalną. Podstawa tego pola jest równa przedziałowi (x_1, x_2) (Rys. 18).

Gęstość prawdopodobieństwa można również przedstawić w postaci **dystrybuanty**, którą oznaczamy symbolem $F(x)$. Odpowiada ona roz-

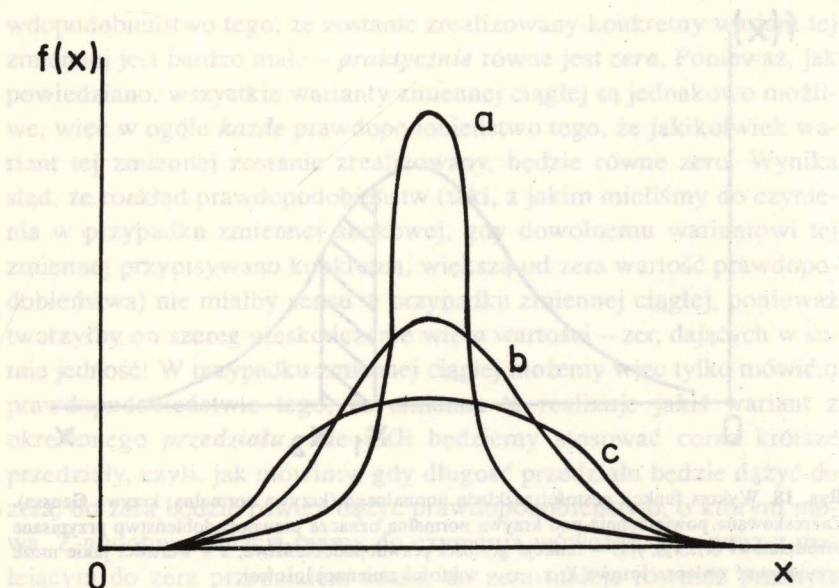


Rys. 18. Wykres funkcji gęstości rozkładu normalnego (krzywa normalna, krzywa Gaussa). Zakreskowana powierzchnia pod krzywą normalną oznacza prawdopodobieństwo przypisane przedziałowi (x_1, x_2) ; $f(x)$ – funkcja gęstości prawdopodobieństwa; x – wartości jakie może przyjmować zmienna losowa X ; x_1, x_2 – wartości zmiennej losowej.

kładowi częstości skumulowanej, czyli dystrybuancie empirycznej, którą oznaczamy symbolem $G(x)$.

– współczynnik asymetrii, współczynnik kurtozy

Rozkład normalny jest rozkładem symetrycznym, o ściśle określonym rozproszeniu wartości liczbowych zmiennej losowej wokół średniej arytmetycznej. Rozproszenie to oceniają dwa współczynniki. **Współczynnik asymetrii** (tzw. moment trzeci) jest miarą asymetrii (skośności, ekscesu). Ocenia on symetryczność rozproszenia wyników wokół średniej arytmetycznej. Wartość dodatnia tego współczynnika wskazuje na asymetrię prawostronną, wartość ujemna na asymetrię lewostronną. Gdy rozkład jest symetryczny, wówczas współczynnik ów równy jest zero. **Współczynnik kurtozy** (tzw. moment czwarty) jest miarą kurtozy (spłaszczenia). Ocenia on stopień zagęszczenia wyników wokół średniej. Wartość dodatnia tego współczynnika wskazuje, że zagęszczenie wyników wokół średniej jest większe niż w rozkładzie normalnym. Mamy wtedy do czynienia z rozkładem bardziej wysmukłym (**rozkład leptokurtyczny**) niż rozkład normalny (**rozkład mezkurtyczny**). Jego wartość ujemna wskazuje, że badany rozkład jest bardziej spłaszczony



Rys. 19. Rodzaje kurtozy. A – rozkład wysmukły (leptokurtyczny); B – rozkład normalny (mezokurtyczny); C – rozkład spłaszczony (platykurtyczny); $f(x)$ – funkcja gęstości prawdopodobieństwa; x – wartości jakie może przyjmować zmienna losowa X .

(**rozkład platykurtyczny**) niż rozkład normalny (Rys. 19). Do opisanego rozkładu normalnego wystarczą tylko dwa parametry: średnia arytmetyczna i odchylenie standardowe (patrz niżej). Aby opisać rozkład leptokurtyczny i platykurtyczny z reguły potrzeba więcej parametrów.

Jak widać z powyższego, **każdy** rozkład normalny jest rozkładem symetrycznym, ale **nie każdy** rozkład symetryczny jest rozkładem normalnym. Współczynniki, o których mowa, służą do oceny tego, jak dalece dany rozkład odbiega od rozkładu normalnego.

– standardowy rozkład normalny

Rozkład normalny wyznaczany jest za pomocą dwóch parametrów: średniej arytmetycznej μ i odchylenia standardowego σ . Zmiana wartości średniej μ powoduje zmianę położenia całego rozkładu poprzez przesunięcie go wzdłuż osi x (Rys. 20). Zmiana wartości σ powoduje zmianę kształtu krzywej normalnej: im wyższa jest wartość odchylenia standardowego, tym bardziej spłaszczony jest kształt tej krzywej

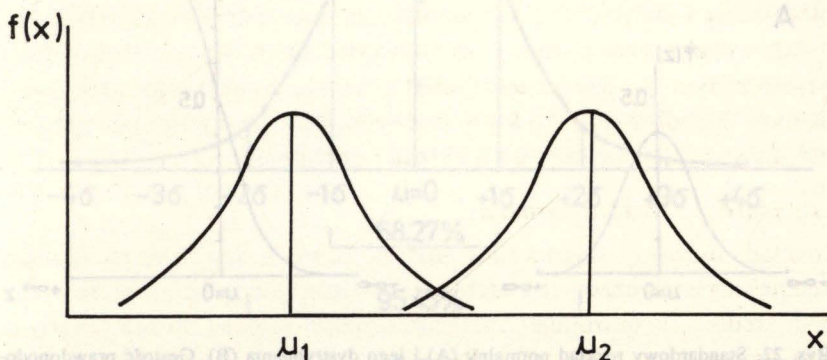
(Rys. 21). W tej sytuacji, jak widać, mamy do czynienia nie z jednym, lecz z wieloma rozkładami normalnymi.

Na szczęście mamy do dyspozycji jeden *standardowy rozkład normalny* o stałych parametrach $\mu = 0$ i $\sigma = 1$. Rozkład ten jest rozkładem tzw. *zmiennej losowej z*. Zmienna ta przyjmuje wartości od minus nieskończoności do plus nieskończoności (Rys. 22). Gęstość prawdopodobieństwa tego rozkładu ma najwyższą wartość dla $\mu = 0$.

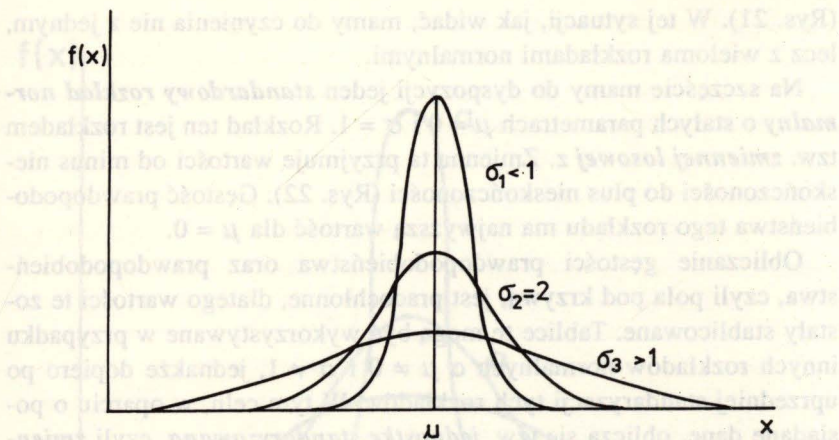
Obliczanie gęstości prawdopodobieństwa oraz prawdopodobieństwa, czyli pola pod krzywą, jest pracochłonne, dlatego wartości te zostały stabilizowane. Tablice te mogą być wykorzystywane w przypadku innych rozkładów normalnych o $\mu \neq 0$ i $\sigma \neq 1$, jednakże dopiero po uprzedniej standaryzacji tych rozkładów. W tym celu, w oparciu o posiadane dane, oblicza się tzw. *jednostkę standaryzowaną*, czyli *zmienną u*, która ma rozkład normalny o $\mu = 0$ i $\sigma = 1$, tzn. taki sam jak zmienna z . W takiej zestandaryzowanej postaci różne rozkłady normalne mogą być także porównywane ze sobą.

– reguła trzech odchyień standardowych (trzech sigm)

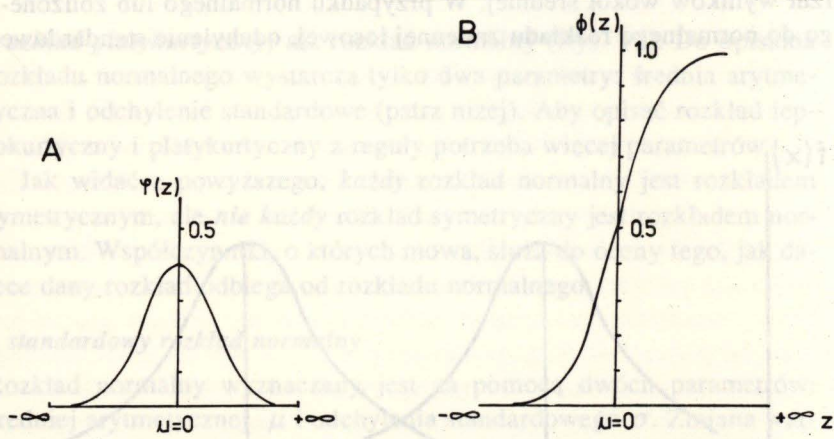
Odchylenie standardowe jest miarą rozproszenia wyników wokół średniej arytmetycznej. Im jego wartość jest większa, tym większy jest rozrzut wyników wokół średniej. W przypadku normalnego lub zbliżonego do normalnego rozkładu zmiennej losowej, odchylenie standardowe



Rys. 20. Rozkłady normalne dla różnych wartości średniej μ i dla $\sigma_1 = \sigma_2$; $f(x)$ – funkcja gęstości prawdopodobieństwa; x – wartości jakie może przyjmować zmienna losowa X .

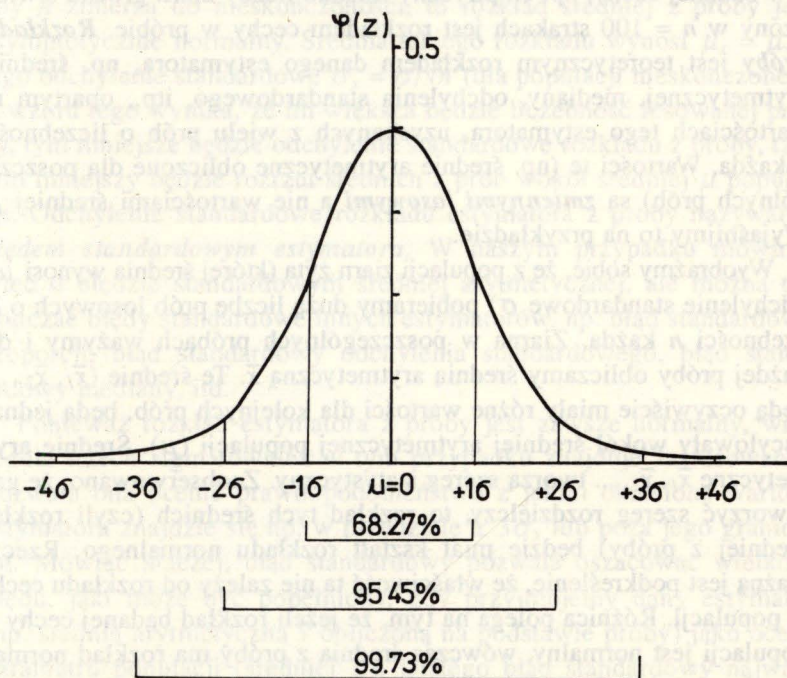


Rys. 21. Rozkłady normalne dla różnych wartości σ i dla tej samej wartości średniej μ ; $f(x)$ – funkcja gęstości prawdopodobieństwa; x – wartości jakie może przyjmować zmienna losowa X .



Rys. 22. Standardowy rozkład normalny (A) i jego dystrybuanta (B). Gęstość prawdopodobieństwa standardowego rozkładu normalnego ma najwyższą wartość w punkcie 0 i wynosi ona 0.3989. Rozkład ten symbolicznie zapisujemy w postaci $N(\mu, \sigma)$ lub $N(0, 1)$. $\phi(z)$ – funkcja gęstości prawdopodobieństwa; $\Phi(z)$ – dystrybuanta.

wykazuje interesującą i ważną właściwość, zwaną **regułą trzech odchyłeń standardowych** lub **regułą trzech sigm**. Reguła ta mówi, że w przedziale od $\mu - \sigma$ do $\mu + \sigma$ ($-\sigma < \mu < +\sigma$) mieści się 68.27% wszystkich wyników, w przedziale od $\mu - 2\sigma$ do $\mu + 2\sigma$ ($-2\sigma < \mu < +2\sigma$) mieści się 95.45% wszystkich wyników, a w przedziale od $\mu - 3\sigma$ do $\mu + 3\sigma$ ($-3\sigma < \mu < +3\sigma$) mieści się 99.73% wszystkich wyników, czyli, praktycznie biorąc, przedział ten obejmuje wszystkie wyniki (Rys. 23). W oparciu o regułę trzech sigm możemy ocenić prawdopodobieństwo tego, że określona wartość zmiennej wystąpi np. w przedziale $\pm 3\sigma$, czy też znajdzie się poza nim. W pierwszym przypadku prawdopodobieństwo to wynosi 0.9973, w drugim tylko 0.0027. To daje nam podstawę do wnioskowania, że z prawdopodobieństwem **praktycznie** równym 1, wartość wylosowana z populacji o rozkładzie normalnym, trafi do przedziału $\pm 3\sigma$, natomiast jest **praktycznie** niemożli-



Rys. 23. Ilustracja reguły trzech odchyłeń standardowych (trzech sigm). Szczegóły w tekście.

we, by wartość ta trafiła poza granice tego przedziału. Gdyby jednak taki fakt zaistniał, moglibyśmy podejrzewać, że wartość ta została wylosowana z innej populacji. Taki z grubsza schemat rozumowania tkwi u podstawy wnioskowania statystycznego. Na ogół nie interesujemy się jednak zmiennością indywidualnych wartości cechy, lecz zmiennością estymatorów, czyli statystyk służących do estymacji nieznanych parametrów populacji. Dlatego reguła trzech sigm znajduje zastosowanie przede wszystkim do oceny błędu standardowego estymatora.

– rozkład z próby, błąd standardowy estymatora, centralne twierdzenie graniczne

Termin “rozkład estymatora z próby” lub krótko “rozkład z próby” jest najbardziej podstawowym pojęciem wnioskowania statystycznego. Terminu tego nie należy mylić z terminem “rozkład w próbie”.

Rozkład w próbie to empiryczny rozkład zaobserwowanych wariantów badanej cechy, np. empiryczny rozkład liczby ziarn grochu stwierdzony w $n = 100$ strąkach jest rozkładem cechy w próbie. **Rozkład z próby** jest teoretycznym rozkładem danego estymatora, np. średniej arytmetycznej, mediany, odchylenia standardowego, itp., opartym na wartościach tego estymatora, uzyskanych z wielu prób o liczebności n każda. Wartości te (np. średnie arytmetyczne obliczone dla poszczególnych prób) są **zmiennymi losowymi** a nie wartościami średniej μ . Wyjaśnijmy to na przykładzie.

Wyobraźmy sobie, że z populacji ziarn żyta (której średnia wynosi μ a odchylenie standardowe σ) pobieramy dużą liczbę prób losowych o liczebności n każda. Ziarna w poszczególnych próbach ważymy i dla każdej próby obliczamy średnią arytmetyczną \bar{x} . Te średnie ($\bar{x}_1, \bar{x}_2, \dots$) będą oczywiście miały różne wartości dla kolejnych prób, będą jednak oscylowały wokół średniej arytmetycznej populacji (μ). Średnie arytmetyczne $\bar{x}_1, \bar{x}_2, \dots$ tworzą szereg statystyczny. Zaobserwowano, że gdy utworzyć szereg rozdzielczy, to rozkład tych średnich (czyli rozkład średniej z próby) będzie miał kształt rozkładu normalnego. Rzeczą ważną jest podkreślenie, że właściwość ta nie zależy od rozkładu cechy w populacji. Różnica polega na tym, że jeżeli rozkład badanej cechy w populacji jest normalny, wówczas średnia z próby ma rozkład normalny **niezależnie** od wielkości próby. Jeżeli, natomiast, rozkład badanej

cechy w populacji nie jest normalny, wtedy rozkład średniej z próby zbliża się do rozkładu normalnego w miarę jak wzrasta liczebność prób n .

Prawidłowość przybliżenia rozkładu średniej z próby do rozkładu normalnego wraz ze wzrostem liczebności próby została uogólniona w twierdzeniu A. M. Lapunowa, w tzw. *centralnym twierdzeniu granicznym*. Twierdzenie to uchodzi za najważniejsze z wielu ważnych tzw. twierdzeń granicznych. Ma ono podstawowe znaczenie w statystyce, uzasadnia bowiem, z punktu widzenia teorii, stosowanie metod opartych na krzywej normalnej.

Ogólnie biorąc, twierdzenie to można sformułować słownie w następujący sposób: Jeżeli dana zmienna losowa jest sumą dużej liczby niezależnych zmiennych losowych o *dowolnych rozkładach*, przy czym żadna z tych zmiennych nie ma dominującego wpływu na wielkość tej sumy, to rozkład tej zmiennej zmierza do rozkładu normalnego, gdy liczebność próby n zmierza do nieskończoności. Można udowodnić, że gdy n zmierza do nieskończoności, to rozkład średniej z próby jest asymptotycznie normalny. Średnia takiego rozkładu wynosi $\mu_x = \mu$, a jego odchylenie standardowe $\sigma_x = \sigma/\sqrt{n}$ (dla populacji nieskończonej). Z wzoru tego wynika, że im większa będzie liczebność losowanej próby, tym mniejsze będzie odchylenie standardowe rozkładu z próby, tzn. tym mniejszy będzie rozrzut średnich z prób wokół średniej μ populacji. Odchylenie standardowe rozkładu estymatora z próby nazywamy *błędem standardowym estymatora*. W naszym przypadku mówimy więc o błędzie standardowym średniej arytmetycznej, ale można też obliczać błędy standardowe innych estymatorów, np. błąd standardowy proporcji, błąd standardowy odchylenia standardowego, błąd standardowy mediany, itd.

Ponieważ rozkład estymatora z próby jest zawsze normalny, więc reguła trzech sigm również w tym przypadku znajduje zastosowanie. Pozwala ona ocenić prawdopodobieństwo z jakim określona wartość estymatora znajdzie się np. w przedziale $\pm 3\sigma_x$ lub poza jego granicami. Mówiąc inaczej, błąd standardowy pozwala oszacować wielkość błędu, jaki może być popełniony, gdy przyjmujemy dany estymator (np. średnią arytmetyczną \bar{x} obliczoną na podstawie próby) jako ocenę parametru populacji (średniej μ). Dlatego błąd standardowy najważniejsze znaczenie posiada w \rightarrow *estymacji przedziałowej*.

4. DOŚWIADCZALNICTWO

PODSTAWOWE TERMINY

Obserwacja, doświadczenie

W powszechnym użyciu są dwie metody badawcze: obserwacja i doświadczenie. *Obserwacją* nazywamy stwierdzenie, jakie wartości przybierają cechy badanego obiektu. Wartość badanej cechy nazywamy *wynikiem*. Zbiór wyników to zbiór *danych*. Obserwacją jest na przykład zbadanie, jaki jest ciężar kłosa żyta. Wiadomo, że każda roślina żyta ma określoną wartość tej cechy, tzn. ma odmienny ciężar kłosa. Co powoduje tę zmienność, tego powiedzieć z całą pewnością nie możemy, gdyż w grę wchodzi bardzo wielka liczba *czynników* takich, jak np. wilgotność, nasłonecznienie, temperatura, skład chemiczny gleby, itp. Przyczynę zaobserwowanej zmienności można ustalić w różny sposób. Na drodze obserwacji można wytypować kilka czynników, o których przypuszcza się, że mogą mieć wpływ na ciężar kłosa i zbadać związek między tymi czynnikami a badaną cechą. Można też zbadać wpływ tych czynników na ciężar kłosa żyta za pomocą *doświadczenia* (eksperymentu). Nie ma ścisłego rozgraniczenia między doświadczeniem a obserwacją, lecz tutaj przyjmiemy, że w doświadczeniu obserwator zmienia niektóre warunki obserwacji i w ten sposób wpływa na jej wynik. W dalszej części niniejszego opracowania zapoznamy się ze sposobami zakładania doświadczeń. Zanim to jednak nastąpi, należy przyswoić sobie kilka niezbędnych pojęć związanych z doświadczalnictwem.

Jednostka eksperymentalna

Próba wylosowana z populacji składa się z części składowych tej populacji. Te części składowe mogą być *jednoelementowe* (np. przedmiot, człowiek, roślina, zwierzę, poletko) lub *wieloelementowe* (np. grupy przedmiotów, grupy ludzi – rodziny, grupy zwierząt – mioty, grupy roślin – kępy, grupy poletek). Część składową próby, wobec której stosu-

je się zabieg nazywamy *jednostką eksperymentalną*. Tak więc próba, to zbiór jednostek eksperymentalnych.

Zabieg

Zabieg (traktowanie, obiekt eksperymentalny), to dowolna operacja wykonana na jednostkach eksperymentalnych próby. Przykładem zabiegu może być rodzaj stosowanego nawożenia, rodzaj użytej do skarmiania paszy, metoda uprawy roli, metoda hodowli, metoda nauczania, rodzaj stosowanego leku, oddziaływanie określonych temperatur na wzrost roślin, itp. Można również za rodzaj zabiegu uznać stosowany w badaniach gatunek, odmianę, rasę rośliny lub zwierzęcia, typ gleby, troficzność wody, ekspozycję, itp. Są to jak gdyby zabiegi przeprowadzone przez przyrodę, niezależnie od człowieka. Rola badacza sprowadza się w tym wypadku do zebrania wyników tych naturalnych eksperymentów.

W najprostszym przypadku zabiegi odpowiadają różnym *natężeniom* (poziomom, wariantom) badanego czynnika. Na przykład, jeżeli badamy wpływ nawożenia azotowego (jeden czynnik) na plon żyta i w tym celu rośliny te zasilamy nawozem azotowym w ilości 0 g, 0,5 g i 1 g nawozu na wazon (trzy natężenia badanego czynnika), to znaczy, że wobec tych roślin stosujemy trzy zabiegi. Doświadczenie polega więc na tym, że wobec jednostek eksperymentalnych stosujemy jeden lub wiele zabiegów, badając wpływ jednego lub wielu czynników na daną jednostkę eksperymentalną. Czynniki, które nie są przedmiotem eksperymentu, mogą być *kontrolowane* (stałe) lub *niekontrolowane* przez badacza.

Przykład 3. Badamy wpływ nawożenia NPK na ciężar kłosów żyta. Rośliny żyta hodujemy w 3 skrzynkach. Każda skrzynka reprezentuje więc jedną próbę, natomiast rośliny pełnią rolę jednostek eksperymentalnych. Do każdej skrzynki dodajemy inną dawkę nawozu. Pozostałe czynniki, takie jak: wilgotność, temperatura, fotoperiod, itp., których wpływu na ciężar kłosów nie badamy, mogą być kontrolowane, jeżeli doświadczenie przeprowadzamy w warunkach ściśle kontrolowanych (np. w komorze klimatyzowanej) lub nie kontrolowane, jeżeli doświadczenie przeprowadzamy w warunkach polowych.

Model stały, losowy, mieszany

Jeżeli poziomy badanych czynników zostały wybrane przez eksperymentatora w sposób arbitralny (tzn. jest on zainteresowany właśnie tymi, a nie innymi poziomami badanych czynników), wówczas mówimy, że doświadczenie zostało zaplanowane według *modelu stałego*. W takim przypadku statystyczne wnioski dotyczą tylko wybranych przez badacza poziomów i nie mogą być uogólniane na inne poziomy tych czynników.

Jeżeli poziomy badanych czynników zostały wylosowane ze zbioru wielu możliwych poziomów, wówczas mówimy, że doświadczenie zostało założone według *modelu losowego*. Ponieważ w tym przypadku badane poziomy stanowią losowy wybór wielu możliwych poziomów, więc wnioski statystyczne mogą być uogólnione na wszystkie poziomy, spośród których wylosowane zostały poziomy badane w danym doświadczeniu.

Jeżeli poziomy jednego czynnika zostały wybrane w sposób arbitralny, a drugiego w sposób losowy, wówczas mamy do czynienia z *modelem mieszanym*. Również ten fakt determinuje zakres wniosków statystycznych.

Doświadczenie jednoczynnikowe i wieloczynnikowe

Eksperyment można zakładać jako tzw. doświadczenie jednoczynnikowe lub doświadczenie wieloczynnikowe. *Doświadczenie jednoczynnikowe* (klasyczne), to takie doświadczenie, w którym badamy wpływ tylko jednego czynnika wobec jednostek eksperymentalnych, przy czym badany czynnik może występować w różnych natężeniach (poziomach, wariantach). *Doświadczenie wieloczynnikowe* (czynnikowe, faktorowane, kombinowane), to takie doświadczenie, w którym jednocześnie bada się wpływ dwu lub więcej czynników wobec jednostek eksperymentalnych, przy czym badane czynniki mogą występować w różnych natężeniach. W doświadczeniu czynnikowym badane są wszelkie możliwe *kombinacje* poziomów testowanych czynników. Kombinacje te pełnią, w tym wypadku, rolę zabiegów.

