

Wojciech Rajpold\*

## Statystyka w archeologii, czyli dlaczego nie trzeba bać się liczb<sup>1</sup>

### *Statistics in archaeology. So why shouldn't you be afraid of numbers?*

The great motorway research and construction investments have brought and are still bringing a huge set of new data. In 2019 alone, one million new archaeological artefacts were sourced. Therefore, there is a problem of systematic and comprehensive development of the obtained sources, in which statistics may be helpful. The article introduces selected statistical methods and shows examples of their use. It focuses on their usefulness in archaeological research, and thus it may become a boost for their wider use in the development of archaeological sources.

KEY WORDS: statistical analysis, archaeology, mathematics, metric data

Submission: 23.09.2021; Acceptance: 26.11.2021

Archeologia powszechnie wiązana jest z naukami społecznymi, humanistycznymi. W wielu krajach ściśle łączy się z historią (tradycja środkowoeuropejska) lub antropologią i etnologią (kraje anglosaskie). Jest to w pełni zrozumiałe, gdyż jej celem jest odtwarzanie społeczno-kulturowej przeszłości człowieka. Z uwagi na fakt, iż jej narodziny wiążą się z romantyczną wizją odkrywania przeszłości przez XVIII- i XIX-wiecznych antykwariuszy (Bahn (red.) 2019, 522) oraz żądnych mitycznych skarbów poszukiwaczy przygód (Fagan 2021, 13–14), w dzisiejszej popkulturze obraz archeologa budują takie postaci jak Indiana Jones czy Lara Croft. Bardzo często wizję archeologa-podróżnika i tropiciela przysłowiowego Świętego Graala podtrzymują media, które w programach popularnonaukowych w roli prowadzących obsadzają osoby wpisujące się w ten schemat (Holtorf 2007, 29–33; Maczyńska 2011, 324–325). W takim ujęciu archeolog jawi się, jako człowiek zafascynowany przeszłością, skupiony na poszukiwaniu cennych artefaktów, do swoich badań wykorzystujący spryt oraz mitologiczne opowieści, podania i legendy. Taki wizerunek odbiegał jednak od postaci archeologa już z końca XIX i początków XX wieku. W tym czasie archeolodzy powszechnie korzystali ze zdobyczy nauk ścisłych, a poszukiwacza przygód zastąpił wnikliwy badacz (Piotrowski 2005). Wraz z narodzinami w latach pięćdziesiątych i sześćdziesiątych ubiegłego wieku tzw. nowej archeologii i badaczami w osobach Gordona Childa, Davida Clarke'a, Levisa Biforda (Johnson 2013, 33–53) nastąpiło odejście od pogoni za najcenniejszymi i najpiękniejszymi artefaktami, a w zamian do arsenału badawczego archeologa wprowadzono zdobycze nauk ścisłych. Obecnie można pokusić

się o stwierdzenie, że archeologia jest najbardziej ścisłą spośród wszystkich nauk społecznych. Najdobitniejszym chyba tego przykładem jest datowanie radiowęglowe (uhonorowane nagrodą Nobla w 1960 r.) (Ławecka 2003, 153), a także wiele innych fizycznych metod datowania, m.in. uranowo-torowa czy termoluminescencyjna (Gąssowski 1983, 74–81; Ławecka 2003 160–164). Do badań składu chemicznego artefaktów wykorzystuje się spektroskopię Ramana (Edwards 2009; Gediga et al. 2017, 57–63). Prospekcja przy użyciu metod geomagnetycznych jest niezastąpiona w nieinwazyjnych badaniach stanowisk (Bahn (red.) 2019, 524–527). Z tego względu nie powinno budzić zdziwienia, że w archeologii powinno znaleźć się miejsce również dla matematyki. Zwłaszcza, że pozyskiwane w trakcie wykopalisk artefakty są masowe, policzalne – możemy je i zmierzyć, i zważyć. Dzięki temu są to dane idealnie nadające się do wykorzystania w analizach statystycznych.

W tym miejscu warto poświęcić nieco uwagi ilości, a ściślej – masowości artefaktów archeologicznych. Dla zobrazowania tego problemu przyjrzyjmy się bliżej ciągle trwającym pracom na grodzisku w Chotyńcu, stan. 1, pow. jarosławski. Tylko z jednego obiektu, tzw. *zolnika* (Czopek et al. 2017; Czopek 2020) w roku 2018 pozyskano 39000 gródek polepy o łącznej masie ponad 1200 kg, ponad 8000 fragmentów ceramiki i prawie 5000 kości. Są to ogromne liczby, a mówimy tu o jednym stanowisku i tylko jednym sezonie badań. Nie sposób nie wspomnieć o badaniach autostradowych, które przyniosły szereg nowych materiałów (Czopek, Pelisiak 2014; Czopek 2017). Podczas tych prac przebadano tysiące arów powierzchni kilkudziesięciu stanowisk. Przyjmuje się, że co roku do magazynów muzealnych trafia od 1 do 2 milionów zabytków archeologicznych (<https://naukawpolsce.pl/aktualnosci/news%2C33541%2Cniec-magazy-now-pomoze-rozwiazac-problem-przechowywania-zabytkow>). Stanowi to gigantyczne wyzwanie logistyczne, wymagające bu-

<sup>1</sup> Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC – 2016/23/N/H53/00666.

dowy nowych magazynów (Bugaj 2018). W ten trend wpisuje się składnica materiałów archeologicznych przy Uniwersytecie Rzeszowskim oraz Regionalna Składnica Zabytków Archeologicznych w Zgłobniu, oddział Muzeum Okręgowego w Rzeszowie (por. artykuł w tym tomie). Obok problemu logistycznego, związanego z przechowywaniem ciągle przybywających materiałów, mamy znacznie poważniejsze zagadnienie, związane z ich opracowywaniem. I właśnie tu z pomocą przychodzi statystyka – dziedzina matematyki, która służy do organizowania wielkich zbiorów. Trzeba jednak stwierdzić, że archeolodzy ze statystyki wciąż nie korzystają w stopniu wystarczającym. Zwłaszcza, jeżeli zwrócimy uwagę na wspomnianą wcześniej problematykę odkryć archeologicznych, które stanowią idealną bazę danych dla analiz statystycznych.

Wydana w 1991 roku przez Mike'a Flatchera i Gary'ego Locka pozycja pt. *Digging Numbers: Elementary statistics for archaeologist* (Flatcher, Lock 1991), będąca podręcznikiem statystyki dla archeologów, nie spowodowała gwałtownego wzrostu liczby prac, bazujących na tych analizach. Co istotne, nie jest to jedynie polski problem. Dobrą ilustracją są spostrzeżenia badaczy zainteresowanych tą problematyką na temat konferencji pt. *Proceedings of the 43rd Annual Conference on Computer Applications and Quantitative Methods in Archaeology* (Shaus et al. 2017, 398). Konferencja ta dotyczyła wykorzystywania oprogramowania komputerowego w archeologii, tymczasem ilość wystąpień/artykułów odnoszących się do analiz statystycznych była tu śladowa. Ten pesymistyczny obraz powoli jednak ewoluuje i wypada tu, chociaż wspomnieć o niektórych pracach polskich badaczy, wykorzystujących analizy statystyczne.

Tego rodzaju analiz używa się najczęściej przy próbach budowania chronologii. Dobrymi przykładami są materiały kultury trzcinieckiej z Małopolski i Polski południowo-wschodniej, dla której – w celu uszczegółowienia faz jej rozwoju – wykorzystano matryce korelacji R-Spearmana oraz testy  $\chi^2$  (Górski 2007; Adamik-Proksa 2019). Analiza korespondencji oraz seriacja znalazły zastosowanie w badaniach materiałów z kurhanów kultury komorowskiej z dorzecza górnego Dniestru. Pomogło to w uszeregowaniu danych oraz pokazało

różniczasowość tych założeń (Makarowicz et al. 2016, 156). Do analiz chronologicznych materiałów z II fazy rozwoju tar-nobrzesckiej kultury łużyckiej (Rajpold 2020) i z cmentarzyska kultury łużyckiej w Kietrzu (Mrówka 2011) wykorzystano tablice korespondencji. Oprócz analiz dotyczących chronologii, należy wspomnieć o badaniach Marcina S. Przybyły (2016), który wykorzystując tablice korelacji i oprogramowanie do analiz sieci społecznych, badał relacje pomiędzy ornamentyką naczyń ze środkowego okresu epoki brązu w Kotlinie Karpackiej a tym, z jakich regionów pochodziły. Dla badań przestrzennych osady w Janowicach, pow. tarnowski Marta Korczyńska (2014) wykorzystwała analizy korespondencji. Ta sama badaczka z powodzeniem pokazała, że można wykorzystywać analizę korespondencji i centralistyczną sieć powiązań, aby pokazać zależności pomiędzy badaniami Marka Gedla a eksploracjami badaczy z niemieckich ośrodków (Korczyńska 2016). Dla potrzeb określenia wieku i płci zmarłych z cmentarzyska w Kłyżowie, pow. stalowowolski wykorzystano wykresy dendrytowe oraz korelacje R-Spearmana (Bobak, Trybała-Zawiślak 2016). Dla dociekań dotyczących płci oraz pozycji społecznej zmarłych ze scytyjskich kurhanów Marcin Burghardt wykorzystywał tablice korespondencji oraz wykresy dendrytowe (Burghardt 2016). W nurt rozważań dotyczących wieku i płci wpisuje się także niedawno wydana monografia autora niniejszego tekstu (Rajpold 2021), z której część przykładów omówionych zostanie w dalszej części.

Powyżej zaprezentowany wybór pokazuje bogactwo oraz różnorodność możliwości wykorzystywania tej dziedziny wiedzy w archeologii. W dalszej części pracy skupimy się na prezentacji wieloaspektowości metod wnioskowania statystycznego w badaniach archeologicznych.

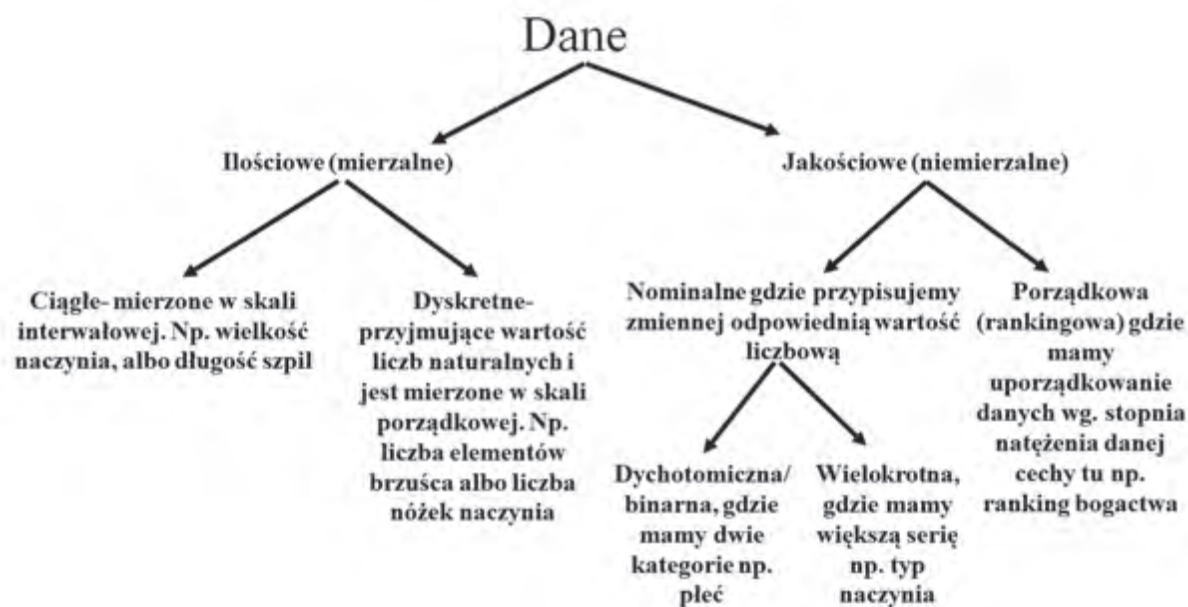
Warto jeszcze zanotować, że do swoich badań autor wykorzystał program Statistica 13.3, który w sposób prosty i intuicyjny daje możliwość wykonywania obliczeń statystycznych na dużym zestawie danych. Ponieważ znaczącą większość wyliczeń wykonuje program, autor postanowił, że przy omawianiu poszczególnych metod skupi się na przedstawieniu ich zastosowań w archeologii, a omawianie ich działania pod kątem matematyki ograniczy do minimum.

## DANE I SKALE POMIAROWE

Absolutnie podstawową kwestią, którą należy poruszyć jest omówienie tzw. cechy statystycznej. Najprostsza definicja mówi, że jest to funkcja, przypisująca badanej bazie danych elementy zbioru wartości (Waliszewski et al. 1998, 26). Można je podzielić (ryc. 1) na dwa typy: ilościowe (mieralne) i jakościowe (niemierzalne) (Stanisz 2006, 22–24). Pierwsze, można mierzyć na skali ciągłej (interwałowej) – są to np. waga lub długość artefaktu (dane ciągłe). Do elementów ilościowych zalicza się również tzw. dane dyskretne, które mierzy się na skali porządkowej – może to być np. liczba odcisków palcowych na naczyniu. Drugi rodzaj dotyczy informacji, których nie jesteśmy w stanie zmierzyć lub zważyć. Wyróżniamy tu dane nominalne, gdzie interesującej nas zmiennej przyporządkujemy odpowiednią wartość liczbową. Mogą to być dane binarne (np. płeć), jak i wielokrotne (np. forma naczynia). Mamy także dane porządkowe, gdzie możemy posegregować artefakty

według natężenia konkretnej cechy. Co istotne, te same dane możemy wyrazić na więcej niż jeden sposób. Tu przykładem może być długość szpil – jest to zmienna mierzalna, którą możemy wyrazić w skali ciągłej, ale możemy również przyjąć skalę jakościową. Wówczas szpilę o długości do 5 cm określimy jako krótką, do 10 cm – średnią, a mającą ponad 10 cm – długą.

W zależności od rodzaju danych, a także przyjętych metod, stosujemy różnego rodzaju skale pomiarowe (Walesiak 1996; Stanisz 2006, 23–24; Bielicka 2011, 21). Najślabszą i dającą najmniej informacji jest skala nominalna. W jej przypadku wszystkie wartości są równe i są to jedynie etykiety identyfikujące te przedmioty; może nią być na przykład forma naczynia (waza, dzban itp.). W skali tej stosuje się liczbę jako nazwę (Bielicka 2011, 21), czyli np. 1 – waza, 2 – dzban, a 3 – amfora. W takim ujęciu cechą jest forma naczynia, a etykietą konkretne formy. Dla tej skali możemy jedynie wyliczyć liczbność



Ryc. 1. Schematyczne przedstawienie rodzajów danych stosowanych w statystyce  
 Fig. 1. Schematic representation of the types of data used in statistics

i procent oraz wskazać najczęściej pojawiającą się wartość, czyli dominantę. Można także wykonać test  $\chi^2$ .

Kolejną, korzystniejszą dla analiz statystycznych, jest tzw. skala porządkowa, gdzie dane układamy w ciągu, przy użyciu konkretnej relacji porządkującej (Stanisz 2006, 23). W skali tej możemy mierzyć dane przymiotnikowe, które można stopniować – np. wielkość. Tu możemy posłużyć się następującym zapisem: 1 – szpila mała, 2 – szpila średnia, 3 – szpila duża. Z pojęciem skali porządkowej łączy się jeszcze tzw. rangi wiązane. Jeżeli w danej próbie mamy kilka tych samych wartości, można im nadać rangę wiązaną, czyli średnią arytmetyczną z ewentualnych rang. Można ją wykorzystywać przy testach  $\chi^2$  i korelacjach rangowych, jak Spearmana czy tau Kendalla.

Najpowszechniejszą jest jednak tzw. skala ciągła (Stanisz 2006, 23), w której dane ułożone są w równych przedziałach, a 0 to wartość umowna (możliwe są wartości poniżej zera). Ważne, by przedziały były sobie równe, czyli odległość między 10 a 20 jest taka sama jak między 30 a 40. Skala ta jest najbardziej uniwersalna i można na niej wykonać większość analiz statystycznych. Co warte odnotowania istnieje pewien szczególny rodzaj skali ciągłej, tzw. skala Likierca (Likierc 1932). Likierc zastosował w niej pięć wypowiedzi, którym przyporządkował wartości liczbowe: -2 – zdecydowanie się nie zgadzam, -1 – raczej się nie zgadzam, 0 – nie mam zdania, 1 – raczej się zgadzam i 2 – zgadzam się. Jak nietrudno się domyślić naj-

częściej jest ona stosowana w naukach socjologicznych oraz psychologii.

Chociaż skala ta jest powszechna, w przypadku archeologii jej rola jest marginalna, gdyż trudno tu wskazać dane o ujemnych wartościach. Można oczywiście próbować je wartościować i na przykład uznać, że pochówek bardzo biedny, nie zawierający żadnego przedmiotu otrzymuje wartość -2, z jednym przedmiotem -1, z trzema przedmiotami nie jest ani biedny ani bogaty i otrzymuje 0, itd. Problem zaczyna się jednak przy 6 i więcej przedmiotach. Można to oczywiście przemodelować, ale znacznie lepiej jest przyjąć kolejny rodzaj skali, a mianowicie ilorazową, w której zero nie stanowi wartości umownej (Bielecka 2011, 22–23). Przykładem takim jest np. waga przedmiotu, gdyż coś nie może ważyć mniej niż 0. Jest to najlepszy rodzaj skali, na której możemy wykonać wszystkie analizy statystyczne. Warto jednak mieć na uwadze, o czym wspomniano wyżej, że różnica między tymi dwoma skalami jest umowna. Dużo czytelniejszym przykładem jest wiek – jeżeli określimy go w latach, mamy skalę ilorazową, jeśli podamy go jako konkretną datę, wówczas przedziałową.

