



<i>Rodzaj dokumentu:</i>	Materiał dodatkowy
<i>Zagadnienie:</i>	Elementy analizy statystycznej w biologii
<i>Egzamin:</i>	Egzamin maturalny
<i>Przedmiot:</i>	Biologia
<i>Poziom egzaminu:</i>	Rozszerzony
<i>Adresaci dokumentu:</i>	Nauczyciele biologii Uczniowie szkół ponadpodstawowych
<i>Data publikacji dokumentu:</i>	7 października 2022 r.

Zespół redakcyjny:

dr Łukasz Banasiak (CKE)
Katarzyna Grudziąż
dr Takao Ishikawa (CKE)
Jadwiga Filipowska (CKE)

Recenzenci:

dr Tomasz Wyszomirski (UW)
dr Tomasz Karpowicz (recenzja językowa)
Grzegorz Zagulski (OKE w Gdańsku)
Aleksandra Szkutnik-Stokłosa (OKE w Jaworznie)
dr Inga Bator (OKE w Krakowie)
Dorota Mościcka (OKE w Łomży)

Materiał został opracowany przez Centralną Komisję Egzaminacyjną we współpracy z okręgowymi komisjami egzaminacyjnymi.

Centralna Komisja Egzaminacyjna
ul. Józefa Lewartowskiego 6, 00-190 Warszawa
tel. 22 536 65 00
sekretariat@cke.gov.pl

Okręgowa Komisja Egzaminacyjna w Gdańsku

ul. Na Stoku 49, 80-874 Gdańsk
tel. 58 320 55 90
komisja@oke.gda.pl

Okręgowa Komisja Egzaminacyjna w Jaworznie

ul. Adama Mickiewicza 4, 43-600 Jaworzno
tel. 32 616 33 99
oke@oke.jaworzno.pl

Okręgowa Komisja Egzaminacyjna w Krakowie

os. Szkolne 37, 31-978 Kraków
tel. 12 683 21 99
oke@oke.krakow.pl

Okręgowa Komisja Egzaminacyjna w Łomży

al. Legionów 9, 18-400 Łomża
tel. 86 473 71 20
sekretariat@oke.lomza.pl

Okręgowa Komisja Egzaminacyjna w Łodzi

ul. Ksawerego Praussa 4, 94-203 Łódź
tel. 42 634 91 33
sekretariat@lodz.oke.gov.pl

Okręgowa Komisja Egzaminacyjna w Poznaniu

ul. Gronowa 22, 61-655 Poznań
tel. 61 854 01 60
sekretariat@oke.poznan.pl

Okręgowa Komisja Egzaminacyjna w Warszawie

pl. Europejski 3, 00-844 Warszawa
tel. 22 457 03 35
info@oke.waw.pl

Okręgowa Komisja Egzaminacyjna we Wrocławiu

ul. Tadeusza Zielińskiego 57, 53-533 Wrocław
tel. 71 785 18 94
sekretariat@oke.wroc.pl

Spis treści

1. Wstęp	5
2. Statystyka w podstawie programowej kształcenia ogólnego	5
3. Podstawowe miary statystyczne	8
4. Opis wyników badań a wnioskowanie	17
5. Ćwiczenia	19
6. Statystyka w zadaniach z <i>Informatora</i> oraz arkusza pokazowego	26

1. Wstęp

Materiał dodatkowy ze statystyki na egzaminie maturalnym z biologii jest przeznaczony dla nauczycieli biologii szkół ponadpodstawowych oraz dla uczniów przygotowujących się do egzaminu maturalnego z biologii na poziomie rozszerzonym. W tym materiale wyjaśniono podstawowe pojęcia stosowane w statystyce oraz w analizie danych uzyskanych w wyniku przeprowadzenia obserwacji czy doświadczenia, a także zamieszczono ćwiczenia, które przybliżą znaczenie najważniejszych miar statystycznych w kontekście biologicznym. W tym materiale znajdują się również przykładowe zadania egzaminacyjne z *Informatora* i arkusza pokazowego z obszernymi komentarzami.

Ten materiał powinien ułatwić nauczycielom zilustrowanie podstawowych zagadnień dotyczących statystyki w biologii, a uczniom przygotowującym się do egzaminu maturalnego na poziomie rozszerzonym sprawdzenie swoich umiejętności z prostej analizy statystycznej na przykładowych zadaniach egzaminacyjnych.

2. Statystyka w podstawie programowej kształcenia ogólnego

Nauczanie elementów statystyki wpisano w podstawę programową matematyki zarówno szkoły podstawowej, jak i ponadpodstawowej, natomiast w pozostałych przedmiotach ogólnych uwzględniono wybrane umiejętności wykorzystania elementów statystyki do opisywania i interpretacji badanych zjawisk.

Poniżej przedstawiono wymagania podstawy programowej w zakresie nauczania statystyki.

A. Matematyka

Szkoła podstawowa

Klasy IV–VI

XIII. Elementy statystyki opisowej. Uczeń:

- 1) gromadzi i porządkuje dane;
- 2) odczytuje i interpretuje dane przedstawione w tekstach, tabelach, na diagramach i na wykresach, na przykład: wartości z wykresu, wartość największą, najmniejszą, opisuje przedstawione w tekstach, tabelach, na diagramach i na wykresach zjawiska przez określenie przebiegu zmiany wartości danych, na przykład z użyciem określenia „wartości rosną”, „wartości maleją”, „wartości są takie same” („przyjmowana wartość jest stała”).

Klasy VII i VIII

XIII. Odczytywanie danych i elementy statystyki opisowej. Uczeń:

- 1) interpretuje dane przedstawione za pomocą tabel, diagramów słupkowych i kołowych, wykresów, w tym także wykresów w układzie współrzędnych;
- 2) tworzy diagramy słupkowe i kołowe oraz wykresy liniowe na podstawie zebranych przez siebie danych lub danych pochodzących z różnych źródeł;
- 3) oblicza średnią arytmetyczną kilku liczb.

Szkoła ponadpodstawowa (zakres podstawowy)

1. Preambuła

„Matematyka jest nauką, która stanowi istotne wsparcie dla innych dziedzin, zwłaszcza dla nauk przyrodniczych [...]”.

2. Wymagania szczegółowe

Podstawa programowa z matematyki wprowadza z zakresu statystyki miary tendencji centralnej oraz rozproszenia.

XII. Rachunek prawdopodobieństwa i statystyka.

Zakres podstawowy. Uczeń:

- 3) oblicza **średnią arytmetyczną** i **średnią ważoną**, znajduje **medianę** i **dominantę**;
- 4) oblicza **odchylenie standardowe** zestawu danych (także w przypadku danych odpowiednio pogrupowanych), interpretuje ten parametr dla danych empirycznych.

B. Biologia

Podstawa programowa z biologii podkreśla rolę samodzielnego planowania i przeprowadzania obserwacji i doświadczeń. Warto zwrócić szczególną uwagę na opracowywanie, analizowanie i interpretowanie wyników badań oparte na prostych analizach statystycznych.

Szkoła ponadpodstawowa (zakres podstawowy)

Wymagania ogólne i szczegółowe w zakresie podstawowym odnoszące się do prostych analiz statystycznych są zawarte w przedstawionych poniżej wymaganiach z zakresu rozszerzonego.

Szkoła ponadpodstawowa (zakres rozszerzony)

1. Cele kształcenia – wymagania ogólne

II. Rozwijanie myślenia naukowego; doskonalenie umiejętności planowania i przeprowadzania obserwacji i doświadczeń oraz wnioskowania w oparciu o wyniki badań. Zdający:

- 1) określa problem badawczy, formułuje hipotezy, planuje i przeprowadza oraz dokumentuje **obserwacje** i proste **doświadczenia** biologiczne;
- 2) określa warunki doświadczenia, rozróżnia próbę kontrolną i badawczą;
- 3) opracowuje, analizuje i interpretuje wyniki badań w oparciu o **proste analizy statystyczne**;
- 5) [...] formułuje wnioski.

2. Wymagania szczegółowe

Wśród wymagań szczegółowych podstawy programowej z biologii w zakresie rozszerzonym znalazło się 30 obowiązkowych obserwacji lub doświadczeń, które powinny zostać zrealizowane podczas nauki w czteroletnim liceum i pięcioletnim technikum. Poniżej przedstawiono dwa przykłady wymagań odwołujących się bezpośrednio do obserwacji lub doświadczenia.

IX. Różnorodność roślin.

3. Odżywianie się roślin. Zdający:

- 5) [...] planuje i przeprowadza **doświadczenie** wykazujące wpływ temperatury i natężenia światła na intensywność fotosyntezy.

XVII. Ekologia.

2. Ekologia populacji. Zdający:

- 2) charakteryzuje populację, określając jej cechy (liczebność, zagęszczenie, struktura przestrzenna, wiekowa i płciowa); dokonuje **obserwacji** cech populacji wybranego gatunku.

3. Podstawowe miary statystyczne

Podczas prowadzenia obserwacji lub doświadczeń wykonuje się pomiary badanej cechy na wielu osobnikach. Aby zwięźle podsumować uzyskane wyniki, stosuje się różne miary tendencji centralnej i zmienności.

Miary tendencji centralnej – pojedyncze wartości stanowiące centrum zbioru wyników

- średnia
- mediana
- dominanta

Miary zmienności – zwane też miarami rozproszenia, określają, jak bardzo zróżnicowane są wyniki

- rozstęp
- odchylenie standardowe

Poniżej przedstawiono krótką charakterystykę poszczególnych miar statystycznych.

Uwaga:

Wzory matematyczne i przykłady rachunkowe podano w celu lepszego zrozumienia omawianych miar statystycznych. Na egzaminie maturalnym z biologii nie przewiduje się ich obliczania, ale zakłada się wyłącznie interpretację danych przy użyciu podstawowych miar statystycznych.

Średnia arytmetyczna – suma wszystkich pomiarów podzielona przez liczbę tych pomiarów:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n}$$

\bar{x} – średnia arytmetyczna

$x_1, x_2, \dots, x_i, \dots, x_n$ – kolejne wyniki pomiarów

n – liczba pomiarów

Przykład 1.