Przykład 5. Badamy wpływ azotu i gęstości siewu na plon żyta. Stosujemy więc dwa czynniki. Pierwszy o trzech poziomach, które umownie oznaczamy jako poziom 0, 1 i 2, oraz drugi o dwóch poziomach: 0, 1. Poziomy obu zabiegów łączymy ze sobą według zasady "każdy z każ-

dym", tak jak to przedstawiono w Tab. 7. Gdy badamy 3 poziomy nawożenia azotowego i 2 poziomy gęstości siewu, otrzymuje się $3 \times 2 = 6$ kombinacji, a doświadczenie czynnikowe tego typu nazywa się doświadczeniem typu 3×2 .

Tab. 7. Przykład doświadczenia czynnikowego typu 3×2 . Badany jest wpływ dwóch czynników (azotu i gęstości siewu) na plon żyta. Uwzględniono trzy poziomy nawożenia azotowego (0, 1, 2) oraz dwa poziomy gęstości siewu (0, 1).

Poziomy nawożenia azotowego	Poziomy gęstości siewu	Kombinacje (zabiegi)
0	0	00
1	0	10
2	0	20
0	1	01
1	1	11
2	1	21

Doświadczenie kontrolne

Aby stwierdzić, czy stosowane zabiegi mają pewien wpływ na badaną cechę czy też nie, w każdym doświadczeniu musimy przewidzieć tzw. **kontrolę**. Kontrola lub doświadczenie kontrolne, to nic innego jak wydzielone losowo jednostki eksperymentalne, wobec których nie stosuje się żadnych zabiegów, a raczej – stosuje się zabieg o natężeniu zero.

Ważności doświadczenia kontrolnego nie sposób przecenić. Wystarczy, jak pisze Wilson (1968), zastanowić się nad następującym twierdzeniem: "Setki doświadczeń wykazały dowodnie, że bicie w tam-tamy przywraca świecenie słońca po zaćmieniu".

Warto jeszcze podkreślić, że zaplanowanego eksperymentu nie wolno przeprowadzać "na raty". To znaczy, nie wolno np. badać wpływu wybranych zabiegów w innym terminie niż przeprowadzony został eksperyment kontrolny: **całe** doświadczenie musi być przeprowadzone w takich samych warunkach i w jednym czasie.

Replikacja a powtórzenie

Jeżeli będziemy ważyć kłosa żyta, to okaże się, że ciężar danego kłosa (jednostki eksperymentalnej) będzie zawsze nieco inny niż ciężar dru-

giego kłosa. Każdy człowiek inaczej reaguje na tabletkę od bólu głowy. U jednego ból ustąpi po 15 minutach, u drugiego po pół godzinie, u innego w ogóle nie ustąpi. Z tego widać, że badaną cechę (ciężar kłosa, czas reakcji na tabletkę przeciwbólową, itp.) cechuje pewna zmienność, o której już wcześniej była mowa. Ta zmienność cechy *musi* być uwzględniona w analizie statystycznej, aby można było właściwie, tzn. z odpowiednim prawdopodobieństwem, ocenić uzyskane wyniki. Oczywiście jest więc, że nie można ograniczyć się do obserwacji jednego okazu danej rośliny, ani przeprowadzić doświadczenia na jednym króliku, gdyż w ten sposób nie uzyskamy *żadnej* informacji na temat zmienności badanej cechy i, co za tym idzie, wysnute wnioski będą bez wartości. Zabezpieczamy się przed taką ewentualnością stosując *replikacje*. Replikacja to wielokrotne wykonanie tego samego zabiegu lub wielokrotne przeprowadzenie tej samej obserwacji względem kolejnych jednostek eksperymentalnych, stanowiących daną próbę. Jeżeli więc próba liczy pięć jednostek eksperymentalnych i wobec nich (w tych samych warunkach) stosuje się ten sam zabieg, to mamy do czynienia z pięcioma replikacjami.

Oprócz replikacji wyróżnia się również *powtórzenia*. Powtórzenie polega na *ponownym* przeprowadzeniu całego eksperymentu w innym terminie i/lub w innym miejscu. Rozróżnienie między replikacją i powtórzeniem jest bardzo ważne w przypadku opracowywania wyników niektórych typów doświadczeń (por. doświadczenia wielokrotne, Okta-
ba 1971).

Przykład 4. W doświadczeniu wazonowym badamy wpływ nawożenia azotowego (jeden czynnik) na plon żyta. Jeżeli jednostką eksperymentalną jest jeden wazon z 10 roślinami i wobec niego zastosujemy nawożenie, to uzyskamy jeden wynik. Jeżeli doświadczenie zaplanujemy z 5 wazonami (5 jednostek eksperymentalnych), każdy po 10 roślin i wobec nich zastosujemy to samo nawożenie (ten sam zabieg), to otrzymamy 5 wyników, odpowiadających pięciu replikacjom. Gdyby drugi wariant doświadczenia przeprowadzić w jednym roku, a w następnym roku przeprowadzić identyczny eksperyment, to będziemy mieli do czynienia z powtórzeniem. Żeby uzyskać powtórzenia, doświadczenie to można też przeprowadzić w jednym roku, ale na przykład w czterech różnych stacjach doświadczalnych.

SCHEMATY DOŚWIADCZALNE

Wyróżnić można dwie główne metody zakładania doświadczeń – *metodę zmiennych połączonych* (zależnych, związanych) oraz *metodę zmiennych niepołączonych* (niezależnych, niezwiązanych)⁸. Metody te będą omówione szczegółowo niżej. W tym miejscu zwrócimy tylko uwagę na zasadnicze, różniące je cechy.

W przypadku *metody zmiennych połączonych*, oprócz głównego źródła zmienności, jakim jest badany czynnik, uwzględnia się *dotatkowe* źródło zmienności – *niejednorodność* materiału doświadczalnego. Zakładamy, mianowicie, że w materiale tym można wydzielić różniące się między sobą grupy, z których każda złożona jest z takich samych (identycznych) jednostek eksperymentalnych. Grupy te nazywa się *blokami*. Na przykład, ze względu na płeć, materiał doświadczalny można podzielić na dwa bloki: samice i samce. Bloki te wydzielamy dlatego, ponieważ podejrzewamy, że osobniki odmiennej płci będą w różny sposób reagowały na badany czynnik, co może zniekształcić wynik eksperymentu. Wydzielając wspomniane wyżej bloki możemy wyeliminować zakłócający czynnik, którym – jak podejrzewamy – jest płeć.

Uogólnieniem metody zmiennych połączonych na więcej niż dwa zabiegi jest *układ bloków kompletnie zrandomizowanych*, natomiast jego, z kolei, rozwinięciem jest *układ kwadratu tacińskiego*. W przypadku tego układu zakładamy, że oprócz głównego źródła zmienności istnieją jeszcze dwa dodatkowe źródła, sprawiające, że materiał eksperymentalny jest niejednorodny. Uwzględniając je w schemacie eksperymentalnym możemy wyeliminować ich zakłócający efekt na wynik doświadczenia.

W przypadku *metody zmiennych niepołączonych* zakłada się, że mamy do dyspozycji *jednorodny* materiał eksperymentalny. Jedyńm źródłem zmienności jest więc tylko badany czynnik. Uogólnieniem tej metody na więcej niż dwa zabiegi jest *układ kompletnej randomizacji*.

⁸ Zakładając zależność lub niezależność zmiennych, w konsekwencji zakłada się zależność lub niezależność prób. Można więc również mówić o *próbach połączonych* (zależnych) i *niepołączonych* (niezależnych).

Doświadczenia jednoczynnikowe i wieloczynnikowe, założone według układu kompletnej randomizacji, bloków kompletnie zrandomizowanych lub kwadratu łacińskiego jako model stały, losowy lub mieszany, analizowane są za pomocą tzw. *analizy wariancji*. Mówiąc bardzo ogólnie, jest to metoda służąca do porównywania więcej niż dwóch prób (poddanych różnym zabiegom) ze względu na średnią wartość badanej cechy. W niniejszym opracowaniu zaprezentowane będą wersje analizy wariancji przeznaczone do opracowania wyników doświadczenia jednoczynnikowego, zaplanowanego według schematu kompletnej randomizacji, bloków kompletnie zrandomizowanych i kwadratu łacińskiego jako model stały (patrz rozdział 7.C.). Więcej informacji na temat innych schematów doświadczalnych niż tu wspomniane, doświadczeń wieloczynnikowych, analizy wariancji dla modelu stałego, losowego i mieszanego można znaleźć gdzie indziej (np. Ahrens 1970, Elandt 1964, Gawęcki i Wagner 1984, Oktaba 1971, 1980, Steczkowski i Zeliaś 1982, Żuk 1989).

Metoda zmiennych połączonych

Metodę zmiennych połączonych stosujemy wówczas, gdy mamy do dyspozycji *niejednorodny materiał* eksperymentalny, tzn. gdy jednostki eksperymentalne, lub ich grupy, różnią się między sobą. Mówiąc inaczej, jednostki eksperymentalne podlegają klasyfikacji ze względu na cechę, o której przypuszczamy, że może mieć wpływ na wynik stosowanych zabiegów. Taką cechą może być na przykład gatunek, rasa, płeć, rodzaj gleby, marka samochodu, kolor skóry, masa ciała, pochodzenie społeczne, itp. Jak już wyżej wspomniano, wyróżnione w ten sposób kategorie (grupy) jednostek eksperymentalnych nazywamy *blokami*. Jeżeli blok składa się z dwu lub więcej (≥ 2) jednostek eksperymentalnych, to w obrębie danego bloku jednostki te powinny być bliźniaczo podobne. Między blokami, natomiast, powinny zaznaczać się wyraźne różnice. Ponieważ drugą cechą, według której klasyfikuje się jednostki eksperymentalne są *zabiegi*, (które mogą być również traktowane jako próby) mówimy więc, że materiał eksperymentalny podlega *klasyfikacji podwójnej* (krzyżowej).

Jeżeli blok składa się z *jednej* jednostki eksperymentalnej, to *tej samej* jednostce przyporządkowuje się *różne* zabiegi (Tab. 8). W ten spo-

Tab. 8. Schemat doświadczenia założonego metodą zmiennych połączonych: każdy blok składa się z jednej jednostki eksperymentalnej, a , b , lub c , która najpierw poddawana jest zabiegowi A , a następnie zabiegowi B . Na przykład: krowy a , b , i c , pochodzące z różnych ras, karmione są najpierw paszą A , potem paszą B . Za każdym razem badana jest reakcja krowy, np. dzienna produkcja mleka, na daną paszę. Na podstawie uzyskanych wyników ocenia się, czy badane pasze mają istotnie różny wpływ na mleczność krów czy też nie.

Blok	Jednostki eksperymentalne	Zabiegi	
		A	B
1	a	aA	aB
2	b	bA	bB
3	c	cA	cB

Tab. 9. Schemat doświadczenia założonego metodą zmiennych połączonych: każdy blok składa się z dwóch jednostek eksperymentalnych, z których jedna poddawana jest zabiegowi A , a druga zabiegowi B . Jednostki w bloku są względem siebie bliźniaczo podobne, natomiast między parami jednostek (blokami) występują różnice. Na przykład: krowy a_1 i a_2 pochodzą z tej samej rasy (tak samo krowy b_1 i b_2 oraz c_1 i c_2), ale krowy oznaczone jako a , b , i c pochodzą z różnych ras. Zabiegi przyporządkowywane są losowo krowom danej pary.

Blok	Jednostki eksperymentalne		Zabiegi	
			A	B
1	a_1	a_2	a_1A	a_2B
2	b_1	b_2	b_2A	b_1B
3	c_1	c_2	c_1A	c_2B

sób jest ona zarazem swoją własną kontrolą. Jest to **zasadnicza cecha metody zmiennych połączonych**.

Jeżeli blok składa się z **dwóch** jednostek eksperymentalnych (Tab. 9), to, jak wspomniano wcześniej, jednostki te powinny być bliźniaczo do siebie podobne. W praktyce wygląda to w ten sposób, że doбира się je **parami**, na przykład, osobniki jednego wieku, kondycji, płci, rozmiarów, gatunku, itd. W podobny sposób można dobrać pary identycznych poletek, wazonów, pożywek, itp. Danej parze bliźniaczych jednostek przyporządkowuje się, w sposób losowy, odpowiednio dwa zabiegi. W tym przypadku, kontrolą jest jedna z jednostek ekspe-

rymentalnych w bloku (parze). Ponieważ są to zawsze bliźniacze jednostki eksperymentalne, więc zasada wykorzystania danej jednostki eksperymentalnej jako własnej kontroli jest zachowana. Tym niemniej, tam gdzie jest to możliwe należy preferować metodę wykorzystywania jednostki eksperymentalnej jako własnej kontroli, nad metodą łączenia jednostek w pary, ponieważ nigdy nie jesteśmy w stanie tak dobrać dwóch osobników, aby były one podobne do siebie w 100%. Taką gwarancję może dać tylko jeden osobnik wykorzystany zarówno jako przedmiot zabiegu oraz jako jego własna kontrola.

- bloki kompletnie zrandomizowane

Uogólnieniem metody zmiennych połączonych na więcej niż dwa zabiegi (próby) jest schemat doświadczalny określany jako **metoda bloków kompletnie zrandomizowanych** (metoda bloków losowych). W tym wypadku blok musi składać się z tylu bliźniaczych jednostek eksperymentalnych, ile przewidujemy zabiegów (Tab. 10). Każdej jed-

Tab. 10. Schemat doświadczenia założonego metodą bloków kompletnie zrandomizowanych, czyli metodą zmiennych połączonych, uogólnioną na więcej niż dwa zabiegi. Każdy blok składa się z takiej samej liczby jednostek eksperymentalnych, które są bliźniaczo podobne względem siebie, ale różnią się od jednostek eksperymentalnych z sąsiednich bloków. Na przykład, krowy w danym bloku pochodzą z tej samej rasy, natomiast krowy tworzące sąsiednie bloki zostały wylosowane z różnych ras. Zabiegi przyporządkowane są w sposób losowy krowom danego bloku. Liczba zabiegów równa jest liczbie krow w bloku.

Blok	Jednostki			Zabiegi		
				A	B	C ...
1	a_1	a_2	$a_3 \dots$	a_1A	a_2B	$a_3C \dots$
2	b_1	b_2	$b_3 \dots$	b_3A	b_2B	$b_1C \dots$
3	c_1	c_2	$c_3 \dots$	c_2A	c_1B	$c_3C \dots$

nostce eksperymentalnej przypisuje się losowo jeden z zabiegów. Losowanie (randomizację) wykonuje się oddzielnie dla każdego bloku. Jeżeli dysponujemy dużą liczbą identycznych jednostek eksperymentalnych, to dla każdego bloku losujemy tyle grup jednostek, ile jest zabiegów.

Liczba jednostek eksperymentalnych w poszczególnych grupach

Tab. 11. Przykłady danych ortogonalnych. A – Dla każdego bloku wylosowano 5 grup jednostek eksperymentalnych (tyle ile jest zabiegów), każda po 5 jednostek eksperymentalnych. W ten sposób liczba obserwacji (wyników) w poszczególnych kratkach tabeli, utworzonych na przecięciu zabiegów z blokami jest zawsze taka sama i wynosi 5 obserwacji. B – Liczby jednostek eksperymentalnych, a tym samym i liczby wyników w poszczególnych kratkach tabeli, są wielokrotnościami liczb zabiegu A w kierunku poziomym. C – Liczby jednostek eksperymentalnych, a tym samym i liczby wyników w poszczególnych kratkach tabeli, są wielokrotnościami bloku 1 w kierunku pionowym.

		A				
		Zabiegi				
Bloki		A	B	C	D	E
1		5	5	5	5	5
2		5	5	5	5	5
3		5	5	5	5	5

		B				
		Zabiegi				
Bloki		A	B	C	D	E
1		1	3	2	4	7
2		4	12	8	20	16
3		2	6	8	4	10

		C				
		Zabiegi				
Bloki		A	B	C	D	E
1		1	2	6	5	4
2		4	4	18	15	8
3		3	6	12	10	12

musi być jednakowa lub musi być proporcjonalną krotnością w kierunku poziomym lub pionowym (Tab. 11).

O takim materiale eksperymentalnym mówimy, że jest *ortogonalny*. We wszystkich pozostałych przypadkach mamy do czynienia z materiałem eksperymentalnym *nieortogonalnym*. Informacja o tym, czy dane są czy też nie są ortogonalne ma istotne znaczenie, ponieważ od tego zależy dalsze opracowywanie danych liczbowych i wnioskowanie.

Podstawowe metody statystyczne analizowania danych podlegających klasyfikacji podwójnej zostały opracowane dla danych ortogonalnych.⁹

Czasami zdarza się, że jeden lub kilka wariantów doświadczenia założonego według schematu bloków losowych ulegnie zniszczeniu i będzie brakować wyników. W takim wypadku brakujące wyniki można szacować za pomocą metody opisanej np. przez Puchalskiego (1980).

Metoda bloków kompletnie zrandomizowanych bardzo często stosowana jest w doświadczeniach polowych. Jest ona typowym schematem doświadczalnym, który pozwala wyeliminować zmienność gleby w kierunku bloków. Za pomocą tej metody można też zakładać doświadczenia laboratoryjne, na przykład wazonowe. W doświadczeniu polowym rolę bloków grają pasy ziemi dobrane tak, aby zmienność gleby między nimi była jak największa, natomiast między poletkami (jednostkami eksperymentalnymi) w obrębie bloku jak najmniejsza (Rys. 24).

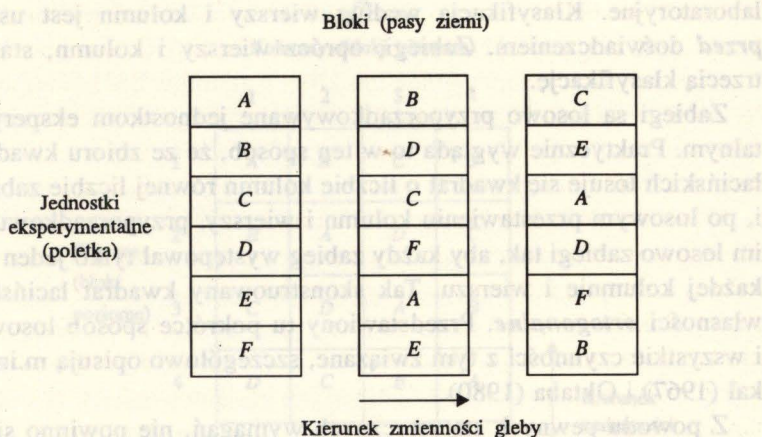
W doświadczeniu wazonowym rolę bloków grać mogą ławy, na których ustawia się jednakowe wazon-y. Jeżeli wszystkie wazon-y wypełnione są jednakową glebą, wtedy zmienność między blokami (ławami) można uzyskać lokując je na różnych stanowiskach, warunkujących np. różne warunki świetlne czy termiczne. Zmienność tę można również uzyskać różnicując glebę stosowaną do napełniania wazonów stojących na różnych ławach, tzn. wazon-y stojące na danej ławie napełnia się inną glebą niż wazon-y stojące na innych ławach.

Metodą bloków kompletnie zrandomizowanych można zakładać doświadczenia czynnikowe typu $2 \times 2 = 2^2$ (dwa czynniki, każdy o dwóch poziomach = 4 kombinacje), $2 \times 2 \times 2 = 2^3$ (trzy czynniki, każdy o dwóch poziomach = 8 kombinacji) i $2 \times 2 \times 2 \times 2 = 2^4$ (cztery czynniki, każdy o dwóch poziomach = 16 kombinacji). Doświadczenie typu 2^5 (5 czynników, każdy o 2 poziomach) ma już 32 kombinacje i powinno być zakładane według innych metod.

- kwadrat łaciński

Dalszym rozwinięciem układu bloków losowych jest *metoda kwadratu łacińskiego*. Metodę tę stosuje się zwykle w doświadczeniach polowych wówczas, gdy nie mamy informacji na temat zmienności gleby danego

⁹ Statystyczne metody analizowania danych nieortogonalnych omawia m. in. Oktaba (1971) i Żuk (1989).



Rys. 24. Przykład doświadczenia polowego założonego według metody bloków kompletnie zrandomizowanych. Badamy wpływ odmiany A, B, ..., F (zabiegi) na plon żyta. W tym celu wyznaczono trzy powierzchnie doświadczalne (bloki), ponieważ przewidziane są trzy replikacje. Bloki uporządkowano zgodnie z wyróżnionym kierunkiem zmienności gleby, np. zgodnie ze wzrastającą żyznością gleby. W obrębie każdego bloku wydzielono tyle poletek, ile jest odmian (zabiegów), a następnie odmiany w blokach rozmieszczono losowo na poszczególnych poletkach. Według tej zasady, można też badać wpływ innych zabiegów, np. wpływ różnych nawozów na plon żyta, itp.

poła i w związku z tym nie możemy uporządkować bloków we właściwy sposób, tzn. zgodnie z kierunkiem tej zmienności.

Aby założyć doświadczenie według schematu kwadratu łacińskiego należy najpierw wytyczyć dwa kierunki – zazwyczaj kierunek orki i kierunek prostopadły do niego. Ten sposób pozwoli nam wyrugować zróżnicowanie gleby w tych kierunkach. Następnie wyznaczamy prostokątną powierzchnię doświadczalną o bokach równoległych do wyróżnionych kierunków i dzielimy ją w każdym z tych dwóch kierunków na bloki poziome zwane *wierszami* (poziome pasy poletek) i na bloki pionowe zwane *kolumnami* (pionowe pasy poletek) (Rys. 25). Liczba wierszy i kolumn w kwadracie łacińskim *musi być jednakowa* i *musi być równa* liczbie zabiegów. Zabiegi te oznacza się literami łacińskimi A, B, C, D, E – stąd nazwa “kwadrat łaciński”. Rolę zabiegów mogą grać odmiany, terminy siewu, gęstość siewu, rodzaje nawożenia, terminy pokosów, sposoby nawadniania, itp. Oprócz doświadczeń polowych, tą metodą można też zakładać doświadczenia produkcyjne i

laboratoryjne. Klasyfikacja według wierszy i kolumn jest ustalana **przed** doświadczeniem. Zabiegi, oprócz wierszy i kolumn, stanowią trzecią klasyfikację.

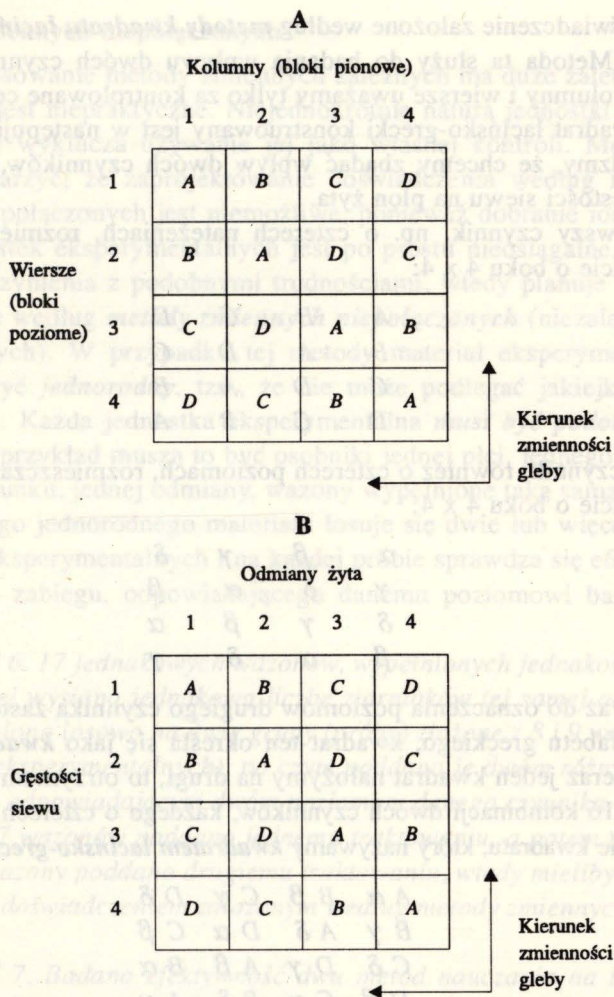
Zabiegi są losowo przyporządkowywane jednostkom eksperymentalnym. Praktycznie wygląda to w ten sposób, że ze zbioru kwadratów łacińskich losuje się kwadrat o liczbie kolumn równej liczbie zabiegów i, po losowym przestawieniu kolumn i wierszy, przyporządkowuje się im losowo zabiegi tak, aby każdy zabieg występował tylko jeden raz w każdej kolumnie i wierszu. Tak skonstruowany kwadrat łaciński ma własności *ortogonalne*. Przedstawiony tu pokrótce sposób losowania, i wszystkie czynności z tym związane, szczegółowo opisują m.in. Perkal (1967) i Oktaba (1980).

Z powodu pewnych statystycznych wymagań, nie powinno się stosować kwadratów mniejszych niż 5×5 , ani większych niż 12×12 . Jeżeli stosujemy kwadraty mniejsze niż 5×5 , wtedy doświadczenie należy replikować kilkakrotnie. W przypadku kwadratu 4×4 doświadczenie należy replikować przynajmniej dwa razy, a kwadraty 3×3 muszą być replikowane przynajmniej pięć razy. Kwadratów 2×2 nie stosuje się w ogóle.