Reasumując, analizowane przez nas informacje możemy zapisywać na wiele różnych sposobów oraz mierzyć na różnych skalach. W zależności od przyjętych metod oraz tego, co chcemy zbadać, dopasowujemy odpowiednią skalę i jej przekształcenia.

Jednym z najbardziej istotnych zagadnień, które bada statystyka, jest to, czy różnice pomiędzy analizowanymi grupami są istotne. Najpowszechniej używany jest test chi<sup>2</sup> opracowany przez Karla Pearsona (1900). Polega on na porównaniu częstości zaobserwowanych z częstościami oczekiwanymi, przy założeniu hipotezy zerowej o braku związku między tymi zmiennymi. Przyjmuje się, że gdy poziom istotności<sup>2</sup> tego testu jest niższy niż 0,05, różnice między badanymi grupami są istotne. Aby zobrazować jego działanie można wyobrazić sobie rzut monetą w trzech seriach po 10 rzutów (tabela 1).

**Tabela 1. Przykład różnic w istotności między seriami rzutów monetą, wykazany przez test chi<sup>2</sup>**

**Table 1. Example of the significance of the difference between series of coin tosses, as demonstrated by the chi<sup>2</sup> test**

Rzut monetą – 1	Reszka	Orzeł	chi <sup>2</sup>	p
Seria 1	5	5	0	1
Seria 2	5	5		
Seria 3	5	5		
Rzut monetą – 2	Reszka	Orzeł	chi <sup>2</sup>	p
Seria 1	5	5	20	0,00005
Seria 2	10	0		
Seria 3	0	10		

W każdej z nich możemy wyrzucić po tyle samo orłów i reszek. W takim przypadku nie ma różnicy między tymi grupami. To też pokazuje test, którego wynik to 0, a istotność wynosi 1. Jednak, jeżeli w pierwszej serii wyrzucimy po tyle samo orłów i reszek, w kolejnej 10 orłów, a w trzeciej 10 reszek to te różnice będą duże. I na to też wskaże test (jego wynik to 20, a p = 0,00005). W naszym przypadku, jako przykład posłużyła zależność obecności metalu w grobach względem wieku zmarłych (tabela 2) oraz tego, czy pochówek był pojedynczy, czy zbiorowy (tabela 3) (dane z cmentarzysk tarnobrzeskiej kultury łużyckiej). W pierwszym przypadku test wykazał, że nie ma różnic ze względu na wiek, a w drugim wskazał, że są różnice ze względu na liczbę zmarłych w grobie. W obydwu egzemplarzach porównaliśmy tylko dwie grupy, ale – jak pokazał przykład z monetą – porównywać można większą grupę. Trzeba jednak mieć na uwadze, że im większy zestaw badanych zespołów, tym różnice będą większe i w takim przypadku szczególne znaczenie ma wynik istotności różnic statystycznych. Niestety metoda ta nie pokazuje, co dokładnie różni analizowane zbiory; wskazuje jedynie, że są one niehomogeniczne. Stanowi to jednak wskazówkę, że należy w tej grupie czegoś szukać.

<sup>2</sup> Dalej będziemy używać skrótu „p”.

**Tabela 2. Tabela liczebności, obrazująca zależności pomiędzy obecnością metalu a wiekiem zmarłego z cmentarzysk tarnobrzeskiej kultury łużyckiej (Rajpold 2021, 444, tabl. 587)**

**Table 2. Table of quantities, showing the relationship between the presence of metal and the age of the deceased from the cemeteries of the Tarnobrzeg Lusatian culture (Rajpold 2021, 444, tabl. 587)**

Obecność metalu	Dzieci	Dorośli	chi <sup>2</sup>	p
Jest	93	234	0,006	0,93
Nie ma	413	1028		

**Tabela 3. Tabela liczebności, obrazująca zależności pomiędzy obecnością metalu a tym, czy pochówek był pojedynczy, czy zbiorowy dla cmentarzysk tarnobrzeskiej kultury łużyckiej (Rajpold 2021, 444, tabl. 589)**

**Table 3. Table of quantities, illustrating the relationship between the presence of metal and whether the burial was single or collective for the cemeteries of the Tarnobrzeg Lusatian culture (Rajpold 2021, 444, tabl. 589)**

Obecność metalu	Jest	Nie ma	chi <sup>2</sup>	p
Pojedyncze	328	1473	22,41	0
Zbiorowe	94	222		

W sytuacji, gdy dane wyrażamy na skalach mierzalnych, do sprawdzenia istotności możemy wykorzystać dwa rodzaje testów. Pierwszy, U Manna-Whitneya (Stanisz 2006, 370–372), stosujemy, gdy mamy dane binarne. Ma on za zadanie sprawdzić hipotezę zerową o tym, że badane grupy pochodzą z populacji o takich samych rozkładach mediany. Jeżeli wynik istotności jest mniejszy niż 0,05 przyjmujemy, że istnieje istotna różnica w rozkładzie mediany między badanymi grupami. Przykładem są badania (tabela 4) dla cmentarzyska w Kłyżowie, stan. 2, dotyczące związku między wielkością popielnicy (suma szerokości brzuśca, wylewu, dna i wysokości naczynia podzielona przez 4) a tym, czy grób znajdował się w części wschodniej lub zachodniej. W części wschodniej naczynia uzyskały 18,43, a w zachodniej 17,84. Tu posłużymy się wspomnianym testem, którego wynik to 0,99, a istotność to 0,32. Tym samym otrzymujemy sugestię, że między tymi dwiema strefami nie ma istotnych różnic, a na całym cmentarzysku naczynia przybierały w miarę zbliżone rozmiary.

Tam, gdzie mamy więcej niż dwie grupy, wykorzystujemy test Kruskala-Wallisa (Stanisz 2006, 372–374). Jego zasada jest podobna jak w przypadku testu U Manna-Whitneya i ma na celu sprawdzenie, czy badane próbki pochodzą z populacji o takim samym rozkładzie mediany. Przykładem jego zastosowania są dane z osady kultury trzcinieckiej w Terliczce,

**Tabela 4. Przykład testu U Manna-Whitneya, dotyczący wielkości popielnic względem części cmentarzyska, z której pochodziły (Kłyżów, stan. 2)**

**Table 4. Example of Mann-Whitney's U test concerning the size of urns in relation to the part of the cemetery they came from (Kłyżów, site 2)**

Kierunek	Średnia	N	Mediana	UM-W	p
Wschód	18,43	75	19,18	0,99	p
Zachód	17,84	49	18,38		
Ogół	18,2	124	18,83		

**Tabela 5. Przykład testu Kruskala-Wallisa dotyczący grubości ścianek naczyń z osady kultury trzcinieckiej w Terliczce stan. 4-5**

**Table 5. An example of the Kruskal-Wallis test concerning the wall thickness of vessels from the settlement of the Trzciniec culture in Terliczka, site 4-5**

Strefa	Średnia	N	Mediana	KW	p
I	8,37	49	7	34,19	0
II	9,15	204	7		
III	9,34	829	8		
IV	8,63	144	7		
V	8,21	75	7		
Ogólne	9,13	1301	7		

pow. rzeszowski, stan. 4-5 (tabela 5), dotyczące związku między grubością naczyń a częścią stanowiska. Joanna Adamik-Proksa (2019, 52-58) wyróżniła na tej osadzie 5 stref, które miały reprezentować jej 5 faz rozwoju. Jak można zauważyć, różnice w średniej grubości naczyń między tymi strefami nie są duże, jednak wynik testu informuje (wynik -34,19 przy  $p = 0$ ), że są istotne statystycznie. Oznacza to, że faktycznie istniały strefy, gdzie wytwarzano grubsze naczynia. Co istotne, test ten daje dodatkowo wyniki odnośnie stopnia podobieństwa i różnic między badanymi strefami - w takim ujęciu widać między innymi, że strefy I, IV i V są identyczne, a strefy II i III wykazują podobieństwa do siebie (tabela 6).

**Tabela 6. Prezentacja różnic między strefami z osady w Terliczce stan. 4-5 pod względem grubości ścianek naczyń na podstawie testu Kruskala-Wallisa**

**Table 6. Presentation of the differences between the zones from the settlement in Terliczka, site 4-5 in terms of vessel wall thickness based on the Kruskal-Wallis test**

	I	II	III	IV	V
I		0,200136	0,029431	1	1
II	0,200136		1	0,204641	0,01054
III	0,029431	1		0,004038	0,000239
IV	1	0,204641	0,004038		1
V	1	0,01054	0,000239	1	

## KORELACJE

Tym, co najmocniej kojarzy się z analizami statystycznymi, są różnego rodzaju korelacje (łac. correlation-, correlatio od „com-” – łącznie i „relations-, relatio” – związek). W najpowszechniejszym rozumieniu służą one do określania istotności wpływu jednej cechy na drugą (Stanisz 2006, 289). Oczywiście, warto mieć na uwadze, że nie każda korelacja oznacza związek. Popularnym przykładem, do jakich absurdów może doprowadzić ślepe czytanie korelacji, jest wykres stworzony przez Tylera Vigena (<http://www.tylervigen.com/spurious-correlations>). Wykazał on istnienie korelacji pomiędzy liczbą utonięć i liczbą filmów z Nicolasem Cage'em. Z całą pewnością są to dwie, zupełnie niepowiązane ze sobą kwestie, ale przykład ten pokazuje, do jakich absurdów można dojść przy nieumiejętnym czytaniu korelacji. Z tego względu, przy wykorzystaniu korelacji (a także innych metod statystycznych) trzeba zachować pewną dozę ostrożności i zdrowego rozsądku.

Analizując liczbę zmiennych, możemy stwierdzić 4 rodzaje zależności, które można badać przy użyciu analiz korelacji (za Stanisz 2006, 289):

1. Zmienna zależna jednowymiarowa oraz jedna zmienna niezależna

2. Zmienna zależna jednowymiarowa oraz wiele zmiennych niezależnych

3. Zmienna zależna wielowymiarowa oraz jedna zmienna niezależna

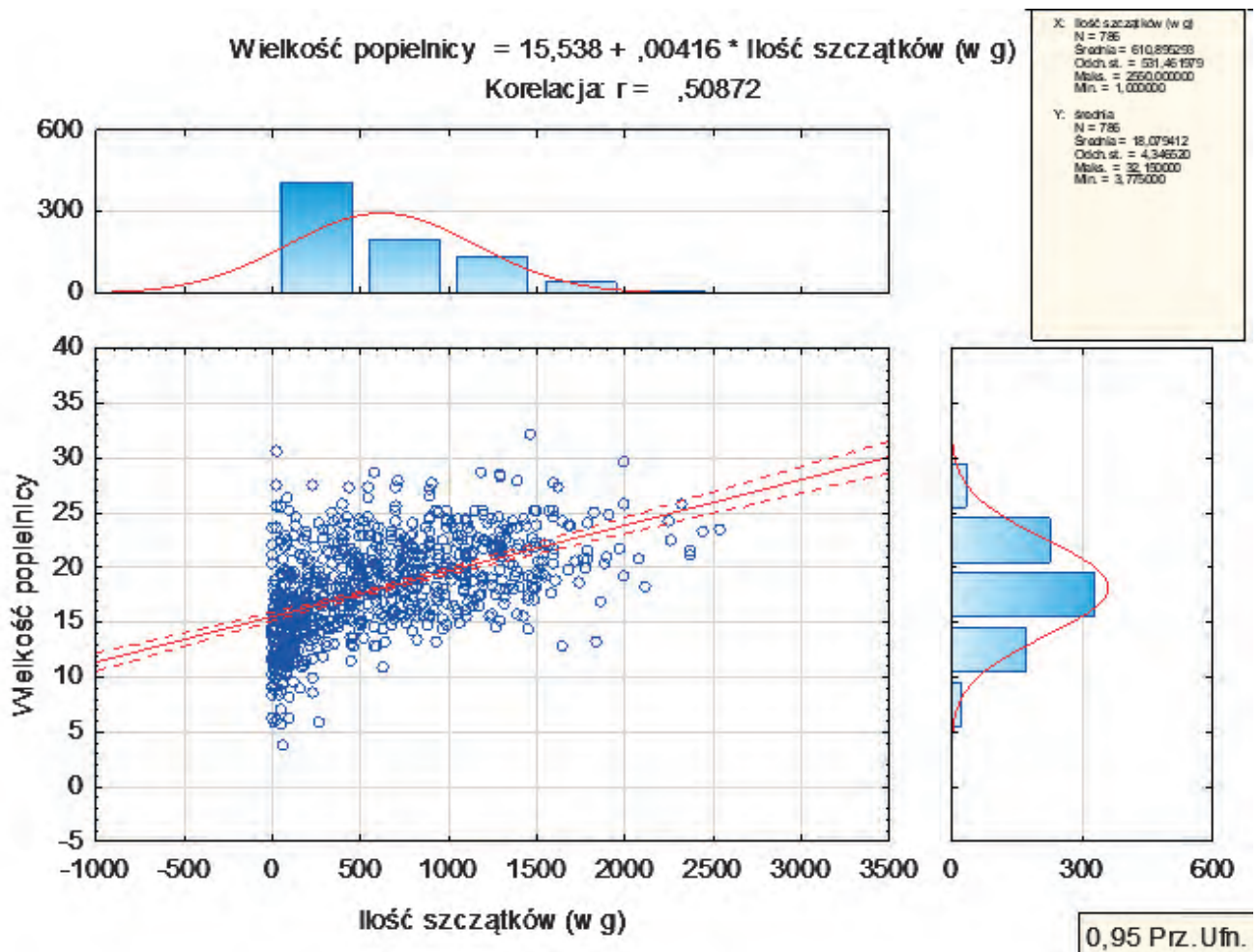
4. Zmienna zależna wielowymiarowa oraz wiele zmiennych niezależnych.

W tym miejscu należy wyjaśnić kilka terminów. Zmienna zależna to współczynnik, który badamy, a niezależna to ta, której związek z interesującą nas kwestią sprawdzamy (Zimbardo, Gerrig 2012, 36). Na przykład, jeżeli weryfikujemy związek między wiekiem zmarłego a wymiarami popielnicy, zmienną zależną jest wiek, a niezależną wymiar urny. W przypadku badań korelacyjnych taki podział nie jest regułą – możemy go odwrócić i badać związek między wymiarami popielnicy a wiekiem i wtedy odwrotne cechy będą zależnymi i niezależnymi. Ważne, by pamiętać, że niezależna jest ta, której wpływ chcemy dopiero wykazać. O wiele bardziej skomplikowana jest kwestia zmiennych jednowymiarowych i wielowymiarowych. W pierwszym przypadku są to cechy, które określamy tylko jednym wymiarem jak wzrost, wiek, waga, mogące przybierać formę ciągłą i skokową (Major 2007, 31-32). Zmienne ciągłe to takie, które mogą przybierać dowolne wartości z pewnego

przedziału liczbowego np. waga lub wzrost. Zmienne skokowe mają za to wartości skończone lub dające się przeliczyć, np. ranga bogactwa pochówku. To ten rodzaj danych najczęściej spotykamy w archeologii. W przypadku wartości wielowymiarowych te również mogą przybierać formę zmiennych ciągłych i skokowych, ale tu następuje połączenie dwóch i większej liczby zmiennych (Czaja, Preweda 2000), co prowadzi do powstania osobnej funkcji określającej np. związek wagi i wielkości metalowego przedmiotu z wiekiem zmarłego. Z tym typem częściej spotykamy się przy o wiele bardziej skomplikowanych analizach, jak np. modelowaniu matematycznym.

Kolejną kwestią związaną z korelacją jest graficzna prezentacja wyników, a mianowicie kształt, jaki przyjmuje krzywa na wykresie. Może być ona liniowa, gdzie dane tworzą linię prostą oraz krzywoliniowa – w formie linii tworzącej łuk (Stanisz 2006, 289–291, rys. 12.1). Jej wartość może być dodatnia, gdy wzrostowi wartości jednej zmiennej towarzyszy wzrost drugiej i ujemna, gdy spadkowi wartości jednej zmiennej towarzyszy spadek wartości drugiej. Wartości korelacji będą się zamykać w przedziale od -1 do 1, gdzie 0 oznacza brak związku między badanymi czynnikami, a 1 i -1 – silny związek.

Najpowszechniejsza jest tzw. korelacja Pearsona, która może być stosowana tylko dla zmiennych mierzalnych, minimum na skali przedziałowej, których związek musi być przynajmniej liniowy (Stanisz 2006, 295–296). Oznacza to, że przy jej pomocy możemy sprawdzać np. związek między wymiarami naczynia a ilością szczątków. Jednak, jeżeli chcemy zbadać relacje między formą popielnicy a jej wymiarami, musimy zastosować inny rodzaj korelacji. Przykładem tego rodzaju korelacji są wyniki związku pomiędzy wielkością urny a ilością szczątków dla materiałów tarnobrzeskiej kultury łużyckiej (ryc. 2). Związek między tymi dwoma kwestiami jest duży, w dodatku istotny statystycznie (p dla tej korelacji wynosiło 0). Co także istotne, korelacja Pearsona daje nam wzór, obrazujący współzależność między badanymi czynnikami. W prezentowanym przykładzie otrzymaliśmy:  $wielkość\ popielnicy = 0,004 * ilość\ szczątków + 15,53$ . Mając na uwadze tę regułę i wiedzę o ilości szczątków, możemy określić potencjalne wymiary popielnicy tam, gdzie uległa ona zniszczeniu. Pamiętać jednak należy o ostrożnym podchodzeniu do wyników tych analiz, które mogą okazać się zwodnicze. Wpływ na wynik może mieć chociażby fakt, iż ilość szczątków mogła ulec rozproszeniu (np. ze względu



Ryc. 2. Krzywa korelacji Pearsona pomiędzy wielkością popielnicy a ilością szczątków kostnych z cmentarzysk tarnobrzeskiej kultury łużyckiej (za Rajpold 2021, 249, ryc. 4.145)

Fig. 2. Pearson's correlation curve between the size of urns and the amount of bone remains from the cemeteries of the Tarnobrzeg Lusatian Culture (after Rajpold 2021, 249, Fig. 4.145)

na destrukcję urny), ale pozwala na przybliżone odtwarzanie tych informacji. Jest to również użyteczne w modelowaniu matematycznym, o czym szerzej w dalszej części.