Na początek rozpatrzmy prosty przykład. W badaniu otrzymano następujące wielkości lęgu¹ kanadyjskiego gatunku gęsi (*Branta canadensis*): 8, 6, 5, 7, 5, 9, 3, 5.

$$\bar{x} = \frac{8 + 6 + 5 + 7 + 5 + 9 + 3 + 5}{8} = 6$$

Średnia wielkość lęgu w badaniu wyniosła 6 jaj. Średnia arytmetyczna zawsze leży na osi liczbowej pomiędzy wartościami najmniejszą i największą: $x_{min} < \bar{x} < x_{max}$, czyli w tym przypadku $3 < \bar{x} < 9$.

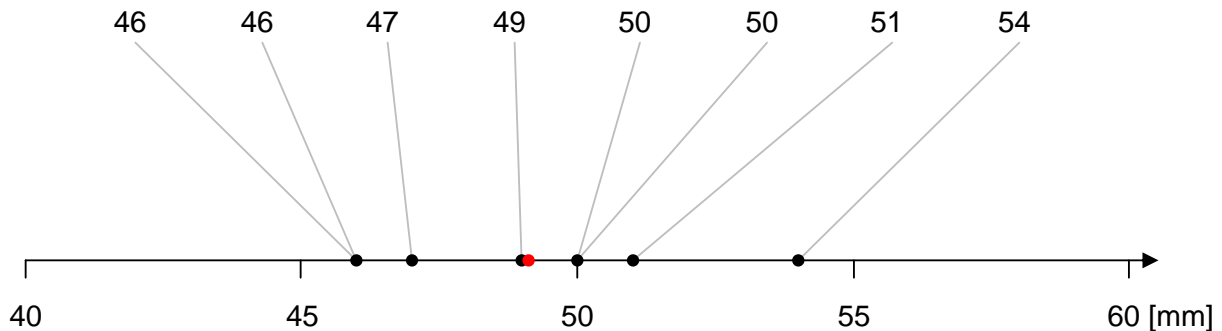
¹ Wielkość lęgu to liczba jaj wysiadywanych jednocześnie.

Przykład 2.

Pomiary długości działek zewnętrznego okółka okwiatu kosańca szczecinkowego (*Iris setosa*) dały następujące wyniki w milimetrach: 51, 49, 47, 46, 50, 54, 46, 50. **Średnia arytmetyczna** pomiarów wynosi:

$$\bar{x} = \frac{51 + 49 + 47 + 46 + 50 + 54 + 46 + 50}{8} \text{ mm} = \mathbf{49,125} \text{ mm}$$

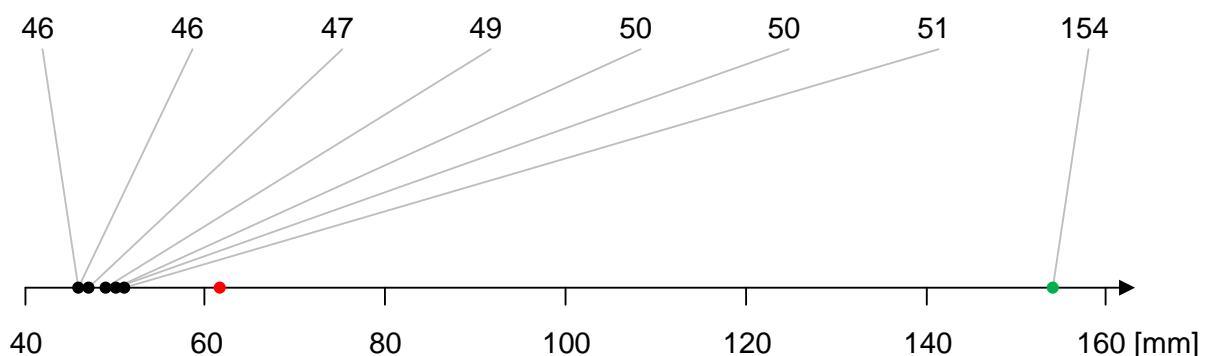
Poniżej przedstawiono **zbiór danych** oraz **średnią arytmetyczną** na osi liczbowej:



Średnia arytmetyczna jest bardzo wrażliwa na tzw. obserwacje odstające – wartości pomiarów znacznie różniące się od typowych wyników². Jeden pomiar znacznie odbiegający od pozostałych może w dużym stopniu zawyżyć lub zaniżyć wartość średnią. Przykładowo: **omyłkowe dopisanie jedynek** przed szóstym wynikiem pomiaru zwiększy wartość średnią o 12,5 mm:

$$\bar{x} = \frac{51 + 49 + 47 + 46 + 50 + \mathbf{154} + 46 + 50}{9} \text{ mm} = \mathbf{61,625} \text{ mm}$$

Średnia arytmetyczna wynosząca 61,625 mm jest większa od każdego z siedmiu niezaburzonych wyników pomiarów, a więc nie jest dobrą miarą tendencji centralnej w tym przypadku – **większość wyników** leży na osi liczbowej na lewo od **średniej**:



Obserwacje odstające mogą być spowodowane naturalnymi procesami, np. pomiar temperatury ciała człowieka podczas infekcji bakteryjnej będzie zawyżony. Obserwacje odstające najczęściej świadczą jednak o błędach popełnionych podczas wykonywania

² Używa się też terminów obserwacja skrajna lub ekstremalna.

pomiarów lub przygotowywania dokumentacji. Do takich błędów zalicza się m.in. pomylenie jednostek, omyłkowe przesunięcie przecinka dziesiętnego, nieprawidłowe użycie sprzętu badawczego oraz niewłaściwe oznaczenie badanego organizmu. Jeżeli nie można poprawić błędnych wyników pomiarów, to należy je bezwzględnie usunąć ze zbioru danych przed obliczeniem wartości średniej.

Podczas dołączania kolejnych wyników doświadczenia lub obserwacji do zbioru danych średnia arytmetyczna może się zwiększyć lub zmniejszyć. Te wahania są jednak zasadniczo coraz mniejsze wraz ze wzrostem liczby wykonanych pomiarów. W przypadku doświadczeń lub obserwacji z odpowiednio dużą liczbą pomiarów można w pewnym momencie zauważyć, że kontynuowanie badań ma nieznaczny wpływ na podsumowanie wyników i dalsze wnioskowanie.

Średnia arytmetyczna ważona uwzględnia różne wagi (znaczenie) poszczególnych wyników pomiarów i jest określona wzorem:

$$\bar{x} = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i} = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n}$$

\bar{x} – średnia arytmetyczna

$x_1, x_2, \dots, x_i, \dots, x_n$ – kolejne wyniki pomiarów

$w_1, w_2, \dots, w_i, \dots, w_n$ – wagi przyporządkowane kolejnym pomiarom

n – liczba pomiarów

Przykład 1.

Średnia arytmetyczna ważona jest wykorzystywana, jeżeli dane zostały pogrupowane. W poniższej tabeli przedstawiono podsumowanie wyników obserwacji 1 046 porodów marmozety zwyczajnej (*Callithrix jacchus*)³, należącej do małp szerokonosych. „d

Wielkość miotu (wynik pomiaru)	Liczba porodów (waga)
1	52
2	649
3	335
4	10
Suma	1046

Aby obliczyć średnią wielkość miotu, należy skorzystać z powyższego wzoru i zastosować liczbę porodów o danej wielkości jako wagi:

$$\bar{x} = \frac{52 \cdot 1 + 649 \cdot 2 + 335 \cdot 3 + 10 \cdot 4}{52 + 649 + 335 + 10} = \frac{2395}{1046} \approx 2,29 \text{ (wartość przybliżona)}$$

³ C.P. Windle, H.F. Baker, R.M. Ridley, A.K. Oerke, R.D. Martin, *Unreared litters and parental reduction of litter size in the common marmoset (Callithrix jacchus)*, „Journal of Medical Primatology” 28, 1999.

Dzięki wykorzystaniu wzoru na średnią ważoną w liczniku znalazła się całkowita liczba noworodków, a w mianowniku – łączna liczba porodów.

Nieuwzględnienie wag doprowadziłoby w tym przypadku do wyniku znajdującego się na osi liczbowej dokładnie w połowie odległości między wartościami najmniejszą i największą:

$$\bar{x} \neq \frac{1 + 2 + 3 + 4}{4} = 2,5$$

Obliczenia prowadzone w ten sposób nie uwzględniają tego, że porody o różnej wielkości zdarzają się z różną częstością. Błędny wynik 2,5 sugeruje, że porody małe (1 lub 2 noworodki) oraz duże (3 lub 4 noworodki) są tak samo prawdopodobne, kiedy w rzeczywistości mały poród jest bardziej prawdopodobny od dużego.

Średnia arytmetyczna jest wartością abstrakcyjną – średnia wielkość miotu marmozety zwyczajnej wynosi w przybliżeniu 2,29, choć podczas porodu może na świat przyjść tylko jedno, dwa, trzy albo cztery młode.

Przykład 2.

W poniższej tabeli przedstawiono wyniki monitoringu populacji wilka w kompleksach leśnych północno-wschodniej Polski w październiku 2018 r.⁴

Monitorowany obszar	Zagęszczenie populacji, $\frac{\text{os.}}{100 \text{ km}^2}$	Monitorowana powierzchnia, km^2
Puszcza Augustowska	3,0	1 583
Puszcza Białowieska	2,6	769
Puszcza Knyszyńska	2,0	1 793

Aby obliczyć średnie zagęszczenie wilka w lasach północno-wschodniej Polski, należy wziąć pod uwagę wielkości powierzchni monitorowanych obszarów potraktowane jako wagi poszczególnych obserwacji:

$$\bar{x} = \frac{3,0 \frac{\text{os.}}{100 \text{ km}^2} \cdot 1\,583 \text{ km}^2 + 2,6 \frac{\text{os.}}{100 \text{ km}^2} \cdot 769 \text{ km}^2 + 2,0 \frac{\text{os.}}{100 \text{ km}^2} \cdot 1\,793 \text{ km}^2}{1\,583 \text{ km}^2 + 769 \text{ km}^2 + 1\,793 \text{ km}^2}$$

$$\bar{x} = \frac{103,344 \text{ os.}}{4145 \text{ km}^2} = \frac{103,344}{41,45} \frac{\text{os.}}{100 \text{ km}^2} \approx 2,49 \frac{\text{os.}}{100 \text{ km}^2}$$

Podczas obliczeń należy uwzględnić jednostkę zagęszczenia ($\frac{\text{os.}}{100 \text{ km}^2}$) i jednostkę pola powierzchni (km^2). Końcowy wynik przedstawiono z dokładnością do dwóch miejsc po przecinku.