Metoda kwadratów łacińskich przeznaczona jest zasadniczo do doświadczeń *jednoczynnikowych*. Doświadczenia te można przeprowadzać nie kontrolując czynników ubocznych (w takim przypadku kolumny i wiersze reprezentują tylko pasy pionowych i poziomych poletek, por. Rys. 25 A) lub kontrolując wybrane czynniki uboczne (wówczas warianty jednego czynnika są reprezentowane przez kolumny a warianty drugiego czynnika są reprezentowane przez wiersze, por. Rys. 25 B). W tym drugim przypadku doświadczenie założone według metody kwadratu łacińskiego można również traktować jako doświadczenie *dwuczynnikowe* z jedną cechą kontrolowaną, lub nawet jako doświadczenie *trójczynnikowe*.

– kwadrat łacińsko-grecki

Należy pamiętać, że założone metodą kwadratu łacińskiego doświadczenie dwu- lub trójczynnikowe traktujemy tylko jako doświadczenie wstępne, pilotowe. Potem należy jednak przeprowadzić właściwe *doświadczenie wieloczynnikowe*, przeznaczone do badania wpływu kilku czynników na rozważaną cechę. Przykładem takiego doświadczenia



Rys. 25. Przykład doświadczenia założonego według metody kwadratu łacińskiego.

A – przykład doświadczenia, którego celem jest porównanie plonu czterech odmian żyta (A, B, C, D). Badany jest więc jeden czynnik, przy czym żaden czynnik nie jest kontrolowany. **B** – przykład doświadczenia, którego celem jest zbadanie wpływu czterech różnych nawozów na plon czterech odmian żyta, wysianego w czterech różnych zagęszczeniach. W tym przypadku badany jest również jeden czynnik (nawożenie), ale oprócz tego kontrolowane są dwa czynniki: odmiany żyta oraz gęstość siewu. To doświadczenie można jednak traktować również jako doświadczenie dwuczynnikowe (nawożenie w czterech natężeniach i żyto w czterech odmianach) z jedną cechą kontrolowaną (gęstość siewu), albo nawet jako doświadczenie trójczynnikowe – bez cech kontrolowanych.

jest doświadczenie założone według *metody kwadratu łacińsko – greckiego*. Metoda ta służy do badania wpływu dwóch czynników, przy czym kolumny i wiersze uważamy tylko za kontrolowane cechy uboczne. Kwadrat łacińsko-grecki konstruowany jest w następujący sposób. Powiedzmy, że chcemy zbadać wpływ dwóch czynników, nawożenia oraz gęstości siewu na plon żyta.

Pierwszy czynnik, np. o czterech natężeniach, rozmieszczamy w kwadracie o boku 4 x 4:

A	B	C	D
B	A	D	C
C	D	A	B
D	C	B	A

Drugi czynnik, również o czterech poziomach, rozmieszczamy także w kwadracie o boku 4 x 4:

α	β	γ	δ
γ	δ	α	β
δ	γ	β	α
β	α	δ	γ

Ponieważ do oznaczenia poziomów drugiego czynnika zastosowano litery alfabetu greckiego, kwadrat ten określa się jako *kwadrat grecki*. Jeżeli teraz jeden kwadrat nałożymy na drugi, to otrzymamy rozmieszczenie 16 kombinacji dwóch czynników, każdego o czterech poziomach, na planie kwadratu, który nazywamy *kwadratem łacińsko-greckim*:

A α	B β	C γ	D δ
B γ	A δ	D α	C β
C δ	D γ	A β	B α
D β	C α	B δ	A γ

W kwadracie tym każdy poziom obu czynników występuje tylko jeden raz w każdej kolumnie i wierszu, żadna kombinacja (=zabieg) nie powtarza się dwukrotnie, każdy poziom jednego czynnika spotyka się tylko jeden raz z każdym poziomem drugiego czynnika. Oprócz wspomnianych tutaj, istnieją jeszcze kwadraty bardziej złożone. Więcej informacji na ten temat można znaleźć np. u Perkała (1967).

Metoda zmiennych niepołączonych

Chociaż stosowanie metody zmiennych zależnych ma duże zalety, często jednak jest niepraktyczne. Niejednokrotnie natura jednostki eksperymentalnej wyklucza używanie jej jako własnej kontroli. Może się również zdarzyć, że zaprojektowanie doświadczenia według metody zmiennych połączonych jest niemożliwe, ponieważ dobranie identycznych jednostek eksperymentalnych jest po prostu nieosiągalne. Jeżeli mamy do czynienia z podobnymi trudnościami, wtedy planuje się doświadczenie według **metody zmiennych niepołączonych** (niezależnych, niezwiązanych). W przypadku tej metody materiał eksperymentalny powinien być **jednorodny**, tzn., że nie może podlegać jakiegokolwiek klasyfikacji. Każda jednostka eksperymentalna **musi być podobna** do drugiej, na przykład muszą to być osobniki jednej płci, jednego miotu, jednego gatunku, jednej odmiany, wazony wypełnione taką samą glebą, itp. Z takiego jednorodnego materiału losuje się dwie lub więcej prób jednostek eksperymentalnych i na każdej próbie sprawdza się efekt tylko jednego zabiegu, odpowiadającego danemu poziomowi badanego czynnika.

Przykład 6. 17 jednakowych wazonów, wypełnionych jednakową glebą, do której wysiano jednakową liczbę ziarniaków tej samej odmiany żyta, podzielono losowo na dwie grupy (próby) złożone z 8 i 9 wazonów (jednostek eksperymentalnych), po czym poddano je dwóm różnym traktowaniom, odpowiadającym dwóm poziomom danego czynnika. Gdyby wszystkie 17 wazonów poddano jednemu traktowaniu, a potem te same wszystkie wazony poddano drugiemu traktowaniu, wtedy mielibyśmy do czynienia z doświadczeniem założonym według metody zmiennych połączonych.

Przykład 7. Badano efektywność dwu metod nauczania na tym samym kursie (materiał jednorodny ze względu na przeprowadzoną wstępną selekcję słuchaczy). Wszystkich słuchaczy uczestniczących w kursie ponumerowano i potem jedną ich część losowo przyporządkowano jednemu wykładowcy, a drugą część drugiemu wykładowcy.

Jak widać z przytoczonych przykładów, przy stosowaniu tej metody, próby jednostek eksperymentalnych **nie muszą** być równoliczne.

Fakt zastosowania metody zmiennych niepołączonych sprawia, że dane eksperymentalne reprezentują **klasyfikację pojedynczą**, tzn. jed-

nostki eksperymentalne klasyfikowane są tylko ze względu na traktowania odpowiadające poziomom badanego czynnika.

– układ kompletnej randomizacji

Uogólnienie metody zmiennych niepołączonych na więcej niż dwa zabiegi (czyli na więcej niż dwie próby) określane jest jako **układ kompletnej randomizacji**. Metodę randomizacji (czyli losowego przyporządkowywania zabiegów) pomija się wtedy, gdy eksperymentalny materiał z natury podlega klasyfikacji pojedynczej. Dotyczy to na przykład wyników uzyskiwanych dla różnych ras zwierząt, dla różnych miotów, odmian lub gatunków roślin, itp., które w tym wypadku odgrywają rolę poziomów właściwego czynnika.

5. WNISKOWANIE STATYSTYCZNE

W statystyce matematycznej wyróżnia się dwa podstawowe typy wnioskowania statystycznego. Jednym z nich zajmuje się *teoria weryfikacji hipotez statystycznych*. Ten sposób wnioskowania polega na tym, że na podstawie próby rozstrzygamy, czy wysunięta *hipoteza statystyczna* (przypuszczenie dotyczące parametrów lub kształtu rozkładu populacji) jest słuszna, czy też nie. Zespół reguł pozwalających przyjąć lub odrzucić sprawdzaną hipotezę nazywamy *testem statystycznym*. Drugi typ wnioskowania dotyczy sposobu szacowania określonych parametrów populacji, w oparciu o estymatory próby. Problemem tym zajmuje się *teoria estymacji parametrów*.

WERYFIKACJA HIPOTEZ STATYSTYCZNYCH

Obecnie, na prostym przykładzie, zademonstrujemy, na czym polega *weryfikacja hipotezy statystycznej* za pomocą *testu statystycznego*. Na początku zostanie przedstawiony pewien schemat analizy statystycznej, a następnie zostaną omówione poszczególne jej etapy.

Przykład 8. Zamierzamy zbadać, czy średni ciężar kłosa żyta odmiany A różni się istotnie (jest lżejszy lub cięższy) od średniego ciężaru kłosa odmiany B. Ponieważ próby będą pochodziły z dwóch różnych odmian (populacji), będziemy mieli do czynienia z próbami niezależnymi. Próby te będą podlegały klasyfikacji pojedynczej, ponieważ dadzą się sklasyfikować tylko ze względu na jedną cechę – odmiany żyta. Rozważane odmiany żyta można traktować jak dwa zabiegi. Odpowiadają one dwu poziomom jednego czynnika, którego wpływ na ciężar kłosa żyta chcemy zbadać. W takiej sytuacji jak wyżej przedstawiono, do porównania średnich ciężarów kłosa odmiany A i B możemy zastosować albo test z, albo test t. W pierwszym przypadku musimy pobrać dwie losowe próby o liczebności większej niż 30 elementów, ale wówczas, bez dodatkowego badania, możemy założyć, że rozkład cechy w obu populacjach, a tym samym i w obu próbach, jest rozkładem normalnym (prawo wielkich liczb!). W drugim przypadku możemy pobrać małe

próby losowe, ale wtedy, powinniśmy mieć pewność, że próby te zostały wylosowane z populacji o rozkładach normalnych.

Gdyby próby były odpowiednio liczne, za pomocą odpowiednich testów moglibyśmy sprawdzić hipotezę, że rozkład cechy w badanych populacjach jest normalny. Gdy próba jest nieliczna, na ogół zrobić tego nie można, dlatego albo powinniśmy informację na ten temat zaczerpnąć z innego źródła, np. z wcześniej przeprowadzonych badań, lub przyjmujemy założenie o rozkładzie normalnym cechy w populacji, ryzykując, że wnioski wyciągnięte na podstawie wyniku testu t będą fałszywe.

Założmy, że wiemy iż rozkład cechy w populacji jest normalny i że zdecydowaliśmy się na test t . Z odmiany A i B (czyli populacji A i B) pobraliśmy proste próby losowe o liczebności $n_1 = n_2 = 10$ kłosów. Obliczamy średni ciężar kłosa w obu próbach i za pomocą testu t sprawdzamy, czy różnica między średnim ciężarem kłosa odmiany A (\bar{x}_1) i średnim ciężarem kłosa odmiany B (\bar{x}_2) jest istotna (znaczy to, że różnicę między średnimi tych prób można przypisać odmianom), czy też jest nieistotna (znaczy to, że zaobserwowana różnica jest wynikiem przypadku – można ją przypisać na przykład przypadkowemu doborowi takich a nie innych kłosów). Stawiamy więc hipotezę zerową, że różnica między średnimi jest nieistotna. Hipotezę tę będziemy próbowali obalić na korzyść hipotezy alternatywnej, że różnica między średnimi jest istotna. Przyjmujemy poziom istotności $\alpha = 0.05$. Założmy, że w wyniku obliczeń otrzymaliśmy wartość statystyki $t = 0.953$. Liczba stopni swobody dla tego przykładu wynosi $n_1 + n_2 - 2 = 10 + 10 - 2 = 18$. Z tablic rozkładu t Studenta, dla 18 stopni swobody, na poziomie istotności 0.05, odczytujemy wartość krytyczną statystyki $t_{0.05} = 2.101$. Ponieważ wartość empiryczna statystyki t jest mniejsza od jej wartości krytycznej, czyli $t = 0.953 < t_{0.05} = 2.101$, nie ma podstaw do odrzucenia hipotezy zerowej. To pozwala nam na wyciągnięcie wniosku, że obie średnie arytmetyczne, wyliczone dla badanych prób, nie różnią się między sobą w sposób statystycznie istotny, a więc: średni ciężar kłosa żyta odmiany A (populacji A) jest taki sam jak średni ciężar kłosa żyta odmiany B (populacji B).

Etapy analizy statystycznej

Przedstawiony powyżej schemat postępowania jest **zawsze taki sam**, niezależnie od stosowanego testu i testowanych hipotez zerowych.

Składa się on z następujących operacji:

1. sformułowanie problemu badawczego oraz wybór schematu doświadczalnego i przeprowadzenie eksperymentu lub obserwacji;
2. wybór testu statystycznego;
3. postawienie hipotezy zerowej i założenie poziomu istotności;
4. dokonanie niezbędnych obliczeń – obliczamy statystykę testu;
5. odczytanie wartości krytycznej statystyki;
6. porównanie wartości krytycznej statystyki z wartością empiryczną;
7. przyjęcie hipotezy zerowej lub jej odrzucenie na korzyść hipotezy alternatywnej.

Zaplanowanie doświadczenia i, co za tym idzie, wybór odpowiedniego schematu doświadczalnego jest sprawą bardzo ważną, ponieważ w sposób jednoznaczny określa, jaki rodzaj testu będzie mógł być zastosowany do opracowania danych eksperymentalnych lub pochodzących z obserwacji.

Hipoteza zerowa

Kiedy mamy już za sobą przeprowadzenie eksperymentu (lub dokonanie obserwacji) i wybór testu, formułujemy odpowiednią **hipotezę zerową**, którą oznaczamy symbolem H_0 . Hipotezę tę można określić jako hipotezę braku różnicy, zgodności lub związku. Oznacza to, że hipotezę zerową zawsze formułujemy w następujący sposób:

- **nie ma różnicy** między rozważaną statystyką próby i parametrem populacji, z której ta próba została pobrana;
- **nie ma różnicy** między badanymi statystykami;
- **nie ma zgodności** między rozkładem empirycznym i teoretycznym;
- **nie ma zgodności** między badanymi rozkładami empirycznymi;
- **nie ma związku** między badanymi cechami.

W **Przykładzie 8** hipoteza zerowa głosiła, że nie ma istotnej różnicy między średnimi arytmetycznymi, \bar{x}_1 i \bar{x}_2 . Zwróćmy jednak uwagę, że porównując średnie z próby, de facto porównujemy średnie μ_1 i μ_2 , (charakterystyczne dla odmian, czyli populacji), których średnie wyliczone dla prób są jedynie szacunkami (ocenami). Dlatego tak sformułowaną hipotezę zerową notuje się zwykle w następujący sposób:

$$H_0 : \mu_1 - \mu_2 = 0$$

(różnica obu średnich wynosi zero)

lub

$$H_0 : \mu_1 = \mu_2$$

(wartość jednej średniej równa się wartości drugiej średniej),

gdzie μ_1 oznacza średnią arytmetyczną *populacji* odmiany A, z której pobrano próbę n_1 , a μ_2 oznacza średnią arytmetyczną *populacji* odmiany B, z której pobrano próbę n_2 .

Hipoteza alternatywna

Hipotezą przeciwną w stosunku do hipotezy zerowej, jest *hipoteza alternatywna*. Tę hipotezę można określić jako hipotezę istnienia różnicy, zgodności lub związku. Przyjmujemy ją dopiero *po odrzuceniu* hipotezy zerowej.

Alternatywą hipotezy zerowej, sformułowanej wcześniej, jest hipoteza, że obie średnie różnią się między sobą istotnie. Hipotezę tę zapisujemy w sposób następujący:

$$H_1 : \mu_1 \neq \mu_2$$

(wartość jednej średniej jest różna od wartości drugiej średniej).

Ponieważ w tym przypadku stwierdzamy tylko, że średnia μ_1 jest różna od średniej μ_2 , tzn. może być mniejsza lub większa, więc wartości krytyczne statystyki testu odczytuje się z tzw. tablicy *testu dwustronnego*. Jeżeli hipoteza alternatywna brzmi:

$$H_1 : \mu_1 > \mu_2$$

(średnia μ_1 jest większa od średniej μ_2)

lub

$$H_1 : \mu_1 < \mu_2$$

(średnia μ_1 jest mniejsza od średniej μ_2),

wtedy wartości krytyczne odczytujemy z tablicy *testu jednostronnego*. A więc dwustronność czy jednostronność testu zdeterminowana jest rodzajem hipotezy alternatywnej, tzn. czy pytamy tylko o istotność różnicy, czy też jeszcze o jej kierunek.

Testy parametryczne, hipotezy parametryczne

Wiele testów statystycznych opartych jest na założeniu, że rozkład badanej cechy w populacji jest rozkładem normalnym. Aby testy te można było stosować warunek ten *musi być* spełniony. Jeżeli dane nie spełniają tego warunku, wtedy testów tych stosować *nie można*. Wybrnąć z tej sytuacji można w dwojaki sposób:

1. można przekształcić zmienną o rozkładzie nienormalnym na zmienną o rozkładzie normalnym – odnośnie transformacji danych patrz np. Norcliffe (1986: 69), Parker (1978: 106), Steczkowski i Zeliaś (1982: 113);

2. można zastosować test, który nie wymaga założenia normalności rozkładu badanej cechy.

Testy stosowane do badania danych spełniających warunek normalności rozkładu nazywa się *testami parametrycznymi*. Biorąc ściśle jednak, testy parametryczne to wszystkie testy, przy konstrukcji których wymagane jest założenie o postaci rozkładu badanej cechy, np. że rozkład ten jest normalny, dwumianowy czy Poissona. Test parametryczny służy do weryfikowania tzw. *hipotezy parametrycznej*. Jest to hipoteza, która precyzuje wartość parametru populacji o znanym rozkładzie. Hipoteza parametryczna głosi więc na przykład, że średnia arytmetyczna \bar{x} próby jest równa średniej arytmetycznej μ populacji, lub że dwie niezależne próby losowe o średnich \bar{x}_1 i \bar{x}_2 pochodzą z dwóch populacji o identycznych średnich.

Testy nieparametryczne, hipotezy nieparametryczne

Testy stosowane do opracowywania prób pobranych z populacji o dowolnym rozkładzie badanej cechy nazywa się *testami nieparametrycznymi*. Testy te służą do weryfikowania tzw. *hipotez nieparametrycznych*. Są to hipotezy głoszące, że rozkład cechy w populacji jest rozkładem określonego typu, np. rozkładem dwumianowym, Poissona, normalnym, równomiernym itp. Niektóre testy nieparametryczne są alternatywami testów parametrycznych. Moc testów nieparametrycznych (\rightarrow moc testu) jest mniejsza niż moc testów parametrycznych, dlatego aby obalić hipotezę zerową za pomocą testu nieparametrycznego należy posłużyć się próbą o większej liczebności niż w przypadku testu parametrycznego.

Stopnie swobody

Po wyborze testu należy przystąpić do obliczania statystyki tego testu zgodnie z właściwą procedurą. Oprócz statystyki testu, na tym etapie opracowania statystycznego danych, oblicza się również tzw. **stopnie swobody**. Stopnie swobody to liczba niezależnych pomiarów dostępnych do obliczenia statystyki testu. Sens stopni swobody, za przykładem Srba i Owena (1959), można wyjaśnić przez porównanie z sytuacją, kiedy mały chłopiec zakłada buty. "Ma on dwa buty, ale tylko jeden stopień swobody, ponieważ gdy włoży jeden but, właściwy czy niewłaściwy, to drugi but automatycznie staje się także albo pasującym, albo nie pasującym na daną nogę. Podobnie w tabeli o dwóch kolumnach cyfr: jedną wartość można wypełnić dowolnie, ale druga jest uwarunkowana faktem, że ogólna suma musi wynosić określoną liczbę obserwacji objętych danym eksperymentem, a odchylenia w dwóch klasach muszą się równoważyć wzajemnie. Jeżeli natomiast mamy cztery klasy, dowolne trzy są zwykle swobodne, lecz jakiegokolwiek byłyby one, czwarta jest już ustalona. Jeżeli więc mamy cztery klasy, zwykle są tylko trzy stopnie swobody". Mówiąc inaczej, liczba stopni swobody równa jest liczbie wyników w próbie, które mogą przyjąć **dowolną** wartość. W najprostszym przypadku liczba ta wynosi $n - 1$. Nie zawsze tak bywa. Zależy to od stosowanego testu, dlatego przy każdym teście statystycznym zawsze jest podany sposób obliczania stopni swobody.

Poziom istotności

Wartość krytyczną statystyki odczytujemy przy takim **poziomie istotności** α , jaki sobie wcześniej założyliśmy. W badaniach przyrodniczych najczęściej przyjmuje się poziom istotności $\alpha = 0.05$. W naszym przykładzie wartość krytyczna statystyki t odczytana przy tym poziomie istotności i 18 stopniach swobody wynosi $t_{0,05} = 2.101$. Przyjmijmy, że w wyniku obliczeń otrzymano dla statystyki t wartość 2.358. Tym razem empiryczna wartość t jest większa od wartości krytycznej, więc hipotezę zerową odrzucamy i przyjmujemy hipotezę alternatywną, która głosi, że średnie arytmetyczne obu badanych populacji różnią się między sobą w sposób istotny, a stwierdzoną różnicę należy przypisać badanym odmianom żyta, A i B .

– błąd pierwszego rodzaju

Wniosek powyższy wypowiadamy na poziomie istotności $\alpha = 0.05$. Poziom istotności, zwany też *ryzykiem błędu*, określa prawdopodobieństwo popełnienia tzw. *błędu pierwszego rodzaju*. Jest to błąd polegający na odrzuceniu prawdziwej hipotezy zerowej i przyjęciu fałszywej hipotezy alternatywnej. W odniesieniu do rozpatrywanego przykładu oznacza to, że w 5 przypadkach na 100 zbadanych prób (czyli w 5%) możemy się pomylić stwierdzając, że odmiany *A* i *B* różnią się istotnie jeżeli chodzi o średni ciężar kłosa żyta (przyjęcie fałszywej hipotezy alternatywnej), podczas gdy ta różnica faktycznie jest nieistotna (odrzucenie prawdziwej hipotezy zerowej). Błąd ewentualnej pomyłki można zmniejszyć, jeżeli założymy wyższy poziom istotności, na przykład $\alpha = 0.01$ lub $\alpha = 0.001$. Wtedy prawdopodobieństwo pomyłki będzie wynosiło odpowiednio 1% lub 0.1%.

Wybór poziomu istotności zależy od charakteru badań. Jeżeli rozważany problem dotyczy zagadnień przyrodniczych, to jak już powiedziano, zadowalamy się ryzykiem błędu wynoszącym 5% ($\alpha = 0.05$). Inaczej jest jeżeli testujemy na przykład nowy rodzaj leku, od którego zależy zdrowie czy życie ludzkie. Gdy na podstawie analizy statystycznej mamy wydać sąd o przydatności tego specyfiku, to oczywiście poziom istotności $\alpha = 0.05$ już nam nie wystarcza, bo to oznaczałoby w praktyce, że na każdych 100 leczonych tym lekarstwem pacjentów, pięciu, w najlepszym razie, nie odczułoby żadnej poprawy. Z tego powodu, przy tego rodzaju badaniach, staramy się jak tylko można zmniejszyć ryzyko błędu pierwszego rodzaju.

– błąd drugiego rodzaju

Należy zaznaczyć, że poziom istotności (α) ocenia prawdopodobieństwo (procent pomyłki) *tylko* w przypadku *odrzucenia* hipotezy zerowej. Jeżeli hipotezę zerową *przyjmujemy*, musimy liczyć się z możliwością popełnienia tzw. *błędu drugiego rodzaju* β . Błąd ten polega na przyjęciu fałszywej hipotezy zerowej i odrzuceniu prawdziwej hipotezy alternatywnej. W rozważanym przykładzie będzie to oznaczało, że uznamy iż różnica między badanymi średnimi jest nieistotna (przyjęcie fałszywej hipotezy zerowej), podczas gdy w rzeczywistości różnica ta jest istotna (odrzucenie prawdziwej hipotezy alternatywnej). W prze-

ciwieństwie do błędu pierwszego rodzaju, prawdopodobieństwo popełnienia błędu drugiego rodzaju jest *nieokreślone*. Dlatego, gdy brak jest podstaw do odrzucenia hipotezy zerowej, badacz ma prawo przyjąć tę hipotezę, ale czyni to na *własną* odpowiedzialność, ponieważ w takim przypadku procent błędnych decyzji jest nieznanym.

Dla dowolnego testu, prawdopodobieństwa błędów pierwszego i drugiego rodzaju są wzajemnie przeciwstawne, tzn. że dla *danej* liczebności próby n , im mniejsze jest ryzyko popełnienia błędu pierwszego rodzaju, tym większe jest prawdopodobieństwo popełnienia błędu drugiego rodzaju. Jak z tego widać, nie można *jednocześnie* minimalizować ryzyka popełnienia obu błędów. Zmniejszenie prawdopodobieństw α i β można uzyskać zwiększając liczebność próby.

Moc testu

Przy danym poziomie istotności α , tę samą hipotezę zerową można testować za pomocą różnych testów statystycznych. W takim przypadku wybieramy taki test, który charakteryzuje się najmniejszym błędem drugiego rodzaju β . Chodzi o to, aby prawdopodobieństwo odrzucenia hipotezy zerowej było jak największe wówczas, gdy hipoteza alternatywna jest prawdziwa. To prawdopodobieństwo właśnie nazywamy *mocą testu* i wynosi ono $1 - \beta$. Reasumując, spośród różnych testów najmocniejszy jest ten, dla którego (przy ustalonym α) wartość $1 - \beta$ jest największa.

UWAGA: Test statystyczny pozwala tylko na odrzucenie hipotezy zerowej lub na stwierdzenie, że brak jest podstaw do odrzucenia hipotezy zerowej. Wniosek o nie odrzuceniu hipotezy zerowej *nie jest* jednoznaczny z wnioskiem o jej prawdziwości. Również odrzucenie hipotezy zerowej *nie oznacza*, że hipoteza alternatywna została *udowodniona*, ponieważ: W STATYSTYCE NIE MA HIPOTEZ UDOWODNIONYCH, SĄ TYLKO LEPIEJ LUB GORZEJ ZWERYFIKOWANE.

ESTYMACJA PARAMETRÓW STATYSTYCZNYCH

Estymacja punktowa, estymacja przedziałowa

Postępowanie polegające na szacowaniu nieznanego parametru populacji za pomocą estymatora (statystyki próby), nazywamy *estymacją parametru*. W przeciwieństwie do parametru, estymator jest zmienną losową.