W sytuacji, gdy przynajmniej jedna zmienna jest mierzona na skali porządkowej, wykorzystuje się korelację rang Spearmana (Stanisz 2006, 337). Przy jej pomocy można sprawdzać dane, w których obydwie zmienne będą wyrażone na skali rangowej. Można także porównywać dane w systemie rangowym z mierzonymi na skali ciągłej, np. ilość szczątków a głębokość pochówku wyrażona w skali rangowej (0–10, 10–20, 30–40 cm). Jako przykład ponownie sięgniemy po materiały z cmentarzysk tarnobrzeskiej kultury łużyckiej i zweryfikujemy związek bogactwa pochówku (wyrażonego w 5 rangach) z wymiarami popielnicy, wyrażonymi w skali ciągłej. W takim przypadku (ryc. 3, tabela 6) okazuje się, że istniała istotna statystycznie korelacja między tymi czynnikami i można wykazać związek większych naczyń z bogatszymi grobami.

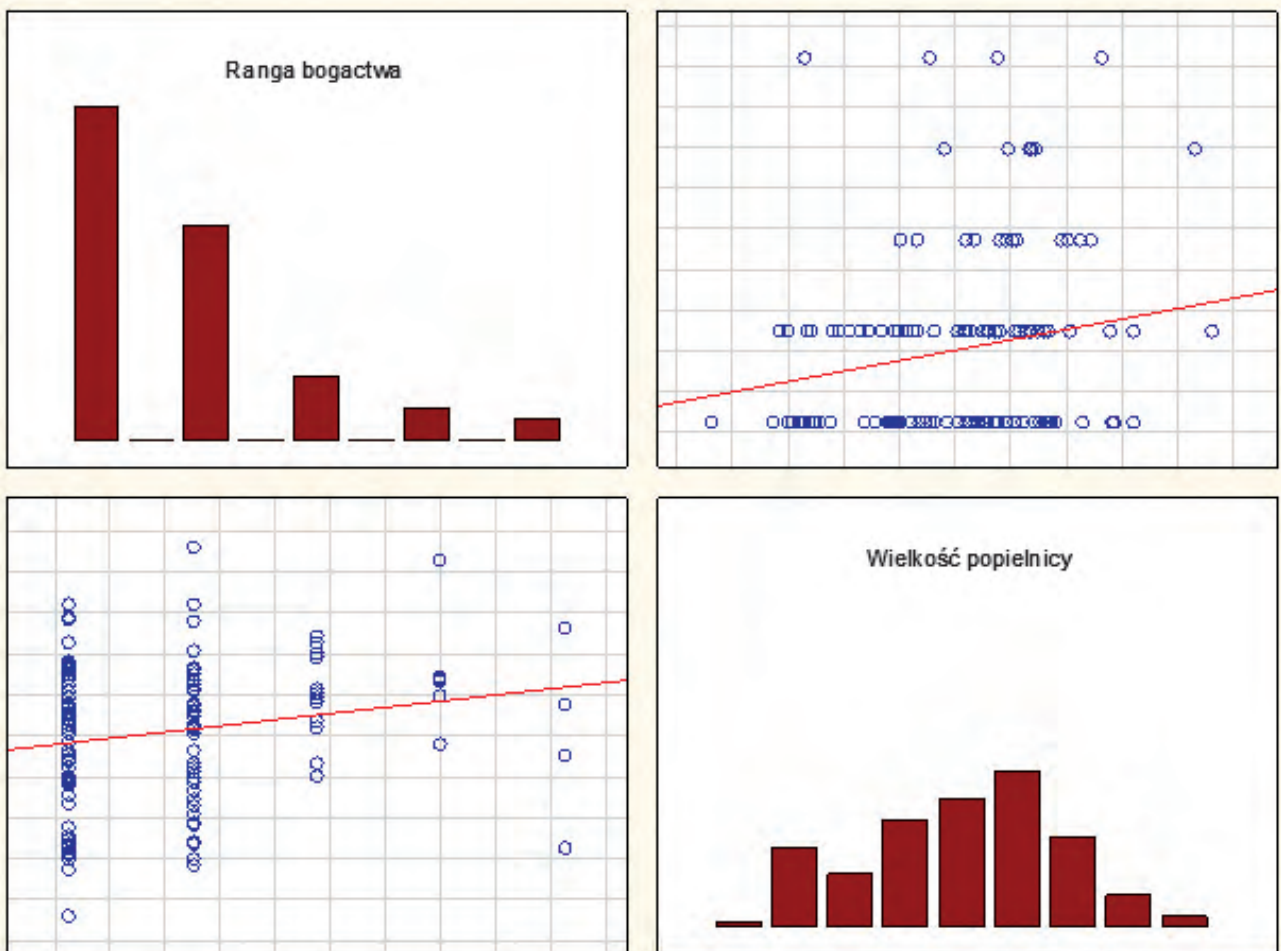
Pozostała jeszcze do omówienia korelacja Tau Kendalla (Stanisz 2006, 337), która pozwala na ocenę podobieństwa zmiennych, mierzonych na skali porządkowej. Można ją wykorzystywać zamiennie względem korelacji rang Spearmana. Dane możemy zapisać poprzez oznaczenia binarne, gdzie 0 przypisujemy brak cechy, a 1 jej występowanie (tabela 8). Moż-

na jednak wprowadzić przymiotniki typu duży, mały, średni. Przykładem (tabela 9) są materiały z osad kultury trzcinieckiej z Polski południowo-wschodniej (za Adamik-Proksa 2019). Korelowano tu współwystępowanie ze sobą (w tych samych obiektach) rozmaitych form naczyń i ich ornamentyki. Na tej podstawie i w oparciu o daty C14 można było zbudować system chronologiczny dla kultury trzcinieckiej, uwzględniający przemiany w ceramice.

**Tabela 7. Korelacja r-Spearmana między wymiarami popielnic a rangą bogactwa pochówków tarnobrzeskiej kultury łużyckiej**

**Tabela 7. R-Spearman's correlation between the dimensions of urns and the rank of the richness of the burials of the Tarnobrzeg Lusatian culture**

	Ranga bogactwa	Wielkość popielnicy	P
Ranga bogactwa	1	0,19	0,03
Wielkość popielnicy	0,19	1	



Ryc. 3. Korelacja r-Spearmana pomiędzy bogactwem pochówku a wielkością popielnic z cmentarzysk tarnobrzeskiej kultury łużyckiej  
Fig. 3. R-Spearman's correlation between the wealth of burials and the size of urns from the cemeteries of the Tarnobrzeg Lusatian culture

**Tabela 8. Przykład zapisu danych do korelacji Tau-Kendala, dotyczący sposobu ornamentyki naczyń kultury trzcinieckiej z osad z Polski południowo-wschodniej, gdzie 1 – cecha występuje, 0 – cechy brak**

**Table 8. An example of data recording for the Tau-Kendall rank correlation coefficient, concerning the ornamentation of the Trzciniec culture vessels from settlements in south-eastern Poland, where 1 – feature is present, 0 – the feature is absent**

Obiekt	E	B	C	D	F	A	J	O	M	N	K	P	Brzegi proste	I	H	G	S	Brzegi zaokrąglone	L	U	T
67	0	0	0	0	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0
155	0	0	0	0	0	0	0	1	0	0	0	0	1	1	1	0	0	0	0	0	0
159	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
215	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
269	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1
283	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	0	0	0	0	0	0

## METODY GRUPOWANIA – TABLICA KORESPONDENCJI, WYKRESY DENDRYTOWE

Statystyka oferuje szereg metod, pozwalających na grupowanie danych. Najpopularniejsze są wykresy dendrytowe oraz tablice korespondencji. Te pierwsze mają na celu stworzenie grup obiektów najbardziej zbliżonych (Stanisz 2007b, 114–122). Najczęściej stosuje się wykresy aglomeracyjne. W schematach tych z początku każdy obiekt to osobne skupisko i następnie łączy się je w większe grupy. Są także diagramy podziałowe, w których notujemy odwrotne podejście, wychodzimy od dużych grup i stopniowo następuje ich rozbijanie na mniejsze. Te jednak stosuje się stosunkowo rzadko, a w wielu pakietach statystycznych są wręcz pomijane.

Pierwszą kwestią, którą należy omówić przy wykresach dendrytowych, jest tzw. miara podobieństwa (Stanisz 2007b, 114–115). Jest to liczba rzeczywista, którą przypisuje się dwóm parom obiektów ( $O_i, O_j$ ). Mieści się ona w przedziale 0–1, a współczynnik podobieństwa obiektu  $O_i$  do  $O_j$  jest taki sam, jak  $O_j$  do  $O_i$ . Obiekty najbardziej podobne będą tworzyć jedną grupę, a najbardziej oddalone osobną. W tym miejscu dochodzi termin miary niepodobieństwa, czyli tzw. odległości (metryka), który bada, jak bardzo dane obiekty się od siebie różnią. To ta miara służy do grupowania obiektów. Jest ona określana na parach obiektów, przyjmuje wartości w zbiorze liczb rzeczywistych i musi spełniać 3 warunki. Pierwszy – zwrotności, który zakłada, że odległość obiektu od samego siebie wynosi 0. Kolejny – symetrii – stwierdza, że odległość obiektu  $O_i$  od  $O_j$  jest równa odległości  $O_j$  od  $O_i$ . Trzecie założenie, nazywane warunkiem trójkąta, oznacza, że odległość obiektu  $O_i$  od  $O_k$ , jest mniejsza lub równa sumie odległości między  $O_i, O_j$  i  $O_j, O_k$ .

Aby zmierzyć odległości między obiektami (Stanisz 2007b, 115–117) wykorzystuje się funkcje odległości. Dysponujemy szeregiem tego typu metryk. Jedną z najpopularniejszych jest funkcja pełnych wiązań, którą wykorzystuje się, gdy dane mają charakter dyskretny. Możemy je oznaczyć w następujący sposób: 1 – cecha pojawia się, 0 – nie występuje. Bywa przydatna na przykład przy ocenianiu podobieństwa między zdobieniem na naczyniach. Jako przykład niech posłużą dwie formy ornamentu – otworki pod krawędzią wylewu i guzki. Naczynie A, na którym były otworki zostałyby zapisane, jako (0.1), a naczynie B, na którym byłyby guzki, jako (1.0). Odległość wy-

siłaby w takim przypadku 1 – dwie pary i obydwie niezgodne. Kolejny typ metryki 1-r Pearsona, jako miarę odległości wykorzystuje korelację Pearsona. Można ją wykorzystać na przykład, gdy chcemy porównać grupę stanowisk pod kątem jednej badanej cechy, np. wymiarów naczyń. Niewątpliwie najczęściej wykorzystuje się wówczas tzw. odległość Euklidesową. Ogólnie jest ona uznawana za dającą najbardziej „naturalne” wyniki. Oblicza się ją poprzez odejmowanie współrzędnych kolejnych punktów i podnoszenie do kwadratu, a następnie wyciąganie z ich sumy pierwiastka. W tej odległości, jeżeli punkt A będzie miał współrzędne 4, 2, 3, a punkt B – 3, 2, 1 to ich odległość będzie wynosiła 2,23, ponieważ

$$\sqrt{(4-3)^2 + (2-2)^2 + (3-1)^2} = \sqrt{5} = 2,23.$$

Jeżeli chcemy przypisać większą wagę obiektom bardziej oddalonym, wykorzystujemy tzw. kwadrat odległości euklidesowej. W tej metryce odległość euklidesową podnosi się do kwadratu. Jeżeli obiekty różnią się tylko w jednym wymiarze, stosuje się odległość Czebyszewa, gdzie obliczamy różnicę między współrzędnymi punktów i bierzemy jako odległość najwyższą wartość. Czyli – przy przykładowych wartościach współrzędnych dla obiektów A i B – odległość wynosiłaby 2. Jeszcze jednym rodzajem odległości jest odległość Manhattan, inaczej miejska. Jest ona sumą różnic mierzonych wzdłuż wymiarów. W przeciwieństwie do odległości Euklidesowej w przypadku tej miary duże różnice między obiektami (przypadki odstające) są stłumione, ponieważ nie podnosi się ich do kwadratu. Dla naszych danych to będzie 3, ponieważ  $1 + 0 + 2 = 3$ .

W zależności od przyjętej metody, odległości między badanymi punktami (grobami, stanowiskami, formami popielnic, itp., itd.) będą większe lub mniejsze, zawsze będą jednak prowadzić do wyróżnienia grup najbardziej do siebie zbliżonych. I tu dochodzi kwestia grupowania i metod oceniania różnic i podobieństw między grupami. Również i w tym zakresie statystyka oferuje szereg metod (Stanisz 2007b, 120–123), pozwalających ocenić odległość między skupiskami. Najbardziej podstawową jest metoda pełnych wiązań (najdalszego sąsiedztwa). W niej odległość między skupiskami wyznaczają punkty najbardziej oddalone. Dobre efekty daje, gdy tworzą one naturalnie

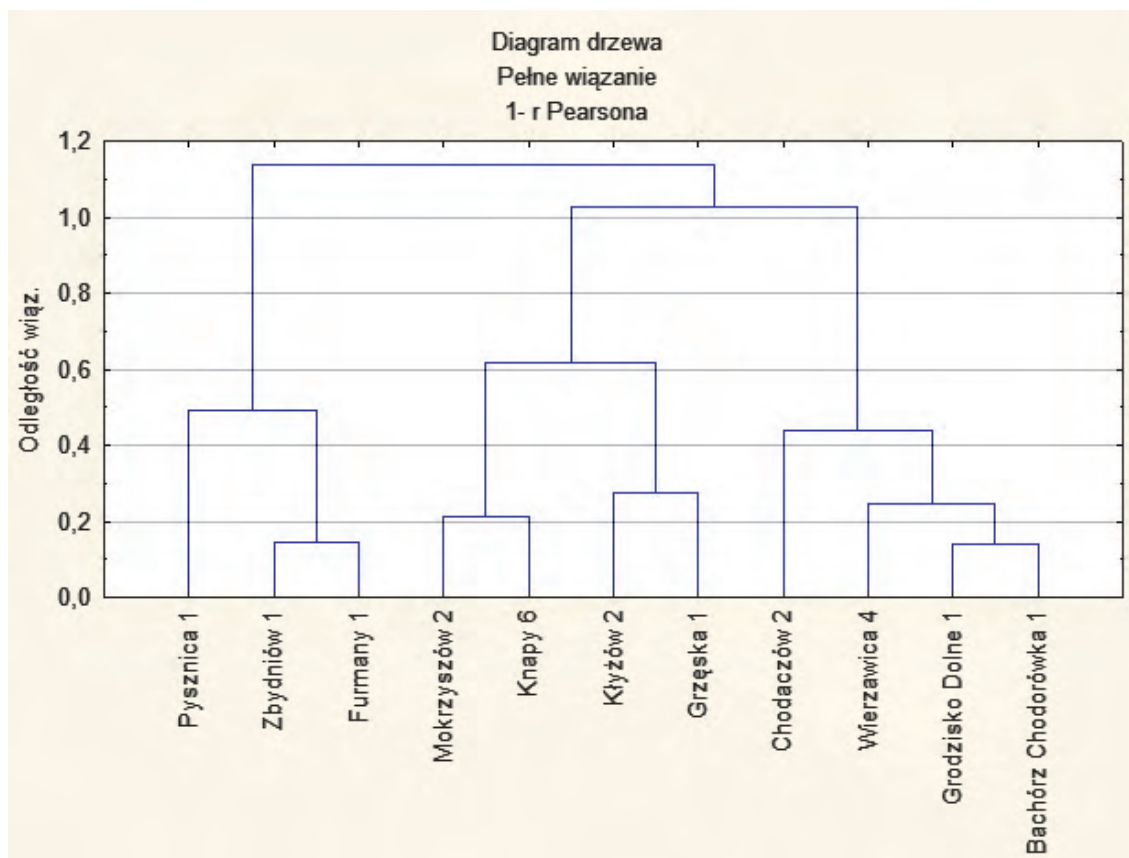


Tabela 9. Wynik korelacji Tau-Kendalla w wybranych zespołach osadowych i grodowych (za Adamik-Proksa 2019, 77, tab. 2.6)  
 Table 9. The result of the Kendall's Tau correlation in selected settlement and stronghold complexes (after Adamik-Proksa 2019, 77, Tab. 2.6)