⁴ <https://www.gios.gov.pl/pl/poiis-monitoring-wilka-i-rysia>

Dominanta (moda, modalna) – najczęstsza wartość pomiaru w danym zbiorze danych.

Przykład 1.

Wróćmy do naszego prostego przykładu wielkości lęgu u gęsi: 8, 6, 5, 7, 5, 9, 3, 5. Sprawdzamy, która liczba występuje najczęściej: 8, 6, **5**, 7, **5**, 9, 3, **5**. Dominantą jest tu liczba **5** (ona jako jedyna powtórzyła się trzykrotnie, pozostałe liczby wystąpiły tylko raz). Na tej podstawie możemy stwierdzić, że podczas prowadzenia dalszych badań najczęściej będziemy obserwować w gniazdach 5 lub zbliżoną liczbę jaj.

Przykład 2.

Wróćmy teraz do naszego przykładu dotyczącego wielkości miotu marmozety. Przeanalizujmy tabelę. Znajdźmy wynik pomiaru, który pojawia się najczęściej.

Wielkość miotu (wynik pomiaru)	Liczba porodów (waga)
1	52
2	649
3	335
4	10
Suma	1046

W przypadku wielkości miotu marmozety dominanta wynosi **2**, ponieważ z powyższej tabeli można odczytać, że najczęściej rodzą się bliźnięta, a więc jest to najbardziej prawdopodobna wielkość miotu przy każdym kolejnym porodzie. Innymi słowy spodziewamy się przyjścia na świat dwójki młodych, ale z mniejszym prawdopodobieństwem może się zdarzyć także poród mniejszy albo większy.

Dla małych zbiorów danych może się okazać, że więcej niż jedna wartość pomiaru występuje z największą częstością – taki zbiór danych ma więcej niż jedną dominantę. W przypadku pomiarów długości działek okwiatu u kosańca dominantami są 46 mm i 50 mm – obydwa wyniki pomiarów występują dwukrotnie w tym zbiorze danych:

51, 49, 47, **46**, **50**, 54, **46**, **50**

W małych próbach dominanta może się zmieniać z próby na próbę. Gdybyśmy powtórzyli pomiary na innych ośmiu okazach kosańca, to z dużym prawdopodobieństwem taki zbiór danych charakteryzowałby się inną wartością modalną.

Jeżeli wartości pomiarów się nie powtarzają (każdy pomiar ma inną wartość), nie da się wyznaczyć dominanty. Jest to typowe dla cech ciągłych mierzonych za pomocą precyzyjnych urządzeń pomiarowych. Gdybyśmy wykonali pomiary długości działek kosańca za pomocą suwmiarki z dokładnością do 0,1 mm, to najprawdopodobniej każdy z ośmiu pomiarów miałby inną wartość.

Dominanta jest zatem użyteczną miarą tendencji centralnej jedynie dla dużych prób i cech o charakterze nieciągłym. Z myślą o cechach ciągłych można stosować grupowanie wyników w przedziały i wyznaczyć przedział modalny – przedział, w którym znalazło się najwięcej wyników.

Mediana (wartość środkowa) – taka wartość badanej cechy, która dzieli zbiór danych na dwie równoliczne części: połowa obserwacji ma wartości mniejsze lub równe medianie, a druga połowa – równe lub większe od mediany.

Przykład 1. – parzysta liczba wyników

Aby wyznaczyć medianę, najwygodniej zacząć od uporządkowania zbioru danych od wartości najmniejszych do największych. To krok pierwszy. Przywołajmy nasz prosty przykład wielkości lęgu u gęsi: 8, 6, 5, 7, 5, 9, 3, 5. Po uporządkowaniu otrzymujemy taki zestaw danych: 3, 5, 5, 5, 6, 7, 8, 9.

W następnym kroku liczymy, ile mamy danych – jest 8 liczb. Dla zbiorów danych o parzystej liczbie wyników medianę oblicza się jako średnią z **dwóch środkowych wartości**.

$$3, 5, 5, \mathbf{5, 6}, 7, 8, 9$$

W ostatnim kroku musimy zatem dodać do siebie dwie sąsiadujące liczby – czwartą i piątą z kolei⁵ – i podzielić na dwa. W naszym przypadku mediana wynosi $\frac{5+6}{2} = 5,5$. Średnia arytmetyczna z dwóch środkowych wartości stanowi jednocześnie środek przedziału wyznaczonego przez te dwie wartości.

Dzięki obliczeniu mediany możemy powiedzieć, że połowa gęsi złożyła co najwyżej 5 jaj, a druga połowa – 6 lub więcej jaj. Możemy też stwierdzić, że gęś, która złożyła 8 jaj, znajduje się w tej połowie próby, która miała wyższe wyniki lęgu. Badana próba jest stosunkowo mała, ale w dużym przybliżeniu możemy stwierdzić, że przy prowadzeniu dalszych badań z takim samym prawdopodobieństwem będziemy odnajdywać lęgi małe – liczące do 5 jaj, oraz lęgi duże – co najmniej 6 jaj.

Przykład 2. – nieparzysta liczba wyników

W przypadku zbiorów danych o nieparzystej liczbie wyników medianę stanowi **pojedyncza wartość środkowa** uporządkowanego zbioru, a więc jest równa wartości jednego z uzyskanych wyników. Poszukajmy mediany w następującym zestawie 11 danych z obserwacji lęgowych wykonanych w innej lokalizacji:

$$2, 3, 4, 5, 5, \mathbf{8}, 9, 9, 11, 11, 12$$

Środkowa – szоста z kolei⁶ – liczba to 8. Mediana wynosi zatem 8.

⁵ Pozycję dwóch środkowych liczb w szeregu uporządkowanym o parzystej liczbie pomiarów można określić ze wzorów: $\frac{n}{2}$ oraz $\frac{n}{2} + 1$, gdzie n to liczba pomiarów.

⁶ Pozycję środkowej liczby w szeregu uporządkowanym o nieparzystej liczbie pomiarów można określić ze wzoru $\frac{n+1}{2}$, gdzie n to liczba pomiarów.

Przykład 3. – wpływ obserwacji odstających

Uporządkowanie zbioru danych z wynikami pomiarów długości działek okwiatu kosańca prowadzi do następującego wyniku w milimetrach:

46, 46, 47, **49, 50**, 50, 51, 54

Liczba wyników jest parzysta, a **dwie środkowe wartości** to 49 mm i 50 mm (odpowiednio: czwarta i piąta wartość po uporządkowaniu). Mediana jest zatem równa **49,5 mm**, ponieważ $\frac{49+50}{2}$ mm = 49,5 mm.

Mediana jest nieczuła na obserwacje odstające. Przykładowo: **omyłkowe dopisanie jedynki** do najwyższego wyniku pomiaru nie ma żadnego wpływu na wartość mediany w dalszym ciągu równej 49,5 mm:

46, 46, 47, **49, 50**, 50, 51, **154**

Warto przypomnieć, że po takim samym zaburzeniu wyników średnia arytmetyczna wzrosła o 12,5 mm.

Rozstęp – różnica między maksymalnym (największym) i minimalnym (najmniejszym) wynikiem pomiaru. Znalezienie maksimum i minimum jest łatwiejsze po uporządkowaniu zbioru danych, ale nie okazuje się to konieczne.

Przykład 1.

W naszym prostym przykładzie wielkości lęgu u gęsi: 8, 6, 5, 7, 5, **9, 3**, 5, rozstęp wynosi $9 - 3 = 6$.

Przykład 2.

W zbiorze wyników pomiarów długości działek kosańca wartość maksymalna wynosi 54 mm, a wartość minimalna jest równa 46 mm. Rozstęp wynosi zatem 8 mm, ponieważ $54 \text{ mm} - 46 \text{ mm} = 8 \text{ mm}$.

Rozstęp, podobnie jak średnia arytmetyczna, jest wrażliwy na obserwacje odstające, które zawsze zwiększają jego wartość. Podczas dołączania do zbioru danych kolejnych wyników, rozstęp nie może maleć, tzn. pozostaje on albo na poprzednim poziomie, jeżeli kolejny wynik pomiaru mieści się w przedziale wyznaczonym przez wartości minimalną i maksymalną, albo rośnie, jeżeli kolejny wynik pomiaru wykracza poza dotychczasowy zbiór danych. Z tego powodu rozstęp jest rzadko używaną miarą rozproszenia, a w praktyce dużo częściej stosuje się odchylenie standardowe.

Odchylenie standardowe – to miara skupienia wyników wokół wartości średniej określona następującym wzorem:

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

s – odchylenie standardowe

$x_1, x_2, \dots, x_i, \dots, x_n$ – kolejne wyniki pomiarów

\bar{x} – średnia arytmetyczna

n – liczba pomiarów

Wartość odchylenia standardowego wyznacza typowe wartości badanej cechy zawierające się w przedziale $\bar{x} \pm s$. Jednocześnie przyjmuje się, że wystąpienie wartości cechy spoza przedziału $\bar{x} \pm 3s$ jest mało prawdopodobne.

Jeżeli zmienność badanej cechy jest dobrze przybliżona rozkładem normalnym⁷ i pobrano próbę liczącą co najmniej kilkanaście obserwacji, to można przewidzieć, że około 2/3 przyszłych wartości pomiarów będzie mieściło się w przedziale $\bar{x} \pm s$, a około 99% – w przedziale $\bar{x} \pm 3s$.

Odchylenie standardowe, podobnie jak średnia arytmetyczna, ma tę samą jednostkę, co mierzona cecha organizmu lub populacji, i jest wrażliwe na obserwacje odstające.

Przykład

Podczas eksperymentu badano siłę kiełkowania nasion. W każdej z czterech doniczek wysiano 100 nasion. Na koniec eksperymentu policzono liczbę skiełkowanych nasion w każdej z doniczek i otrzymano następujące wyniki: 66, 72, 51, 78.