Istnieją dwa sposoby estymacji: estymacja punktowa i przedziałowa. **Estymacja punktowa** polega na podaniu jednej liczby odpowiadającej przypuszczalnej wartości parametru. Liczba ta, to wartość estymatora otrzymana z n -elementowej próby. Wartość tę traktuje się jako najlepsze przybliżenie nieznaney wartości parametru. **Estymacja przedziałowa** polega na znajdowaniu przedziału liczbowego, w którym, z zadanyym z góry prawdopodobieństwem $1 - \alpha$, zwanym **poziomem ufności** lub współczynnikiem ufności, znajduje się nieznaną wartość danego parametru populacji, np. średniej arytmetycznej. Przedział ten nazywa się **przedziałem ufności** (\rightarrow szacowanie przedziałów ufności, str. 105).

Poziom ufności zależy od wielkości błędu pierwszego rodzaju α . Jeżeli $\alpha = 0.05$, to poziom ufności wynosi $1 - 0.05 = 0.95$. Oznacza to, że na każde 100 przedziałów wyznaczonych dla 100 prób, w 95 przypadkach prawdziwa wartość parametru populacji znajduje się wewnątrz przedziału, natomiast w 5 przypadkach znajduje się poza przedziałem.

Im mniejsze jest α , tym mniejsze jest prawdopodobieństwo błędnego wyznaczenia przedziału ufności, ale tym większy jest ten przedział i tym mniej precyzyjny.

Między poziomem ufności, przedziałem ufności i błędem standardowym istnieje ścisła zależność. Ponieważ błąd standardowy maleje w miarę jak rośnie n , więc poziom ufności i przedział ufności można kształtować poprzez zmianę liczebności próby n .

6. PRZEGLĄD ZAGADNIENI, KTÓRE MOGĄ BYĆ PRZEDMIOTEM ANALIZY STATYSTYCZNEJ

Testowanie losowości próby

Jak wspomniano wcześniej, sądy o populacji wydaje się na podstawie próby. Aby sądy te były prawdziwe, próba musi być pobrana w sposób losowy. Taka próba jest próbą reprezentatywną dla danej populacji. Za pomocą odpowiednich testów można zbadać, czy pobrana próba ma rzeczywiście charakter losowy. Testy weryfikujące losowy charakter próby mogą być stosowane w przypadku cech mierzalnych (ilościowych) i niemierzalnych (jakościowych). Spełniają one swoje zadanie tylko wtedy, gdy porządek elementów w próbie jest niezmienny, tzn. taki, w jakim próbę pobierano. Niektóre z tych testów mogą być stosowane do weryfikacji hipotezy, że ciąg obserwacji jest wolny od \rightarrow trendu. Testy te są szczególnie użyteczne do wykrywania tendencji i zmian cyklicznych (\rightarrow autokorelacja).

Testowanie jednorodności próby

Równie ważną jak losowość jest też kwestia jednorodności próby. Mówiąc o jednorodności próby mamy na myśli to, że różnice między danymi nie powinny być zbyt duże. Wyobraźmy sobie próbę złożoną z $n = 10$ pomiarów (cm): 5, 30, 31, 35, 41, 43, 44, 47, 72. Jak widać, dwie wartości w tej próbie, największa i najmniejsza, wyraźnie odbiegają od pozostałych. Może to być rezultatem błędnego pomiaru lub spowodowane tym, że zostały zmierzone elementy wylosowane z innej (lub innych) populacji. Fakt ten może mieć istotny wpływ na wartość statystyki próby, co z kolei spowoduje, że wnioski wyciągnięte na podstawie takich niejednorodnych danych będą błędne. W konsekwencji zostaną podjęte niewłaściwe decyzje. Test jednorodności pozwala odpowiedzieć na pytanie, czy dana *próba* jest *jednorodna*, czy też kontrolowane dane należy odrzucić.

Należy pamiętać, że jednorodność jest pojęciem względnym i zależy od rozpatrywanej cechy. Na przykład próba może być jednorodna pod

względem wysokości osobników, lecz niejednorodna ze względu na kolor skóry, wiek, płeć, itp.

Testowanie istotności miar położenia

Najczęściej testowane miary położenia to: kategoria modalna, wartość medialna i średnia arytmetyczna.

Testy istotności miar położenia służą do:

1. porównania miary położenia danej cechy w próbie z miarą położenia w populacji lub, mówiąc inaczej, do porównania miary położenia w próbie ze standardem, np. sprawdzamy istotność różnicy między średnim plonem żyta uzyskanym w rozpatrywanym gospodarstwie indywidualnym (\bar{x} w próbie) a średnim plonem żyta jaki osiągany jest w województwie, do którego należy to gospodarstwo (μ w populacji);

2. porównania miar położenia danej cechy w dwu lub więcej próbach, np. porównanie średniego plonu żyta w dwóch lub więcej gospodarstwach indywidualnych – próbach.

Testowanie istotności miar rozproszenia

Najczęściej testowane miary rozproszenia to: dyspersja względna klasyfikacji, wariancja i odchylenie standardowe. Testy istotności dotyczące miar rozproszenia stosuje się w następujących przypadkach:

– gdy zachodzi potrzeba porównania dokładności co najmniej dwu metod pomiaru (klasyfikacji). O metodzie wykazującej mniejszą wariancję wnioskujemy, że jest metodą dokładniejszą;

– przy badaniu zjawisk manifestujących się różnicowaniem rozproszenia wyników w różnych sytuacjach eksperymentalnych w stosunku do kontroli;

– gdy zastosowanie danego testu uzależnione jest od tego, aby wariancje prób były jednakowe (np. test t Studenta, analiza wariancji). Wówczas, przed zastosowaniem takiego testu, należy porównać wariancje w próbach i zweryfikować hipotezę zerową o braku różnic istotnych.

Reasumując, testy istotności miar rozproszenia służą do:

1. porównania wariancji empirycznej ze standardem;

2. porównania empirycznych wartości wariancji w dwu lub więcej próbach.

Testowanie istotności proporcji (frakcji, procentu, wskaźnika struktury)

Jeżeli interesuje nas, jak często (z jakim nasileniem) występuje wariant danej cechy w próbie, to obliczamy w tym celu **proporcję**. Na przykład, gdy w próbie liczącej 100 osobników znajduje się 60 blondynów, 30 brunetów i 10 rudych, to proporcja blondynów będzie wynosić $60/100 = 0.6$, brunetów $30/100 = 0.3$ a rudych $10/100 = 0.1$. Suma proporcji jest zawsze dokładnie równa jedności: $0.6 + 0.3 + 0.1 = 1.0$. Jeżeli proporcję pomnożymy przez 100%, to otrzymamy frakcję w **procentach** (odsetki).

W zasadzie nie ma sensu obliczać odsetek z próby liczącej mniej niż 100 elementów, gdyż wówczas przypadkowa zmiana liczby obserwacji o jednostkę, wywoła zmianę większą niż 1%. Na przykład, zmiana o jednostkę w przypadku próby liczącej 20 elementów, w procentach oznacza zmianę o 5%. Proporcja jest względną miarą częstości nasilenia przypadków zaliczanych do danego wariantu badanej cechy. Ta właściwość proporcji umożliwia dokonywanie porównań nasilenia zdarzeń właściwych dla określonego wariantu badanej cechy w różnych, różnicznych próbach.

Jeżeli elementy próby podlegają klasyfikacji dychotomicznej (rodzaj skali pomiarowej jest w tym przypadku bez znaczenia), to poszczególne elementy (wyniki obserwacji lub doświadczenia) przyporządkowuje się dwu rozłącznym kategoriom badanej cechy, np. mężczyźni – kobiety, orzeł – reszka, osobnicy poniżej 40 lat – osobnicy powyżej 40 lat, chorzy – zdrowi, itp. W ten sposób otrzymujemy tzw. rozkład dwupunktowy (zero–jedynekowy), a zmienną tego typu (jak opisano) określa się jako zmienną zero–jedynekową (por. str. 53). Tę z dwu proporcji, która ma wyższą wartość można traktować jako szczególny rodzaj średniej takiego rozkładu – kategorię modalną, o której często mówi się jako o **wskaźniku struktury**.

Testy proporcji, obliczonej dla rozkładu zero–jedynekowego, służą do:

1. porównania proporcji empirycznej ze standardem. Na przykład porównujemy proporcję chłopców ($p = 0.65$) urodzonych w gminie A (próba) z oczekiwaną proporcją chłopców ($\pi = 0.5$) w populacji. Proporcja ta została oszacowana na podstawie wielkiej liczby obserwacji dotyczących stosunku płci wśród nowonarodzonych dzieci.

2. porównania proporcji empirycznych w dwu lub więcej próbach. Na przykład porównujemy trzy proporcje ($p_1 = 0.3$, $p_2 = 0.35$, $p_3 = 0.41$) obliczone dla trzech prób o liczebności n_1 , n_2 , n_3 .

Testowanie zgodności rozkładów

Testy zgodności służą do:

1. porównania rozkładu empirycznego badanej cechy w danej próbie z rozkładem oczekiwanym (teoretycznym), czyli ze standardem. Najczęściej testuje się zgodność rozkładu empirycznego z teoretycznym rozkładem dwumianowym, Poissona, normalnym i z rozkładami hipotetycznymi, innego kształtu niż tu wymienione;

2. porównania rozkładu empirycznego badanej cechy w dwu lub więcej próbach. Ten rodzaj testu jest nieparametryczną alternatywą parametrycznych testów istotności miar położenia i rozproszenia. Nie wymaga zatem założenia o typie rozkładu.

Testowanie niezależności cech

Testy niezależności stosuje się wówczas, gdy badana jest współzależność (współzmiennność) co najmniej dwóch cech, charakteryzujących daną populację i należy rozstrzygnąć, czy współzależność ta istnieje czy też nie. Analizę współzależności można przeprowadzić dla próby danych wyrażonych za pomocą skali nominalnej, porządkowej lub przedziałowej.

W przypadku cech wyrażonych za pomocą skali nominalnej i porządkowej, do pomiaru siły współzależności tych cech stosuje się nieparametryczne miary korelacji, odpowiednio: współczynniki siły związku oraz współczynniki korelacji rang. W pierwszym przypadku są to np.: wskaźnik współzmienności Yule'a, współczynnik Q Yule'a, współczynnik Bykowskiego, współczynnik siły związku r_p , współczynnik kontyngencji C , współczynnik V Cramera, czy współczynnik zbieżności T Czuprowa. W drugim przypadku są to: współczynnik korelacji rang Spearmana, współczynnik korelacji rang Kendalla, współczynnik konkordancji w Kendalla, wskaźnik korelacji medianowej, czy też wskaźnik korelacji Tukey'a.

Najlepiej jednak rozwinięte są techniki analizy współzależności cech dla danych wyrażonych za pomocą skali przedziałowej. W przy-

padku danych tego typu mamy do dyspozycji dwie metody analizy: **metodę regresji** i **metodę korelacji**. Bardzo ważną sprawą jest właściwe stosowanie obu tych metod, dlatego naszą uwagę skupimy przede wszystkim na tej kwestii. Zanim zdecydujemy o wyborze jednej z wyżej wspomnianych metod, musimy najpierw zastanowić się, z którym z dwóch przedstawionych niżej modeli współzależności mamy do czynienia.

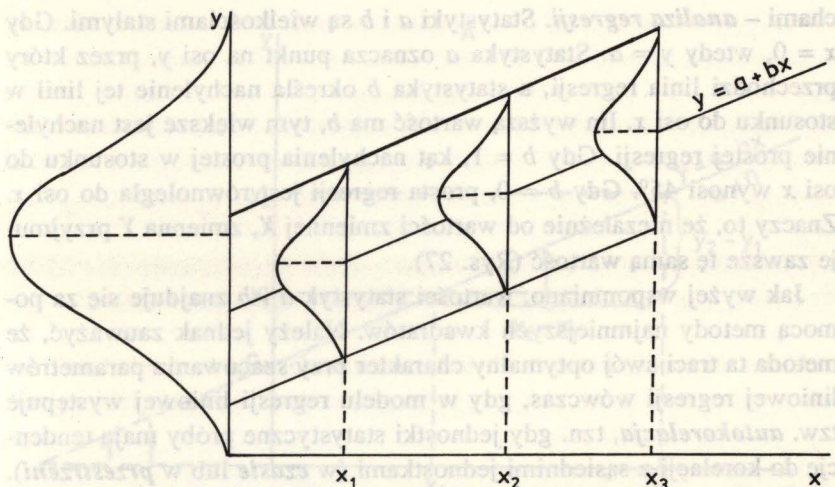
– Model 1: regresja

W najprostszym przypadku model ten dotyczy zależności między dwoma cechami, z których jedna (Y) jest **zmienną zależną**, a druga (X) **zmienną niezależną**. Cecha Y jest **zmienną losową** o rozkładzie normalnym, natomiast cecha X jest **zmienną kontrolowaną**. Przybiera ona wartości ściśle zdeterminowane w tym sensie, że są one ustalane z góry przez eksperymentatora (Rys. 26). Ilustracją tego modelu może być na przykład sytuacja, gdy badamy zależność tempa wzrostu kultury roślin (zmienna Y) od czasu (zmienna X), lub gdy badamy zależność między intensywnością fotosyntezy (zmienna Y) a temperaturą (zmienna X). W obu przypadkach wartości zmiennej X są ustalane przez eksperymentatora, który decyduje, że co trzeci dzień będzie notował liczbę roślin w badanej kulturze oraz, że efekt fotosyntezy w postaci wydzielanego tlenu będzie mierzony w temperaturze 5° , 10° , 15° , 20° , 25° , 30° , 35° i 40° . W ten sposób zmienna niezależna X nie jest obciążona naturalną zmiennością ani błędem pomiaru. W sytuacji jak wyżej opisano nie ma więc problemu z wyróżnieniem zmiennej zależnej i zmiennej niezależnej, ponieważ wynika to już z samego charakteru badań. Eksperymentator dokonując pomiarów zmiennej Y , z góry zakłada, że zmienna ta jest **liniową funkcją** zmiennej X . Zakładając, że zmienna Y jest zależna od zmiennej X , a nie na odwrót, zakłada więc, że istnieje tylko jednostronna – **asymetryczna** zależność funkcyjna. Na przykład zakłada, że wzrost populacji jest funkcją czasu, ale nie na odwrót. Zauważmy jednak, że twierdzenie takie ma sens z biologicznego punktu widzenia, ale nie z matematycznego. W tym drugim przypadku rozgraniczenie takie jest nieistotne.

Funkcję, o której mowa, można zapisać za pomocą równania:

$$y = a + bx \quad (1).$$

Dla populacji zapis ten przyjmie postać: $Y = \alpha + \beta X$. Statystyki a i b



Rys. 26. Współzależność między dwiema zmiennymi, z których jedna (Y) jest zmienną zależną, a druga (X) jest zmienną niezależną. Cecha Y jest zmienną losową o rozkładzie normalnym, natomiast cecha X jest zmienną kontrolowaną. Zmienna X przybiera wartości ustalone z góry przez badacza. Każdej ustalonej wartości zmiennej kontrolowanej (np. x_1 , x_2 , x_3) odpowiada zbiór wartości zmiennej losowej Y , który ma rozkład normalny. Zakłada się, że między zmienną Y i zmienną X istnieje zależność asymetryczna: zmienna Y zależy od zmiennej X , ale nie na odwrót. Zależność ta ma charakter liniowej zależności funkcyjnej (regresja prostoliniowa), którą zapisujemy za pomocą równania: $y = a + bx$. y – wartości jakie może przyjmować zmienna losowa Y ; x – wartości zmiennej X ustalone przez badacza.

to szacunki (oceny) nieznanych parametrów α i β . Wartości tych statystyk znajduje się za pomocą *metody najmniejszych kwadratów*. Po wstawieniu do równania (1) wartości a i b możemy wyliczyć oczekiwane wartości zmiennej Y dla określonych wartości zmiennej X , a następnie, w oparciu o te wyliczone wartości, wykreślić prostą obrazującą liniową funkcję zmiennej Y względem zmiennej X . Jest to zarazem linia najlepiej pasująca do empirycznych danych. Tę linię nazywa się *linią regresji* lub *linią regresji drugiego rodzaju*¹⁰, równanie opisujące tę linię – *równaniem regresji*, a tę metodę badania współzależności mię-

¹⁰ *Linie regresji pierwszego rodzaju*, obrazującą związek między zmiennymi, wykreśla się w oparciu o tzw. średnie warunkowe. W ten sposób powstaje linia łamana. Linii tej nie można opisać za pomocą równania tak jak linii regresji drugiego rodzaju. Z tego względu ma ona mniejsze zastosowanie w analizie współzależności cech.

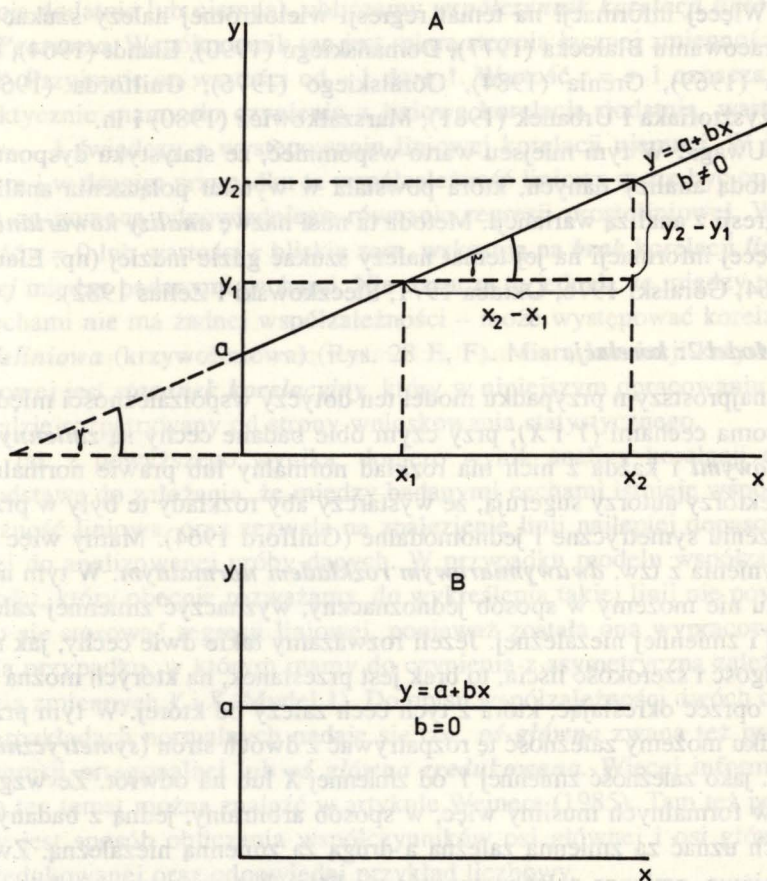
chami – **analizą regresji**. Statystyki a i b są wielkościami stałymi. Gdy $x = 0$, wtedy $y = a$. Statystyka a oznacza punkt na osi y , przez który przechodzi linia regresji, a statystyka b określa nachylenie tej linii w stosunku do osi x . Im wyższą wartość ma b , tym większe jest nachylenie prostej regresji. Gdy $b = 1$, kąt nachylenia prostej w stosunku do osi x wynosi 45° . Gdy $b = 0$, prosta regresji jest równoległa do osi x . Znaczy to, że niezależnie od wartości zmiennej X , zmienna Y przyjmuje zawsze tę samą wartość (Rys. 27).

Jak wyżej wspomniano, wartości statystyk a i b znajduje się za pomocą metody najmniejszych kwadratów. Należy jednak zauważyć, że metoda ta traci swój optymalny charakter przy szacowaniu parametrów liniowej regresji wówczas, gdy w modelu regresji liniowej występuje tzw. **autokorelacja**, tzn. gdy jednostki statystyczne próby mają tendencję do korelacji z sąsiednimi jednostkami (w *czasie* lub w *przestrzeni*). Ma to na przykład miejsce wówczas, gdy występuje tendencja do cyklicznego następstwa pewnych zjawisk, np.: 7 lat tłustych – 7 lat chudych – 7 lat tłustych – 7 lat chudych itd. W takim przypadku należy zastosować inną metodę estymacji parametrów liniowej regresji.

Statystyka b pozwala określić średni przyrost zmiennej zależnej, związany ze wzrostem o jednostkę wartości zmiennej niezależnej. W ten sposób, znajomość równania regresji umożliwia prognozowanie wartości zmiennej zależnej, gdy znamy wartość zmiennej niezależnej.

Prognozowanie, o którym mowa, będzie tym dokładniejsze, im silniejszy związek będzie istnieć między obu zmiennymi. Oceną siły związku (korelacji) między współzależnymi cechami zajmuje się metoda analizy zwana **metodą korelacji** (\rightarrow Model 2: korelacja, str. 102). Miarą korelacji jest **współczynnik korelacji liniowej, r** . Współczynnik ten mierzy siłę tylko prostoliniowej zależności. Dlatego dyskusowanie istotności współczynnika korelacji r w sytuacji, gdy poprawne było zastosowanie analizy regresji nie ma sensu, ponieważ fakt istnienia liniowego związku między badanymi zmiennymi stanowi założenie a priori tej metody. Ważne natomiast zastosowanie ma w tym wypadku tzw. **współczynnik determinacji r^2** , który informuje, jaką część zmienności Y można przypisać związkowi tej zmiennej ze zmienną X .

Analizę regresji można uogólnić na dowolną liczbę cech, z których jedną można uważać za zmienną zależną lub objaśnianą (zmienna losowa), a pozostałe za zmienne niezależne (zmienne objaśniające).



Rys. 27. Linia regresji obrazująca zależność zmiennej Y od zmiennej X . a – oznacza punkt na osi y , przez który przechodzi linia regresji; b – oznacza nachylenie linii regresji w stosunku do osi x . Statystyka b to, mówiąc inaczej, $\operatorname{tg} \gamma = \frac{y_2 - y_1}{x_2 - x_1}$. Jeżeli b , czyli $\operatorname{tg} \gamma$, równy jest jedności ($b = 1$), wówczas kąt $\gamma = 45^\circ$. Gdy $b = 0$, $\gamma = 0^\circ$. y – wartości jakie może przyjmować zmienna losowa Y ; x – wartości zmiennej X ustalane przez badacza; A – regresja prostoliniowa dla b różnego od zera ($b \neq 0$); B – regresja prostoliniowa dla b równego zero ($b = 0$).

Aby określić związek między tymi zmiennymi, stosuje się **metodę regresji wielokrotnej** (wielorakiej), przy czym, jeśli zakładamy, że zależność między zmienną objaśnianą i zespołem zmiennych objaśniających ma charakter funkcji liniowej, wówczas będziemy mieć do czynienia z **metodą prostoliniową regresji wielokrotnej**.

Więcej informacji na temat regresji wielokrotnej należy szukać w opracowaniu Blalocka (1977), Domańskiego (1990), Elandt (1964), Fisz (1969), Grenia (1984), Góralskiego (1976), Guilforda (1964), Krzysztofiaka i Urbanek (1981), Marszałkowicz (1980) i in.

Uwaga: W tym miejscu warto wspomnieć, że statystyka dysponuje metodą analizy danych, która powstała w wyniku połączenia analizy regresji z analizą wariancji. Metoda ta nosi nazwę *analizy kowariancji*. Więcej informacji na jej temat należy szukać gdzie indziej (np. Elandt 1964, Góralski 1976, Oktaba 1971, Steczkowski i Zeliaś 1982).

– Model 2: korelacja

W najprostszym przypadku model ten dotyczy współzależności między dwoma cechami (Y i X), przy czym obie badane cechy są *zmiennymi losowymi* i każda z nich ma rozkład normalny lub prawie normalny. Niektórzy autorzy sugerują, że wystarczy aby rozkłady te były w przybliżeniu symetryczne i jednomodalne (Guilford 1964). Mamy więc do czynienia z tzw. *dwuwymiarowym rozkładem normalnym*. W tym modelu nie możemy w sposób jednoznaczny, wyznaczyć zmiennej zależnej i zmiennej niezależnej. Jeżeli rozważamy takie dwie cechy, jak np. długość i szerokość liścia, to brak jest przesłanek, na których można by się oprzeć określając, która z tych cech zależy od której. W tym przypadku możemy zależność tę rozpatrywać z dwóch stron (*symetrycznie*), tzn. jako zależność zmiennej Y od zmiennej X lub na odwrót. Ze względów formalnych musimy więc, w sposób arbitralny, jedną z badanych cech uznać za zmienną zależną a drugą za zmienną niezależną. Zwyczajowo, zmienną zależną oznacza się literą Y , a zmienną niezależną literą X .