	A	B	D	E	F	C	G	M	J	K	I	N	H	L
A	1	0,53	0,18	0,24	0,04	0,12	0,02	-0,19	-0,26	-0,31	-0,29	-0,4	-0,21	-0,19
B	0,53	1	0,44	0,62	-0,02	-0,05	0,11	-0,28	-0,39	-0,46	-0,43	-0,48	-0,28	-0,28
D	0,18	0,44	1	0,27	0,32	0,23	0,28	-0,42	-0,15	-0,42	-0,36	-0,65	-0,05	-0,24
E	0,24	0,62	0,27	1	0,08	-0,03	0,07	-0,17	-0,24	-0,28	-0,26	-0,22	-0,17	-0,17
F	0,04	-0,02	0,32	0,08	1	0,12	0,71	-0,19	-0,26	-0,31	-0,29	-0,26	-0,06	-0,19
C	0,12	-0,05	0,23	-0,03	0,12	1	0,13	-0,16	-0,03	-0,26	-0,24	-0,34	-0,3	-0,16
G	0,02	0,11	0,28	0,07	0,71	0,13	1	-0,27	-0,37	-0,3	-0,4	-0,44	-0,12	-0,27
M	-0,19	-0,28	-0,42	-0,17	-0,19	-0,16	-0,27	1	0,27	0,2	0,66	0,29	-0,04	-0,13
J	-0,26	-0,39	-0,15	-0,24	-0,26	-0,03	-0,37	0,27	1	0,36	0,41	0,22	-0,02	-0,17
K	-0,31	-0,46	-0,42	-0,28	-0,31	-0,26	-0,3	0,2	0,36	1	0,46	0,5	0,03	0
I	-0,29	-0,43	-0,36	-0,26	-0,29	-0,24	-0,4	0,66	0,41	0,46	1	0,43	0,08	-0,19
N	-0,4	-0,48	-0,65	-0,22	-0,26	-0,34	-0,44	0,29	0,22	0,5	0,43	1	0,25	0,47
H	-0,21	-0,28	-0,05	-0,17	-0,06	-0,3	-0,12	-0,04	-0,02	0,03	0,08	0,25	1	0,34
L	-0,19	-0,28	-0,24	-0,17	-0,19	-0,16	-0,27	-0,13	-0,17	0	-0,19	0,47	0,34	1

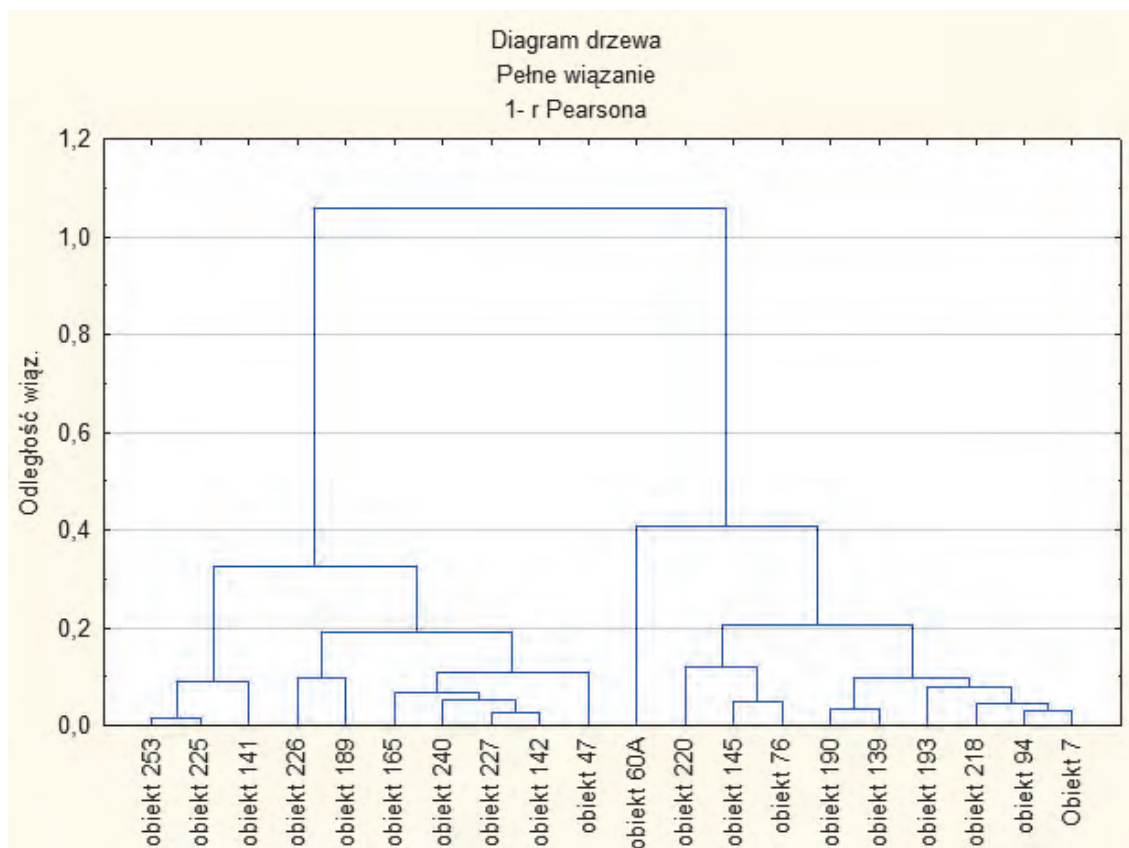
oddzielone zwarte zgrupowania. Weźmy na przykład różnice między stanowiskami z różnych okresów chronologicznych, gdzie każde skupisko będzie wyznaczać różne fazy rozwoju. Podobna jest tzw. metoda ważonych środków ciężkości (mediany). Ta metryka dodatkowo uwzględnia liczbę obiektów w danym skupisku. Sprawdza się więc lepiej, jeżeli pomiędzy badanymi grupami notujemy duże różnice w liczbie obiektów. W metodzie środków ciężkości, odległość między skupiskami wyznacza różnica między ich średnimi punktami ciężkości. Jeżeli badane obiekty tworzą skupiska formujące się w osobne „kępki” lub mające charakter „łańcucha”, wykorzystuje się metodę średnich połączeń. W metodzie tej odległość między dwoma skupieniami oblicza się, jako średnią odległość między wszystkimi parami obiektów, należących do dwóch różnych skupień. Ostatnia – metoda Warda – jest najbardziej skomplikowana i najmocniej różni się od pozostałych. Wykorzystuje się w niej analizę wariacji. Na każdym etapie sprawdza się odległości między wszystkimi parami skupisk i wybiera się te, które w rezultacie łączenia dają najmniejsze różnice. Jest to bardzo efektywna metoda, jednak prowadzi do tworzenia skupień o małej wielkości. W zależności od przyjętej metody wykres może się nieco różnić, co jednak naprawdę istotne, każda z tych metod pozwala na uporządkowanie badanych przez nas obiektów i daje szansę na wyróżnienie skupisk, które można ze sobą porównywać i próbować wychwytywać przyczyny, które wpływały na ich zróżnicowanie.

Pierwszy przykład ich zastosowania (ryc. 4) dotyczy związku między formą urny a cmentarzyskiem, z którego pochodziła. Pogrupowano tu cmentarzyska tarnobrzeskiej kultury łużyckiej w 3 zbiory (Rajpold 2021, 275). W podziale tym Zbydniów, stan. 1 i Pysznicza, stan. 1, pow. stalowowolski oraz Furmany stan. 1, pow. tarnobrzeski ze względu na stylistykę nadszańską zostały przyporządkowane do jednej grupy, która odpowiadała II fazie rozwoju wspomnianej jednostki kulturowej. Co też istotne w takim podziale – Pysznicza, jako najmłodsza z tych trzech (Czopek 2001, 167–186) – odbiega od pozostałych. Druga grupa to cmentarzyska w Bachorzu-Chodorówce, stan. 1, pow. rzeszowski, Chodaczowie, stan. 2, Wierzawicach, stan. 4 i Grodzisku Dolnym, stan. 1, pow. leżajski, połączone za sprawą garnków tulipanowatych i reprezentujące najwcześniejszą fazę. Tu warto zwrócić uwagę na Chodaczów. Na cmentarzysku tym stwierdzono również pewne związki z III fazą, co potwierdza obserwacje Sylwestra Czopka, który sugerował rozwój tej nekropolii aż po schyłek rozwoju tarnobrzeskiej kultury łużyckiej (Czopek 1996, 67). Bez wątpliwości najciekawsze są wyniki dla ostatniej grupy, łączącej cmentarzyska z III fazy rozwoju tej kultury. Mamy tu dwie podgrupy: Knapy, stan. 6, pow. tarnobrzeski i Mokrzyszów, stan. 2, pow. tarnobrzeski oraz Kłyżów, stan. 2, pow. stalowowolski z Grzęską, stan. 1, pow. przeworski. W przypadku tych pierwszych podobieństwo może być efektem ich bliskości. Już jednak w przypadku Kłyżowa i Grzęski o podobnych relacjach nie można mówić. Niemniej, być



Ryc. 4. Przykład wykresu dendrytowego opracowany dla cmentarzysk tarnobrzeskiej kultury łużyckiej porównujący je pod względem formy popielnicy (za Rajpold 2021, 264 ryc. 4.157)

Fig. 4. An example of a dendrite diagram developed for the cemeteries of the Tarnobrzeg Lusatian culture comparing them in terms of the urn forms (from Rajpold 2021, 264, Fig. 4.157)



Ryc. 5. Przykład wykresu dendrytowego opracowany dla materiałów ceramicznych (ich technologii) z obiektów tarnobrzeskiej kultury łużyckiej z osady w Białobrzegach stan. 8 (dane do utworzenia wykresu zawdzięczam mgr. Joannie Ligodzie z Muzeum Okręgowego w Rzeszowie)

Fig. 5. An example of a dendrite diagram developed for ceramic materials (their technology) from features of the Tarnobrzeg Lusatian culture uncovered at the settlement in Białobrzegi, site 8 (the data to create the chart have been collected thanks to Joanna Ligoda from the Regional Museum in Rzeszów)

może dodanie kolejnych cmentarzysk, pozwoliłoby na jakieś studia dotyczące różnic o charakterze regionalnym. Tu warto podkreślić, że za wydzieleniem się Knap i Mokrzeszowa stał wysoki udział garnków jajowatych.

Drugi przykład dotyczył technologii ceramiki pochodzącej z osady tarnobrzeskiej kultury łużyckiej w Białobrzegach<sup>3</sup> (ryc. 5). Można tu wyróżnić dwie grupy i co istotne, po zmapowaniu tych materiałów, okazuje się, że jedna to ary w centrum osady, a druga – ary na jej obrzeżach.

Szczególnie przydatna jest tablica korespondencji (Stanisz 2007b, 307–355). Często stosuje się ją w badaniach dotyczących prób łączenia zmian stylistyki naczyń z chronologią. Przyjmuje się, że jeżeli dane przybierają kształt paraboli, można mówić o ewolucyjnych zmianach. Sposób powstawania tego rodzaju wykresów znacząco wykracza poza ramy tego tekstu, ale warto odnotować, iż bierze się tu pod uwagę związek między wierszami i kolumnami w tabelach wielodzielczych. Oznacza to, że możemy na przykład badać współzależności między grobami z danego cmentarzyska a elementami popielnic. Szczególnie istotną cechą

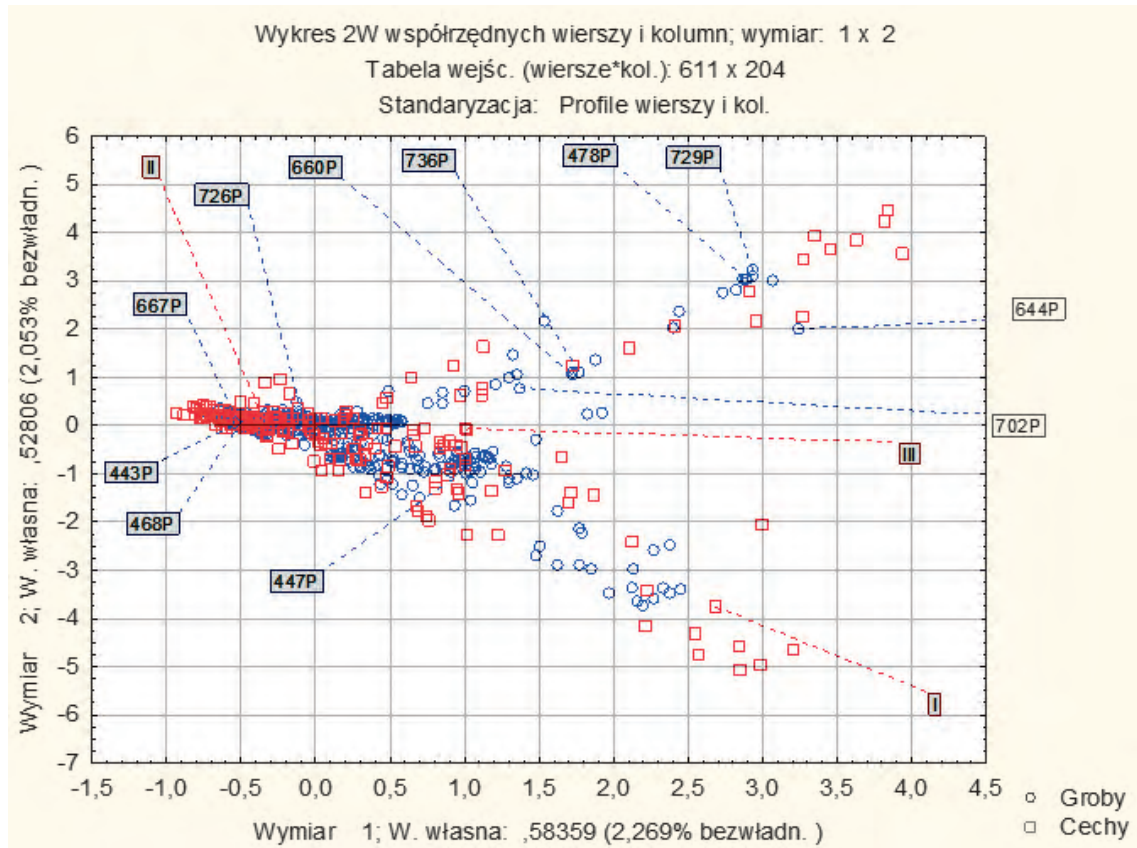
związaną z tablicami korespondencji, którą warto poruszyć, jest tzw. miara bezwładności. Jest ona związana z metryką  $\chi^2$  i określa miarę rozproszenia profili wokół odpowiednich przeciętnych profili. Informuje nas przede wszystkim o różnicach między profilami a profilem przeciętnym. Jeżeli jest niewielka, to oznacza niewielkie rozproszenie wokół profilu przecięcia, jeżeli wysoka – to duże. Innymi słowy – im mniejsza, tym lepsze odwzorowanie rzeczywistości. Służy także przy obliczaniu, jak wiele informacji tracimy, podczas tworzenia wykresu, a także dodawaniu do niego kolejnych wymiarów. Wymiary te pozwalają, co prawda stworzyć ciekawsze skupienie, jednak zawsze wiąże się to z utratą części danych, co obniża ich wartość.

Jeden z przykładów zastosowania tej metody (ryc. 6) dotyczył zmian w stylistyce naczyń na cmentarzyskach tarnobrzeskiej kultury łużyckiej (Rajpold 2020). Wykorzystano tu groby, które zawierały informacje o datach C14 oraz fazach chronologicznych. W układzie tym dane przyjęły kształt paraboli, a daty ułożyły się w sekwencji od najstarszych do najmłodszych. Tym samym otrzymaliśmy potwierdzenie o słuszności wykorzystania tej metody przy badaniach dotyczących chronologii i zmian w stylistyce naczyń. Warto jeszcze zanotować, że w wykresie tym widać znaczne zagęszczenie wyników w centrum paraboli; jej skrzydła są słabiej czytelne. Jest to efekt wykorzy-

<sup>3</sup> Dane pozyskane od mgr. Joanny Ligody z Muzeum Okręgowego w Rzeszowie.

stania materiałów, które reprezentują środkową fazę rozwoju tarnobrzeskiej kultury łużyckiej. Pozostałe fazy były słabo reprezentowane na badanych nekropolach, niemniej jednak

wykres zdołał je wychwycić. Wykorzystanie wyłącznie cmentarzysk z II fazy rozwojowej wynikało z prób wydzielenia podfaz w rozwoju stylistyki nadszańskiej (Rajpold 2020).



Ryc. 6. Przykład tablicy korespondencji dla materiałów z cmentarzysk (w Pysznicy, Zbydniowie i Furmanach) pomiędzy sposobem ornamentowania naczyń oraz ich formą a grobem, z którego pochodziły. Oznaczono także groby zawierające daty C14 (Rajpold 2020, 224 ryc. 7)

Fig. 6. An example of a correspondence board for materials from cemeteries (in Pysznica, Zbydniów and Furmanach) between the way of ornamenting the vessels and their form and the grave they came from. The graves containing the dates C14 were also marked (Rajpold 2020, 224 fig. 7)

## WYKRES SZANS ODDS RATIO ORAZ PUNKT ODCIĘCIA KRZYWEJ ROC

Bardzo przydatną metodą statystyczną pomocną przy tworzeniu modeli decyzyjnych jest punkt odcięcia krzywej ROC. Analiza ta ma na celu wyznaczenie optymalnego punktu odcięcia i oceny, jakości tego klasyfikatora (Harańczyk 2010). Najprościej ujmując, wskazuje punkt, powyżej którego szanse na interesujące nas zjawisko będą największe. W kwestii metodyki opiera się ona na tzw. błędnych decyzjach. Każdy badany punkt to pewna liczba wskazań, które nie spełniają naszych założeń. Celem prezentowanej metody jest wybranie punktu, w którym będzie najmniej błędnych wskazań. Aby to zbadać, powstaje tzw. macierz klasyfikacji (tabela 8), w której TP to prawidłowe wskazania wyróżnionej klasy, TN – prawidłowe niewskazania wyróżnionej klasy, FP – niepoprawne wskazanie wyróżnionej klasy i FN – niewskazanie wyróżnionej klasy, kiedy powinno się ją wskazać (Harańczyk 2010, 81). Tabela taka powstaje dla każdego punktu i służy wyszukaniu tego, który będzie wskazywał najmniej FP i FN, a najwięcej TN i TP. Do tego, aby okre-

ślić ten punkt, stosuje się pewne reguły decyzyjne, najczęściej miary czułości i specyficzności (Harańczyk 2010, 82). Pierwsza to iloczyn TP przez sumę TP i FN, a druga iloczyn TN przez sumę TP i FP. W ten sposób otrzymujemy dwie współrzędne, które możemy oznaczyć na osi X i Y, gdzie oś X to specyficzność, a Y czułość. Punkty łączymy, a ten, który wypadnie najbardziej współrzędnej 0,1, będzie poszukiwanym, ponieważ będzie miał czułość równą 1 (wszystkie obiekty wybranej klasy wykryto) i swoistości równej 1 (nie wskazano błędnie żadnego obiektu za obiekt wyróżnionej klasy) (Harańczyk 2010, 82). W tym miejscu należy jeszcze wspomnieć o tzw. polu pod wykresem krawędzi ROC, czyli tzw. AUC. Przyjmuje ono wartość w przedziale od 0–1 i im jest ona wyższa, tym model lepiej odwzorowuje rzeczywistość. Jest to podstawowy sposób oceny wartości tej statystyki.