Średnia liczba skiełkowanych nasion wynosi:

$$\bar{x} = \frac{66 + 72 + 51 + 78}{4} = 66,75$$

Odchylenie standardowe liczby skiełkowanych nasion jest równe:

$$s = \sqrt{\frac{(66 - 66,75)^2 + (72 - 66,75)^2 + (51 - 66,75)^2 + (78 - 66,75)^2}{4}}$$

$$s = \sqrt{\frac{(-0,75)^2 + (5,25)^2 + (-15,75)^2 + (11,25)^2}{4}}$$

$$s = \sqrt{\frac{0,5625 + 27,5625 + 248,0625 + 126,5625}{4}}$$

⁷ Rozkład normalny ma kształt krzywej dzwonowej. Wiele cech organizmów ma rozkład zbliżony do normalnego. Dobrym przykładem jest rozkład wzrostu kobiet i mężczyzn w populacji ludzkiej.

$$s = \sqrt{\frac{402,75}{4}} = \sqrt{100,6875} \approx 10,03432$$

Do tego samego wyniku można dojść z wykorzystaniem alternatywnego wzoru na odchylenie standardowe:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} = \sqrt{\overline{x^2} - (\bar{x})^2}$$

s – odchylenie standardowe

$x_1, x_2, \dots, x_i, \dots, x_n$ – kolejne wyniki pomiarów

\bar{x} – średnia arytmetyczna wyników pomiarów

\bar{x}^2 – kwadrat średniej arytmetycznej

$\overline{x^2}$ – średnia arytmetyczna kwadratów wyników pomiarów

n – liczba pomiarów

Po podstawieniu do wzoru otrzymujemy:

$$s = \sqrt{\frac{66^2 + 72^2 + 51^2 + 78^2}{4} - 66,75^2}$$

$$s = \sqrt{\frac{4356 + 5184 + 2601 + 6084}{4} - 4455,5625}$$

$$s = \sqrt{\frac{18225}{4} - 4455,5625}$$

$$s = \sqrt{4556,25 - 4455,5625}$$

$$s = \sqrt{100,6875} \approx 10,03432$$

Wyznamy teraz przedział $\bar{x} \pm s$ z dokładnością do drugiego miejsca po przecinku:

$$(56,72; 76,78)$$

Pamiętamy, że w omawianym eksperymencie w każdej doniczce było 100 nasion, a więc tak wyznaczony przedział obejmuje typowe wartości siły kiełkowania badanych nasion wyrażone w procentach. Do uzyskanych wyników należy jednak podchodzić z dużą rezerwą, ponieważ wartości średniej arytmetycznej i odchylenia standardowego zostały obliczone tylko dla czterech powtórzeń eksperymentu. Takie uproszczenie zostało wprowadzone w celu ułatwienia rachunków, ale podczas prowadzenia badań powinno się wykonywać, jeżeli to możliwe, większą liczbę powtórzeń.

4. Opis wyników badań a wnioskowanie

Opis wyników badań za pomocą miar tendencji centralnej i rozproszenia przekłada się na łatwiejsze wyciągnięcie wniosków z przeprowadzonych badań. Wnioskowanie polega na udzieleniu odpowiedzi na postawione pytanie badawcze lub zweryfikowaniu sformułowanych hipotez na podstawie analizy uzyskanych wyników doświadczenia lub obserwacji. Taka analiza często polega na porównaniu uzyskanych wyników między badanymi grupami.

Poniżej przedstawiono plan obserwacji uwzględniający: problem badawczy, pomiary badanych cech, opis wyników oraz analizę możliwych wyników wraz z wnioskowaniem.

Przykład:

Zależności pokarmowe między motylem bielinkiem kapustnikiem a ostróżeczką polną.

Dorośla forma motyla bielinka kapustnika (*Pieris brassicae*) żywi się nektarem kwiatów. Jest to możliwe dzięki aparatowi gębowemu typu ssącego w postaci odpowiednio długiej rurki. Ostróżeczka polna (*Consolida regalis*) produkuje nektar w miodnikach znajdujących się na dnie wyciągniętej w ostrogę górnej działki kielicha.

Problem badawczy: Czy populacja bielinka może korzystać z nektaru produkowanego przez ostróżeczkę?

Mierzone cechy: długość ssawki 30 osobników bielinka oraz długość ostrogi 30 okazów ostróżeczki.

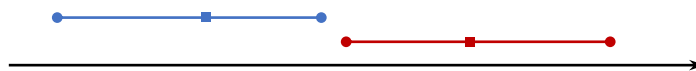
Opis wyników: przedstawienie zmienności obydwu cech na osi liczbowej w postaci minimum, mediany i maksimum.

Analiza wyników i wnioskowanie: porównanie rozkładów badanych cech i udzielenie odpowiedzi na postawione pytanie badawcze⁸.

Poniżej przedstawiono cztery różne możliwe wyniki obserwacji, analizę tych wyników i wnioski wyciągnięte na podstawie analizy. Kolorem **czzerwonym** przedstawiono wyniki dla długości ssawki, a kolorem **niebieskim** – dla długości ostrogi kwiatu. Koła oznaczają minimum i maksimum, a kwadraty – medianę⁹.

Przykład 1.

Wyniki:



Analiza: Rozkłady badanych cech w próbie są rozłączne. Ssawki motyli w próbie były zawsze dłuższe od ostróg kwiatów. Badania na większej próbie mogłyby jednak wykazać, że zmienność badanych cech w populacji częściowo na siebie nachodzi.

Wniosek: Dostęp populacji motyli *P. brassicae* do nektaru *C. regalis* jest nieograniczony lub ograniczony w niewielkim stopniu.

⁸ W tym prostym przykładzie podczas analizy i wnioskowania zakładamy, że na dostępność nektaru nie mają wpływu inne zmienne nieujęte w tych badaniach, takie jak lepkość nektaru.

⁹ Oś liczbową nie została wyskalowana, ponieważ rozpatrujemy cztery hipotetyczne scenariusze, w których ważne jest jedynie wzajemne położenie na osi liczbowej uwzględnionych w analizie statystyk.

Przykład 2.

Wyniki:



Analiza: Rozkłady badanych cech w próbie częściowo się pokrywają. Najkrótsze ssawki motyli były krótsze od najdłuższych ostróg kwiatów.

Wniosek: Większość motyli w populacji *P. brassicae* ma nieograniczony dostęp do nektaru *C. regalis*, ale niewielka część z nich nie może żerować na największych okazach kwiatów.

Przykład 3.

Wyniki:

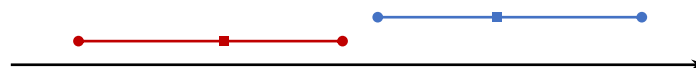


Analiza: Rozkłady badanych cech w próbie częściowo się pokrywają. Większość motyli miała ssawki krótsze od najkrótszych ostróg kwiatów. W przybliżeniu połowa kwiatów miała ostrogi dłuższe od najdłuższych ssawek motyli.

Wniosek: Tylko subpopulacja motyli *P. brassicae* o najdłuższych ssawkach ma dostęp do części nektaru produkowanego przez *C. regalis*.

Przykład 4.

Wyniki:



Analiza: Rozkłady badanych cech w próbie są rozłączne. Ssawki motyli w próbie były zawsze krótsze niż ostrogi kwiatów. Badania na większej próbie mogłyby jednak wykazać, że zmienność badanych cech w populacji częściowo na siebie nachodzi.

Wniosek: W populacji *P. brassicae* motyle nie mają w ogóle dostępu do nektaru *C. regalis* lub ten dostęp jest ograniczony w znacznym stopniu – tylko niektóre motyle mogą mieć dostęp do nektaru.

Cztery powyższe przykłady pokazują, że prawidłowe wnioski można wyciągnąć jedynie na podstawie uważnej analizy wyników, polegającej na porównaniu podstawowych miar statystycznych. Wnioski zawsze muszą odnosić się do postawionego pytania badawczego.

5. Ćwiczenia

Ćwiczenie 1.

Czy na podstawie pomiarów wzrostu uczniów w jednej klasie można przewidzieć wyniki pomiarów wzrostu uczniów z innych klas?

- Zmierz wzrost co najmniej kilkunastu chłopców lub dziewcząt w klasie. Na podstawie uzyskanych wyników pomiarów oblicz wartości średniej i odchylenia standardowego.**

Średnia arytmetyczna: $\bar{x} = \dots\dots\dots$

Odchylenie standardowe: $s = \dots\dots\dots$

- Podaj granice przedziału $\bar{x} \pm s$.**

Dolna granica przedziału: $\dots\dots\dots$

Górna granica przedziału: $\dots\dots\dots$

- Zmierz wzrost co najmniej kilkunastu osób wybranej wcześniej płci uczęszczających do innych klas. Określ, jaka część nowych wyników pomiarów zawiera się w przedziale $\bar{x} \pm s$ wyliczonym na podstawie pierwszej serii pomiarów.**

Odsetek nowych wyników znajdujących się w wyznaczonym przedziale: $\dots\dots\dots$ %

Komentarz

To ćwiczenie wymaga zaangażowania uczniów z różnych klas. Ze względu na to, że intensywny proces wzrostu chłopców trwa aż do 17. roku życia, ćwiczenie powinno być wykonywane w ramach jednego rocznika (intensywny wzrost dziewcząt kończy się wcześniej – około 13. roku życia, a więc przed rozpoczęciem nauki w szkole ponadpodstawowej). Pomiarów wzrostu mogą być przeprowadzane na bieżąco podczas kolejnych etapów ćwiczenia, ale możliwe jest także ze strony nauczyciela udostępnienie uczniom wcześniej wykonanych pomiarów.

W ćwiczeniu 1.1. i 1.2. podczas obliczania średniej arytmetycznej, odchylenia standardowego i granic przedziału należy pamiętać o podaniu odpowiedniej jednostki długości.

Ze względu na to, że zmienność wzrostu w populacji ludzkiej jest dobrze opisana rozkładem normalnym, w granicach przedziału $\bar{x} \pm s$ należy spodziewać się około 2/3 przyszłych obserwacji. Jakość wyznaczonego przedziału zależy przede wszystkim od liczby pomiarów. Im będzie ich więcej, tym wyznaczony przedział będzie lepszej jakości, tzn. odsetek przewidzianych wyników pomiarów w innych klasach będzie bliższy 2/3.

Po wykonaniu przez uczniów z pomocą nauczyciela ćwiczenia dla jednej płci zalecane jest samodzielne wykonanie ćwiczenia dla drugiej płci.

Ćwiczenie 2.

Z kolekcji Muzeum Brytyjskiego wybrano losowo po dziesięć czaszek samic i samców szakala złocistego (*Canis aureus*). Poniżej przedstawiono wyniki pomiarów długości szczęki wyrażone w milimetrach.