W przypadku tego modelu, analizę współzależności przeprowadza się za pomocą metody korelacji. Metoda ta pozwala sprawdzić, czy badane cechy są współzależne. Analizę korelacji między dwoma cechami najlepiej zacząć od tzw. *wykresu korelacyjnego*. W przestrzeni dwuwymiarowej, w układzie współrzędnych x i y , odnotowujemy wartości badanych cech. W ten sposób otrzymujemy obraz istniejącego między nimi związku, który określany bywa jako "chmura punktów". Na podstawie takiego wykresu można się wstępnie zorientować o istnieniu zależności i jej kierunku (Rys. 28). Jeżeli są ku temu podstawy (chmura punktów sugeruje, że między badanymi cechami istnieje liniowa kore-

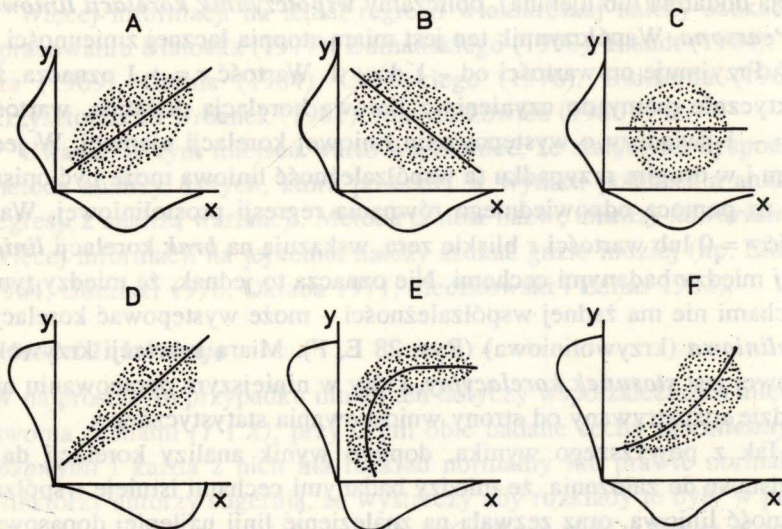
lacja dodatnia lub ujemna), obliczamy **współczynnik korelacji liniowej r Pearsona**. Współczynnik ten jest miarą stopnia łącznej zmienności X i Y . Przyjmuje on wartości od -1 do $+1$. Wartość $r = +1$ oznacza, że faktycznie mamy do czynienia z liniową korelacją dodatnią, wartość $r = -1$ świadczy o występowaniu liniowej korelacji ujemnej. W jednym i w drugim przypadku ta współzależność liniowa może być opisana za pomocą odpowiedniego równania regresji prostoliniowej. Wartość $r = 0$ lub wartości r bliskie zera, wskazują na **brak korelacji liniowej** między badanymi cechami. Nie oznacza to jednak, że między tymi cechami nie ma żadnej współzależności – może występować korelacja **nieliniowa** (krzywoliniowa) (Rys. 28 E, F). Miarą korelacji krzywoliniowej jest **stosunek korelacyjny**, który w niniejszym opracowaniu nie będzie rozpatrywany od strony wnioskowania statystycznego.

Jak z powyższego wynika, dopiero wynik analizy korelacji daje podstawę do założenia, że między badanymi cechami istnieje współzależność liniowa, oraz zezwala na znalezienie linii najlepiej dopasowanej do analizowanej próby danych. W przypadku modelu współzależności, który obecnie rozważamy, do wykreślenia takiej linii nie powinno się stosować regresji liniowej, ponieważ została ona wypracowana dla przypadku, w którym mamy do czynienia z asymetryczną zależnością zmiennych Y i X (Model 1). Do opisu współzależności dwóch cech o rozkładach normalnych nadaje się tzw. **oś główna** zwana też prostą regresji ortogonalnej lub **oś główna zredukowana**. Więcej informacji na ten temat można znaleźć w artykule Weinerja (1985). Tam też podany jest sposób obliczania współczynników osi głównej i osi głównej zredukowanej oraz odpowiedni przykład liczbowy.

Analizę korelacji można uogólnić na dowolną liczbę cech, z których jedną przyjmujemy za zmienną zależną a pozostałe za zmienne niezależne. Analiza taka służy do badania siły związku między:

- zmienną zależną a wszystkimi zmiennymi niezależnymi, działającymi łącznie. Ten typ korelacji nazywamy **korelacją wielokrotną** (wieloraką). Jej miarą jest **współczynnik liniowej korelacji wielokrotnej**;
- parą zmiennych przy wyłączeniu pozostałych zmiennych. W tym przypadku mówimy o **korelacji cząstkowej** (częściowej). Jej miarą jest **współczynnik liniowej korelacji cząstkowej**.

Więcej informacji na temat korelacji wielokrotnej i cząstkowej można znaleźć u Białocka (1977), Domańskiego (1990), Elandt (1964), Fi-



Rys. 28. Hipotetyczne formy wykresów rozrzutu "chmury punktów" na wykresie korelacyjnym, gdy postacie rozkładu wartości zmiennej Y i X są różne (wg Guilforda 1964, nieco zmienione). Kształt rozkładu zmiennej Y i X zaznaczono schematycznie przy odpowiedniej osi – y lub x . Linia przerywana oznacza kształt zależności (regresję). y – wartości jakie może przyjmować zmienna losowa Y ; x – wartości jakie może przyjmować zmienna losowa X ; A – korelacja prostoliniowa dodatnia; B – korelacja prostoliniowa ujemna; C – brak korelacji (dla tej chmury punktów nie ma również współzależności między zmiennymi Y i X); D – korelacja prostoliniowa dodatnia (rozkłady obu zmiennych mają asymetrię ujemną!); E – korelacja krzywoliniowa; F – korelacja krzywoliniowa. Najważniejszym warunkiem stosowania współczynnika korelacji liniowej r Pearsona jest, by zależność między zmiennymi była prostoliniowa, tzn. wymagana jest regresja prostoliniowa. Warunek ten jest spełniony, gdy chmura punktów przyjmuje kształt elipsy (przypadek A i B). W pozostałych przypadkach należy stosować inne miary korelacji, np. stosunek korelacyjny (korelacja krzywoliniowa).

sza (1969), Góralskiego (1976), Grenia (1984), Guilforda (1964), Krzysztofiaka i Urbanek (1981), Marszałkiewicz (1980) i in.

UWAGA: Należy mocno podkreślić, że stwierdzenie korelacji między badanymi cechami nie jest jednoznaczne z istnieniem **zależności przyczynowej** między tymi cechami. Zależność taka może istnieć, lecz nie musi! Stwierdzenie występowania korelacji jest jedynie **sygnałem**, że pomiędzy badanymi cechami (zmiennymi) **może** istnieć jakiś bezpośredni lub pośredni związek przyczynowy.

Testy niezależności cech, ze względu na charakter rozwiązywanych

zagadnień, można podzielić na testy dotyczące analizy regresji i testy dotyczące korelacji cech.

Testy dotyczące analizy regresji służą do:

1. wykrywania autokorelacji;
2. sprawdzenia liniowości regresji;
3. sprawdzenia czy dwie linie regresji są identyczne;
4. oceny istotności współczynnika regresji liniowej;
5. oceny istotności różnicy dwóch współczynników regresji liniowej.

Testy dotyczące korelacji cech służą do:

1. oceny istotności miary korelacji;
2. oceny istotności różnic dwóch miar korelacji cech.

Szacowanie przedziałów ufności

Testy istotności stanowią pierwszy krok we wnioskowaniu statystycznym, dotyczącym porównywania statystyki próby z parametrem populacji (standardem). Na podstawie wyniku testu istotności możemy wywnieść wniosek o charakterze *jakościowym*: wiemy tylko, że badana statystyka nie różni się od parametru populacji lub, że ma wartość od niego różną. Czasami wiemy też, że jest większa lub mniejsza od standardu. Postępowanie takie dotyczy estymacji punktowej. Zdarza się jednak, że chcemy znać przedział, w którym, z określonym prawdopodobieństwem, mieści się wartość parametru populacji. Do rozwiązania takiego zagadnienia stosuje się tzw. *oszacowanie przedziałowe*. Jest to metoda właściwa estymacji przedziałowej (→ Estymacja parametrów statystycznych, str. 92). Za pomocą odpowiednich wzorów wyznacza się przedział ufności, w którym, powinna się znajdować prawdziwa wartość szacowanego parametru populacji.

Przedziały ufności można szacować dla miar położenia, rozproszenia i korelacji. Przedziały ufności mogą grać rolę testów istotności.

Testowanie szeregów czasowych

Szeregiem czasowym nazywamy próbę, w której wartości badanej cechy zostały uszeregowane według następstwa w czasie. Mamy tu więc do czynienia z dwoma cechami, z których jedna jest *zmienną losową* (zależną), charakteryzującą poziom pewnego zjawiska, a druga jest *zmienną czasową* (niezależną). Analiza szeregów czasowych polega

więc na badaniu zmian, jakim podlega pewne zjawisko w czasie. Zmiany te dotyczą tendencji rozwojowych (trendu), wahań okresowych i wahań sezonowych.

Trend, to powolne i systematyczne zmiany, jakim podlega zjawisko w czasie.

Wahania okresowe (cykliczne), to zmiany zjawiska, które powtarzają się regularnie w analogicznych jednostkach czasu kolejnych cykli. Przykładem zjawisk cyklicznych mogą być zjawiska związane np. z cyklem dobowym, rocznym, wieloletnim.

Wahania sezonowe, to wahania okresowe o cyklu rocznym. Wywoływane są one zmianami pór roku, lub pewnymi czynnikami zwyczajowymi, które powtarzają się lub nasilają w analogicznych porach roku.

Testy dla szeregów czasowych służą do:

1. badania trendu;
2. badania wahań okresowych (cyklicznych);
3. badania równoczesnych wahań w dwu szeregach czasowych.

Testy dla szeregów czasowych mogą być użyteczne do wykrywania tendencji i zmian cyklicznych w badaniach przyrodniczych. W niniejszym opracowaniu nie będą one jednak rozważane. Informację na temat szacowania parametrów i weryfikacji hipotez statystycznych w szeregach czasowych można znaleźć gdzie indziej (Krzysztofiak i Urbanek 1981, Luszniwicz 1980, Marszałkowicz 1980 i in.).

7. WYKAZ TESTÓW STOSOWANYCH DO STATYSTYCZNEJ ANALIZY DANYCH

Wraz z coraz łatwiejszym dostępem do sprzętu komputerowego, do analizy statystycznej danych coraz powszechniej wykorzystuje się komputerowe programy statystyczne (np. BMDP, Statgraphics, itp.). Niezależnie jednak od tego, czy stosuje się te programy, czy też obliczenia przeprowadza się za pomocą kalkulatora, problem wyboru odpowiedniego testu jest zawsze aktualny – *żadna maszyna za nas tego nie robi!* Informacje zamieszczone w tej części pracy mają właśnie pomóc nam w wyborze odpowiedniego testu statystycznego.

Wszystkim zaprezentowanym dalej testom wspólne są dwa warunki, a mianowicie: 1. liczebność próby jest bardzo mała w porównaniu z liczebnością populacji generalnej; 2. próba jest prostą próbą losową. Testy, o których mowa, zgodnie z przyjętym wcześniej podziałem, zostały podzielone na:

- A – testy sprawdzające losowość próby;
- B – testy sprawdzające jednorodność próby;
- C – testy istotności miar położenia;
- D – testy istotności miar rozproszenia;
- E – testy istotności proporcji;
- F – testy zgodności rozkładów;
- G – testy niezależności cech;
- H – szacowanie przedziałów ufności.

W obrębie wyróżnionych w ten sposób grup, tam gdzie tego wymagała sytuacja, wszystkie testy zostały sklasyfikowane według: 1. schematu doświadczalnego (zmienne połączone – zmienne niepołączone), 2. liczby badanych prób (testy dla 1 próby, dla 2 prób, dla ≥ 2 prób), 3. w zależności od skali pomiaru (skala nominalna, porządkowa, przedziałowa) oraz 4. ze względu na rodzaj testu (test parametryczny – nieparametryczny). Wydaje się, że ten ujednoczony, pod względem formalnym, przegląd testów statystycznych ułatwi szybką orientację w ich wielkim i różnorodnym bogactwie. Zamieszczona przy każdym teście szczegółowa notatka bibliograficzna umożliwi dotarcie do źródła.

Celowo nie ograniczono się do jednego przykładu, ilustrującego zastosowanie danego testu. Tam, gdzie było to możliwe, cytowano wiele źródeł omawiających ten test. Zapoznanie się z różnymi przykładami stosowania danego testu ułatwia zrozumienie jego możliwości i ograniczeń. Może być również wzorem i twórczym natchnieniem, gdy przystępujemy do opracowywania statystycznego własnych danych.

Noty bibliograficzne, prezentowane w poniższym wykazie testów, nie wyczerpują wszystkich, dostępnych na krajowym rynku, źródeł. Nie jest to zresztą konieczne. Po zapoznaniu się z niniejszym opracowaniem uważny czytelnik z pewnością zrozumie od czego zależy wybór danego testu i nie powinien mieć trudności w korzystaniu z podręczników statystyki nie cytowanych w tej pracy.

Na koniec jeszcze jedna uwaga. Kiedy już, rozważając wyszczególnione wyżej warunki, dotrzemy do danego testu, należy **koniecznie** zapoznać się z dodatkowymi założeniami, które **powinny** być spełnione, aby ten test można było zastosować. Między innymi trzeba zaczerpnąć informacji na temat minimalnej liczebności próby lub prób.

W tej części pracy stosowane będą następujące oznaczenia:

m – oznacza liczbę prób lub liczbę cech;

k, w – oznaczają liczbę kategorii danej cechy;

$m \times k$ – oznacza tablicę wielopolową dla $m \geq 2$ prób (wiersze) i $k \geq 2$ kategorii rozważanej cechy (kolumny);

$w \times k$ – oznacza tablicę wielopolową dla $m = 1$ próby, gdzie w reprezentuje ≥ 2 kategorii jednej cechy (wiersze), a k reprezentuje ≥ 2 kategorii drugiej cechy (kolumny). Ten typ tabeli określane bywa jako tabela kontyngencji.

A. Testy sprawdzające losowość próby

Problem: Hipoteza zerowa, że próba jest losowa (uporządkowanie elementów w próbie jest losowe) testowana jest wobec alternatywy, że próba ta nie jest losowa (uporządkowanie elementów w próbie nie jest losowe).

A.1. Skala porządkowa

A.1.1. Testy nieparametryczne

A.1.1.1. Kwantylový test serii

Źródło: DOMAŃSKI 1979: 62–64

A.1.1.2. Medianowy test serii

Źródło: DOMAŃSKI 1979: 64–67; FREUND 1971: 320–321; GREŃ 1984: 139 (Model I, przykład 1); PERKAL 1967: 135

A.1.1.3. Kwantylový test serii

Źródło: DOMAŃSKI 1979: 67; PERKAL 1967: 135

A.1.1.4. Test oparty na liczbie serii trzech rodzajów elementów

Źródło: DOMAŃSKI 1979: 68–70

A.1.1.5. Test oparty na długości serii

Źródło: DOMAŃSKI 1979: 70–74

A.1.1.6. Test oparty na ogólnej liczbie serii znaków

Źródło: DOMAŃSKI 1979: 75–76; FREUND 1971: 318–320

A.1.1.7. Test Moora-Wallisa

Źródło: DOMAŃSKI 1979: 76–77

A.1.1.8. Test χ^2 oparty na oczekiwanych i zaobserwowanych seriach monotonicznych różnej długości (test Moora-Wallisa)

Źródło: DOMAŃSKI 1979: 77–78

A.1.1.9. Test χ^2 oparty na liczbie serii znaków

Źródło: DOMAŃSKI 1979: 79–80

A.1.1.10. Test kolekcjonera

Źródło: DOMAŃSKI 1979: 80–82

A.1.1.11. Test permutacji

Źródło: DOMAŃSKI 1979: 82–84

A.1.1.12. Test pokerowy

Źródło: DOMAŃSKI 1979: 84–86

A.1.1.13. Test ostatniej cyfry (test χ^2)

Źródło: PERKAL 1967: 136

B. Testy sprawdzające jednorodność próby

Problem: Hipoteza zerowa głosząca, że wszystkie elementy próby pobrane zostały z populacji generalnej, będącej przedmiotem badania, testowana jest wobec alternatywy, że w próbie znajdują się elementy pobrane z innej populacji generalnej.

B.1. Skala przedziałowa

B.1.1. Testy parametryczne

B.1.1.1. Eliminowanie jednej obserwacji, największej lub najmniejszej

B.1.1.1.1. Test jednorodności próby

Źródło: GÓRALSKI 1976: 175–176

B.1.1.1.2. Test T_n

Źródło: PUCHALSKI 1980: 314–315

B.1.1.2. Eliminowanie dwóch obserwacji skrajnych, największej i najmniejszej

Źródło: PUCHALSKI 1980: 315–316

B.1.1.3. Eliminowanie dwóch obserwacji skrajnych, największych lub najmniejszych

Źródło: PUCHALSKI 1980: 315–316

C. Testy istotności miar położenia

C.1. Testy dla $m = 1$ próby

C.1.1. Porównanie kategorii modalnej ze standardem

Problem: Testowana jest hipoteza zerowa, że kategoria modalna zaobserwowana w próbie nie różni się istotnie od kategorii modalnej w populacji (standard).

Hipoteza ta może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– kategoria modalna w próbie nie różni się istotnie od standardu;

2. test jednostronny:

– kategoria modalna w próbie jest wyższa niż standard;

– kategoria modalna w próbie jest niższa niż standard.

C.1.1.1. Skala nominalna

C.1.1.1.1. Testy parametryczne

C.1.1.1.1.1. Test z

Uwaga: Cecha przyjmuje więcej niż dwie wzajemnie wykluczające się kategorie.

Źródło: GÓRALSKI 1976: 148–150; DOMAŃSKI 1990: 97–99

C.1.2. Porównanie wartości medialnej (mediany) ze standardem

Problem: Testowana jest hipoteza zerowa, że mediana zaobserwowana w próbie nie różni się w sposób istotny od mediany w populacji (standard). Hipoteza ta może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane mediany różnią się istotnie;

2. test jednostronny:

– mediana w próbie jest większa niż mediana w populacji;

– mediana w próbie jest mniejsza niż mediana w populacji.

C.1.2.1. Skala przedziałowa

C.1.2.1.1. Testy parametryczne

C.1.2.1.1.1. Test z

Źródło: GÓRALSKI 1976: 152

C.1.3. Porównanie średniej arytmetycznej ze standardem

Problem: Testowana jest hipoteza zerowa, że średnia arytmetyczna w próbie nie różni się w sposób istotny od średniej arytmetycznej w populacji (standard). Hipoteza ta może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane średnie arytmetyczne różnią się istotnie;

2. test jednostronny:

– średnia arytmetyczna w próbie jest większa niż średnia w populacji;

– średnia arytmetyczna w próbie jest mniejsza niż średnia w populacji.

C.1.3.1. Skala przedziałowa

C.1.3.1.1. Testy parametryczne

C.1.3.1.1.1. Test z (test u)

Źródło: ASKENAS, SAWICKI 1970: 57–58; BŁALOCK 1977: 162–166; FREUND 1971: 247–250; GREŃ 1984: 57 (Model I), 59 (przykład 1), 59 (Model III), 59 (przykład 1); PUCHALSKI 1980: 251–253; SAWICKI 1982: 90–91

C.1.3.2.1.2. Test t Studenta

Uwaga: Test t ma zastosowanie w sytuacjach, gdy mamy do czynienia z małą próbą i gdy można założyć, że rozkład zmiennej w populacji jest normalny. Niestety, gdy próby są małe, mamy zwykle poważne wątpliwości co do kształtu rozkładu w populacji. Dlatego test t nie jest specjalnie użyteczny, mimo, że jest raczej popularny (Błalock 1977).

Źródło: BŁALOCK 1977: 166–172; FREUND 1971: 250–251; GÓRALSKI 1976: 150–151; GREŃ 1984: 58 (Model II), 60 (przykład 2); PUCHALSKI 1980: 253–255; SAWICKI 1982: 91–93

C.2. Testy dla $m = 2$ prób

C.2.1. Testy porównujące kategorie modalne

Problem: Hipoteza zerowa, że różnica między porównywanymi kategoriami modalnymi jest nieistotna (tzn. porównywane próby pochodzą z tej samej populacji lub z różnych populacji o identycznej tendencji centralnej) może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane kategorie modalne różnią się istotnie (tzn. porównywane próby pochodzą z populacji o różnych tendencjach centralnych);

2. test jednostronny:

– kategoria modalna w próbie n_1 jest wyższa niż w próbie n_2 ;

– kategoria modalna w próbie n_1 jest niższa niż w próbie n_2 .

C.2.1.1. Skala nominalna

C.2.1.1.1. Zmienne niepołączone

C.2.1.1.1.1. Testy parametryczne

C.2.1.1.1.1.1. Test z

Uwaga: Cecha przyjmuje więcej niż dwie wzajemnie wykluczające się kategorie

Źródło: GÓRALSKI 1976: 163–164; DOMAŃSKI 1990: 123–124

C.2.2. Testy porównujące rozkłady częstości w próbach

Problem: Hipoteza zerowa, że porównywane rozkłady częstości są zgodne (tzn. porównywane próby pochodzą z tej samej populacji lub z różnych populacji o takim samym rozkładzie cechy, czyli o identycznej tendencji centralnej), może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane rozkłady częstości nie są zgodne i różnią się istotnie (tzn. porównywane próby pochodzą z populacji o różnych rozkładach cechy);

2. test jednostronny:

– wartość miary położenia jednego z rozkładów jest większa (mniejsza) od wartości tej samej miary położenia drugiego rozkładu.

C.2.2.1. Skala nominalna

Uwaga: Do testowania hipotezy zerowej można wykorzystać testy zgodności (F.2.2.1.).

C.2.3. Testy porównujące zgodność uporządkowania elementów w próbach

Problem: Należą tutaj testy porównujące rozkłady uporządkowania wartości badanej cechy w dwu próbach. Testowana jest hipoteza zerowa, że w porównywanych próbach rozkłady te są zgodne (tzn. obie próby pochodzą z jednej populacji lub z dwóch populacji o takich samych rozkładach, czyli o identycznej tendencji centralnej). Hipoteza ta może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane rozkłady nie są zgodne (próby pochodzą z populacji o różnych tendencjach centralnych);

2. test jednostronny:

– wartość miary położenia jednego z rozkładów jest większa (mniejsza) od wartości tej samej miary położenia drugiego rozkładu.

C.2.3.1. Skala porządkowa

Uwaga: Do weryfikowania hipotezy zerowej mogą być wykorzystane testy zgodności dla $m = 2$ prób (F.2.3.1.)

C.2.4. Testy oceniające istotność różnicy między średnimi arytmetycznymi

Problem: Testowana jest hipoteza zerowa, że różnica między porównywanymi średnimi arytmetycznymi jest statystycznie nieistotna. Hipoteza ta może być testowana wobec jednej z następujących hipotez alternatywnych:

1. test dwustronny:

– porównywane średnie różnią się istotnie

2. test jednostronny:

– wartość średniej w próbie n_1 jest wyższa niż wartość średniej w próbie n_2 ;

– wartość średniej w próbie n_1 jest niższa niż wartość średniej w próbie n_2 .

C.2.4.1. Skala przedziałowa

C.2.4.1.1. Zmienne połączone

C.2.4.1.1.1. Testy parametryczne

C.2.4.1.1.1.1. Test z (test u)

Źródło: FREUND 1971: 249–250; KRZYSZTOFIAK, URBANEK 1981: 241–243

C.2.4.1.1.1.2. Test t

Źródło: ASKENAS, SAWICKI 1970: 60–62; BLALOCK 1977: 207–209; FREUND 1971: 250–251; GÓRALSKI 1976: 182–184; OKTAB 1980: 127–129; PUCHALSKI 1980: 259–264; SAWICKI 1982: 105–107

C.2.4.1.2. Zmienne niepołączone

C.2.4.1.2.1. Testy parametryczne

C.2.4.1.2.1.1. Test z (test u) dla jednakowych wariancji

Źródło: ASKENAS, SAWICKI 1970: 59–60; FREUND 1971: 251–252; GREŃ 1984: 65 (Model I), 67 (Model III, przykład 1); KRZYSZTOFIAK, URBANEK 1981: 239–241; SAWICKI 1982: 105

C.2.4.1.2.1.2. Test t dla jednakowych wariancji

Źródło: ASKENAS, SAWICKI 1970: 59–60; BLALOCK 1977: 198–201; FREUND 1971: 252–254; GÓRALSKI 1976: 176–177; GREŃ 1984: 65 (Model II), 68 (przykład 2); KRZYSZTOFIAK, URBANEK 1981: 249–252; OKTAB 1980: 122–125; PUCHALSKI 1980: 255–259; SAWICKI 1982: 103–104

C.2.4.1.2.1.3. Test t dla różnych wariancji i różnych liczebności prób

Źródło: BLALOCK 1977: 201–203; GÓRALSKI 1976: 177–178

C.2.4.1.2.1.4. Test t dla różnych wariancji i równych liczebności prób

Źródło: OKTAB 1980: 132

C.2.4.1.2.1.5. Test C dla różnych wariancji i różnych liczebności prób (test przybliżony C Cochra i Coxa)

Źródło: OKTAB 1980: 132–133

C.3. Testy dla $m \geq 2$ prób

C.3.1. Testy porównujące kategorie modalne

Problem: Testowana jest hipoteza zerowa, że różnice między porównywanymi kategoriami modalnymi są nieistotne (tzn. próby pochodzą z tej samej populacji lub z różnych populacji o tych samych rozkładach, czyli o tej samej tendencji centralnej). Hipoteza ta jest testowana względem alternatywy, że badane kategorie modalne nie są zgodne, tzn. że przynajmniej jedna próba została wylosowana z populacji o innej tendencji centralnej (test dwustronny).

C.3.1.1. Skala nominalna

C.3.1.1.1. Zmienne niepołączone

C.3.1.1.1.1. Testy parametryczne

C.3.1.1.1.1. Test F

Uwaga: Cecha przyjmuje więcej niż dwie wzajemnie wykluczające się kategorie. Test ten jest specyficzną formą analizy wariancji.

Źródło: GÓRALSKI 1976: 165–166; DOMAŃSKI 1990: 125

C.3.2. Testy porównujące rozkłady częstości w próbach

Problem: Testowana jest hipoteza zerowa, że porównywane rozkłady częstości są zgodne (tzn. próby pochodzą z tej samej populacji lub z różnych populacji o tych samych rozkładach, czyli o tej samej tendencji centralnej). Hipoteza ta jest testowana względem alternatywy, że badane rozkłady częstości nie są zgodne – przynajmniej jedna próba została wylosowana z populacji o innym rozkładzie, czyli innej tendencji centralnej (test dwustronny).