Nasz przykład (tabela 11, ryc. 7) dotyczy danych o wymiarach popielnic z cmentarzyska w Kłyżowie. Jak widzimy, opty-

malnym punktem odcięcia jest 18,43 – powyżej tego punktu szanse na obecność szczątków dziecka w grobie są minimalne. Ponadto, pole pod krzywą AUC pokazuje, że jest to bardzo mocne statystycznie zestawienie. Tu warto zwrócić uwagę na sposób zapisu danych (tabela 12), gdzie poprzez 1 i 2 kodujemy dorosłych i dzieci, a obok przypadku zapisujemy wymiary popielnicy.

Punkt odcięcia – oprócz wskazania, gdzie przebiega granica, powyżej której szansa na dane zjawisko jest minimalna – znajduje także zastosowanie w tzw. ilorazie szans (tzw. Odds Ratio). Iloraz ten daje możliwość sprawdzenia szansy zaistnienia danej sytuacji w grupie A względem B (Stanisz 2007a, 222–223). Wykorzystując wspomniany wcześniej punkt odcięcia krzywej ROC, można określić, jak duże są szanse na to, że w grobie z popielnicą o wymiarach powyżej tego punktu, znajdować się będą szczątki dziecka czy osoby dorosłej. Zapis danych do tego wykresu jest bardzo podobny do tego z ROC, z tym że tu zamieniamy wymiary popielnicy na liczbę: 1 to dane poniżej interesującego nas punktu; 2 – powyżej. Jak widać na przykładzie dotyczącym materiałów z Kłyżowa (tabela 13) szanse na to, że na tym cmentarzysku pochówek dziecka znajduje się w naczyniu mającym więcej niż 18,43 punktu, są prawie 20-krotnie mniejsze, niż to, iż w takim naczyniu znajdzie się osoba dorosła. Oczywiście, przy wykresie szans Odds Ratio ten punkt jest dowolny i można ustawić wskaźnik na mniejsze lub większe naczynie. Można także obliczyć szanse na pochówek dorosłego względem dziecka w naczyniu mającym mniej niż 12 punktów. Wykres ten nie ogranicza nas jedynie do danych mierzonych na skali ciągłej, zwłaszcza, że i tak następuje prze-

kształcenie i wprowadzenie dwóch zakresów. Nie musimy się także ograniczać do analizowania dwóch podziałów – w wykresie tym można wprowadzić, na przykład, szanse na pojawienie się szczątków kobiety, mężczyzny i dziecka w naczyniu powyżej określonego wymiaru. Istotne jest jednak, że możemy mieć równoległe jedynie dwie zmienne główne. Tym samym możemy na przykład badać obecność lub brak metalu, obecność lub brak zdobienia, relacje dorosły–dziecko, itp. względem dowolnej ilości innych zmiennych. W takim ujęciu przykładem jest wykres szans Odds Ratio dla obecności popielnicy mającej więcej niż 17,7 punktu na jednym z 11 cmentarzysk tarnobrzeskiej kultury łużyckiej (ryc. 8).

Warto podkreślić, że metoda ta bierze poprawkę na wielkość badanej próbki. Jest to ważne, jeżeli badane grupy są nierówne liczebnie. Ten problem obrazuje przykład zastosowania tego wykresu dla danych dotyczących badań polepy z zolnika w Chotyńcu (ryc. 9, tabela 13). Wokół zolnika znaleziono ogromną ilość polepy, która tworzyła kolistą strukturę (ryc. 10). Do analizy wykorzystano fragment z miejsca, gdzie polepa grupowała się w największej ilości, tworząc tzw. płaszczatkę, w której można wyróżnić 2 warstwy. Interesowało nas, o ile większe są szanse na pojawienie się polepy noszącej ślady konstrukcji względem zwykłej polepy. W warstwie II mieliśmy o 200 polep konstrukcyjnych więcej, mogłoby się wydawać więc, że to w tej warstwie te szanse powinny być większe. Wykres szans Odds Ratio pokazuje jednak, że jest odwrotnie – to w warstwie I szanse są większe. Jest to wynik niemal dwukrotnie większej liczby niekonstrukcyjnej polepy w warstwie II względem I.

**Tabela 10. Tabela 10. Macierz klasyfikacji – stan faktyczny i wskazanie modelu (Harańczyk 2010, tab. 1)**

**Table 10. Classification matrix – the actual state and indication of the model (Harańczyk 2010, Tab. 1)**

	Zaobserwowano stan wyróżniony	Nie zaobserwowano stan wyróżniony
Przewidywano stan wyróżniony	TP <sup>a</sup>	FP <sup>b</sup>
Nie przewidywano stanu wyróżnionego	FN <sup>c</sup>	TN <sup>d</sup>

a – prawidłowe wskazania wyróżnionej klasy; b – niepoprawne wskazanie wyróżnionej klasy; c – niewskazanie wyróżnionej klasy, kiedy powinno się ją wskazać; d – prawidłowe niewskazania.

a – Correct indications of the distinguished class; b – Incorrect indication of the highlighted class; c – Failure to indicate the distinguished class when it should be indicated; d – Correct non-indications.

**Tabela 11. Przykład danych dotyczących relacji między wiekiem a wymiarami popielnicy do krzywej ROC, opracowany na podstawie materiałów z cmentarzyska w Kłyżowie (za Rajpold 2021, 61, tabl. 4.13)**

**Table 11. An example of data presenting the relationship between the age and dimensions of the urn to the ROC curve, prepared on the basis of materials from the cemetery in Kłyżów (after Rajpold 2021, 61, tabl. 4.13)**

Wiek dorośli dzieci						
Wielkość popielnicy						
	AUC	SE	AUC Dolny 95%	AUC Górny 95%	z	p
Dorośli/dzieci	0,86	0,034	0,793	0,927	10,51	0

Tabela 12. Przykładowy sposób zapisu danych w celu uzyskania krzywej ROC (gdzie 1 to dzieci, a 2 – dorośli) (Rajpold 2021, 40, tabl. 3.2)

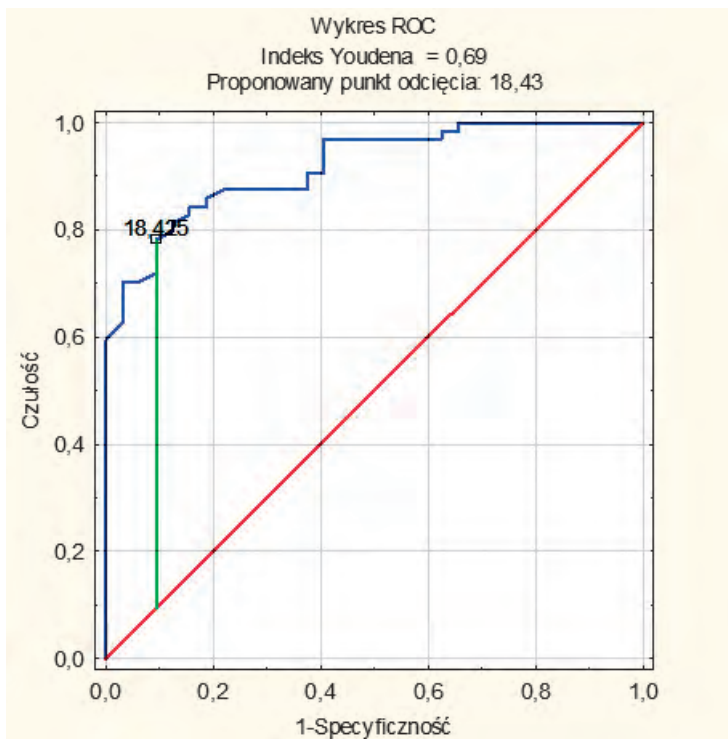
Table 12. Example of data records in order to obtain the ROC curve (where 1 is children, and 2 – adults) (Rajpold 2021, 40, tabl. 3.2)

Wiek	Wymiary naczyń	Wiek	Wymiary naczyń	Wiek	Wymiary naczyń
1	11	1	14	2	16
1	12	1	15	2	23
1	13	1	15	2	24
1	13	1	16	2	24
1	13	1	17	2	25
1	13	1	21	2	25
1	14	2	21	2	26
1	14	2	21		
1	14	2	22		

Tabela 13. Przykład tablicy z danymi do szans Odds Ratio. Prezentuje ona szanse na pojawianie się pochówku dziecka względem osoby dorosłej w popielnicy mającej więcej niż 18,43 punktu (dane: cmentarzysko w Kłyżowie – Rajpold 2021, 61, tabl. 4.14)

Table 13. Example of a table showing data with Odds Ratio. It presents the chances of a child's burial occurrence in relation to an adult in the urn with more than 18.43 points (data: cemetery in Kłyżów – Rajpold 2021, 61, tabl. 4.14)

Szanse Odds Ratio			
Średnia naczyń	Or	CI	p
Dorośli	4,4	(2,48–7,8)	0
Dzieci	0,23	(0,13–0,4)	0



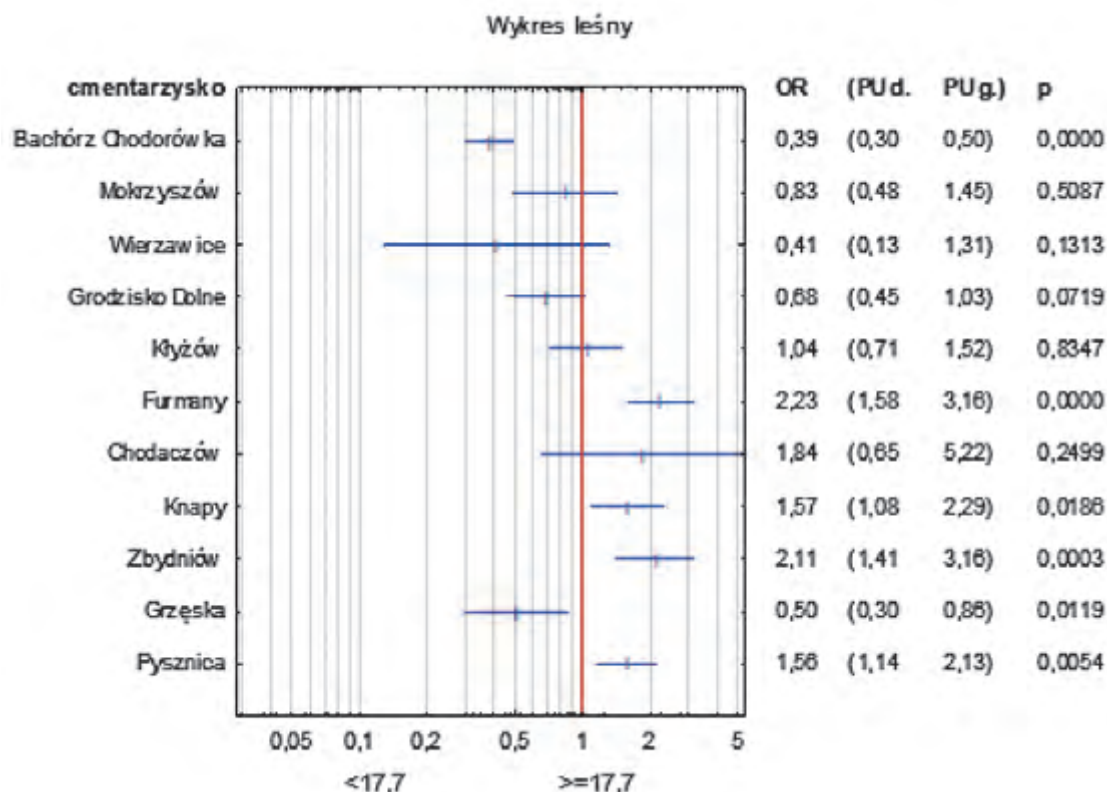
Ryc. 7. Przykład wykresu odcięcia punktu krzywej ROC opracowany na podstawie materiałów (dotyczących relacji między wiekiem a wymiarami popielnicy) z cmentarzyska w Kłyżowie

Fig. 7. An example of a chart of cut-off point for the ROC curve point developed on the basis of materials (concerning the relationship between age and dimensions of the urn) from the cemetery in Kłyżów

## MODELOWANIE

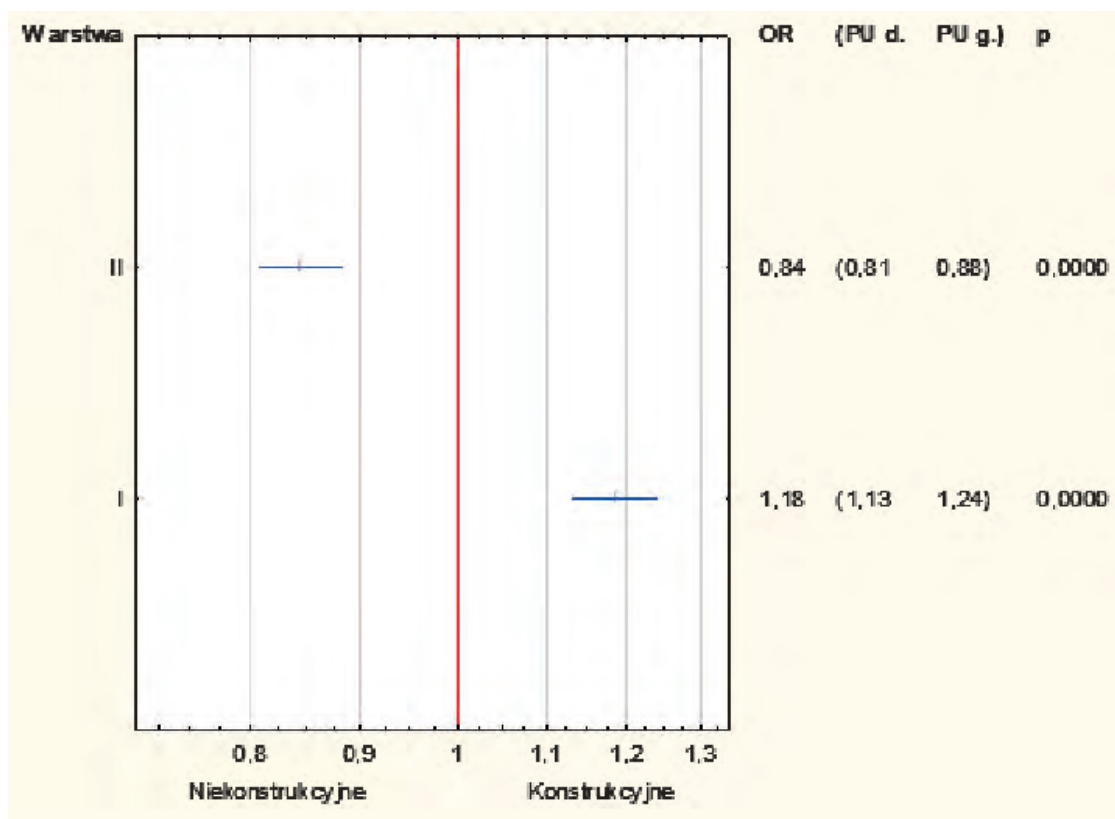
Największą mocą statystyki jest możliwość tworzenia modeli. Na podstawie wyników badań pozwalają one stworzyć formułę, opisującą analizowaną przez nas cechę. Na tej podstawie można przewidywać przyszłe wyniki. Znamy wiele typów modeli – np. logistyczne, liniowe i drzewa decyzyjne. Są one szeroko wykorzystywane, między innymi, w medycynie (Schmitz 1986; Spitznagel 2008; Hilbe, Greene 2008; Petrie, Sabin 2013: 88–104), socjologii (Pampel 2000) czy ekonomii (Cramer 2001). Co ważne, już od lat 80 są one stosowane także

w pracach, dotyczących archeologii osadnictwa. W badaniach tych mają za zadanie zapobiegać niszczeniu potencjalnych stanowisk archeologicznych oraz wskazywać optymalne miejsca do prowadzenia badań wykopaliskowych (Warren, Asch 2000; Wheatley, Gillings, 2002; Kvamme, Kenneth 2006; Gilda 2016). W ostatnim czasie autor niniejszego artykułu przedstawił także modele matematyczne dla wieku oraz płci zmarłych z grobów tarnobrzeskiej kultury łużyckiej (Rajpold 2021, 485–516). I to właśnie te dane posłużą, jako przykład zastosowania mo-



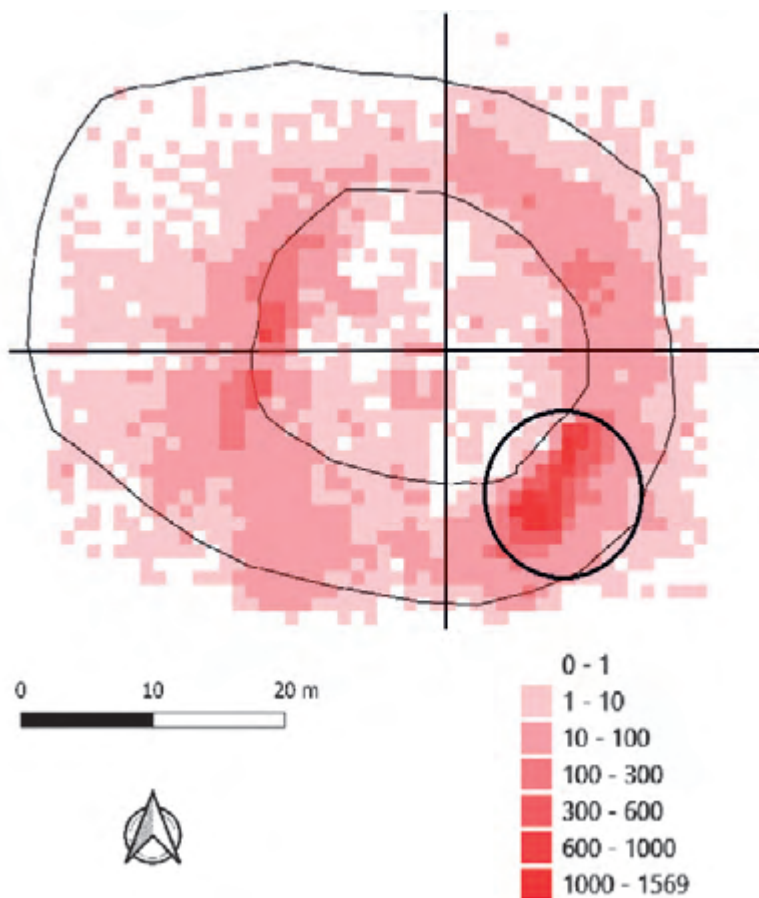
Ryc. 8. Przykładowy wykres szans Odds Ratio dla obecności popielnicy mającej więcej niż 17,7 punktów na jednym z 11 cmentarzyk tarnobrzeskiej kultury lużyckiej (za Rajpold 2021, 248, tabl. 4.143)

Fig. 8. An example of Odds Ratio for the presence of an urn with more than 17.7 points in one of the 11 cemeteries of the Tarnobrzeg Lusatian culture (after Rajpold 2021, 248, table 4.143)



Ryc. 9. Przykład wykresu szans Odds Ratio dla relacji pomiędzy ilością polepy konstrukcyjnej i niekonstrukcyjnej z zolnika w Chotyńcu 1

Fig. 9. An example of Odds Ratio for the relationship between the amount of constructional and non-constructional burnt daub from the zolnik in Chotyńiec 1



Ryc. 10. Gęstość występowania polepy na zolniku w Chotyńcu wraz z zaznaczoną częścią przyjętą do analiz

Fig. 10. Density of occurrence of the burnt daub on the zolnik in Chotyńiec with the marked part accepted for analysis

delowania. Problematyka ta jest niezwykle szeroka i nie sposób omówić ją tu w całości, należy jednak – chociaż pokrótce – przedstawić kilka podstawowych typów modeli.