Samce:

120 107 110 116 114 111 113 117 114 112

Samice:

110 111 107 108 110 105 107 106 111 111

1. Uzupełnij tabelę – znajdź minimum, maksimum oraz medianę wyników pomiarów długości szczęki szakali w podziale na płeć zwierzęcia. Wyniki zapisz w milimetrach.

	Minimum	Mediana	Maksimum
Samce			
Samice			

2. Zinterpretuj uzyskane wyniki badań – określ, czy płeć szakala złocistego można bezbłędnie rozpoznać wyłącznie na podstawie wyników pomiarów długości szczęki. Odpowiedź uzasadnij.

.....

.....

.....

.....

Przykładowe rozwiązanie

1.

Zestaw danych po uporządkowaniu z zaznaczonymi na żółto wartościami środkowymi:

Samce:

107 110 111 112 **113 114** 114 116 117 120

Samice:

105 106 107 107 **108 110** 110 111 111 111

	Minimum	Mediana	Maksimum
Samce	107	$(113 + 114)/2 = 113,5$	120
Samice	105	$(108 + 110)/2 = 109,0$	111

2.

- Nie, ponieważ mediany długości szczęki są przesunięte tylko o 4,5 mm, a rozstęp długości szczęki u samców wynosi nawet kilkanaście milimetrów.
- Nie jest to możliwe, ponieważ część samic ma dłuższe szczęki od niektórych samców, choć z reguły to szczęki samców są dłuższe.

Ćwiczenie 3.

W ciągu miesiąca zawodnik wykonał cztery treningi biegowe na dystansie siedmiu kilometrów. W każdym dniu treningu zostało zmierzone tętno spoczynkowe oraz tętno wysiłkowe sportowca. Wyniki pomiarów w postaci liczby uderzeń serca na minutę przedstawiono w poniższej tabeli.

Data	Tętno spoczynkowe	Tętno wysiłkowe
5 marca	58	160
6 marca	59	162
13 marca	66	174
29 marca	64	161

Ćwiczenie 3.1.

Uzupełnij tabelę – na podstawie przedstawionych wyników pomiarów oblicz wartość średnią oraz odchylenie standardowe tętna spoczynkowego oraz tętna wysiłkowego. Odpowiedzi zapisz w liczbie uderzeń serca na minutę z dokładnością do drugiego miejsca po przecinku.

	Tętno spoczynkowe	Tętno wysiłkowe
Średnia arytmetyczna		
Odchylenie standardowe		

Zasady oceniania

2 pkt – za poprawne obliczenie wartości średnich i odchyleń standardowych.

1 pkt – za poprawne obliczenie wartości średnich.

0 pkt – za odpowiedź niespełniającą wymagań za 1 pkt albo za brak odpowiedzi.

Rozwiązanie

$$\bar{x}_s = \frac{58 + 59 + 66 + 64}{4} = \frac{247}{4} = 61,75$$

$$s_s = \sqrt{\frac{(58 - 61,75)^2 + (59 - 61,75)^2 + (66 - 61,75)^2 + (64 - 61,75)^2}{4}}$$

$$s_s = \sqrt{\frac{14,0625 + 7,5625 + 18,0625 + 5,0625}{4}} = \sqrt{\frac{44,75}{4}} = \sqrt{11,1875} \approx 3,34$$

$$\bar{x}_w = \frac{160 + 162 + 174 + 161}{4} = \frac{657}{4} = 164,25$$

$$s_w = \sqrt{\frac{(160 - 164,25)^2 + (162 - 164,25)^2 + (174 - 164,25)^2 + (161 - 164,25)^2}{4}}$$

$$s_w = \sqrt{\frac{18,0625 + 5,0625 + 95,0625 + 10,5625}{4}} = \sqrt{\frac{128,75}{4}} = \sqrt{32,1875} \approx 5,67$$

	Tętno spoczynkowe	Tętno wysiłkowe
Średnia arytmetyczna	61,75 (wartość dokładna)	164,25 (wartość dokładna)
Odchylenie standardowe	3,34 (wartość przybliżona)	5,67 (wartość przybliżona)

Komentarz

Podczas wykonywania obliczeń należy stosować jak największą dokładność. Większość kalkulatorów prostych wykonuje obliczenia z dokładnością do 7–9 miejsc po przecinku, ale w tym przykładzie najdłuższe rozwinięcia dziesiętne wyników częściowych są czterocyfrowe. Błędem jest zaokrąglanie wyników częściowych. Zaokrąglić do wymaganej precyzji należy tylko końcowy wynik – wynik pierwiastkowania.

Ćwiczenie 3.2.

Określ, który z mierzonych parametrów – tętno spoczynkowe czy tętno wysiłkowe – wykazywał większą zmienność między treningami. Odpowiedź uzasadnij, porównując wartości liczbowe rozstępów obydwu parametrów.

.....

.....

.....

.....

Przykładowe rozwiązanie

Różnica między wartościami minimalną i maksymalną pomiaru wynosi tylko 8 dla tętna spoczynkowego, a dla tętna wysiłkowego – aż 14 uderzeń serca na minutę. Zatem to tętno wysiłkowe miało większą zmienność.

Ćwiczenie 4.

Nadciśnienie tętnicze jest przyczyną wielu chorób układu krążenia. Nadciśnienie tętnicze rozpoznaje się, kiedy ciśnienie skurczowe przekracza 140 mmHg, a ciśnienie rozkurczowe – 90 mmHg. W zależności od stopnia przekroczenia normy rozpoznaje się różne stopnie nadciśnienia tętniczego.

Aby ocenić skuteczność terapii nadciśnienia tętniczego, piętnastu pacjentom dwukrotnie zmierzono ciśnienie krwi: bezpośrednio przed podaniem oraz dwie godziny po podaniu dawki 25 mg kaptoprylu. W poniższej tabeli przedstawiono wyniki pomiarów ciśnienia rozkurczowego.

L.p.	Ciśnienie tętnicze rozkurczowe, mmHg		Zmiana (efekt terapeutyczny)
	Przed podaniem kaptoprylu	Dwie godziny po podaniu kaptoprylu	
1.	130	125	-5
2.	122	121	-1
3.	124	121	-3
4.	104	106	+2
5.	112	101	-11
6.	101	85	-16
7.	121	98	-23
8.	124	105	-19
9.	115	103	-12
10.	102	98	-4
11.	98	90	-8
12.	119	98	-21
13.	106	110	+4
14.	107	103	-4
15.	100	82	-18

Kaptopryl stosowany jest w dawce maksymalnej 150 mg na dobę w trzech dawkach podzielonych. Nie należy przekraczać maksymalnej dawki dobowej, ponieważ nie zwiększa to skuteczności działania leku, a może zaszkodzić życiu lub zdrowiu pacjenta.

- Oblicz średnią zmianę ciśnienia krwi po podaniu kaptoprylu oraz odchylenie standardowe tej zmiany. Odpowiedzi zapisz z dokładnością do drugiego miejsca po przecinku.**

Uwaga: ze względu na dużą ilość danych zalecane jest prowadzenie obliczeń z wykorzystaniem komputera, np. w arkuszu kalkulacyjnym.

Średnia arytmetyczna: \bar{x} = mmHg

Odchylenie standardowe: s = mmHg

2. Uzupełnij poniższe zdania tak, aby powstała poprawna interpretacja uzyskanych wyników badań. Podkreśl właściwe określenie w każdym nawiasie.

Wśród badanych pacjentów efekt terapeutyczny podania dawki 25 mg kaptoprylu był (zbliżony / zróżnicowany). Z tego powodu różni pacjenci o tym samym stopniu nadciśnienia powinni dostawać (taką samą / indywidualnie ustalaną) dawkę leku, aby ich ciśnienie tętnicze krwi utrzymywało się na prawidłowym poziomie. Wśród pacjentów z ciśnieniem rozkurczowym na poziomie 105 mmHg dawka 25 mg kaptoprylu jest z reguły (wystarczająca / niewystarczająca), aby obniżyć ciśnienie rozkurczowe poniżej 90 mmHg.

Rozwiązanie

1.

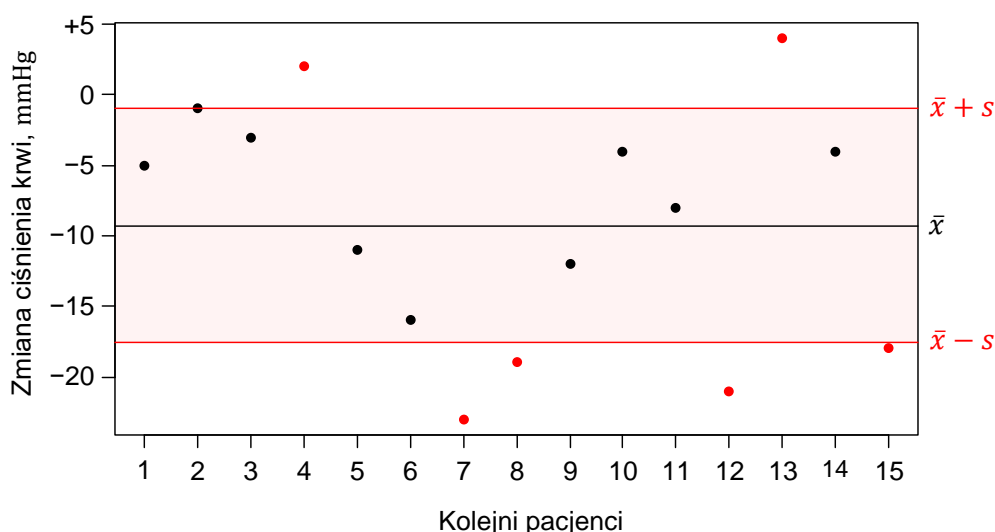
Średnia arytmetyczna: $\bar{x} = -9,27$ mmHg
 Odchylenie standardowe: $s = 8,32$ mmHg

2.

Wśród badanych pacjentów efekt terapeutyczny podania dawki 25 mg kaptoprylu był (zbliżony / zróżnicowany). Z tego powodu różni pacjenci o tym samym stopniu nadciśnienia powinni dostawać (taką samą / indywidualnie ustalaną) dawkę leku, aby ich ciśnienie tętnicze krwi utrzymywało się na prawidłowym poziomie. Wśród pacjentów z ciśnieniem rozkurczowym na poziomie 105 mmHg dawka 25 mg kaptoprylu jest z reguły (wystarczająca / niewystarczająca), aby obniżyć ciśnienie rozkurczowe poniżej 90 mmHg.