C.3.2.1. Skala nominalna

Uwaga: Do weryfikacji hipotezy zerowej mogą być wykorzystane testy zgodności (F.3.1.1.)

C.3.3. Testy porównujące zgodność uporządkowania elementów w próbach

Problem: Należą tutaj testy porównujące rozkłady uporządkowania wartości badanej cechy w $m \geq 2$ próbach. W zależności od tego, czy stosujemy metodę zmiennych połączonych, czy też niepołączonych, testowane są, odpowiednio, dwie lub jedna hipoteza zerowa. W przypadku zmiennych zależnych (klasyfikacja podwójna), weryfikowane są dwie hipotezy zerowe:

1. obserwacje uporządkowane według rzędów (bloków) pochodzą z tej samej lub z różnych populacji o takim samym rozkładzie cechy, czyli o identycznej tendencji centralnej;
2. obserwacje uporządkowane według kolumn (zabiegów, prób) pochodzą z tej samej lub z różnych populacji o takim samym rozkładzie cechy, czyli o identycznej tendencji centralnej.

Hipotezy te weryfikowane są względem alternatywy, że przynajmniej jedna z badanych prób została wylosowana z populacji o odmiennym rozkładzie.

W przypadku zmiennych niezależnych (klasyfikacja pojedyncza), weryfikowana jest tylko hipoteza zerowa dla kolumn (zabiegów).

C.3.3.1. Skala porządkowa

Uwaga: Do weryfikowania hipotezy zerowej mogą być wykorzystane testy zgodności (F.3.2.1.)

C.3.4. Testy oceniające istotność różnic między średnimi arytmetycznymi: analiza wariancji

Problem: W przypadku zmiennych połączonych (klasyfikacja podwójna), testowane są dwie hipotezy zerowe:

1. średnie arytmetyczne kolumn są sobie równe;
2. średnie arytmetyczne rzędów są sobie równe.

Hipotezy te można również sformułować w następujący sposób:

1. wpływ k zabiegów jest jednakowy (podział wg kolumn nie jest istotny);
2. wpływ w bloków jest jednakowy (podział wg rzędów nie jest istotny).

Hipotezy te weryfikowane są względem alternatywy, że porównywane średnie arytmetyczne różnią się między sobą istotnie.

W przypadku zmiennych niepołączonych (klasyfikacja pojedyncza), testowana jest tylko hipoteza zerowa dla kolumn (zabiegów, prób).

Analiza wariancji nie daje odpowiedzi na pytanie, które mianowicie średnie różnią się między sobą istotnie, a które nie. Aby odpowiedzieć na to pytanie, należy zastosować dodatkowe testy, opracowane specjalnie w tym celu (por. C.3.5.).

C.3.4.1. Skala przedziałowa

C.3.4.1.1. Zmienne połączone (klasyfikacja podwójna)

C.3.4.1.1.1. Testy parametryczne

C.3.4.1.1.1.1. Analiza wariancji dla bloków kompletnie zrandomizowanych bez replikacji

Uwaga: Rozpatrujemy tabelę z pewną liczbą komórek, przy czym w każdej komórce tabeli jest tylko jeden wynik.

Źródło: FREUND 1971: 304–306; GREŃ 1984: 106–110; OKTABA 1980: 180–187; PAWŁOWSKI 1980: 198–203; PUCHALSKI 1980: 274–280; SAWICKI 1982: 155–162

C.3.4.1.1.1.2. Analiza wariancji dla bloków kompletnie zrandomizowanych z replikacjami

Uwaga: Rozpatrujemy tabelę z pewną liczbą komórek i taką samą liczbą wyników (obserwacji) w każdej komórce tabeli, przy czym liczba wyników w komórce > 1 .

Źródło: BLALOCK 1977: 284–294; PUCHALSKI 1980: 281–285

C.3.4.1.1.1.3. Analiza wariancji dla bloków kompletnie zrandomizowanych przy brakujących wynikach: wzór Yatesa

Źródło: OKTABA 1980: 197–199; PUCHALSKI 1980: 308–310

C.3.4.1.1.1.4. Analiza wariancji dla bloków kompletnie zrandomizowanych: test χ^2

Źródło: GÓRALSKI 1976: 184–186

C.3.4.1.1.1.5. Analiza wariancji dla kwadratu łacińskiego

Źródło: OKTABA 1980: 211–214; PAWŁOWSKI 1980: 203–206; PUCHALSKI 1980: 310–314

C.3.4.1.1.1.6. Analiza wariancji dla kwadratu łacińskiego przy brakujących wynikach

Źródło: OKTABA 1980: 221–222; PUCHALSKI 1980: 310 (313)–314

C.3.4.1.2. Zmienne niepołączone (klasyfikacja pojedyncza)

C.3.4.1.2.1. Testy parametryczne

C.3.4.1.2.1.1. Analiza wariancji dla układu kompletnej randomizacji

Uwaga: Jest to uogólnienie testu różnicy między średnimi arytmetycznymi na więcej niż dwie próby.

Źródło: BLALOCK 1977: 272–284; FREUND 1971: 299–302, 303 (ćw. 8); GÓRALSKI 1976: 178–180; GREŃ 1984: 99–103; OKTABA 1980: 157–167; PAWŁOWSKI 1980: 190–198; PUCHALSKI 1980: 266–274; SAWICKI 1982: 126–133

C.3.5. Testy istotności stosowane w analizie wariancji po odrzuceniu hipotezy zerowej

Problem: Testy te służą do badania istotności różnic między poszczególnymi średnimi arytmetycznymi w badanych próbach, jeżeli w wyniku przeprowadzonej analizy wariancji odrzucono hipotezę zerową, tzn. stwierdzono, że przynajmniej jedna z badanych prób pochodzi z innej populacji niż pozostałe.

C.3.5.1. Testy parametryczne

C.3.5.1.1. Metoda Tukeya (test luki, test skrajnego odchylenia, test F Snedecora) dla prób równolicznych

Uwaga: Metoda ta wymaga dość żmudnych obliczeń.

Źródło: OKTABA 1980: 188–189 (bloki kompletnie zrandomizowane), 214–215 (kwadrat łaciński), 168–172 (układ kompletnie zrandomizowany)

C.3.5.1.2. Metoda Duncana (test Duncana, wielokrotny test rozstępu) dla prób równo- i różnolicznych

Uwaga: Jest to metoda prosta i niezbyt pracochłonna, dlatego jest ona często stosowana.

Ostatnio jednak test ten został poddany krytyce i niektórzy autorzy przestrzegają przed stosowaniem go jako jedyne go sprawdzianu (Steczkowski i Zeliaś 1982).

Źródło: GÓRALSKI 1976: 181–182 (układ kompletnej randomizacji); OKTABA 1980: 189–191 (bloki całkowicie zrandomizowane), 172–177 (układ kompletnej randomizacji)

C.3.5.1.3. Test d Halperina dla prób równo- i różnicowych

Uwaga: Jest to test przeznaczony dla układu kompletnej randomizacji.

Źródło: GÓRALSKI 1976: 180–181

D. Testy istotności miar rozproszenia

D.1. Testy dla $m = 1$ próby

Problem: Testowana jest hipoteza zerowa, że rozpatrywana miara rozproszenia wyników w próbie jest taka sama jak miara rozproszenia w populacji (standard). Hipoteza ta może być testowana względem jednej z następujących alternatyw:

1. test dwustronny:

– zaobserwowana miara rozproszenia różni się istotnie od standardu;

2. test jednostronny:

– zaobserwowana miara rozproszenia jest większa od standardu;

– zaobserwowana miara rozproszenia jest mniejsza od standardu.

D.1.1. Porównanie względnej dyspersji klasyfikacji ze standardem

D.1.1.1. Skala nominalna

D.1.1.1.1. Testy nieparametryczne

D.1.1.1.1.1. Test z

Źródło: GÓRALSKI 1976: 153–154, patrz też 31–33

D.1.2. Porównanie wariancji ze standardem

D.1.2.1. Skala przedziałowa

D.1.2.1.1. Testy parametryczne

D.1.2.1.1.1. Test z

Źródło: FREUND 1971: 263–264, 267 (ćw. 4); GÓRALSKI 1976: 154–155; PAWŁOWSKI 1980: 176–178

D.1.2.1.1.2. Test χ^2

Źródło: GREŃ 1984: 83–86

D.2. Testy dla $m = 2$ prób

Problem: Testowana jest hipoteza zerowa, że rozpatrywane miary rozproszenia w porównywanych próbach nie różnią się w sposób istotny. Hipoteza ta może być testowana względem jednej z następujących alternatyw:

1. test dwustronny:

– porównywane miary rozproszenia różnią się istotnie;

2. test jednostronny:

– miara rozproszenia w próbie n_1 ma wyższą wartość niż w próbie n_2 (próba n_1 ma większe rozproszenie niż próba n_2);

– miara rozproszenia w próbie n_1 ma niższą wartość niż w próbie n_2 (próba n_1 ma mniejsze rozproszenie niż próba n_2).

D.2.1. Test porównujący rozproszenia względne dwu klasyfikacji

D.2.1.1. Skala nominalna

D.2.1.1.1. Zmienne niepołączone

D.2.1.1.1.1. Testy nieparametryczne

D.2.1.1.1.1.1. Test z

Źródło: GÓRALSKI 1976: 190–192

D.2.2. Testy porównujące rozproszenie wyników w próbach

D.2.2.1. Skala porządkowa

D.2.2.1.1. Zmienne niepołączone

D.2.2.1.1.1. Testy nieparametryczne

D.2.2.1.1.1.1. Test sumy kolejności W

Źródło: PUCHALSKI 1980: 303–304

D.2.2.1.1.1.2. Test sumy kolejności R

Źródło PUCHALSKI 1973: 205–208

D.2.2.1.1.1.3. Test U Manna-Whitneya (test Wilcozona, test sumy rang Wilcozona-Manna-Whitneya)

Uwaga: Test ten jest słabszy przy porównywaniu rozproszenia w próbach niż test Walda-Wolfowitza (D.2.2.1.1.1.4.)

Źródło: BLALOCK 1977: 225–231; DOMAŃSKI 1979: 148–150; FREUND 1971: 313–316, 317, 318 (ćw. 4 i 5); PAWŁOWSKI 1980: 238–240

D.2.2.1.1.1.4. Test serii Walda-Wolfowitza (Test Walda-Wolfowitza, test serii u , test serii)

Uwaga: Test ten nadaje się szczególnie do badania różnic między rozproszeniami wyników w próbach lub postaciami rozkładu. Por. uwagę przy teście D.2.2.1.1.1.3.

Źródło: BLALOCK 1977: 221–225; DOMAŃSKI 1979: 140–141; KRZYSZTOFIAK, URBANEK 1981: 265–266; PAWŁOWSKI 1980: 235–238; PERKAL 1967: 91–92; PUCHALSKI 1980: 293–294

D.2.2.1.1.1.5. Test Sadowskiego

Źródło: PERKAL 1967: 107–108

D.2.3. Testy oceniające istotność różnicy między wariancjami

Problem: Testowana jest hipoteza zerowa, że porównywane wariancje nie różnią się istotnie. Hipoteza ta może być testowana względem jednej z następujących alternatyw:

1. test dwustronny:

– porównywane wariancje różnią się istotnie;

2. test jednostronny:

– wariancja w próbie n_1 jest większa niż w próbie n_2 ;

– wariancja w próbie n_1 jest mniejsza niż w próbie n_2 .

D.2.3.1. Skala przedziałowa

D.2.3.1.1. Zmienne połączone

D.2.3.1.1.1. Testy parametryczne

D.2.3.1.1.1.1. Test Morgana

Źródło: ASKENAS, SAWICKI 1970: 63–64; GÓRALSKI 1976: 196–197

D.2.3.1.2. Zmienne niepołączone

D.2.3.1.2.1. Testy parametryczne

D.2.3.1.2.1.1. Test F Fishera-Snedecora

Źródło: ASKENAS, SAWICKI 1970: 62–64; FREUND 1971: 265–267; GÓRALSKI 1976: 192–193; GREŃ 1984: 88–90; OKTABA 1980: 130–131; PAWŁOWSKI 1980: 179; SAWICKI 1982: 133–135

D.3. Testy dla $m \geq 2$ prób

Problem: Hipoteza zerowa, że porównywane wariancje nie różnią się istotnie testowana jest względem alternatywy, że wariancje te różnią się między sobą w sposób istotny (test dwustronny).

D.3.1. Testy oceniające istotność różnic między wariancjami

D.3.1.1. Skala przedziałowa

D.3.1.1.1. Zmienne połączone

D.3.1.1.1.1. Testy parametryczne

D.3.1.1.1.1.1. Test q_s

Źródło: GÓRALSKI 1976: 197–198

D.3.1.1.2. Zmienne niepołączone

D.3.1.1.2.1. Testy parametryczne

D.3.1.1.2.1.1. Test Hartleya

Uwaga: Test ten zalecany jest szczególnie, zamiast testu Bartletta (D.3.1.1.2.1.2.), wówczas gdy liczebności prób są < 5 .

Źródło: GÓRALSKI 1976: 193–194

D.3.1.1.2.1.2. Test Bartletta

Uwaga: Jest to jeden z najczęściej stosowanych testów oceniających istotność różnic między wariancjami. Por. jednak uwagę przy teście D.3.1.1.2.1.1.

Źródło: GÓRALSKI 1976: 193–194; GREŃ 1984: 92–95; PAWŁOWSKI 1980: 179–181

D.3.1.1.2.1.3. Test Cochra

Źródło: PAWŁOWSKI 1980: 181–182

E. Testy istotności proporcji (frakcji, procentu, wskaźnika struktury)

E.1. Testy dla $m = 1$ próby

E.1.1. Porównanie proporcji ze standardem

Problem: Testowana jest hipoteza zerowa, że proporcja w próbie nie różni się istotnie od standardu. Hipoteza ta może być testowana względem jednej z następujących alternatyw:

1. test dwustronny:

– proporcja różni się istotnie od standardu;

2. test jednostronny:

– proporcja jest większa od standardu;

– proporcja jest mniejsza od standardu.

E.1.1.1. Skala nominalna

E.1.1.1.1. Testy parametryczne

E.1.1.1.1.1. Test z (test u) dla proporcji

Uwaga: Jest to test dla zmiennej zero-jedynkowej

Źródło: ASKENAS, SAWICKI 1970: 65–67; BLALOCK 1977: 172–175; DOMAŃSKI 1990: 94–97; FREUND 1971: 278–280; GREŃ 1984: 76–77; OKTABA 1980: 120–122; SAWICKI 1982: 93–99

E.2. Testy dla $m = 2$ prób

E.2.1. Testy porównujące dwie proporcje

Problem: Hipoteza zerowa, że porównywane proporcje są zgodne (tzn. porównywane próby pochodzą z tej samej populacji lub z różnych populacji o takim samym rozkładzie zmiennej, czyli o identycznej tendencji centralnej) może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane proporcje nie są zgodne i różnią się istotnie (tzn. porównywane próby pochodzą z populacji o różnych tendencjach centralnych);

2. test jednostronny:

– proporcja w próbie n_1 jest wyższa niż w próbie n_2 ;

– proporcja w próbie n_1 jest niższa niż w próbie n_2 .

E.2.1.1. Skala nominalna

E.2.1.1.1. Zmienne połączone

E.2.1.1.1.2. Testy parametryczne

E.2.1.1.1.2.1. Test z (test u): test różnicy między dwoma proporcjami

Uwaga: Jest to test dla zmiennych zero-jedynkowych

Źródło: ASKENAS, SAWICKI 1970: 68–69; SAWICKI 1982: 109–111

E.2.1.1.2. Zmienne niepołączone

E.2.1.1.2.2. Testy parametryczne

E.2.1.1.2.2.1. Test z (test u): test różnicy między dwoma proporcjami

Uwaga: Jest to test dla zmiennych zero-jedynkowych. Freund (1971) podaje modyfikację tego testu stosowaną do testowania H_0 , że różnica między porównywanymi proporcjami *równa się* danej wielkości stałej.

Źródło: ASKENAS, SAWICKI 1970: 67–68; BLALOCK 1977: 203–206; DOMAŃSKI 1990: 122–123; FREUND 1971: 281–283; 286 (ćw. 5); GREŃ 1984: 79–81; KRZYSZTOFIK, URBANEK 1981: 243–245; OKTABA 1980: 125–127; SAWICKI 1982: 108–109

E.3. Testy dla $m \geq 2$ prób

E.3.1. Testy porównujące więcej niż dwie proporcje

Problem: Testowana jest hipoteza zerowa, że różnice między porównywanymi proporcjami są nieistotne (próby pochodzą z tej samej populacji lub z różnych populacji o tych samych rozkładach, czyli o tej samej tendencji centralnej). Hipoteza ta jest testowana względem alternatywy, że różnice między porównywanymi proporcjami są istotne. Znaczy to, że przynajmniej jedna próba została wylosowana z populacji o innym rozkładzie (innej tendencji centralnej) (test dwustronny).

E.3.1.1. Skala nominalna

E.3.1.1.2. Zmienne niepołączone

E.3.1.1.2.1. Testy nieparametryczne

E.3.1.1.2.1.1. Test zgodności χ^2

Uwaga: Jest to test dla zmiennych zero-jedynkowych

Źródło: OKTABA 1980: 275–276

E.3.1.1.2.1.3. Test zgodności χ^2 porównujący frakcje sukcesów

Uwaga: Jest to test dla zmiennej dwumianowej (równocześnie przeprowadza się wiele doświadczeń, przy czym k jest znane):

Źródło: OKTABA 1980: 276–277

E.3.1.1.2.1.4. Test zgodności χ^2 porównujący frakcje sukcesów

Uwaga: Jest to test dla zmiennej Poissona (równocześnie przeprowadza się wiele doświadczeń, przy czym k jest nieznane, a $p < 0.1$).

Źródło: OKTAB 1980: 277

F. Testy zgodności rozkładów

F.1. Testy dla $m = 1$ próby

Problem: Testowana jest hipoteza zerowa, że zaobserwowany rozkład cechy jest zgodny z rozkładem standardowym, to znaczy z rozkładem określonego typu, na przykład: prostokątnym, dwumianowym, Poissona, normalnym, wykładniczym itd., lub dowolnym rozkładem hipotetycznym. Rozkład dwumianowy, Poissona i normalny zostały pokrótce scharakteryzowane wcześniej. Szczegółowych informacji na ich temat należy szukać w podręcznikach statystyki, np. u Góralskiego (1976). Tam też można znaleźć informacje o innych rozkładach zmiennej losowej, o których tutaj nie wspomiano, a które stosowane są do analizy rozmieszczenia obiektów biologicznych w przestrzeni. W tym względzie por. też Kwiatkowska, Symonides (1980). Pojęcie "dowolnego rozkładu hipotetycznego (teoretycznego)" wprowadzono tutaj na oznaczenie rozkładu standardowego, różnego od wszystkich rozkładów sformalizowanych, takich np. jak wymienione wyżej. Tak więc, dowolnym rozkładem hipotetycznym może być, stwierdzony na podstawie wcześniejszych badań, rozkład badanej cechy, z którym porównywane będą rozkłady tej cechy stwierdzone w próbach pobranych w innym czasie i/lub w innym miejscu. Oprócz rozkładów kompletnych, tzn. rozkładów, w których zmienna losowa przyjmuje wszystkie możliwe wartości, np. $x = 0, 1, \dots, k$ (zmienna skokowa), czasami możemy mieć do czynienia z rozkładami, nie zawierającymi żadnej informacji o liczebności klasy zerowej. Są to tzw. rozkłady ucięte lewostronnie. Na przykład rozkład liczby gatunków na zdjęcie fitosocjologiczne¹¹ jest rozkładem tego typu – nie ma bowiem zdjęć "pustych", czyli takich, w których nie byłoby żadnych roślin. Wynika to z charakteru samych danych. Taka sytuacja wymaga specjalnego traktowania. Numeryczny przykład testowania hipotezy zerowej, głoszącej że rozważana zmienna losowa ma rozkład dwumianowy (lub Poissona), w przypadku, gdy liczebność klasy zerowej jest nieznana, przedstawiono w opracowaniu Wołka i Dawidowicza (1991). Tam też można znaleźć informacje o innych rozkładach uciętych. Hipoteza zerowa, o której mowa, testowana jest względem hipotezy alternatywnej, że rozkład zaobserwowany nie jest zgodny z rozkładem standardowym.

F.1.1. Testy zgodności rozkładu zaobserwowanego z rozkładem hipotetycznym

F.1.1.1. Skala nominalna

F.1.1.1.1. Testy nieparametryczne

¹¹ Zdjęcie fitosocjologiczne, to – krótko mówiąc – spis gatunków roślin występujących na wytypowanej powierzchni wraz z określeniem stopnia ich ilościowości na tej powierzchni.

- F.1.1.1.1.1. Test zgodności χ^2 rozkładu zaobserwowanego z dowolnym rozkładem hipotetycznym
 Źródło: OKTABA 1980: 260–262; PAWŁOWSKI 1980: 215–(220) 221; SAWICKI 1982: 122–123
- F.1.1.1.1.2. Test zgodności χ^2 rozkładu zaobserwowanego z rozkładem prostokątnym
 Uwaga: O rozkładzie prostokątnym (równomiernym, jednostajnym) mówimy wówczas, gdy częstość wyników we wszystkich kategoriach cechy jest taka sama.
 Źródło: GÓRALSKI 1976: 224–226 (ćw. 14.2); OKTABA 1980: 280 (ćw. 1, 2, 3)
- F.1.1.1.1.3. Test zgodności χ^2 rozkładu zaobserwowanego z rozkładem dwumianowym
 Źródło: ASKENAS, SAWICKI 1970: 52–53; FREUND 1971: 290–292; OKTABA 1980: 267–268; PERKAL 1967: 76–78 (ćw. 6); SAWICKI 1982: 125 (przykład 16.4)
- F.1.1.1.1.4. Test zgodności χ^2 rozkładu zaobserwowanego z rozkładem Poissona
 Źródło: GÓRALSKI 1976: 224–226; GREŃ 1984: 117 (przykład 2); OKTABA 1980: 266–267; PAWŁOWSKI 1980: 215–(217) 221; PERKAL 1967: 78–79 (ćw. 7); SAWICKI 1982: 125–126 (przykład 16.5 i 16.6)
- F.1.1.1.1.5. Test zgodności χ^2 rozkładu zaobserwowanego z rozkładem normalnym
 Źródło: ASKENAS, SAWICKI 1970: 55; DOMAŃSKI 1979: 88–91; GREŃ 1984: 116 (przykład 1); PAWŁOWSKI 1980: 215–(218) 221; PERKAL 1967: 79–80 (ćw. 8); SAWICKI 1982: 123–124
- F.1.1.1.1.6. Test Hellwiga zgodności rozkładu zaobserwowanego z rozkładem normalnym (dla małej próby)
 Źródło: DOMAŃSKI 1979: 109–111
- F.1.1.2. Skala porządkowa
- F.1.1.2.1. Testy nieparametryczne
- F.1.1.2.1.1. Testy zgodności rozkładu zaobserwowanego z rozkładem prostokątnym
- F.1.1.2.1.1.1. Test Kołmogorowa-Smirnowa
 Źródło: DOMAŃSKI 1979: 91–100 (ćw. 4.2)
- F.1.1.2.1.1.2. Test oparty na różnicach absolutnych wartości dystrybuant
 Źródło: DOMAŃSKI 1979: 107–108
- F.1.1.2.1.2. Testy zgodności rozkładu zaobserwowanego z rozkładem wykładniczym
- F.1.1.2.1.2.1. Test Kołmogorowa (test Kołmogorowa-Smirnowa)
 Źródło: DOMAŃSKI 1979: 91–100 (ćw. 4.3)
- F.1.1.2.1.2.2. Test Cramera -von Misesa
 Źródło: DOMAŃSKI 1979: 101–103 (ćw. 4.3)
- F.1.1.2.1.2.3. Test Kuipera
 Źródło: DOMAŃSKI 1979: 103–104 (ćw. 4.3)
- F.1.1.2.1.2.4. Test Watsona
 Źródło: DOMAŃSKI 1979: 105–106 (ćw. 4.3)
- F.1.1.2.1.3. Testy zgodności rozkładu zaobserwowanego z rozkładem normalnym
- F.1.1.2.1.3.1. Test Kołmogorowa (Test λ Kołmogorowa)
 Źródło: DOMAŃSKI 1979: 91–100 (ćw. 4.4); GÓRALSKI 1976: 228–231; PAWŁOWSKI 1980: 212–214; PERKAL 1967: 81–82 (sposób graficzny), 82–84 (sposób numeryczny)
- F.1.1.2.1.3.2. Test Cramera-von Misesa
 Źródło: DOMAŃSKI 1979: 101–103 (ćw. 4.5)
- F.1.1.2.1.3.3. Test Kuipera
 Źródło: DOMAŃSKI 1979: 103–104 (ćw. 4.4)
- F.1.1.2.1.3.4. Test Watsona

Źródło: DOMAŃSKI 1979: 105–106 (ćw. 4.5)

F.1.1.2.1.3.5. Test Andersona-Darlinga

Źródło: DOMAŃSKI 1979: 106–107

F.1.1.2.1.3.6. Test Pawłowskiego

Źródło: DOMAŃSKI 1979: 122–124; PAWŁOWSKI 1980: 225–(227)229

F.1.1.2.1.3.7. Test ω^2 Smirnowa

Źródło: DOMAŃSKI 1979: 114–117

F.1.1.2.1.3.8. Test Cramera-Smirnowa

Źródło: DOMAŃSKI 1979: 117–120

F.1.1.2.1.3.9. Test Shapiro-Wilka

Źródło: DOMAŃSKI 1979: 120–122

F.1.1.2.1.3.10. Test przybliżony momentu trzeciego (asymetria) i momentu czwartego (kurtosis, spłaszczenie)

Źródło: PAWŁOWSKI 1980: 221–224; PUCHALSKI 1980: 250–251

F.2. Testy dla $m = 2$ prób

F.2.1. Testy porównujące rozkłady częstości w próbach

Problem: Hipoteza zerowa, że porównywane rozkłady częstości są zgodne (tzn. porównywane próby pochodzą z tej samej populacji lub z różnych populacji o takim samym rozkładzie zmiennej, czyli o identycznej tendencji centralnej) może być testowana wobec jednej z następujących alternatyw:

1. test dwustronny:

– porównywane rozkłady częstości nie są zgodne i różnią się istotnie (tzn. porównywane próby pochodzą z populacji o różnych rozkładach zmiennej);

2. test jednostronny:

– wartość miary położenia jednego z rozkładów jest większa (mniejsza) od wartości tej samej miary położenia drugiego rozkładu.