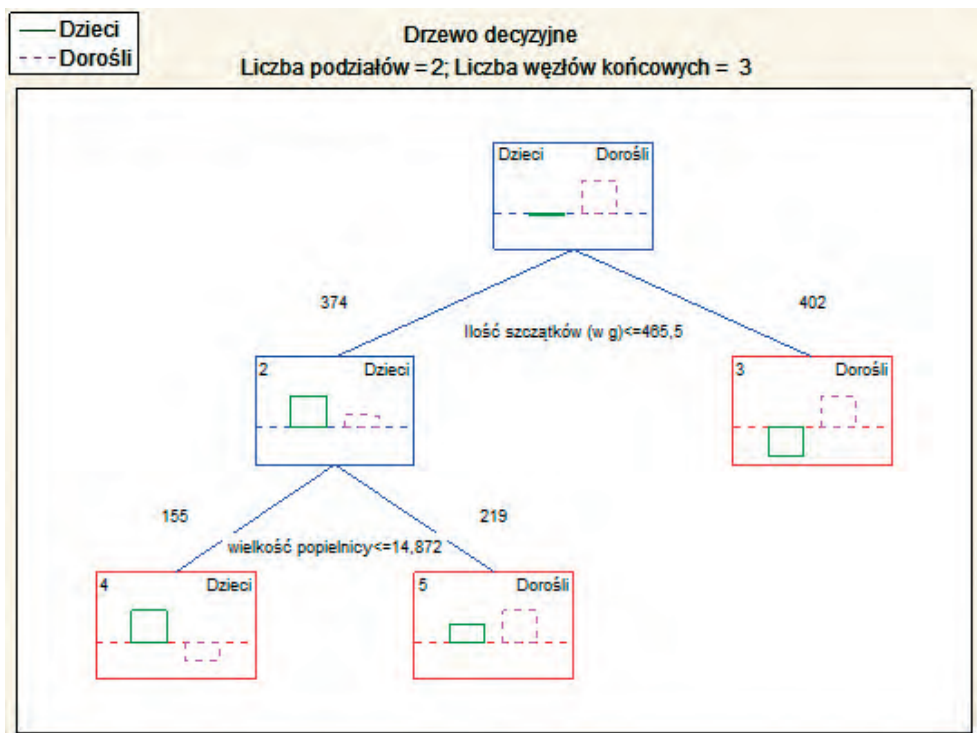
Najprostszą formą modelowania matematycznego są tzw. drzewa decyzyjne. Mają one ustalony kształt, w którym pień to wszystkie możliwe decyzje, kolejne węzły to podejmowane decyzje, a liście to cele. Drzewa te służą do prezentacji decyzji i ich możliwych konsekwencji – w założeniu mają pomagać w ich podejmowaniu (Bujak 2008, 2–3). Metoda ta opiera się o tzw. entropię, czyli liczbę informacji przypadającą na znak zajścia danego zdarzenia (Bujak 2008, 10–11). Charakteryzuje się ona kilkoma cechami, jednak najważniejszą jest, że nie może być ujemna. Jeżeli prawdopodobieństwo zajścia zdarzeń jest takie samo, osiąga maksymalną wartość. Jeżeli zmienne przyjmują wartości 0 i 1, zawsze wynosi 0. Gdy dwa systemy są od siebie niezależne, to jej suma jest równa sumie tych dwóch systemów. W celu obliczenia drzewa decyzyjnego obliczamy entropię dla kolejnych przypadków – aż do momentu, kiedy będzie równać się 0, co zamknie „konar” decyzyjny drzewa. Jako przykład posłużą dane dotyczące dzieci i dorosłych z cmentarzysk tarnobrzeskiej kultury łużyckiej, przy wzięciu pod uwagę wielkości popielnicy, ilości szczątków i wielkości pokryw (ryc. 11). W pierwszym kroku mamy podział ze względu na ilość szczątków. Jeżeli w grobie było więcej niż 516 gramów szczątków kostnych, model wskazuje, że musiał to być grób dorosłego. Drugi krok dotyczył wielkości popielnicy. W tym kroku, jeżeli popielnica miała więcej niż 20 punktów, ponownie model sugeruje osobę dorosłą, nawet mimo niewielkiej ilości szczątków. Brak w tym modelu węzła dla pokryw, chociaż były one wykorzystywane przy jego two-

zeniu. Stało się tak, ponieważ w takim zestawieniu nie miały one znaczenia dla decyzji. Dlatego w sytuacji, gdy mamy dane odnośnie do wielkości popielnicy i ilości szczątków wielkość pokrywy nie jest już nam potrzebna. Jednakże, gdybyśmy mieli tylko dane odnośnie wielkości pokrywy (ryc. 12), wówczas ta zmienna może również stanowić wyznacznik wieku. Taki obraz wynika z tzw. przycinania, gdzie wykres automatycznie wycina dane, które dla podjęcia decyzji mają mniejsze znaczenie. Przy drugim wykresie, gdzie nie było innych danych, taka sytuacja nie zaszła.

Tego typu drzewa – jak nie trudno zauważyć – pomagają w graficznej prezentacji danych. Znajdują także zastosowanie w uczeniu maszynowym i co szczególnie istotne mogą pomóc przy wyborze zmiennych potrzebnych do tworzenia bardziej skomplikowanych i dokładniejszych modeli matematycznych.

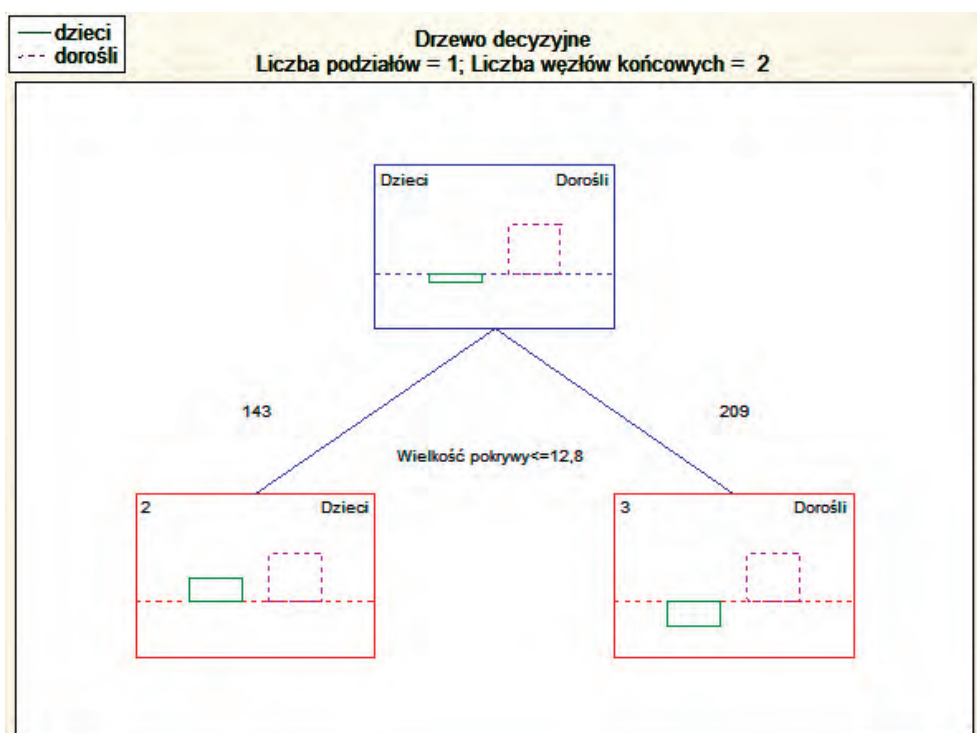
Takim typem jest między innymi model logitowy (Stanisz 2016). Jego podstawowym celem jest stworzenie na podstawie efektów badań empirycznych wzoru matematycznego, który będzie mógł posłużyć do opisanie zmiennej zależnej (Y – badana cecha: wiek, płeć, faza rozwoju) za pomocą zmiennych niezależnych (np. wielkości naczyń, obecność metalu, ilość przedmiotów grobie). W tego typu modelach interesującą nas zmienną (Y) można mierzyć na skalach porządkowej i nominalnej. Całe prawdopodobieństwo dla badanego przedziału będzie zawsze wynosić 1 (czyli 100%), a wyniki modelu dla badanych zmiennych będą się wahać w przedziale od 0 do 1. Co także istotne, liczba równań zawsze będzie o 1 mniejsza niż liczba przyjętych kategorii. Na przykład, jeżeli badamy zmiany form naczyń w kolejnych trzech fazach, otrzymamy 2 równania. Prawdopodobieństwo dla 3 gru-





Ryc. 11. Wykres drzewa decyzyjnego, które do określania wieku wykorzystywało wielkość popielnicy oraz ilość szczątków kostnych

Fig. 11. A decision tree diagram that used the size of the urn and the amount of bone remains to determine the age



Ryc. 12. Wykres drzewa decyzyjnego, które do określania wieku wykorzystywało tylko wymiary pokrywy

Fig. 12. A decision tree diagram that used only the dimensions of the urn to determine the age

py uzyskamy, odejmując sumę dwóch poprzednich. Model przyporządkuje dany przypadek do tej grupy, przy której uzyska najwyższy wynik. Czyli, jeżeli dla fazy I to było 0,1, dla fazy II – 0,3, a dla III – 0,6, to przyporządkuje właśnie do tej ostatniej. Wskazuje to również prawdopodobieństwo danego zdarzenia. Jeżeli model pokazuje 0,6 przy III fazie, możemy powiedzieć, że z 60% prawdopodobieństwem badane naczynie pochodziło z tego okresu.

Aby stworzyć, tego typu model określa się wartości  $a_0, \dots$ , an dla zmiennych, które istotnie wpływają na badaną przez nas cechę. Aby te zmienne wybrać, można wykorzystać, między innymi, wspomniane wcześniej drzewa decyzyjne, można oprzeć się na wykresach szans Odds Ratio; ale jest również tzw. kryterium Akaikego (w skrócie AIC), służące do wyboru najwłaściwszych modeli o różnej liczbie predyktorów (Akaike 1973). Ostatecznie model dla dwóch zmiennych przyjmie poniższą formułę:

$$P(Y = \text{badana zmienna} | X_1 = x_1, \dots, X_n = x_n) = \frac{\exp(a_0 + a_1 \cdot x_1 + \dots + a_n \cdot x_n)}{1 + \exp(a_0 + a_1 \cdot x_1 + \dots + a_n \cdot x_n)}$$

Gdzie: exp to funkcja wykładnicza o podstawie  $e=2,73$ ;

$a_0, a_1, a_2, \dots, a_n$  to współczynniki regresji, obliczone specjalnie dla konkretnej zmiennej niezależnej;

$x_1, x_2, \dots, x_n$  to zmienne niezależne, które mogą być wyrażone na skali porządkowej i nominalnej (np. ilość szczątków, wielkość popielnicy, ranga bogactwa).

Gdy mamy więcej zmiennych otrzymamy większą ilość wzorów. Warto jeszcze zaznaczyć, że statystyka oferuje szereg metod mających na celu oszacowanie istotności takiego modelu. Podstawowymi są: test Walda (Fahremer et al. 2013, 663), odchylenie D (Hosmer, Lemeshow 2000, 145–146), iloraz wiarygodności (LR) (King 1989, 84) oraz test Hosmera-Lemeshowa (Hosmer, Lemeshow 2000, 147–155). Warto także zaznaczyć, że w przypadku pierwszego testu im wyższy wynik, tym większe podstawy do odrzucenia badanej cechy. W przypadku odchylenia D, im wynik ilorazu tego testu z liczbą stopni swobody bliższy 0, tym model jest lepiej dopasowany. Dla ilorazu wiarygodności – im wyższy wynik, tym gorszy model, ponieważ oznacza to, że większa ilość zmienności zmiennej zależnej nie została rozwiązana. I na koniec, w przypadku testu Hosmera-Lemeshowa, badamy licznosci występowania wartości wyróżnionej Og i przewidywane prawdopodobieństwo Eg. Jeżeli obydwie wartości są do siebie zbliżone, model jest dobrze dopasowany.

Pierwszy przykład ich zastosowania (tabela 15) dotyczy prób wykorzystania modeli logitowych do przyporządkowania fragmentów ceramiki do jednej z 5 faz, jakie udało się uchwycić w rozwoju osady kultury trzcinieckiej na stanowisku Pawłosiów 52, pow. jarosławski (Adamik-Proksa 2019, 58–71). Dwoma cechami ceramiki związanymi z chronologią była grubość ścianek i ich barwa. Te właśnie cechy zostały wykorzystane w modelu. Co warto podkreślić, pierwsza cecha jest mierzalna, ponieważ można zmierzyć grubość ścianek. Druga z kolei jest jakościowa – barwom ścianek przypisano odpowiednie cyfry, które je kodowały. Model ten nie jest zbyt dokładny i z nieco ponad 60% prawdopodobieństwem rozpoznaje materiały z fazy II i z ponad 50% fazy V; dla pozostałych jego skuteczność jest 0. Trzeba jednak podkreślić, że jest to bardzo prosty model. Można tu więc dołożyć dodatkowe

cechy, jak choćby forma naczyń czy sposób ornamentyki, tym samym podwyższając jego skuteczność. Co jednak istotne, model ten pokazuje, że wspomniane cechy faktycznie mają znaczenie przy analizach dotyczących chronologii. Warto też zwrócić uwagę, iż po pewnych przekształceniach model ten można przekształcić na prostszy: podział na materiały starsze i młodsze (tabela 16).

Drugi przykład dotyczył pochówków dzieci i osób dorosłych z cmentarzysk tarnobrzeskiej kultury łużyckiej. Autor poniższego tekstu w ramach swojej rozprawy doktorskiej (Rajpold 2021) przedstawił wiele rozmaitych modeli dla każdego z 11 opracowywanych przez niego cmentarzysk, a także dla całości danych. Najlepsze efekty otrzymano przy modelu opartym o ilość szczątków i wymiary popielnicy. W tym miejscu jednak autor chciałby przedstawić dużo ciekawszy – z punktu widzenia archeologii – model, oparty wyłącznie o cechy archeologiczne (tabela 17). Model ten można było skonfrontować z oznaczeniami antropologicznymi i w przypadku osób dorosłych jego skuteczność wynosiła ponad 90%. W przypadku grobów dziecięcych skuteczność okazała się jednak dużo niższa, bo nieco ponad 50%. Model opierający się o ilość szczątków miał efektywniejszą, bo prawie 70%, dokładność we wskazywaniu pochówków dzieci (Rajpold 2021, 486, tabl. 5.2). W tym miejscu należy zaznaczyć, że porównanie grobów dzieci, wskazanych przez model jako osoby dorosłe, wykazało szereg cech wspólnych z pochówkami osób dorosłych. Prawdopodobnie wynika to z wczesnego momentu inicjacji w społeczności tarnobrzeskiej kultury łużyckiej i przyznawania praw do pochówku typowego dla dorosłego osobom, które – według naszych standardów – wciąż byłyby traktowane jako dzieci. To tłumaczy niską skuteczność archeologicznego modelu dla dzieci. Model ten, opierający się wyłącznie o cechy kulturowe, prawdopodobnie – przynajmniej w części przypadków mylnych wska-

**Tabela 14. Tabela liczebności wraz z testem  $\chi^2$  dla danych, dotyczących ilości polepy z dwóch warstw tzw. płoszczadki z grodziska na stanowisku Chotyńiec 1**

**Table 14. Table of quantities with the  $\chi^2$  test for data presenting the amount of burnt daub from two layers, the so-called ploschadka from the fortified settlement at the site of Chotyńiec 1**

				Chi <sup>2</sup>	p
	I	II	Suma		
Konstrukcyjna	883	1094	1977	52,28	0
Niekonstrukcyjna	9788	17011	26799		
Ogół	10671	18105	28776		

Tabela 15. Dane do modelu logitowego dla 5 faz z osady kultury trzcinieckiej w Pawłosiowie, stan. 52, pow. jarosławski

Table 15. Data for the logit model regarding 5 phases from the Trzciniec Culture settlement in Pawłosiów, site 52, Jarosław dist.