Komentarz

Ćwiczenie rozpoczyna się od obliczenia wartości średniej arytmetycznej i wartości odchylenia standardowego. Obliczenia są jednak wykonywane nie na surowych wartościach pomiarów, ale na różnicach dwóch pomiarów: po podaniu leku i przed jego podaniem. Ujemne wartości różnic – określających efekt terapeutyczny leku – oznaczają obniżenie ciśnienia krwi. Poniżej przedstawiono graficzne podsumowanie wyników badań.



W próbie 15 pacjentów ciśnienie rozkurczowe obniżyło się średnio o 9,27 mmHg (**czarna** linia na wykresie). Jednak nie każdy z pacjentów odniósł taką samą korzyść. Typowe efekty terapeutyczne wyrażone jako średnia \pm odchylenie standardowe zawierają się w przedziale od $-0,95$ mmHg do $-17,59$ mmHg (**czerwone** linie na wykresie). Oznacza to, że typowe wyniki terapii są bardzo zróżnicowane: rozciągają się od efektów trudno mierzalnych do efektów mogących przywrócić prawidłowe ciśnienie krwi u pacjentów z ciśnieniem rozkurczowym na poziomie 105 mmHg (nadciśnienie tętnicze drugiego stopnia). Ze względu na duże różnice we wrażliwości pacjentów na działanie leku jego dawka powinna być ustalana indywidualnie dla każdego pacjenta, ale nie można przekraczać jego dawki maksymalnej.

6. Statystyka w zadaniach z *Informatora* oraz arkusza pokazowego

Zadanie 2. (0–4)¹⁰

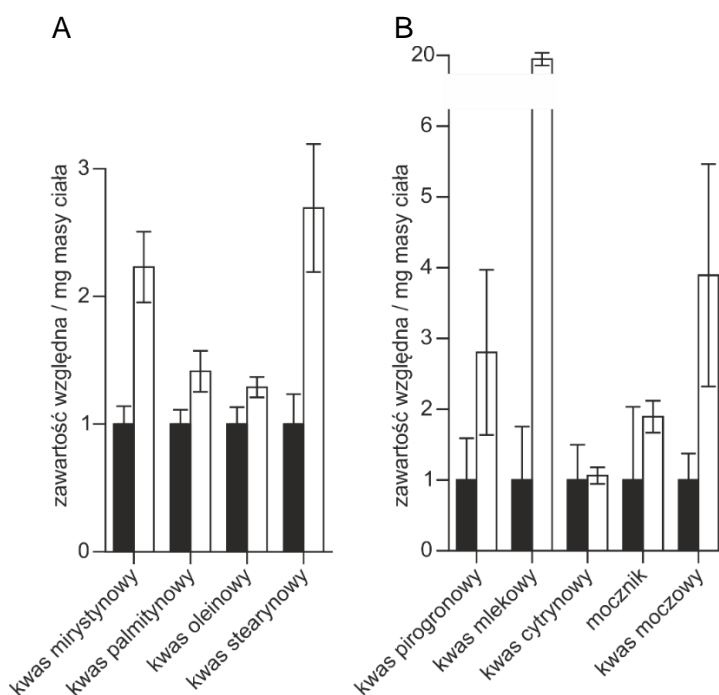
Organizmy modelowe są używane m.in. do badania zależności między genami, szlakami sygnałowymi i metabolizmem. Homeostaza lipidów i glukozy jest w podobny sposób kontrolowana u bezkręgowców i ssaków – za pomocą ścieżki sygnałowej uruchamianej przez insulinę.

W celu określenia, czy dieta wysokotłuszczowa wpływa na metabolizm węglowodanów, lipidów i białek, przeprowadzono badania z wykorzystaniem identycznych genetycznie samic szczepu *w¹¹⁸* wywiltzny karłowatej (*Drosophila melanogaster*) podzielonych na dwie grupy: grupa I – osobniki hodowane na pożywkę standardową, wykorzystywanej w laboratoriach do hodowli wywiltzny i zapewniającej jej wszystkie niezbędne do normalnego rozwoju składniki odżywcze;

grupa II – osobniki hodowane na pożywkę standardową, do której dodano oleju kokosowego (pożywka o dużej zawartości tłuszczów nasyconych).

Po tygodniu oznaczono zawartość określonych metabolitów u osobników należących do każdej z grup.

Na poniższych wykresach przedstawiono zawartość względną kwasów tłuszczowych (A) oraz zawartość względną wybranych metabolitów (B) w ciele much karmionych pożywką o wysokiej zawartości tłuszczów nasyconych, w porównaniu z zawartością tych związków w 1 mg masy ciała much karmionych pożywką standardową.



Informacja do wykresu

- pożywka standardowa □ pożywka o wysokiej zawartości tłuszczów nasyconych

Słupki błędów oznaczają odchylenie standardowe obliczone na podstawie 15 pomiarów.

Na podstawie: E.T. Heinrichsen i inni, *Metabolic and transcriptional response to a high-fat diet in Drosophila melanogaster*, „Molecular Metabolism” 3, 2014, s. 42–54.

¹⁰ Zadania 2., 3., 6. i 11. pochodzą z [Informatora o egzaminie maturalnym z biologii od roku szkolnego 2022/2023](#).

Zadanie 2.1. (0–1)

Oceń, czy poniższe interpretacje przedstawionych wyników badań są prawidłowe. Zaznacz T (tak), jeśli interpretacja wyników jest prawidłowa, albo N (nie) – jeśli jest nieprawidłowa.

1.	W przypadku diety bogatej w tłuszcze nasycone pomiary zawartości kwasu mlekowego miały większą zmienność niż pomiary zawartości kwasu moczowego.	T	N
2.	Wszystkie wyniki pomiarów względnej zawartości kwasu pirogronowego były mniejsze od czterech.	T	N
3.	Najwyższy wynik pomiaru zawartości kwasu pirogronowego w próbie z pożywką standardową był mniejszy od najwyższego wyniku w próbie z pożywką o wysokiej zawartości tłuszczów nasyconych.	T	N

Wymagania ogólne

II. Rozwijanie myślenia naukowego; doskonalenie umiejętności planowania i przeprowadzania [...] doświadczeń oraz wnioskowania w oparciu o wyniki badań.

Zdający:

- 3) [...] analizuje i interpretuje wyniki badań w oparciu o proste analizy statystyczne;
- 4) odnosi się do wyników uzyskanych przez innych badaczy;
- 5) [...] formułuje wnioski.

Wymaganie szczegółowe

III. Energia i metabolizm.

5. Pozyskiwanie energii użytecznej biologicznie. Zdający:

- 2) analizuje [...] przebieg glikolizy, reakcji pomostowej i cyklu Krebsa, wyróżnia substraty i produkty tych procesów.

Zasady oceniania

1 pkt – za trzy poprawne odpowiedzi.

0 pkt – za odpowiedź niespełniającą wymagań za 1 pkt albo za brak odpowiedzi.

Rozwiązanie

NNN

Komentarz

Rozwiązanie zadania 2.1. wymaga umiejętności odczytywania wyników prostych analiz statystycznych. Wszystkie niezbędne informacje można odczytać z wykresów znajdujących się we wstępie do zadania. Wartości średnie z prób wyrażono w jednostkach względnych – grupa kontrolna ma zawsze wystandaryzowaną wartość średnią równą jedności, z kolei wartość średnia w próbie badawczej oznacza, ile razy większa lub mniejsza jest ona w porównaniu z próbą kontrolną. Zmienność wyników w każdej z grup liczących 15 osobników przedstawiono w postaci słupków błędów – w tym przypadku wyrażają one odchylenie standardowe. Im większe odchylenie standardowe, tym większe zróżnicowanie wyników. Odchylenie standardowe dla pomiarów zawartości kwasu mlekowego jest mniejsze niż dla pomiarów zawartości kwasu moczowego, a więc pierwsza interpretacja wyników jest nieprawidłowa. Słupki błędów w postaci odchylenia standardowego określają przedział, w którym znajdują się typowe wyniki pomiarów, ale część wyników badań zawsze leży dalej

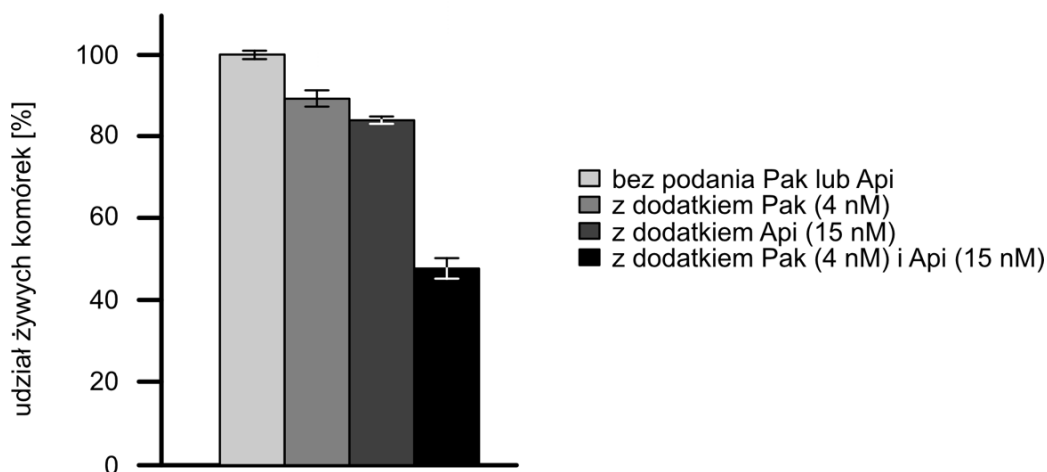
od średniej niż jedno odchylenie standardowe. Zatem drugie i trzecie zdanie także stanowią nieprawidłową interpretację przedstawionych wyników badań. Na podstawie słupków błędów można jedynie powiedzieć, że typowe wyniki względnej zawartości kwasu pirogronowego w próbie badawczej mieszczą się w przedziale od ok. 1,5 do 4,0, ale są też wyniki, które znajdują się poza tym przedziałem.

Zadanie 3. (0–2)

Paklitaxel to związek chemiczny wyizolowany z kory cisa krótkolistnego (*Taxus brevifolia*). Stosuje się go w leczeniu m.in. raka sutka, jajnika i płuc. Działanie paklitaxelu polega na stabilizacji mikrotubul i zahamowaniu ich depolimeryzacji, co uniemożliwia ich skracanie. Apigenina jest flawonem występującym m.in. w kwiatach rumianku, pietruszce i selerze.