F.2.1.1. Skala nominalna

F.2.1.1.1. Zmienne połączone

F.2.1.1.1.1. Testy nieparametryczne

F.2.1.1.1.1.1. Test McNemary (test zgodności χ^2 dla tabeli czteropolowej typu 2×2)

Uwaga: Zmienna przyjmuje dwie wzajemnie wykluczające się kategorie

Źródło: ASKENAS, SAWICKI 1970: 72–73; GÓRALSKI 1976: 166; SAWICKI 1982: 116–117

F.2.1.1.1.2. Zmienne niepołączone

F.2.1.1.2.1. Testy nieparametryczne

F.2.1.1.2.1.1. Test zgodności χ^2 dla tabeli czteropolowej typu 2×2

Uwaga: Jeżeli ogólna liczebność $n_1 + n_2$ jest ≤ 20 lub wynosi $20 < n_1 + n_2 \leq 40$, a którakolwiek z wartości oczekiwanych jest ≤ 5 , to należy stosować test dokładny Fishera (patrz F.2.1.1.2.1.2.).

Źródło: ASKENAS, SAWICKI 1970: 70–72; BŁALOCK 1977: 241–251 (243–248), OKTABA 1980: 271–272; PUCHALSKI 1980: 290–293; SAWICKI 1982: 113–116

F.2.1.1.2.1.2. Test dokładny Fishera dla tabeli czteropolowej typu 2×2

Uwaga: Zapoznaj się z uwagą pod E.2.1.1.2.1.1. Test Fishera posługuje się dokładnymi prawdopodobieństwami, a nie ich przybliżeniami, jak to ma miejsce w przypadku testu χ^2 . Test χ^2 daje zwykle prawdopodobieństwa nieco niższe niż test Fishera, co upoważnia do odrzucenia prawdziwej hipotezy zerowej (błąd pierwszego rodzaju).

Źródło: BŁALOCK 1977: 251–254

F.2.1.1.2.1.3. Test zgodności χ^2 dla tabeli wielopolowej typu $2 \times k$

Źródło: ASKENAS, SAWICKI 1970: 73–74 (omawiają przypadek ogólny, tabela $m \times k$);
 BŁALOCK 1977: 241–251 (248–250); OKTABA 1980: 273–275 (tabela $m \times k$);
 PUCHALSKI 1980: 287–290 (tabela $m \times k$); SAWICKI 1982: 117–118

F.2.2. Testy porównujące zgodność uporządkowania elementów

Problem: Należą tutaj testy porównujące rozkłady uporządkowania wartości zmiennej losowej w dwu próbach. Testowana jest hipoteza zerowa, że w porównywalnych próbach rozkłady te są zgodne (tzn. obie próby pochodzą z jednej populacji lub z dwóch populacji o takich samych rozkładach, czyli o identycznej tendencji centralnej).

Hipoteza ta może być testowana wobec jednej z następujących alteratyw:

1. test dwustronny:

– porównywane rozkłady nie są zgodne (próby pochodzą z populacji o różnych tendencjach centralnych);

2. test jednostronny:

– wartość miary położenia jednego z rozkładów jest większa (mniejsza) od wartości tej samej miary położenia drugiego rozkładu.

F.2.2.1. Skala porządkowa

F.2.2.1.1. Zmienne połączone

F.2.2.1.1.1. Testy nieparametryczne

F.2.2.1.1.1.1. Test znaków

Uwaga: Test znaków uwzględnia tylko kierunek różnicy (+ lub –) między obserwacjami tworzącymi pary. Test ten jest alternatywą testu t różnicy między dwoma średnimi arytmetycznymi dla $m = 2$ prób zależnych (patrz C.2.3.1.1.2.). W przypadku testu znaków hipotezę zerową można sformułować następująco: prawdopodobieństwo uzyskania znaku (–) jest takie samo, jak prawdopodobieństwo uzyskania znaku (+) i wynosi $p = 0.5$. W przypadku testu dwustronnego jej alternatywą jest hipoteza, że $p \neq 0.5$, w przypadku testu jednostronnego hipoteza, że $p > 0.5$ lub $p < 0.5$.

Źródło: BŁALOCK 1977: 150–152; DOMAŃSKI 1979 145–146; FREUND 1971: 309–311; GÓRALSKI 1976: 171–172; GREŃ 1984: 145–147; KRZYSZTOFIAK, URBANEK 1981: 261–262; PAWŁOWSKI 1980: 229–235; PUCHALSKI 1978: 296–297; SAWICKI 1982: 163–165

F.2.2.1.1.1.2. Test Wilcozona dla par (test rangowanych znaków, test znaków kolejności T , test Wilcozona rang różnic dla par)

Uwaga: Test Wilcozona dla par uwzględnia nie tylko kierunek różnicy (+ lub –) między obserwacjami tworzącymi pary, ale również wielkość tych różnic. Uwzględnia więc większą ilość informacji niż test znaków. Test ten jest alternatywą testu t różnicy między dwoma średnimi arytmetycznymi dla $m = 2$ prób zależnych (C.2.3.1.1.2.). Pod względem mocy stoi między testem znaków (C.2.2.1.1.1.) a testem t .

Źródło: BŁALOCK 1977: 235–238; DOMAŃSKI 1979: 146–148; GÓRALSKI 1976: 172–173; GREŃ 1984: 149–152; PUCHALSKI 1980: 297–299; SAWICKI 1982: 165–167

F.2.2.1.2. Zmienne niepołączone

F.2.2.1.2.1. Testy nieparametryczne

F.2.2.1.2.1.1. Test medianowy

Uwaga: Dla obu prób łącznie wyznacza się wyniki mniejsze i większe od mediany. Użyte dane grupuje się w tabelę czteropolową (2×2) i oblicza się statystykę χ^2 . W przypadku testu medianowego H_0 można sformułować w następujący sposób: pra-

wdopodobieństwo uzyskania wyniku mniejszego lub większego od mediany w obu próbach jest takie samo i wynosi $p = 0.5$ albo – obie próby pochodzą z jednej populacji lub z dwóch populacji o takich samych medianach. Dla testu dwustronnego hipoteza alternatywna brzmi: $p \neq 0.5$, dla testu jednostronnego: $p > 0.5$ lub $p < 0.5$.

Źródło: DOMAŃSKI 1979: 153–155; GREŃ 1984: 153–155; KRZYSZTOFIAK, URBANEK 1981: 263–265

F.2.2.1.2.1.2. Test U Manna-Whitneya (test Wilcozona, test sumy rang, test Wilcozona-Manna-Whitneya)

Uwaga: Test ten ma większą moc przy porównywaniu prób pod względem tendencji centralnej niż przy porównywaniu dyspersji lub postaci rozkładu. Dlatego stanowi on bardzo mocną alternatywę testu t różnicy między dwoma średnimi arytmetycznymi dla dwu prób niepołączonych (patrz C.2.3.2.1.2.–4.).

Źródło: BLALOCK 1977: 225–231; DOMAŃSKI 1979: 148–150; FREUND 1971: 313–316; PAWŁOWSKI 1980: 238–240; SAWICKI 1982: 168–169

F.2.2.1.2.1.3. Test Wilcozona (test sumy kolejności)

Źródło: DOMAŃSKI 1979: 150–152; GÓRALSKI 1976: 168–170; PUCHALSKI 1980: 294–296

F.2.2.1.2.1.4. Test serii Walda-Wolfowitza (test Walda-Wolfowitza, test serii u , test serii)

Uwaga: Test ten jest słabszy przy porównaniu prób pod względem tendencji centralnej niż test U Manna-Whitneya, ale jest mocniejszy przy porównywaniu dyspersji lub postaci rozkładu.

Źródło: BLALOCK 1977: 221–225; DOMAŃSKI 1979: 140–141; GREŃ 1984: 140 (Model II), 143 (przykład 2); KRZYSZTOFIAK, URBANEK 1981 265–266; PAWŁOWSKI 1980: 235–238; PERKAL 1967: 91–92; PUCHALSKI 1980: 293–294

F.2.2.1.2.1.5. Test Kołmogorowa-Smirnowa (test Smirnowa, Test D Kołmogorowa-Smirnowa)

Uwaga: Test ten porównuje rozkłady częstości skumulowanych. Ma on moc porównywalną z testem Manna-Whitneya (F.2.2.1.2.1.2.).

Źródło: BLALOCK 1977: 232–235; DOMAŃSKI 1979: 126–131; GÓRALSKI 1976: 227–228; GREŃ 1984: 124 (Model II), 126 (przykład 2); PAWŁOWSKI 1980: 214–215; PERKAL 1967: 89–91; PUCHALSKI 1980: 299

F.2.2.1.2.1.6. Test Gniedienki-Korolyuka

Źródło: DOMAŃSKI 1979: 131–133

F.2.2.1.2.1.7. Test Cramera-von Misesa

Źródło: DOMAŃSKI 1979: 135–137

F.2.2.1.2.1.8. Test Kuipera

Źródło: DOMAŃSKI 1979: 137–138

F.2.2.1.2.1.9. Test Watsona

Źródło: DOMAŃSKI 1979: 138–139

F.2.2.1.2.1.10. Test Stevensa

Źródło: GÓRALSKI 1976: 226–227

F.3. Testy dla $m \geq 2$ prób

F.3.1. Testy porównujące rozkłady częstości w próbach

Problem: Testowana jest hipoteza zerowa, że rozkłady częstości są zgodne (próby pochodzą z tej samej populacji lub z różnych populacji o tych samych rozkładach, czyli, mówiąc inaczej, o tej samej tendencji centralnej). Hipoteza ta jest testowana względem alternatywy, że badane rozkłady częstości nie są

zgodne, tzn. że przynajmniej jedna próba została wylosowana z populacji o innym rozkładzie – innej tendencji centralnej (test dwustronny).

F.3.1.1. Skala nominalna

F.3.1.1.1. Zmienne połączone (klasyfikacja podwójna)

F.3.1.1.1.1. Testy nieparametryczne

F.3.1.1.1.1.1. Test Cochran Q

Uwaga: Test ten jest uogólnieniem testu McNemary (E.2.1.1.1.1.1.). Moc testu jest bliżej nieznaną. Za pomocą tego testu można badać $m \geq 2$ prób ze względu na dwie wykluczające się kategorie: "A" oraz "nie A".

Źródło: GÓRALSKI 1976: 166–168

F.3.1.1.2. Zmienne niepołączone (klasyfikacja pojedyncza)

F.3.1.1.2.1. Testy nieparametryczne

F.3.1.1.2.1.2. Test zgodności χ^2 dla $k \geq 2$ wykluczających się kategorii

Źródło: ASKENAS, SAWICKI 1970: 73–74; BLALOCK 1977: 241–251; FREUND 1971: 283–287; OKTABA 1980: 273–275; PUCHALSKI 1980: 287–290

F.3.2. Testy porównujące zgodność uporządkowania elementów w próbach

Problem: Hipoteza zerowa, że porównywane rozkłady są zgodne (próby pochodzą z tej samej populacji lub z różnych populacji o tych samych rozkładach, czyli o tej samej tendencji centralnej) weryfikowana jest względem alternatywy, że rozkłady te nie są zgodne (test dwustronny). Oznacza to, że przynajmniej jedna z badanych prób została wylosowana z populacji o innym rozkładzie zmiennej, czyli o innej tendencji centralnej. Niektóre z prezentowanych tu testów pozwalają stwierdzić, które właściwie próby różnią się istotnie, a które nie.

F.3.2.1. Skala porządkowa

F.3.2.1.1. Zmienne połączone (klasyfikacja podwójna)

F.3.2.1.1.1. Testy nieparametryczne

Uwaga: Prezentowane tutaj testy są nieparametrycznymi alternatywami analizy wariancji dla klasyfikacji podwójnej.

F.3.2.1.1.1.1. Test Friedmana

Uwaga: Jest to uogólniony test znaków na dowolną liczbę prób. Charakteryzuje się on dużą mocą, która odpowiada mocy testu F Snedecora. Jeżeli wynik testu zmusza do odrzucenia hipotezy zerowej, to aby stwierdzić, które z badanych prób różnią się względem siebie, a które nie, trzeba zastosować odpowiedni test (patrz: Steczkowski, Zeliaś 1982: 213).

Źródło: GÓRALSKI 1976: 174–175

F.3.2.1.1.1.2. Test medianowy

Źródło: PUCHALSKI 1973: 241–244; KRZYSZTOFIAK, URBANEK 1981: 263–265

F.3.2.1.1.1.3. Test sumy kolejności W

Źródło: PUCHALSKI 1980: 304–307

F.3.2.1.1.1.4. Test sumy kolejności dla porównań wielokrotnych: próby równoliczne (test McDonalda-Thompsona)

Uwaga: Po odrzuceniu hipotezy zerowej, test ten pozwala odpowiedzieć na pytanie, które próby różnią się między sobą istotnie, a które nie.

Źródło: PUCHALSKI 1973: 250–251

F.3.2.1.1.1.5. Test wielostopniowy m

Uwaga: Por. uwagę do testu F.3.2.1.1.1.4.

Źródło: PUCHALSKI 1973: 229–232, 252–253

F.3.2.1.2. Zmienne niepołączone (klasyfikacja pojedyncza)

F.3.2.1.2.1. Testy nieparametryczne

Uwaga: Prezentowane tutaj testy są nieparametrycznymi alternatywami analizy wariancji dla klasyfikacji pojedynczej.

F.3.2.1.2.1.1. Test H Kruskala-Wallisa (test sumy rang)

Uwaga: Test ten jest uogólnieniem testu U Manna-Whitneya-Wilcoxon na m prób niezależnych. Jest on bardziej efektywny niż test medianowy, bo wykorzystuje więcej informacji. Pod względem mocy porównuje się go z testem F , a jego wydajność ocenia się na 95%. Test stosuje się zarówno dla uporządkowań zupełnych, jak i w przypadku gdy występują rangi związane. Poprawka na rangi związane uwzględniana jest przez niektórych autorów, patrz np. Steczkowski i Zeliaś (1982) oraz Blalock (1977).

Źródło: BLALOCK 1977: 295–297; DOMAŃSKI 1979: 152–153; FREUND 1971: 316–317; GÓRALSKI 1976: 170–171; KRZYSZTOFIK, URBANEK 1981: 266–267; PUCHALSKI 1980: 300–301; SAWICKI 1982: 169–170

F.3.2.1.2.1.2. Test Q_7^2

Źródło: PAWŁOWSKI 1980: 240–243

F.3.2.1.2.1.3. Test wielostopniowy m dla prób równolicznych (test Conovera)

Uwaga: Test ten ma mniejszą moc niż test H Kruskala-Wallisa. Częściej może zdarzyć się popełnienie błędu drugiego rodzaju. Po odrzuceniu hipotezy zerowej, test ten pozwala odpowiedzieć na pytanie, które próby różnią się między sobą istotnie, a które nie.

Źródło: PUCHALSKI 1980: 301–303

F.3.2.1.2.1.4. Test Smirnowa dla prób równolicznych (dla $m = 3$)

Źródło: DOMAŃSKI 1979: 134–135

F.3.2.1.2.1.5. Test serii wielokrotnych dla prób równolicznych (dla $m \geq 3$)

Źródło: DOMAŃSKI 1979: 142–145

F.3.2.1.2.1.6. Test sumy kolejności dla porównań wielokrotnych dla prób równolicznych.

Uwaga: Po odrzuceniu hipotezy zerowej, test ten pozwala stwierdzić, które próby różnią się między sobą istotnie, a które nie.

Źródło: PUCHALSKI 1973: 228–229

G. Testy niezależności cech

Uwaga: Są to testy badające współzależność (współzmiennność) między cechami, a mianowicie jej kształt (regresję) i siłę (korelację).

G.1. Testy stosowane w analizie regresji

G.1.1. Testy służące do wykrywania autokorelacji¹²

Problem: Hipoteza zerowa, że autokorelacja nie istnieje testowana jest względem alternatywy, że autokorelacja istnieje (test dwustronny).

Uwaga: autokorelacja występuje często wówczas, gdy wątpliwy jest liniowy charakter badanej regresji i model regresji nieliniowej byłby lepszy.

¹² Do wykrywania autokorelacji można wykorzystać niektóre testy sprawdzające losowość próby (Domański 1990). Testy autokorelacji, zarówno parametryczne jak i nieparametryczne, szczegółowo omawia Tomaszewicz (1985).

G.1.1.1. Skala przedziałowa

G.1.1.1.1. Testy parametryczne

G.1.1.1.1.1. Test Durbina-Watsona

Źródło: KRZYSZTOFIAK, URBANEK 1981: 382–387

G.1.1.1.1.2. Test von Neumanna

Źródło: GREŃ 1984: 227–235 (w przykładzie 2 na str. 232 autor demonstruje, kolejno, sposób: 1. testowania hipotezy o braku autokorelacji, 2. estymacji współczynnika autokorelacji oraz 3. estymację parametrów równania regresji liniowej w sytuacji gdy autokorelacja istnieje).

G.1.2. Testy sprawdzające liniowość regresji

Problem: Hipoteza zerowa, że regresja y na x jest liniowa (tzn. jest postaci $y = a + bx$) testowana jest względem hipotezy alternatywnej, że regresja nie jest liniowa (test dwustronny). Sprawdzana jest więc hipoteza o kształcie regresji.

G.1.2.1. Skala porządkowa

G.1.2.1.1. Testy nieparametryczne

G.1.2.1.1.1. Test serii

Źródło: DOMAŃSKI 1979: 172–175; GREŃ 1984: 140 (Model III), 142 (przykład 3); SAWICKI 1982: 170–172

G.1.2.2. Skala przedziałowa

G.1.2.2.1. Testy parametryczne

G.1.2.2.1.1. Test F Fishera

Źródło: BLALOCK 1977: 353–353; GÓRALSKI 1976: 291–292

G.1.3. Testy sprawdzające czy 2 linie regresji są identyczne

Problem: Hipoteza zerowa, że dwie linie regresji są identyczne (test testowana względem alternatywy, że te linie regresji nie są identyczne (test dwustronny)).

G.1.3.1. Skala porządkowa

G.1.3.1.1. Testy nieparametryczne

G.1.3.1.1.1. Test serii

Źródło: DOMAŃSKI 1979: 176–178

G.1.4. Testy istotności współczynnika regresji liniowej (b)

Problem: Hipoteza zerowa, że współczynnik regresji liniowej w populacji równy jest zero (brak zależności między x i y , czyli nachylenie prostej regresji jest nieistotne) testowany jest względem jednej z następujących alternatyw:

1. test dwustronny:

– współczynnik regresji liniowej jest różny od zera;

2. test jednostronny:

– współczynnik regresji jest większy od zera;

– współczynnik regresji jest mniejszy od zera.

Za pomocą prezentowanych testów można również weryfikować H_0 , że współczynnik regresji jest równy określonej wartości (standardowi). Wówczas hipoteza alternatywna będzie brzmieć, że współczynnik ten różni się od standardu (test dwustronny) lub jest większy (mniejszy) od standardu (test jednostronny).

G.1.4.1. Skala porządkowa

G.1.4.1.1. Testy nieparametryczne

G.1.4.1.1.1. Test Theila

Źródło: DOMAŃSKI 1979: 178–180

G.1.4.2. Skala przedziałowa

G.1.4.2.1. Testy parametryczne

G.1.4.2.1.1. Test t

Źródło: FREUND 1971: 339; GREŃ 1984: 181–183; KRZYSZTOFIAK, URBANEK 1981: 285, 380–381; SAWICKI 1982: 138–139 (przykład 18.1)

G.1.5. Testy istotności różnicy 2 współczynników regresji liniowej

Problem: Hipoteza zerowa, że 2 współczynniki regresji liniowej są identyczne (tzn., że dwie proste regresji są równoległe lub nachylenie dwóch prostych regresji jest identyczne) testowana jest względem jednej z następujących alternatyw:

1. test dwustronny:

– porównywane współczynniki nie są identyczne;

2. test jednostronny:

– współczynnik regresji w próbie n_1 jest wyższy niż w próbie n_2 ;

– współczynnik regresji w próbie n_1 jest niższy niż w próbie n_2 .

G.1.5.1. Skala przedziałowa

G.1.5.1.1. Testy parametryczne

G.1.5.1.1.1. Test t

Źródło: GREŃ 1984: 186–189; PERKAL 1967: 120–121 (ćw. 10)

G.2. Testy stosowane w analizie korelacji

G.2.1. Badanie siły współzależności między $m = 2$ cechami

Problem: Hipoteza zerowa, że dwie cechy nie są współzależne (cechy są niezależne), testowana jest względem alternatywy, że cechy te są współzależne (cechy są zależne) – test dwustronny.

G.2.1.1. Skala nominalna

G.2.1.1.1. Testy nieparametryczne

G.2.1.1.1.1. Test niezależności χ^2 dla czteropolowych tablic kontyngencji typu 2×2

Źródło: BŁALOCK 1977: 241–251 (243–248); DOMAŃSKI 1979: 158–162; GÓRALSKI 1976: 207–210; OKTABA 1980: 271–272; PUCHALSKI 1980: 266–269; SAWICKI 1982: 113–116

G.2.1.1.1.2. Test niezależności χ^2 dla wielopolowych tablic kontyngencji typu $w \times 2$ z podziałem stopni swobody

Uwaga: Ta forma testu χ^2 pozwala porównać poszczególne kategorie jednej zmiennej ze względu na częstość występowania kategorii drugiej zmiennej, np. porównanie grup wiekowych ludzi ze względu na częstość objawów badanej choroby. Test pozwala ocenić, między którymi kategoriami wiekowymi różnice są istotne, a między którymi nieistotne.

Źródło: SAWICKI 1982: 118–121

G.2.1.1.1.3. Test χ^2 dla wielopolowych tablic kontyngencji ($w \times k$)

Źródło: ASKENAS, SAWICKI 1970: 73–74; BŁALOCK 1977: 241–251; DOMAŃSKI 1979: 158–162; FREUND 1971: 287–290; GÓRALSKI 1976: 207–210; GREŃ 1984: 130–134; OKTABA 1980: 273–275; PUCHALSKI 1980: 287–290; SAWICKI 1982: 117–118

G.2.1.2. Skala porządkowa

G.2.1.2.1. Testy nieparametryczne

G.2.1.2.1.1. Test niezależności oparty na liczbie serii

Źródło: DOMAŃSKI 1979: 162–164

G.2.2. Badanie współzależności między $m > 2$ cechami

Problem: Hipoteza zerowa, że między m cechami nie ma współzależności (badane cechy są niezależne), testowana jest względem alternatywy, że współzależność ta między badanymi cechami istnieje (badane cechy są zależne) – test dwustronny.

G.2.2.1. Skala nominalna

G.2.2.1.1. Testy nieparametryczne

G.2.2.1.1.1. Test niezależności χ^2 dla $m = 3$ cech, z których każda obejmuje dwie wykluczające się kategorie

Uwaga: Próba n elementowa, wylosowana z populacji ludzi aktywnych zawodowo, rozpatrywana jest ze względu na 3 cechy: 1. wydajność w pracy (wysoka–niska); 2. kondycja fizyczna (dobra–zła); 3. stan cywilny (wolny–żonaty, zamężna). Test ten daje informację o charakterze globalnym. Stwierdza tylko, że cechy są lub nie są związane (zależne).

Źródło: GÓRALSKI 1976: 207–210; OKTABA 1980: 277–290

G.2.2.1.1.2. Metoda interakcji testu χ^2 (Lancastera) dla $m = 3$ cech, z których każda obejmuje dwie kategorie

Uwaga: Ten test jest jakby rozwinięciem testu G.2.2.1.1.1. Informuje nie tylko o związku między badanymi cechami, ale również o tym, czy między badanymi cechami występuje współdziałanie (interakcja), i czy to współdziałanie jest istotne.

Źródło: GÓRALSKI 1976: 210–214

G.2.3. Ustalanie istotności nadwyżek i niedoborów liczebności w przypadku, gdy stwierdzono, że badane cechy są współzależne

Problem: Stosowane tu testy pozwalają wykryć istotne nadwyżki i/lub niedobory liczebności przypadających na poszczególne kategorie badanych cech. Stanowi to punkt wyjścia analizy objaśniającej mechanizm współzależności między cechami, wykrytej wcześniej za pomocą testu niezależności χ^2 (tabele kontyngencji). Hipoteza zerowa, że stwierdzone nadwyżki i/lub niedobory liczebności są nieistotne może być testowana względem alternatywy, że te nadwyżki i/lub niedobory są istotne (test dwustronny).