	Ocena	Standardowe błędy	Walda	95% dolny CI	95% górny CI	p
Wyraz wolny 1	0,98	0,91	1,17	-0,8	2,77	0,28
Barwa ścianek	-0,12	0,31	0,14	-0,73	0,5	0,71
Grubość ścianek	-0,19	0,16	1,42	-0,51	0,12	0,23
Wyraz wolny 2	1,19	0,76	2,42	-0,31	2,69	0,12
Barwa ścianek	-0,27	0,26	1,14	-0,78	0,23	0,29
Grubość ścianek	-0,11	0,13	0,66	-0,37	0,15	0,42
Wyraz wolny 3	-3,3	1,17	7,98	-5,6	-1,01	0
Barwa ścianek	-1,31	0,37	12,42	-2,03	-0,58	0
Grubość ścianek	0,5	0,19	6,75	0,12	0,88	0,01
Wyraz wolny 4	-5,05	1,27	15,8	-7,53	-2,56	0
Barwa ścianek	-1,57	0,38	16,8	-2,33	-0,82	0
Grubość ścianek	0,75	0,2	13,25	0,34	1,15	0
Skala	1	0		1	1	
Dobroć dopasowania						
	Df	Stat.	Stat/Df			
Odchylenie	1880	1339,94	0,71			
Skal.odchylenie	1880	1339,94	0,71			
Chi2 Pearsona	1880	1878,59	1			
Skalow. Chi2 P.	1880	1878,59	1			
AIC		1363,94				
AICC		1364,61				
BIC		1413,84				
Log(il.wiar.)		-669,97				
LR1						
	St. Swobody	Log największej wiarygodności	chi^2	p		
Wyraz wolny	4	-693,91				
Barwa ścianek	4	-684,32	19,18	0		
Faza	4	-669,97	28,7	0		
Poprawne oszacowania dla faz w %						
I	II	III	IV	V		
0	62,59	0	13,95	51,48		

zań – pokazywał groby dzieci, które przez społeczność tarnobrzesckiej kultury łużyckiej uznawane były za osoby dorosłe.

Jeszcze innym typem modeli matematycznych, które można zastosować w archeologii, są tzw. analizy dyskryminacyjne. Ich zasada jest nieco inna niż w przypadku modeli logistycznych, jednak cel jest ten sam – klasyfikacja badanych przypadków pod warunkiem dysponowania informacjami o cechach (predyktorach) tych przypadków (Radkiewicz 2010). Dane wykorzystane w tego typu modelach muszą pochodzić z dwóch lub większej ilości niezwiązanych ze sobą grup, które muszą liczyć przynajmniej dwie obserwacje (Radkiewicz 2010, 144). Badane przez nas cechy wykorzystane, aby różnicować grupy

nazywamy zmiennymi dyskryminującymi. Muszą być one mierzone na skalach przedziałowych, gdyż do ich obliczania potrzebna jest mediana oraz średnia. W jednym zbiorze nie mogą się znaleźć dwie zmienne, które będą ze sobą idealnie skorelowane. Musi być ich zawsze o co najmniej dwie mniej. Ostatnie założenie mówi, że próby pobrano z grup, w których mieliśmy wielowymiarowy rozkład normalny (Radkiewicz 2010, 144). Na podstawie wspomnianych zmiennych powstają tzw. kanoniczne funkcje dyskryminacyjne. Mogą mieć one różne wzory, ale w naszych rozważaniach wykorzystaliśmy funkcje liniowe, które przybierają następujący wzór (za Stanis 2007b, 53):

Tabela 16. Dane do modelu logitowego dla podziału na materiały starsze i młodsze z osady kultury trzcinieckiej w Pawłostowie, stan. 52, pow. jarosławski  
 Table 16. Data for the logit model regarding the division into older and younger materials from the Trzciniec culture settlement in Pawłosiów, site 52, Jarosław dist.

	Wartość parametru	Standardowy błąd	Statystyka Walda	95% dolny koniec przedziału ufności	95% górny koniec przedziału ufności	P	Iloraz szans	95% dolny koniec przedziału ufności dla ilorazu szans	95% górny koniec przedziału ufności dla ilorazu szans
Wyraz wolny	2,29	0,63	13,35	1,06	3,53	0,00			
Barwa	0,27	0,20	1,73	-0,13	0,67	0,19	1,31	0,88	1,95
Grubość ścianek w mm	-0,35	0,11	10,80	-0,57	-0,14	0,00	0,70	0,57	0,87
Dobroć dopasowania									
	Df	Stat.	Stat/Df						
Odchylenie	470,00	630,29	1,34						
Skal.odchylenie	470,00	630,29	1,34						
Chi2 Pearsona	470,00	472,58	1,01						
Skalow. Chi2 P.	470,00	472,58	1,01						
AIC		636,29							
AICC		636,35							
BIC		648,77							
Log(il.wiar.)		-315,15							
LR1									
	St. Swobody	Log największej wiarygodności	chi^2	P					
Wyraz wolny	1	-325,71							
Barwa	1	-320,75	9,94	0,00					
Grubość ścianek w mm	1	-315,15	11,20	0,00					
Poprawne oszacowania dla faz w %									
	Starsze przewidywanie	Młodsze przewidywanie	Procent poprawne						
Starsze	105,00	109,00	49,07						
Młodsze	88,00	171,00	66,02						

Tabela 17. Dane do modelu logitowego służącego określeniu wieku (prawdopodobieństwo pochówku dziecka) przy uwzględnieniu cech popielnicy i innych elementów obrządku pogrzebowego tarnobrzeskiej kultury łużyckiej (Rajpold 2021, 507–508, tabl. 5.31)

Table 17. Data for the logit model used to determine the age (probability of a child's burial) taking into account the features of the urn and other elements of the funeral rites of the Tarnobrzeg Lusatian culture (Rajpold 2021, 507–508, tabl. 5.31)

Efekty	Wartość parametru	Błąd standardowy	Statystyka Walda	95% dolny koniec przedziału ufności	95% górny koniec przedziału ufności	p
Wyraz wolny	6,66	0,70	91,76	5,30	8,02	0,00
Forma popielnicy: waza (1)	-0,01	0,28	0,00	-0,56	0,54	0,97
Forma popielnicy: misa (2)	-1,57	0,72	4,78	-2,98	-0,16	0,03
Forma popielnicy: czerpak lub kubek (3)	-0,54	0,70	0,61	-1,91	0,82	0,44
Forma popielnicy: dzban (4)	1,84	0,69	7,14	0,49	3,20	0,01
Forma popielnicy: waza nadsańska (5)	0,25	0,39	0,40	-0,52	1,01	0,52
Forma popielnicy: garnek jajowaty (6)	-0,01	0,63	0,00	-1,24	1,23	0,99
Forma popielnicy: inne naczynia (7)	0,50	0,45	1,26	-0,37	1,37	0,26
Wysokość szyjki wysoka (1)	0,07	0,27	0,08	-0,45	0,60	0,78
Wysokość szyjki: krótka (2)	0,57	0,26	4,96	0,07	1,08	0,03
Wielkość popielnicy	-0,44	0,04	143,76	-0,51	-0,37	0,00
Brzusiec mocno wydęty- 1	-0,01	0,15	0,00	-0,30	0,28	0,97
Brzusiec smukły-2	0,03	0,16	0,03	-0,29	0,35	0,87
Liczba elementów zdobniczych- 0 (naczynie nie zdobiono)	-0,30	0,29	1,11	-0,86	0,26	0,29
Liczba elementów zdobniczych- 1	-0,32	0,30	1,17	-0,90	0,26	0,28
Liczba elementów zdobniczych- 2	0,04	0,34	0,02	-0,62	0,71	0,90
Liczba elementów zdobniczych- 3	0,68	0,33	4,22	0,03	1,34	0,04
Liczba elementów zdobniczych- 4	0,26	0,43	0,36	-0,59	1,11	0,55
Skręt jest obecny (1)	0,40	0,16	6,38	0,09	0,70	0,01
Skala	1,00	0,00		1,00	1,00	

Efekty	Wartość parametru	Błąd standardowy	Statystyka Walda	95% dolny koniec przedziału ufności	95% górny koniec przedziału ufności	p
LR1						
Efekty	stopnie	Log-	Chi^2			p
Wyraz wolny	1	-515,24				
Podstawowy typ nacznia: waza (1), misa (2), czerpak lub kubek (3), dzban (4), waza nadsańska (5), garnek jajowaty (6), inne (7), garnek tulipanowaty (8)	7	-349,27	28,77	0,00		
Wysokość szyjki – wysoka (1), krótka (2), niewyróżniona (3)	2	-346,34	5,86	0,05		
Brzusiec – mocno wydęty (6) względem smukłego (7)	2	-346,33	0,02	0,99		
Wielkość popielnicy	1	-363,65	303,18	0,00		
Liczba elementów zdobniczych (0-5)	5	-342,54	7,58	0,18		
Obecność skrzydła 1 – jest, 2 – nie ma	1	-339,45	6,18	0,01		
Dobroć dopasowania						
	Df	Stat.	Stat/Df			
odchylenie	890,00	678,91	0,76			
Skal.odchylenie	890,00	678,91	0,76			
Chi^2 Pearsona	890,00	1938,27	2,18			
Skalow. Chi^2 P.	890,00	1938,27	2,18			
AIC		716,91				
BIC		808,34				
R2 Coxa-Snella		0,32				
R2 Nagelkerka		0,47				
Log (il. wiar.)		-339,45				
Poprawność	Dorośli -93,21%	Dzieci -55,4%				

$$D_{kj} = b_0 + b_1x_{1kj} + \dots + b_px_{pkj}$$

$$K_i = c_{i0} + c_{i1}x_1 + \dots + c_{ip}x_p$$

Gdzie:  $p$  – liczba zmiennych dyskryminacyjnych,  $n$  – liczbeność próby,  $g$  – liczba grup (wartości zmiennej grupującej)  $D_{kj}$  – wartość kanonicznej dyskryminacyjnej funkcji dla  $k$ -tego przypadku w  $j$ -tej grupie,  $x_{ikj}$  – wartość  $i$ -tej zmiennej dyskryminującej dla  $k$ -tego przypadku w  $j$ -tej grupie ( $i = 1, \dots, p$ ),  $b_i$  – współczynniki kanonicznej funkcji dyskryminacyjnej, wyznaczonej na podstawie własności tej funkcji ( $b_i$  – określamy, tj. wyliczamy, estymujemy tak, aby średnie grup były jak najbardziej zróżnicowane).

Gdzie:  $c_{ij}, j = 1, \dots, p$  są wyliczonymi z próby specjalnymi współczynnikami zmiennych dyskryminacyjnych. Podstawiając do powyższego wzoru wartości zmiennych (np. ilość szczątków i wielkość popielnicy) dyskryminacyjnych dla badanego przypadku (np. grobu), otrzymujemy dla każdej z grup inną wartość  $K_i$  przyjmujemy się, że badany przypadek będzie należał do grupy (np. wiek z podziałem na dzieci i dorosłych), dla której  $K_i$  osiągnie najwyższy wynik.

Kolejnym istotnym tematem jest kwestia liczby funkcji dyskryminacyjnych, które będą istotne statystycznie. To stwierdzamy przy użyciu tzw. statystyki lambda Wilksa. Aby określić siłę powiązań ze sobą zmiennych i funkcji dyskryminacyjnych stosujemy współczynnik struktury czynnikowej (im bliższy 0, tym gorszy model; optymalny wynik to  $-1$  i  $1$ ) (Stanisz 2007b, 58–59).

Jeżeli już zostaną wybrane funkcje dyskryminacyjne, przechodzi się do kolejnego etapu, a mianowicie klasyfikacji przypadków. Wszystkie procedury tego typu stosują porównanie pozycji (geometrycznej) badanego przypadku względem każdej centroidy (te tworzą zmienne), aby znaleźć ten najbliższy. W przykładowym modelu zastosowano pochodzącą od Fishera klasyfikację, bazującą na linowej kombinacji zmiennych dyskryminujących. Oznacza to, że dla każdej z badanych grup opracowuje się osobne funkcje, które przybierają poniższą formę (za Stanisz 2007b, 64):

Przykładowy model analizy dyskryminacji (tabela 18) oparto na danych z cmentarzyska w Grodzisku Dolnym, stan. 1. Skuteczność tego modelu wynosiła ponad 90% w przypadku dorosłych i w prawie 90% przypadków trafnie wskazywał pochodówki dzieci. Model ten ma więc lepszą skuteczność niż zaprezentowany wcześniej model logitowy, jednakże ten odnosi się tylko do tego jednego cmentarzyska, gdyż na innych jego skuteczność była mniejsza (ponadto, w przeciwieństwie do wcześniejszego modelu, nie można go wykorzystywać do całości danych). Oczywiście możliwe jest stworzenie modelu analizy dyskryminacji dla całości danych, jednak dla materiałów tarnobrzeskiej kultury łużyckiej skuteczność takiego modelu byłaby mniejsza niż wcześniej wspomnianego logitowego. Model ten w teorii można wykorzystać więc tylko do tego jednego cmentarzyska, a to, teoretycznie, znacząco ogranicza jego wartość. Można go jednak odnosić do nekropolii o podobnej chronologii, z naczyniami o zbliżonych wymiarach i podob-

**Tabela 18. Przykład analizy dyskryminacyjnej dla wieku (dzieci i dorosłych) zmarłych z cmentarzyska tarnobrzeskiej kultury łużyckiej w Grodzisku Dolnym, stan. 1, pow. leżajski (Rajpold 2021, 504, tabl. 5.26)**

**Table 18. Example of discriminant analysis used to distinguish the age (children and adults) of the deceased from the cemetery of the Tarnobrzeg Lusatian culture in Grodzisk Dolny, site 1, Leżajsk dist. (Rajpold 2021, 504, tabl. 5.26)**

Zmienne do modelowania						
Efekty	Lambda wilksa	cząstkowe wilksa	F usun (1,23)	p	Toler.	1-Toler. R <sup>2</sup>
Ilość szczątków	0,510816	0,870095	8,51007	0,005045	0,760593	0,239407
Wielkość popielnicy	0,525461	0,845845	10,38827	0,002099	0,760593	0,239407
Wzór funkcji dyskryminacyjnej						
D=4,08-0,001*(il. szczątków)-0,17*(średnia naczynia)						
Funkcja kwalifikacyjna do dzieci	K1= - 10,29 - 0,004*(il. szcząt) + 1,35*(śr. nacz)					
Funkcja kwalifikacyjna do dorosłych	K2= -18,15 - 0,0016*(il. szcząt) + 1,73*(śr. nacz)					
Poprawność klasyfikacji						
Wiek	Macierz klasyfikacji					
	Procent poprawnych	Dzieci	Dorośli			
Obserwacja dzieci	89,47369	17	2			
Obserwacja dorośli	92,68293	3	38			
Razem	91,66666	20	40			

nym stanie zachowania grobów. W takim ujęciu zbudowany na podstawie jednej nekropolii model można próbować dopasować do innej. Zbudowanie bazy algorytmów dla kilku nekropolii z różnych faz rozwoju danej kultury automatycznie daje możliwości ekstrapolowania takich wyników na inne, zbliżone do nich cmentarzyska. W takim ujęciu wartość tego modelu – jako jednego z kilku reperowych – rośnie.

Omówione powyżej analizy nie wyczerpują wszystkich typów modeli matematycznych. Można między innymi wspomnieć jeszcze tzw. regresję wieloraką, którą stosuje się, gdy

mamy dużą ilość zmiennych, mogących wpływać na badaną przez nas cechę. Co istotne, pomaga ona wybrać te zmienne, które mają największe znaczenie (Stanisz 2007a, 59–97; 2016, 78–91). Jednakże z oczywistych przyczyn, omówienie każdego rodzaju modelowania znacząco wykracza poza ramy niniejszego tekstu. Warto jednak jeszcze raz podkreślić to, że cechą wszystkich modeli matematycznych jest stworzenie na podstawie danych empirycznych wzoru, który możliwie jak najdokładniej opisze rzeczywistość i pozwoli przewidzieć na tej podstawie kolejne wyniki.

## PODSUMOWANIE

Podsumowując ten skrótywy przegląd metod statystycznych, należy stwierdzić, że ich znaczenia dla archeologii w sytuacji systematycznego napływu nowych danych nie można nie doceniać. Należy sądzić, że będą one coraz chętniej wykorzystywane. Tym bardziej, że ich możliwości poznawcze wciąż są nie do końca wyeksploatowane przez archeologię. Szczególnie dotyczy to modelowania matematycznego, które – jak pokazano – można wykorzystywać między innymi do badań dotyczących chronologii oraz wskazywania wyznaczników wieku i płci. Co też istotne, obecnie, do wykonywania skomplikowanych obliczeń statystycznych nie potrzeba godzin spędzonych na wyprowadzaniu kolejnych wzorów. Wykorzystując ogólnodostępne programy jak Statistica czy R i RStudio, błyskawicznie możemy wykonywać analizy na zbiorach liczących nawet kilkadziesiąt tysięcy rekordów, nigdy nie korzystając ze wzorów.