Poniżej przedstawiono wykres ilustrujący odsetek komórek nowotworowych, które pozostały żywe po 24 godzinach od podania paklitaxelu (Pak), apigeniny (Api) lub obu związków razem (Pak i Api).

Słupki błędów wskazują odchylenie standardowe.



Na podstawie: Y. Xu i inni, *Synergistic Effects of Apigenin and Paclitaxel on Apoptosis of Cancer Cells*, „PLoS ONE” 6(12), 2011, s. 1–12.

Zadanie 3.2. (0–1)

Uzupełnij poniższe zdania tak, aby powstał poprawny opis interpretacji przedstawionego wykresu. W każdym nawiasie podkreśl właściwe określenie.

Paklitaxel i apigenina podane razem powodują, że po 24 godzinach przeżywa średnio około (50% / 80%) komórek. Na podstawie wykresu zamieszczonego we wprowadzeniu do zadania można stwierdzić, że podanie paklitaxelu i apigeniny (*razem / oddzielnie*) będzie – średnio rzecz biorąc – skuteczniejsze w terapii raka niż podawanie ich (*razem / oddzielnie*).

Wymagania ogólne

II. Rozwijanie myślenia naukowego; doskonalenie umiejętności planowania i przeprowadzania [...] doświadczeń oraz wnioskowania w oparciu o wyniki badań.

Zdający:

- 2) [...] rozróżnia próbę kontrolną i badawczą;
- 3) opracowuje, analizuje i interpretuje wyniki badań w oparciu o proste analizy statystyczne.

Wymagania szczegółowe

IV. Podziały komórkowe. Zdający:

- 3) opisuje cykl komórkowy, z uwzględnieniem zmian ilości DNA w poszczególnych jego etapach; uzasadnia konieczność replikacji DNA przed podziałem komórki;
- 4) opisuje przebieg kariokinezy podczas mitozy [...].

Zasady oceniania

1 pkt – za podkreślenie właściwych określeń we wszystkich nawiasach.

0 pkt – za odpowiedź niespełniającą wymagań za 1 pkt albo za brak odpowiedzi.

Rozwiązanie

Paklitaxel i apigenina podane razem powodują, że po 24 godzinach przeżywa średnio około (50% / 80%) komórek. Na podstawie wykresu zamieszczonego we wprowadzeniu do zadania można stwierdzić, że podanie paklitakselu i apigeniny (razem / oddzielnie) będzie – średnio rzecz biorąc – skuteczniejsze w terapii raka niż podawanie ich (razem / oddzielnie).

Komentarz

Wyniki badań w każdej z czterech grup zostały podsumowane w postaci średniej arytmetycznej i odchylenia standardowego. Te wyniki pokazują, że w każdej z grup wartość odchylenia standardowego jest niewielka w porównaniu z wartością średnią. Dzięki temu wiadomo, że zmienność wyników między powtórzeniami w każdej z grup była niewielka. Zatem średnia precyzyjnie opisuje badane populacje komórek, a różnice między średnimi są dobrym przedstawieniem różnic między badanymi populacjami.

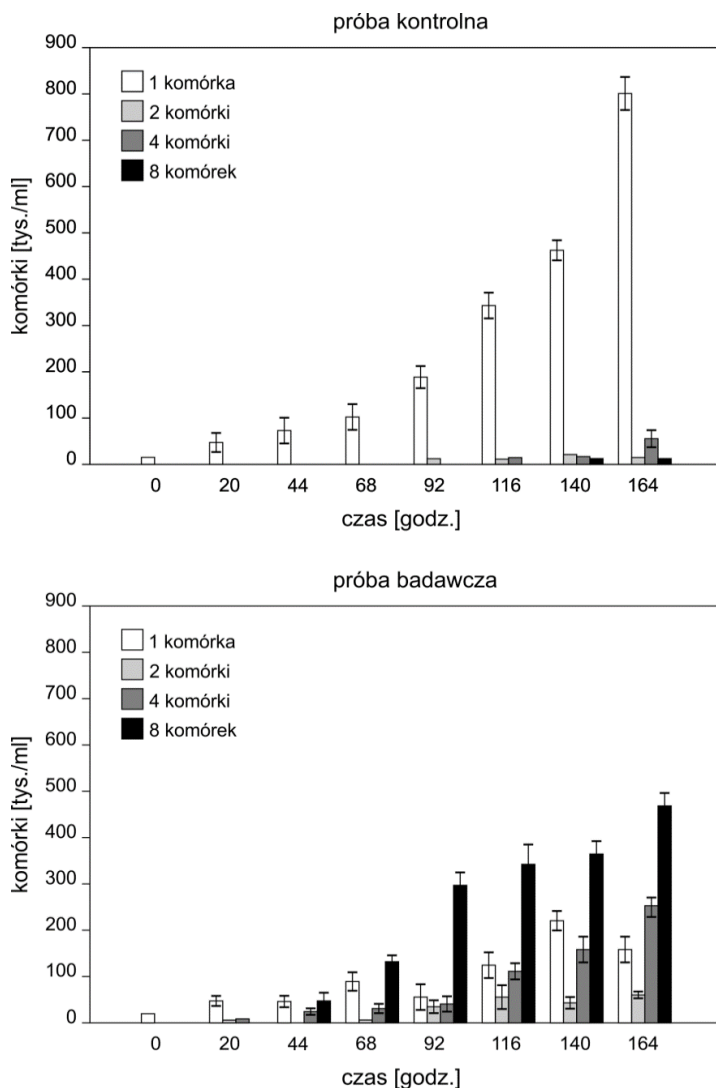
Najważniejszym wnioskiem, jaki można wyciągnąć z przedstawionych wyników badań, jest interakcja między działaniem paklitakselu i działaniem apigeniny. Przy podaniu osobno paklitaxel obniża udział żywych komórek o ok. 12%, a apigenina – o ok. 18%. Można się zatem spodziewać, że sumaryczny efekt terapeutyczny przy podaniu tych dwóch leków razem będzie wynosił ok. 30%. Badania pokazują jednak, że przy podaniu leków razem przeżywa tylko ok. 50% komórek, a więc efekt terapeutyczny jest o 20% silniejszy od spodziewanego. Ten wzrost skuteczności o 20% ponad sumę indywidualnych efektów każdego z leków stanowi miarę interakcji między lekami i określa się jako działanie synergistyczne dwóch leków.

Zadanie 6. (0–4)

Zielenica *Scenedesmus subspicatus* może być jednokomórkowa lub tworzyć 2-, 4- lub 8-komórkowe kolonie. Postawiono hipotezę, że forma morfologiczna glonu wynika z presji roślinożercy.

Aby przetestować tę hipotezę, przeprowadzono następujący eksperyment. Dodano do wody sole mineralne i rozpoczęto hodowlę *S. subspicatus* w warunkach sztucznego oświetlenia i stałej temperatury równej 20 °C. Po sześciu dniach populacja składała się prawie wyłącznie z formy jednokomórkowej o rozmiarach 6–8 × 4–5 μm w zagęszczeniu około 20 tys. komórek/ml. Na tym etapie hodowlę podzielono na cztery części i umieszczono w osobnych naczyniach o objętości 100 ml. Do dwóch z nich dodano po jednym osobniku *Daphnia magna* – skorupiaka o wielkości około 3 mm, który w warunkach naturalnych żywi się fitoplanktonem. Pozostałe warunki hodowli nie uległy zmianie. We wszystkich czterech próbach codziennie sprawdzano gęstość hodowli oraz udział poszczególnych form morfologicznych.

Wyniki eksperymentu przedstawiono na poniższych wykresach w postaci średnich oraz zakresu zmienności w dwóch powtórzeniach.



Na podstawie: D.O. Hessen, E. van Donk, *Morphological changes in Scenedesmus induced by substances released from Daphnia*, „Archiv für Hydrobiologie” 127, 1993, s. 129–140.

Zadanie 6.1. (0–1)

Sformułuj wniosek na podstawie przedstawionych wyników doświadczenia.

.....

.....

Wymaganie ogólne

II. Rozwijanie myślenia naukowego; doskonalenie umiejętności planowania i przeprowadzania [...] doświadczeń oraz wnioskowania w oparciu o wyniki badań.

Zdający:

5) ocenia poprawność zastosowanych procedur badawczych oraz formułuje wnioski.

Wymaganie szczegółowe

VIII. Protisty. Zdający:

1) przedstawia formy morfologiczne protistów.

Zasady oceniania

1 pkt – za poprawnie sformułowany wniosek, odnoszący się do stymulującego wpływu *Daphnia magna* na rozwój kolonii u *Scenedesmus subspicatus*.

0 pkt – za odpowiedź niespełniającą wymagań za 1 pkt albo za brak odpowiedzi.

Przykładowe rozwiązania

- Presja *D. magna* prowadzi do formowania się kolonii u *S. subspicatus*.
- *Scenedesmus subspicatus* odpowiada na presję *D. magna* zwiększeniem rozmiarów kolonii.
- Obecność *Daphnia magna* jest przyczyną grupowania się komórek w kolonie u *S. subspicatus*.

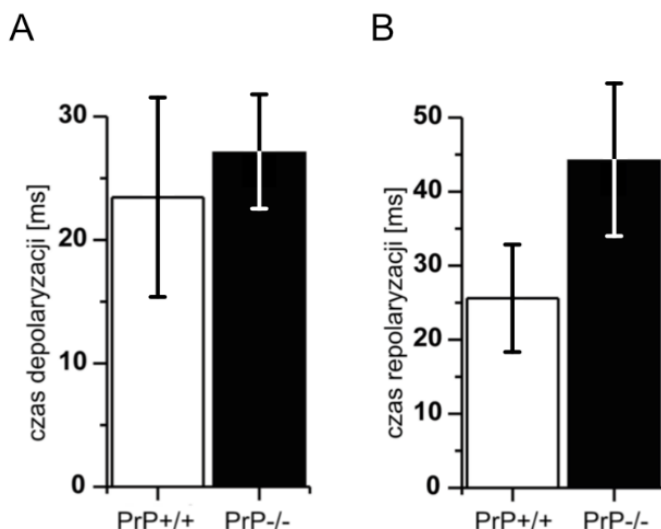
Komentarz

Słupki błędów na wykresach zazwyczaj oznaczają odchylenie standardowe. W tym przypadku jest inaczej – chodzi o „zakres zmienności w dwóch powtórzeniach”, a więc o wartość minimalną i o wartość maksymalną. Pośrodku takiego przedziału znajduje się wartość średnia z dwóch powtórzeń eksperymentu. Badacze zazwyczaj stosują większą liczbę prób (powtórzeń), ale czasami ze względów finansowych lub organizacyjnych badania prowadzi się na mniejszą skalę. Jeżeli – jak w tym przypadku – zmienność między powtórzeniami jest niewielka w porównaniu do efektu wywoływanego przez badany czynnik, to z takich badań można od razu wyciągnąć wnioski. W przeciwnym razie przeprowadzony eksperyment należałoby potraktować jako pilotażowy, a wnioskowanie – odroczyć do momentu uzyskania wyników z udziałem większej liczby powtórzeń.