G.2.3.1. Skala nominalna

G.2.3.1.1. Testy nieparametryczne

G.2.3.1.1.1. Wskaźnik nadwyżki Wankego dla $m = 2$ cech

Źródło: PERKAL 1967: 129

G.2.3.1.1.2. Test ϕ dla $m > 2$ cech

Źródło: GÓRALSKI 1976: 212–214 (podaje przykład dla $m = 3$ cech)

G.3. Miary korelacji

G.3.1. Nieparametryczne miary korelacji

G.3.1.1. Skala nominalna

G.3.1.1.1. Czteropolowe tablice kontyngencji typu 2×2

G.3.1.1.1.1. Wskaźnik współzmienności Yule'a (współczynnik zbieżności, współczynnik Pearsona)

Źródło: KRZYSZTOFIAK, URBANEK 1981: 311–315, (313); PERKAL 1967: 123–124 (przykład 12 i 13); PUCHALSKI 1980: 164–167

G.3.1.1.1.2. Współczynnik Q Yule'a (współczynnik zbieżności, miernik Q Kendalla)

Źródło: BLALOCK 1977: 260–261; KRZYSZTOFIAK, URBANEK 1981: 311–315 (313)

G.3.1.1.1.3. Współczynnik W Bykowskiego

Źródło: KRZYSZTOFIAK, URBANEK 1981: 311–315 (313)

G.3.1.1.1.4. Współczynnik siły związku, r_p

Źródło: GÓRALSKI 1976: 157–159, patrz też 33–35

G.3.1.1.2. Wielopolowe tablice kontyngencji typu $w \times k$ (w kategorii cechy A \times k kategorii cechy B)

G.3.1.1.2.1. Współczynnik kontyngencji C (wskaźnik kontyngencji Pearsona, miernik zbieżności C Pearsona)

Uwaga: Współczynnik ten jest szeroko stosowaną miarą współzależności cech. Tym niemniej należy sobie zdawać sprawę z jego ograniczeń. Współczynnik C nie osiąga jedności nawet wtedy, gdy istnieje pełna współzależność cech: jego wartość rośnie ze wzrostem liczby kolumn i wierszy, zawsze jednak jest mniejsza od jedności. Z tego względu miernik ten jest trudny do interpretacji, chyba, że przeprowadzi się standaryzację, dzieląc jego wartość przez maksymalną wartość możliwą dla danej liczby r i k . W tym względzie porównaj G.3.1.1.2.2. Współczynnika C nie można bezpośrednio porównać z żadną inną miarą korelacji, np. Pearsona r , Spearmana r_s lub Kendalla τ .

Źródło: BLALOCK 1977: 258–260; PERKAL 1967: 124–128; PUCHALSKI 1980: 164–167

G.3.1.1.2.2. Współczynnik siły związku, r_p

Uwaga: Współczynnik ten jest zmodyfikowaną formą współczynnika kontyngencji C (G.3.1.1.2.1.), pozwalającą na standaryzację tego miernika dla wielopolowych tabel kontyngencji. Wypracowany on został przez Góralskiego (1976). Autor nie podaje sposobu testowania istotności tego współczynnika.

Źródło: GÓRALSKI 1976: 33–35

G.3.1.1.2.3. Współczynnik V Cramera (miara współzależności cech)

Źródło: BLALOCK 1977: 258–260

G.3.1.1.2.4. Współczynnik zbieżności T Czuprowa

Źródło: BLALOCK 1977: 258–260; KRZYSZTOFIAK, URBANEK 1981: 275–276

G.3.1.2. Skala porządkowa

G.3.1.2.1. Badanie niezależności $m = 2$ cech porządkujących

Problem: Hipoteza zerowa, że dany współczynnik korelacji jest równy zero (nie różni się istotnie od zera), testowana jest względem alternatywy, że jest on różny od zera (test dwustronny).

G.3.1.2.1.1. Współczynnik korelacji rang r_s Spearmana (wskaźnik korelacji uporządkowania, współczynnik korelacji kolejnościowej)

Uwaga: Współczynnik ten służy do porównywania uszeregowania (uporządkowania) elementów próby ze względu na dwie zmienne (cechy), np. uporządkowanie próby złożonej z 10 osób ze względu na wiek (jedna cecha) i ze względu na stan majątkowy (druga cecha). Robi się to w ten sposób, że osoba (element próby), posiadająca najmniejszą wartość danej cechy otrzymuje rangę 1. Każdy następny, "lepsz" od poprzedniego, wynik otrzymuje kolejną rangę: 2, 3, ..., itd. Współczynnik, o którym mowa, nadaje się do uporządkowań zupełnych, tzn. że wszystkie elementy próby można uporządkować za pomocą relacji $>$ lub $<$, oraz do uporządkowań słabych (rangi powiązane), tzn. że w próbie występują elementy identyczne, np. osoby urodzone w tym samym roku. Jeżeli jednak liczba rang powiązanych jest duża, poleca się stosować wówczas współczynnik korelacji rang Kendalla τ (G.3.1.2.1.2.). Współczynnik r_s należy interpretować analogicznie jak współczynnik korelacji liniowej r Pearsona, bo współczynniki te różnią się nieznacznie, szczególnie przy dużych próbach ($n > 10$).

Źródło: BLALOCK 1977: 354–355; DOMAŃSKI 1979: 168–169; GÓRALSKI 1976: 41–42; 215; KRZYSZTOFIAK, URBANEK 1981: 304–306; PERKAL 1967: 116–118

- G.3.1.2.1.2. Współczynnik korelacji rang τ Kendalla
 Uwaga: Współczynnik ten może być z łatwością stosowany wówczas, gdy występuje duża liczba rang wiązanych.
 Źródło: BLALOCK 1977: 356–362; DOMAŃSKI 1979: 164–168; GÓRALSKI 1976: 40–45, 215
- G.3.1.2.1.3. Wskaźnik korelacji medianowej
 Źródło: PERKAL 1967: 114–115
- G.3.1.2.1.4. Wskaźnik Tukeya (dywanowy) korelacji
 Źródło: PERKAL 1967: 115–116
- G.3.1.2.2. Badanie niezależności $m > 2$ cech porządkujących
- G.3.1.2.2.1. Współczynnik konkordancji w Kendalla
 Źródło: GÓRALSKI 1976: 40–45 (44), 216
- G.3.2. Parametryczne miary korelacji
- G.3.2.1. Skala przedziałowa
- G.3.2.1.1. Korelacja między $m = 2$ cechami
- G.3.2.1.1.1. Korelacja prostoliniowa
- G.3.2.1.1.1.1. Testowanie istotności współczynnika korelacji
- G.3.2.1.1.1.1.1. Test t Studenta
 Źródło: GREŃ 1984: 165 (Model II), 167 (przykład 2); OKTABA 1980: 134; PUCHALSKI 1980: 329–330
- G.3.2.1.1.1.1.2. Test analizy wariancji (test F)
 Źródło: BLALOCK 1977: 337–340
- G.3.2.1.1.1.1.3. Tablice Wallace'a-Snedecora
 Źródło: KRZYSZTOFIAK, URBANEK 1981: 300–301; FREUND 1971: 350–351
- G.3.2.1.1.1.2. Test z (test u) dla porównania współczynnika korelacji ze standardem
 Źródło: GÓRALSKI 1976: 159–160; GREŃ 1984: 166 (Model II), 169 (przykład 3)
- G.3.2.1.1.1.3. Test z Fishera dla porównania 2 współczynników korelacji
 Źródło: BLALOCK 1980: 345–348 (346–347), GÓRALSKI 1976: 203–204; PUCHALSKI 1980 330–332
- G.3.2.1.1.1.4. Test χ^2 dla porównania 2 współczynników korelacji
 Źródło: GÓRALSKI 1976: 204–205

H. Szacowanie przedziałów ufności

- H.1. Szacowanie przedziałowe miar położenia
- H.1.1. Skala nominalna
- H.1.1.1. Przedział ufności dla proporcji
 Źródło: ASKENAS, SAWICKI 1970: 77–78; BLALOCK 1977: 188–190; FREUND 1971: 268–272; GÓRALSKI 1976: 150; GREŃ 1984: 33–38; KRZYSZTOFIAK, URBANEK 1981: 235–237; OKTABA 1980: 144; PAWŁOWSKI 1980: 122–124; PUCHALSKI 1980: 328–329; SAWICKI 1982: 99–100
- H.1.2. Skala przedziałowa
- H.1.2.1. Przedział ufności dla mediany
 Źródło: GÓRALSKI 1976: 152–153; KRZYSZTOFIAK, URBANEK 1981: 245; PUCHALSKI 1980: 327
- H.1.2.2. Przedział ufności dla średniej arytmetycznej
 Źródło: ASKENAS, SAWICKI 1970: 76–77; BLALOCK 1977: 183–186, 188; FREUND 1971:

219–226; GÓRALSKI 1976: 151–152; GREŃ 1984: 21–33; KRZYSZTOFIAK, URBA-
NEK 1981: 230–235; OKTABA 1980: 139–141; PAWŁOWSKI 1980: 111–116; PU-
CHALSKI 1980: 323–324, 324–326; SAWICKI 1982: 95–99

H.1.2.3. Przedział ufności dla różnicy średnich arytmetycznych: zmienne połączone

Źródło: OKTABA 1980: 141–142

H.1.2.4. Przedział ufności dla różnicy średnich arytmetycznych: zmienne niepołączone

Źródło: OKTABA 1980: 142–143

H.1.2.5. Przedział ufności dla średniej λ rozkładu Poissona

Źródło: OKTABA 1980: 144–145

H.2. Szacowanie przedziałowe miar rozproszenia

H.2.1. Skala nominalna

H.2.1.1. Przedział ufności dla względnej dyspersji klasyfikacji

Źródło: GÓRALSKI 1976: 154

H.2.2. Skala przedziałowa

H.2.2.1. Przedział ufności dla wariancji i odchylenia standardowego

Źródło: FREUND 1971: 259–262; GÓRALSKI 1976: 155; GREŃ 1984: 38–45; PAWŁOWSKI
1980: 120–121; PUCHALSKI 1980: 327

H.3. Szacowanie przedziałowe miar korelacji

H.3.1. Skala nominalna

H.3.1.1. Przedział ufności dla współczynnika siły związku

Źródło: GÓRALSKI 1976: 157

H.3.2. Skala przedziałowa

H.3.2.1. Przedział ufności dla współczynnika korelacji liniowej

Źródło: BLALOCK 1977: 340–343; FREUND 1971: 353 (ćw. 10); GÓRALSKI 1976: 160–
161; GREŃ 1984: 164 (Model I), 166 (przykład 1); PUCHALSKI 1980: 332–333

H.4. Szacowanie przedziałowe w analizie regresji

H.4.1. Skala przedziałowa

H.4.1.1. Przedział ufności dla współczynnika regresji liniowej

Źródło: BLALOCK 1977: 343–344; FREUND 1971: 337–341; GREŃ 1984: 176 (Model II,
przykład); SAWICKI 1982: 139 (przykład 18.1)

H.4.1.2. Przedział ufności dla równania regresji

Źródło: BLALOCK 1977: 344–345; GREŃ 1984: 175 (Model I), 176 (przykład)



LITERATURA

- AHRENS H. 1970: Analiza wariancji. PWN, Warszawa.
- ASKENAS Z. SAWICKI F. (red.) 1970: Metody statystyczne w kardiologii. Biblioteka Kardiologiczna T. I. PZWL, Warszawa.
- BARNETT V. 1982: Elementy teorii pobierania prób. PWN, Warszawa.
- BLALOCK H. M. 1977: Statystyka dla socjologów. PWN, Warszawa.
- BOGUCKI Z. 1979: Elementy statystyki dla biologów. Statystyka opisowa. Wyd. Nauk. UAM w Poznaniu, Poznań.
- BYZOW Ł. A. 1951: Graficzne metody w statystyce, planowaniu i ewidencji. PWG, Warszawa.
- ✓ DOMAŃSKI C. 1979: Statystyczne testy nieparametryczne. PWE, Warszawa.
- DOMAŃSKI C. 1990: Testy statystyczne. PWE, Warszawa.
- ELANDT R. 1964: Statystyka matematyczna. PWN, Warszawa.
- FALIŃSKA K. 1990: Osobnik, populacja, fitocenoza. PWN, Warszawa.
- FALIŃSKI J. B. 1990: Kartografia geobotaniczna. T. 1-3. PPWK im. E. Romera, Warszawa-Wrocław.
- FISZ M. 1969: Rachunek prawdopodobieństwa i statystyka matematyczna. PWN, Warszawa.
- FREUND J. E. 1971: Podstawy nowoczesnej statystyki. PWE, Warszawa.
- GAWECKI J., WAGNER W. 1984: Podstawy metodologii badań doświadczalnych w nauce o żywieniu i żywności. PWN, Warszawa-Poznań.
- GŁOWACIŃSKI Z. 1975: Succession of bird communities in the Niepołomice Forest (southern Poland). *Ekol. pol.* 23(2): 231-263.
- GÓRALSKI A. 1976: Metody opisu i wnioskowania statystycznego w psychologii. PWN, Warszawa.
- GREŃ J. 1984: Statystyka matematyczna. Modele i zadania. PWN, Warszawa.
- GRODZIŃSKA K., SZAREK G., GODZIK B., BRANIEWSKI S., CHRZANOWSKA E.: Air pollution mapping in Poland by heavy metal concentration in moss (npbl.)
- GUILFORD J. P. 1964: Podstawowe metody statystyczne w psychologii i pedagogice. PWN, Warszawa.
- HELLWIG Z. 1977: Elementy rachunku prawdopodobieństwa i statystyki matematycznej. PWN, Warszawa.
- JONCKERS L. H. M. 1973: The concept of population in biology. Algal analysis with a suggestion to diminish the confusion to thought. *Acta Biotheoretica* 22: 78-108.
- KRZYSZTOFIAK M. 1971: Liczby, tablice wykresy. PWE, Warszawa.
- KRZYSZTOFIAK M., URBANEK D. 1981: Metody statystyczne. PWN, Warszawa.
- KWIATKOWSKA A. J., SYMONIDES E. 1980: Przegląd metod oceny typu rozkładu przestrzennego populacji roślinnych. *Wiad. ekol.* 26: 25-56.

- LEITNER R., ŻAKOWSKI W. 1975: Matematyka dla kandydatów na wyższe uczelnie. Cz. II. Wyd. Nauk.-Tech., Warszawa.
- LUSZNIWICZ A. 1980: Statystyka ogólna. PWE, Warszawa.
- MARTIN J. 1972: Podstawy matematyki i statystyki dla biologów, lekarzy i farmaceutów. PZWL, Warszawa.
- MARSZAŁKOWICZ T. 1980: Metody statystyki opisowej w badaniach ekonomiczno-rolniczych. PWN, Warszawa.
- NIEDOKOS E. 1990: Zastosowania rachunku prawdopodobieństwa i statystyki matematycznej. PWN, Warszawa.
- NORCLIFFE G. B. 1986: Statystyka dla geografów. PWN, Warszawa.
- OCHRONA ŚRODOWISKA, 1991. GUS, Warszawa.
- OCHYRA R., TOMASZEWICZ H. 1979: Nowe stanowiska *Ricciocarpos natans* (L.) Corda (*Ricciaceae*, *Hepaticopsida*) i przegląd jego rozmieszczenia w Polsce. *Fragm. Flor. Geobot.* 25(3): 429-438.
- OKTABA W. 1971: Metody statystyki matematycznej w doświadczalnictwie. PWN, Warszawa.
- OKTABA W. 1980: Elementy statystyki matematycznej i metodyka doświadczalnictwa. PWN, Warszawa.
- OSIPOW I. 1957: Jak sporządzać wykresy statystyczne. PWG, Warszawa.
- ✓ PARKER R. E. 1978: Wprowadzenie do statystyki dla biologów. PWN, Warszawa.
- PASCHAWER I. 1967: Prawo wielkich liczb i prawidłowości procesu masowego. PWE, Warszawa.
- PAWŁOWSKI Z. 1972: Wstęp do statystycznej metody reprezentacyjnej. PWN, Warszawa.
- PAWŁOWSKI Z. 1980: Statystyka matematyczna. PWN, Warszawa.
- PERKAL J. 1963: Matematyka dla przyrodników i rolników. Cz. II. PWN, Warszawa.
- PERKAL J. 1967: Matematyka dla przyrodników i rolników. Cz. III. PWN, Warszawa.
- PLATT C. 1981: Problemy rachunku prawdopodobieństwa i statystyki matematycznej. PWN, Warszawa.
- POCIECHA J. 1986: Statystyczne metody segmentacji rynku. *Zesz. Nauk. Akad. Ekonom. w Krakowie, Ser. specjalna: Monografie* 71: 1-188.
- PUCHALSKI T. 1973: Wnioskowanie statystyczne (statystyka matematyczna). PWN, Warszawa, Skrypt.
- PUCHALSKI T. 1980: Statystyka. Wykład podstawowych zagadnień. PWN, Warszawa.
- RAPORT O STANIE, ZAGROŻENIU I OCHRONIE ŚRODOWISKA. 1990. GUS, Warszawa.
- SAWICKI F. 1982: Elementy statystyki dla lekarzy. PZWL, Warszawa.
- SRB A. M., OWEN R. D. 1959: Genetyka ogólna. PWRiL, Warszawa.
- STECZKOWSKI J. 1970: Statystyczna procedura określania struktury zbiorowości. *Zesz. Nauk. Akad. Ekonom. w Krakowie, Ser. specjalna: Rozprawy habilitacyjne* 21: 1-131.
- STECZKOWSKI J., ZELIĄŚ A. 1982: Analiza wariancyjna i kowariancyjna w badaniach ekonomicznych. PWN, Warszawa.
- TOMASZEWICZ A. S. 1985: Jednorównaniowe modele ekonometryczne przy nieklasycznych założeniach. *Acta Univ. Lodziensis.*

- WEINER J. 1985: Korelacja i regresja, czyli o szkodliwości kalkulatorów. *Wiad. ekol.* 31(1): 67–84.
- WILSON E., Jr. 1968: Wstęp do badań naukowych. PWN, Warszawa.
- WŁOCZEWSKI T., KĘDZIERSKI Z. 1965: Metodyka leśnych badań hodowlanych. PWRiL, Warszawa.
- WOLEK J., DAWIDOWICZ L. D. 1991: Rozkłady ucięte, dwumianowy i Poissona, w badaniach ekologicznych: zastosowanie praktyczne. *Wiad. ekol.* 37(1): 27–41.
- WÓJCICKI J. J. 1991: Variability of *Prunus fruticosa* Pall. and the problem of an anthropybridization. *Veröff. Geobot. Inst. ETH, Stiftung Rübel, Zürich* 106: 266–272.
- ZAJĄC K. 1971: Zarys metod statystycznych. PWE, Warszawa.
- ZARZYCKI K. 1976: Ecodiagrams of common vascular plants in the Pieniny Mountains (the Polish West Carpathians). *Fragm. Flor. Geobot.* 22(4): 479–497.
- ZASEPA R. 1962: Badania statystyczne metodą reprezentacyjną. Zarys teorii i praktyki. PWN, Warszawa.
- ZIOMEK M. J. 1958: Metody graficzne w statystyce. PWG, Warszawa.
- ŻUK B. 1989: Biometria stosowana. PWN, Warszawa.
- YULE G. U., KENDALL M. G. 1966: Wstęp do teorii statystyki. PWN, Warszawa.

INDEKS

- analiza kowariancji 102
 - regresji 100
 - wariancji 74
- bloki 73, 74
 - kompletnie zrandomizowane 73
- błąd drugiego rodzaju 91
 - gruby 25
 - pierwszego rodzaju 91
 - pomiaru 25
 - przypadkowy 25
 - standardowy estymatora 66
 - systematyczny 25
- cecha statystyczna 24, 25
 - ilościowa 25
 - jakościowa 25
 - mierzalna 25
 - niemierzalna 25
- centralne twierdzenie graniczne 50, 66
- centyle 18
- charakterystyki populacji 16
 - próby 16
 - częstość względna 48
- dane ortogonalne 77
 - nieortogonalne 77
- decyle 18
- długość serii 55
- doświadczenie 68
 - jednoczynnikowe 70
 - kontrolne 71
 - losowe 15, 46
 - wieloczynnikowe 70
- dwuwymiarowy rozkład normalny 102
- dyspersja względna klasyfikacji 21
- dystybuanta 61
 - empiryczna 29
 - ogiwalna (ogiwa) 29
- estymacja parametru 92
 - przedziałowa 92
 - punktowa 92
- estymatory 17
- frakcja 19, 29, 96
- frekwencja 26, 29
- funkcja 48
 - gęstości prawdopodobieństwa 60
 - gęstość prawdopodobieństwa 59
- granice przedziału klasowego 27
- hipoteza alternatywna 88
 - nieparametryczna 89
 - parametryczna 89
 - statystyczna 85
 - zerowa 87
- histogram 28, 31, 32
 - liczebności skumulowanych 29
- jednostka eksperymentalna 68
 - standaryzowana 63
 - statystyczna 13
- kategoria modalna 18, 96
- klasyfikacja podwójna 74
 - pojedyncza 83
- korelacja 102
 - cząstkowa (częściowa) 103
 - wielokrotna (wieloraka) 103
- krzywa Gaussa 60
 - liczebności 28, 31, 32
 - skumulowanych 28, 29
 - normalna 60
- kwadrat łaciński 73, 78
 - grecki 82
 - łacińsko grecki 80
- kwartyle 18
- kwintyle 18
- linia regresji 99
 - drugiego rodzaju 99
 - pierwszego rodzaju 99
- losowanie niezależne 15
 - zależne 15
- metoda najmniejszych kwadratów 99
 - reprezentacyjna 15
 - zmiennych niepołączonych 73, 83
 - połączonych 73, 74
- miary korelacji 17,22
 - położenia 17, 18
 - rozproszenia 17, 21
 - statystyczne 17
- moc testu 92
- model losowy 70
 - mieszany 70
 - stały 70
- nadzieja matematyczna 50
- niepowodzenie 53

- obiekt medialny 19
- obiekty kwartylowe 21
- obserwacja 68
- odchylenie standardowe 22
- odzworowanie 48
- oś główna 103
 - zredukowana 103
- parametry 16
- percentyle 18
- populacja 12
 - dwuwymiarowa 14
 - jednowymiarowa 14
 - nieskończona 14
 - skończona 14
 - wielowymiarowa 14
- powtórzenie 71
- poziom istotności 90
 - ufności 93
- prawdopodobieństwo 48, 51, 59
 - , definicja aksjomatyczna 51
 - - geometryczna 51
 - - klasyczna 51
 - - statystyczna 51
 - zdarzenia 50
- prawidłowość statystyczna 48
- prawo wielkich liczb 48
 - - - Bernoulliego 49
- procent 19, 96
- proporcja 19, 96
- próbna 13, 14
 - duża 16
 - jednorodna 94
 - mała 16
 - prosta losowa 15
 - reprezentatywna 14
- próby niepołączone 73
 - połączone 71
- przedział klasowy 27
 - - otwarty 28
 - - ufności 93
- przekształcanie skal 12
- przeźrzenie zdarzeń elementarnych 46
- ranga 11
- rangi związane 11
- realizacje zmiennej losowej 24
- replicacja 71
- regresja 98
 - wielokrotna 101
- reguła trzech odchyłeń standardowych 63
 - - sigm 63
- rozkład antymodalny 35
 - asymetryczny 33
 - dwumianowy 55
 - dwumodalny 35
 - empiryczny 25
 - - cechy ciągłej 28
 - - cechy skokowej 31
 - jednomodalny 35
 - leptokurtyczny 62
 - liczebności 26, 33
 - - skumulowanych 29
 - mezikurtyczny 62
 - normalny 59
 - platykurtyczny 62
 - Poissona 57
 - prawdopodobieństw 54
 - równomierny 35
 - symetryczny 33
 - teoretyczny 26, 53
 - wielomodalny 35
 - w próbie 66
 - zero-jedynkowy 53
 - z próby 66
- ryzyko błędu 91
- skala pomiaru 10
 - ilorazowa 12
 - interwałowa 12
 - nominalna 10
 - porządkowa 10
 - przedziałowa 12
- schematy doświadczalne 73
 - losowania próby 15
- standardowy rozkład normalny 62
- statystyki 16
- stopnie swobody 90
- stosunek korelacyjny 23, 103
- sukces 53
- szereg rozdzielczy 26
 - statystyczny 26
- szeregi czasowe 105
- średnia arytmetyczna 20
 - - ważona 20
 - geometryczna 20
 - harmoniczna 21
- środek przedziału klasowego 28
- test dwustronny 88
 - jednostronny 88
 - nieparametryczny 89
 - parametryczny 89
 - statystyczny 85
- trend 106
- układ kompletnej randomizacji 73, 84

- uporządkowanie niezupełne (słabe) 11
 - zupełne 11
- wahania okresowe (cykliczne) 106
 - sezonowe 106
- wariancja 21
- wartość medialna 20
 - modalna 18
- weryfikacja hipotezy statystycznej 85
- wielobok liczebności 28, 31, 32
 - - - skumulowanych 29, 31, 33
- wskaźnik struktury 96
- współczynnik asymetrii 61
 - determinacji r^2 100
 - konkordancji 23
 - korelacji liniowej r Pearsona 23, 100, 103
 - - - wielokrotnej 23, 103
 - - - cząstkowej 23, 103
 - - - rang (kolejnościowej) 23
 - kurtozy 61
 - siły związku 23
 - zmienności 22
- wykres korelacyjny 102
 - pasowy 36
 - przestrzenny (bryłowy) 36
 - segmentowy (wiedeński) 36
- wykresy kołowe (cyklogramy) 36
 - liniowe 36
 - mapowe (kartogramy) 36
 - słupkowe 36
- zabieg 69
- zdarzenie elementarne 46
 - losowe 46
 - niemożliwe 47
 - pewne 47
 - prawdopodobne 47
 - złożone 47
- zjawiska (procesy) masowe 48
- zmienna losowa 24, 47, 55
 - - - ciągła 24
 - - - skokowa 24
 - - - u 63
 - - - z 63

NOTATKI

NOTATKI

BIBLIOTEKA
INSTYTUTU BOTANIKI
im. W. Szafera
POLSKIEJ
AKADEMII NAUK
UL. LUBICZ 46
31-512 KRAKÓW

SYGNATURA

2900