Warto zaznaczyć, że wynik analizy statystycznej (korelacji, wykresów dendrytowych, tablicy korespondencji, a nawet modelowania) nie zawsze musi oznaczać wykrycie związku między badanymi czynnikami. Każdą z analizowanych sytuacji należy rozpatrywać ostrożnie – bliżej przyglądać się wszystkim otrzymanym wynikom i pamiętać, że jest to jedynie narzędzie.

Warto przyrzeć się jeszcze sytuacjom, gdy wyniki analiz statystycznych są oddalone od ideału. W takich przypadkach

połączenie rozumowania statystycznego z archeologicznym prowadzi do wyjątkowo ciekawych wniosków. Najdobitniej ukazuje to zasygnalizowany wyżej model logitowy, opierający się wyłącznie o cechy archeologiczne dla wyznaczenia wieku zmarłych. Model ten osiągał gorsze rezultaty niż algorytm, opierający się o czynnik biologiczny, jakim jest ilość szczątków. Na pierwszy rzut oka mamy więc niezadowolające wyniki. Jednakże, gdy uwzględnimy wiedzę o tym, że w wielu społecznościach moment inicjacji dzieci był wczesny (Frazer 1969, 55, 60, 521–523; Nowicka 1991, 342; Genep 2006, 84–126), możemy zrozumieć przyczyny niższej skuteczności tego modelu. Jeden z nich wskazywał bowiem wiek biologiczny, a drugi – kulturowy. W takim ujęciu obydwa modele zaczynają się uzupełniać i już trudniej wskazać ten, który we właściwy sposób obrazuje rzeczywistość.

Nie sposób nie zaznaczyć, że bardzo istotnym jest wykorzystywanie właściwej metody do odpowiednich danych, aby uniknąć zafałszowania obserwacji. Na koniec należy pamiętać, że statystyka jest jedynie pomocnym nam – archeologom – narzędziem i korzystając z jej możliwości, należy zawsze zachowywać zdrowy rozsądek. Łatwo, bowiem wpaść w pułapkę popularnego przysłowia: „jeżeli ja mam dwie nogi, a mój pies cztery, to statystycznie obydwoje mamy trzy”.

## WYKAZ CYTOWANEJ LITERATURY

Adamik-Proksa J. 2019. *Ekumena ludności kultury trzcinieckiej w Polsce południowo-wschodniej studium osadniczo-kulturowe*. Rzeszów: Uniwersytet Rzeszowski.

Akaike H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. W: B.N. Petrov, F. Csaki (red.), *International Symposium on Information Theory*, 267–281. Budapeszt: Akademiai Kiado.

Bahn P. (red.) 2019. *Historia archeologii*. Warszawa: Arkady.

Bobak D., Trybała-Zawiślak K. 2016. Statystyczna analiza zespołów grobowych z cmentarzyska tarnobrzeskiej kultury łużyckiej w Kłyszowie, pow. Stalowowolski. *Materiały i Sprawozdania Rzeszowskiego Ośrodka Archeologicznego* 37, 143–152.

Bielecka A. 2011. *Statystyka dla menedżerów. Teoria i praktyka*. Warszawa: Oficyna Wolters Kluwer Business.

Bugaj A. 2018. Magazyny archeologiczne w Polsce – kwestia przechowywania zabytków archeologicznych, czyli rzecz o stajniach Augiasza i kilka przykładów dobrych praktyk. *Raport* 13, 155–169.

Bujak Ł. 2008. *Drzewa decyzyjne*. Toruń: Uniwersytet Mikołaja Kopernika Wydział Fizyki, Astronomii i Informatyki Stosowanej.

Burghardt M. 2016. Wyznaczniki płci w grobach ludności grup leśnostepowych kultury scytyjskiej. *Materiały i Sprawozdania Rzeszowskiego Ośrodka Archeologicznego* 37, 165–208.

Cramer J.S. 2001. *An introduction to the logit model for economist*. London: Timberlake Consultants.

Czaja J., Preweda E. 2000. Analiza statystyczna zmiennej losowej wielowymiarowej w aspekcie korelacji i predykcji. *Geodezja* 6/2, 129–145.

Czopek S. 1996. *Grupa tarnobrzeska nad środkowym Sanem i dolnym Wisłokiem*. Rzeszów: Muzeum Okręgowe w Rzeszowie.

Czopek S. 2001. *Pysznica, pow. Stalowa Wola, stanowisko 1 – cmentarzysko ciałopalne z przełomu epok brązu i żelaza*. Rzeszów: Uniwersytet Rzeszowski.

Czopek S. 2017. Kilka uwag na marginesie „archeologii autostradowej”. *Raport* 12, 229–237.



- Czopek S. 2020. Znaczenie odkryć w Chotyńcu (południowo-wschodnia Polska) dla interpretacji procesów kulturowo-histerycznych na zachodnim Wołyniu i Podolu (Ukraina) we wczesnej epoce żelaza. *Archeologia Polski* 55, 67–114.
- Czopek S., Trybała-Zawiślak K., Tokarczyk T., Ocadyga-Tokarczyk E., Burghardt M., Adamik-Proksa J., Rajpold W. 2017. Pierwsze sprawozdanie z weryfikacyjnych badań na grodzisku z wczesnej epoki żelaza w Chotyńcu. *Materiały i Sprawozdania Rzeszowskiego Ośrodka Archeologicznego* 38, 291–305.
- Czopek S., Pelisiak A. 2014. Autostrady i co dalej. *Raport* 9, 423–434.
- Edwards H.G. M. 2009. Raman spectroscopic analysis of art and archaeological artefacts of relevance to the preservation of cultural heritage: the illumination of ancient mysteries. *Wiadomości konserwatorskie* 9, 55–64.
- Fagan B. 2021. *Krótką historią archeologii*. Warszawa: Wydawnictwo Read ME.
- Fahrmeir L., Kneib T., Lang S., Marx B. 2013. *Regression: models, methods and applications*. Berlin: Springer.
- Flatcher M., Lock G. 1991. *Digging Numbers: Elementary statistics for archaeologists*. Oxford: University School of Archaeology.
- Frazer J.G. 1969. *Złota Gałąź*. Warszawa: Państwowy Instytut Wydawniczy.
- Gąssowski J. 1983. *Z archeologią za pan brat*. Warszawa: Iskry.
- Gediga B., Łaciak D., Łydzba-Kopczyńska B., Markiewicz M. 2017. *Świat kolorów garncarzy z rejonu Domasławia sprzed około 2800 lat*. Wrocław: Instytut Archeologii i Etnologii PAN.
- Gennep A. 2006. *Obrzędy przejścia. Systematyczne studium ceremonii*. Warszawa: Państwowy Instytut Wydawniczy.
- Gilda L. 2019. *Probability model for archaeological site location, a case study on O'ahu island, Hawai'i*. Master's thesis. Hawai'i, Mānoa.
- Górski J. 2007. *Chronologia kultury trzcinieckiej na lessach Niecki Nidziańskiej*. Kraków: Biblioteka Muzeum Archeologicznego w Krakowie.
- Harańczyk G. 2010. Krzywe ROC, czyli ocena jakości klasyfikatora i poszukiwanie optymalnego punktu odcięcia. [https://media.statsoft.pl/\\_old\\_dnn/downloads/krzywe\\_roc\\_czyli\\_ocena\\_jakosci.pdf](https://media.statsoft.pl/_old_dnn/downloads/krzywe_roc_czyli_ocena_jakosci.pdf).
- Hilbe J., Greene W. 2008. Count response regression models. W: C.R. Rao, J.P. Miller, D.C. Rao (red.), *Epidemiology and medical statistics*, 210–252. Amsterdam, Boston: Elsevier.
- Holtorf C. 2007. *Archaeology is a brand! The meaning of archaeology in contemporary popular culture*. Oxford: Archaeopress.
- Hosmer D.W., Lemeshow S. 2000. *Applied Logistic Regression*. New York: Wiley – Interscience Publication.
- Johnson M. 2013. *Teoria archeologii*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- King G. 1989. *Unifying political methodology: the likelihood theory of statistical inference*. Cambridge: University Press.
- Korczyńska M. 2014. Dynamics of the Depositional Processes: the example of the tree windthrow structure at the graveyard in Janowice, site 44 (AZP 106-65/103). W: T.L. Kienlin, P. Valde-Nowak, M. Korczyńska, K. Cappenberg, J. Ociepa (red.), *Settlement, Communication and Exchange around the Western Carpathians. International Workshop held at the Institute of Archaeology, Jagiellonian University, Kraków October 27–28, 2012*. Oxford: Archaeopress.
- Korczyńska M. 2016. Wpływ archeologii niemieckiej na prace profesora Marka Gedla i recepcja jego dorobku w niemieckim środowisku archeologicznym. W: J. Chochorowski (red.), *Profesor Marek Gedl – życie i dzieło*, 347–382. Kraków: Wydawnictwo Profil – Archoe.
- Kvamme, Kenneth L., 2006. There and Back Again: Revisiting Archaeological Locational Modeling in GIS and Archaeological Site Locational Modeling. W: M.W. Mehrer, K.L. Wescott (red.) *GIS and Archaeological Site Modeling*, 3–38. London: CRC Press.
- Likert R. 1932. A technique for the measurement of attitudes. *Archives of psychology* 140, 55.
- Ławecka D. 2003. *Wstęp do archeologii*. Warszawa–Kraków: Wydawnictwo Naukowe PWN.
- Maczyńska A. 2011. Archeologia i media – małżeństwo z rozsądku. W: A. Marciniak, D. Minta-Tworzowska, M. Pawleta (red.), *Współczesne oblicza przeszłości*, 323–331. Poznań: Wydawnictwo Poznańskie.
- Major M. 2007. *Elementy statystyki. Rachunek prawdopodobieństwa i wnioskowanie statystyczne*. Kraków: Krakowska Szkoła Wyższa im. Andrzeja Frycza Modrzewskiego.
- Makarowicz P., Cwaliński M., Romaniszyn J. 2016. Absolute chronology of the Komarów Culture in the Upper Dniestr Basin in light of research at the Bukivna cemetery. *Analecta Archaeologica Ressoiviensia* 11, 131–164.
- Mrówka Ł. 2011. Development of pottery style on the bronze and Early Iron Age cemetery at Kietrz, Głubczyce district, in the light of statistical analyses. *Recherches Archéologiques* 3, 67–100.
- Nowicka E. 1991. *Świat człowieka – świat kultury*. Warszawa: Wydawnictwo Naukowe PWN.
- Pampel F.C., 2000. *Logistic regression: A primer. Quantitative Applications in the Social Sciences* 132. Sage: Sage Publications.
- Pearson K. 1900. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine* 5.50 (302), 157–175.
- Piotrowski W. 2005. Wykopaliska biskupińskie z lotu ptaka – próba podsumowania. W: J. Nowakowski, A. Prinke, W. Rączkowski (red.), *Biskupin... i co dalej? Zdjęcia lotnicze w polskiej archeologii*. Poznań: Instytut Prahistorii Uniwersytetu im. Adama Mickiewicza, Ośrodek Ochrony Dziedzictwa Archeologicznego, Muzeum Archeologiczne w Biskupinie, Poznańskie Towarzystwo Prehistoryczne.
- Petrie A., Sabin C. 2013. *Medical Statistics at a Glance*. London: John Wiley And Sons Ltd.
- Przybyła M.S. 2016. Middle Bronze Age social networks in the Carpathian Basin. *Recherches Archéologiques* 8, 47–84.
- Radkiewicz P. 2010. Analiza dyskryminacyjna. Podstawowe założenia i zastosowania w badaniach społecznych. *Psychologia Społeczna* 5, 2–3, 142–161.
- Rajpold W. 2020. Zastosowanie analizy korespondencji do klasyfikacji materiałów z późnej epoki brązu na przykładzie obiektów funeralnych z Polski południowo-wschodniej. *Folia Praehistorica Poznaniensia* 25, 209–247.
- Rajpold W. 2021. *Wyznaczniki płci i wieku w grobach tarnobrzeskiej kultury lużyckiej w świetle analiz statystycznych*. Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego.
- Shaus A., Sober B., Faigenbaum-Golovin S., Mendel-Geberovich A., Levin D., Piasezky E., Smith M.A. 2017. A Study in Urnfield Interpretations in Middle Europe. *Zephyrus* VIII, 195–240.
- Schmitz P.I.M. 1986. *Logistic Regression in medical decision making and epidemiology*. Rotterdam: Erasmus University.

Stanisz A. 2006. *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 1. Statystyki podstawowe*. Kraków: StatSoft.

Stanisz A. 2007a. *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 2. Modele liniowe i nieliniowe*. Kraków: StatSoft.

Stanisz A. 2007b. *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny. Tom 3. Analizy wielowymiarowe*. Kraków: StatSoft.

Stanisz A. 2016. *Modele regresji logistycznej. Zastosowania w medycynie, naukach przyrodniczych i społecznych*. Kraków: StatSoft.

Spitznagel E. 2008. Logistic Regression. W: C.R. Rao, J.P. Miller, D.C. Rao (red.), *Epidemiology and medical statistics*, 187–209. Amsterdam, Boston: Elsevier.

Waliszewski W. et al. 1988. *Ecyklopedia szkolna – Matematyka*. Warszawa: Wydawnictwo Szkolne i Pedagogiczne.

Walesiak M. 1996. Dopuszczalne działania na liczbach w badaniach marketingowych z punktu widzenia skal pomiarowych. *Prace naukowe Akademii Ekonomicznej we Wrocławiu*, 718, 103–110.

Warren R.E., Asch D.L. 2000. A predictive model of archaeological site location in the Eastern Prairie Peninsula. W: K.L. Wescott, R.J. Brandon (red.), *Practical Applications of GIS for Archaeologists: a Predictive Modeling Kit*, 5–25. London: Taylor & Francis.

Wheatley D., Gillings M., 2002. *Spatial Technology and Archaeology: the Archaeological Applications of GIS*. London: Taylor & Francis.

Zimbardo G.P., Richard J.G. 2012. *Psychologia i życie*. Warszawa: Wydawnictwo Naukowe PWN SA.

#### Wirtualne źródła informacji

<http://www.tylervigen.com/spurious-correlations>

Wojciech Rajpold

## Statistics in archaeology. So why shouldn't you be afraid of numbers?

### Summary

Archaeology, although it is a human science, draws from other areas of the world of science, especially from the achievements of natural sciences. Physics, chemistry and biology are widely used, for example in determining the chronology (C14, dendrology) or in the study of the chemical composition of artefacts (e.g. Raman spectroscopy). That is why it is not surprising that mathematics is also included in the arsenal of research methods used by archaeologists. The amount of archaeological materials widening the collection from year to year is impressive. However, it creates a huge challenge, including the one associated with the development of such a large number of sources. The artefacts obtained during excavations are massive, countable, therefore we can measure them and weigh them. So this is where statistics comes to the aid – the field of mathematics that organizes large numbers. The possibility of using statistical analyses can be found in many works of Polish researchers, and they show both richness and diversity, as well as usefulness of this field of science in archaeological studies.

The first issue that should be indicated is the type of data surveyed by the statistics. There are two types: a) quantitative (measurable) – e.g. weight or length of the artefact (continuous data) or e.g. the number of coils on the pin head (discrete data); b) nominal (immeasurable; the variable gets a numerical label); they can be binary and multiple. This group also includes ordinal data – otherwise rankable – which arrange materials according to the intensity of the phenomenon.

The type of data, in turn, determines the type of measurement scale which is going to be used. In archaeology, the so-called quotient scale gives the possibility to implement all statistical methods. In that case, we can measure weight and height. You can also use the so-called an ordinal scale that examines rankable data, where we grade the intensity of a specific feature.

However, it is worth emphasizing that statistical data (depending on the methods) can be transformed so that they can be recorded in many ways and measured on various scales.

In statistics, however, the most important issue is to examine whether the differences in the analysed groups are significant. For this purpose, the so-called  $\chi^2$ , Kruskal-Wallis and U-Mann Whitney tests are used. The first test – the most common – compares the observed prevalence with the expected rates. With its help, you can check whether the ceramics obtained from a particular site is technologically homogeneous, for instance. The next two tests are used when data are expressed on measurable scales and are implemented to test the median. They can be used, inter alia, to check whether the differences in the thickness of the vessels from given areas are significant.

What is most, commonly associated with statistics are various types of correlations, or interdependencies (e.g. Pearson correlation, Spearman's Rank correlations, Kendall's Tau correlation). However, it should be considered that not every correlation is a relationship between the studied features, therefore in any statistical method it is so important to exercise a certain degree of caution when reading the results.

Above all, methods that allow archaeologists to group data are very useful in the considerations. The most popular are dendrite diagrams and correspondence boards. Additionally, the cut-off point on the ROC curve and associated with it the Odds Ratio diagram may also be an important method.

Undoubtedly, the greatest power of statistics is mathematical modelling, which allows the researcher, based on the results of empirical research, to create a formula describing the analysed feature. There are many types of models, such as logistic, linear and decision trees. They are widely used, among

others, in medicine, sociology and economics – and they are also implemented in archaeological analyses. The simplest type of model is a decision tree, which is useful for the graphic presentation of data and the possible results of the decisions taken. It is also helpful when selecting the features that have the greatest impact on the topic we are researching. Such an example is the determination of the age of the deceased (divided into adults and children) on the basis of the number of remains and the size of the urn, where two steps were noted that led to the correct indication of the age of the deceased. The logit and discriminant models are much more complicated. They can be created for two variables, e.g. age divided into adults and children, as well as for more variables, e.g. chronological

phases. Basically, in these models we obtain a formula which, based on the examined features, e.g. as for the age – the number of remains and the size of the urn, and for the chronology – the thickness and colour of the walls, shows us (with a certain probability) which age group or chronology can be connected to the given archaeological feature.

Finally, it is worth mentioning that the importance of statistical analyses for archaeology will systematically grow, as constant influx of new materials takes place. Especially, there are some available statistical programs (e.g. Statistica, R and RStudio) which are really useful. However, it is necessary to remember that statistics is only an extremely helpful tool that we can and should use carefully.