Zadanie 11. (0–3)

Białko prionowe (PrP) w neuronach ssaków może występować w dwóch formach przestrzennych: sfałdowanej prawidłowo oraz sfałdowanej nieprawidłowo. Nie poznano dotąd funkcji formy prawidłowej, natomiast forma nieprawidłowa jest przyczyną chorób prionowych, np. choroby Creutzfeldta-Jakoba.

W celu określenia funkcji białka PrP w neuronie hodowano myszy o genotypie dzikim (PrP+/+) oraz myszy z nieaktywnymi obiema kopiami genu kodującego badane białko (PrP-/-). Z myszy wyizolowano neurony i zmierzono czas trwania depolaryzacji (A) i repolaryzacji (B). Wyniki eksperymentu przedstawiono na poniższym wykresie.



Informacja do wykresu

Średni czas depolaryzacji i repolaryzacji wraz z odchyleniem standardowym.

W plemieniu Fore z Papui-Nowej Gwinei dawniej występowała choroba *kuru*, będąca odmianą choroby Creutzfeldta-Jakoba. Choroba rozprzestrzeniła się, ponieważ w tym plemieniu praktykowano rytualny kanibalizm zmarłych. Po zaprzestaniu tych praktyk ustały przypadki nowych zachorowań. U części członków tego plemienia wykryto allel genu kodującego PrP – całkowicie eliminujący ryzyko zachorowania na chorobę Creutzfeldta-Jakoba. Tego allelu nie znaleziono dotychczas w innych ludzkich populacjach.

Na podstawie: H. Khosravani i inni, *Prion protein attenuates excitotoxicity by inhibiting NMDA receptors*, „Journal of Cell Biology” 181, 2008, s. 551–565;

E.A. Asante i inni, *A naturally occurring variant of the human prion protein completely prevents prion disease*, „Nature” 522, 2015, s. 478–481.

Zadanie 11.1. (0–2)

Uzupełnij poniższe zdania tak, aby powstał poprawny wniosek dotyczący wpływu białka PrP na czas depolaryzacji i repolaryzacji neuronu. W każdym nawiasie podkreśl właściwe określenie.

Białko PrP (*skraca / wydłuża*) czas (*depolaryzacji / repolaryzacji*) błony neuronu, ponieważ typowe wyniki pomiarów w próbie badawczej były (*krótsze / dłuższe*) od typowych wyników w próbie kontrolnej. Natomiast ustalenie, czy białko PrP skraca, czy – wydłuża czas (*depolaryzacji / repolaryzacji*), nie jest możliwe, ponieważ wyniki pomiarów w próbach badawczej i kontrolnej w znacznym stopniu się pokrywały.

Wymagania ogólne

- II. Rozwijanie myślenia naukowego; doskonalenie umiejętności planowania i przeprowadzania [...] doświadczeń oraz wnioskowania w oparciu o wyniki badań.
Zdający:
 - 3) [...], analizuje i interpretuje wyniki badań w oparciu o proste analizy statystyczne.
- III. Posługiwanie się informacjami pochodzącymi z analizy materiałów źródłowych. Zdający:
 - 2) odczytuje, analizuje, interpretuje i przetwarza informacje tekstowe, graficzne, liczbowe.

Wymaganie szczegółowe

- XII. Wirusy, wiroidy, priony.
 2. Wiroidy i priony – swoiste czynniki infekcyjne. Zdający:
 - 2) opisuje priony jako białkowe czynniki infekcyjne będące przyczyną niektórych chorób degeneracyjnych OUN (choroba Creutzfeldta-Jakoba [...]).

Zasady oceniania

- 2 pkt – za podkreślenie właściwych określeń w czterech nawiasach.
1 pkt – za podkreślenie właściwych określeń w trzech nawiasach.
0 pkt – za odpowiedź niespełniającą wymagań za 1 pkt albo za brak odpowiedzi.

Rozwiązanie

Białko PrP (skraca / wydłuża) czas (depolaryzacji / repolaryzacji) błony neuronu, ponieważ typowe wyniki pomiarów w próbie badawczej były (krótsze / dłuższe) od typowych wyników w próbie kontrolnej. Natomiast ustalenie, czy białko PrP skraca, czy – wydłuża czas (depolaryzacji / repolaryzacji), nie jest możliwe, ponieważ wyniki pomiarów w próbach badawczej i kontrolnej w znacznym stopniu się pokrywały.

Komentarz

Tekst o prionach wykorzystano do sprawdzenia umiejętności wnioskowania statystycznego [...]. W zadaniu 11.1. zdający powinien zauważyć, że choć w obu eksperymentach zaobserwowano dłuższe średnie czasy depolaryzacji i repolaryzacji w neuronach pozbawionych białka PrP, to tylko w przypadku repolaryzacji typowe wyniki pomiarów w próbie badawczej i kontrolnej są wyraźnie rozgraniczone. Z kolei w drugim eksperymencie, mimo że średnia wartość czasu depolaryzacji była mniejsza u typu dzikiego, to wiele pomiarów w próbie kontrolnej było wyższych niż w próbie badawczej. Z tego powodu wyniki naszego eksperymentu nie pozwalają na wyciągnięcie tak mocnych wniosków.

Zadanie 9.¹¹

Zdjęcie A przedstawia bursztynkę pospolitą (*Succinea putris*) – niewielkiego ślimaka spotykanego w zacienionych wilgotnych zaroślach. Na zdjęciu B widoczny jest osobnik bursztynki, który ma charakterystycznie zgrubiałe czułki, świadczące o obecności w nich larw pasożytniczej przywry *Leucochloridium paradoxum*. Czułki tego ślimaka barwą oraz kształtem przypominają gąsienice motyli i rytmicznie pulsują. Z tego powodu zwracają na nie uwagę ptaki owadożerne, np. sikory, zięby czy muchołówki, w których diecie zwykle nie występują ślimaki. Ptaki te zjadają czułki ślimaka, przez co zarażają się pasożytem.



A

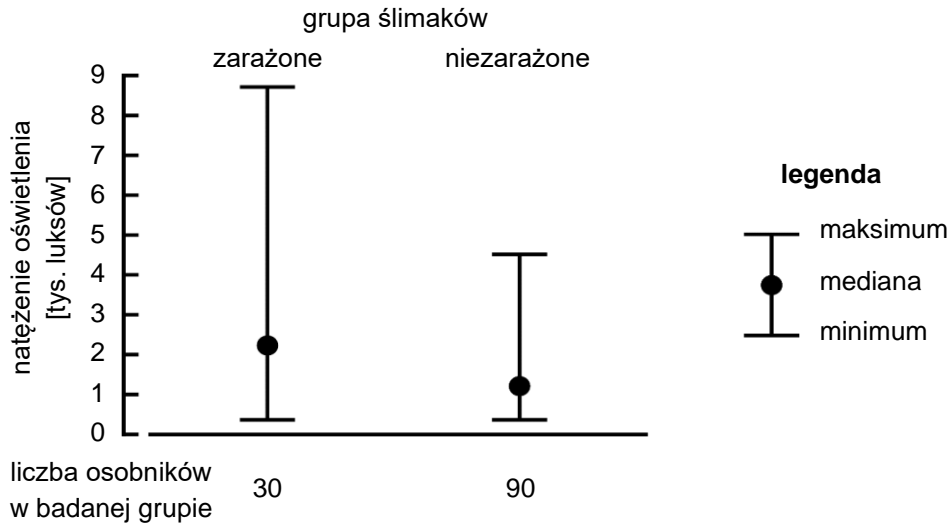


B

Bursztynki zarażają się przywrą, zjadając wraz z roślinami ptasie odchody, w których znajdują się jaja z rozwiniętym w nich pierwszym stadium larwalnym pasożyta (miracidium). W przewodzie pokarmowym ślimaka rozwijają się z nich następne stadia larwalne (sporocysty), które rozmnażają się partenogenetycznie i wytwarzają charakterystyczne worki z larwami kolejnego stadium (metacerkariami), umiejscowione w czułkach bursztynki.

Zbadano wpływ zarażenia przywrą na zachowanie ślimaków. Na wykresie przedstawiono pomiary natężenia oświetlenia w miejscach, które wybierały ślimaki mające czułki zmienione przez larwy *Leucochloridium*, oraz w miejscach, w których gromadziły się bursztynki bez objawów zarażenia tym pasożytem. Obserwacje osobników z obu grup trwały po 45 minut.

¹¹ Zadanie 9. pochodzi z [arkusza pokazowego egzaminu maturalnego z biologii na poziomie rozszerzonym w Formule 2023](#).



Na podstawie: W. Wesołowska, T. Wesołowski, *Do Leucochloridium sporocysts manipulate the behaviour of their snail hosts?*, „Journal of Zoology” 292, 2014; M. Tabin, *Władcy marionetek*, „Wiedza i Życie” 8, 2011; www.flickr.com

Zadanie 9.1. (0–1)

Na podstawie przedstawionych informacji oraz wykresu opisz zmianę w zachowaniu populacji ślimaków wywołaną zarażeniem *L. paradoxum*.

.....

.....

.....

Wymaganie ogólne

- II. Rozwijanie myślenia naukowego; doskonalenie umiejętności planowania i przeprowadzania obserwacji i doświadczeń oraz wnioskowania w oparciu o wyniki badań. Zdający:
- 3) [...] analizuje i interpretuje wyniki badań w oparciu o proste analizy statystyczne.

Wymagania szczegółowe

- XI. Funkcjonowanie zwierząt.
2. Porównanie poszczególnych czynności życiowych zwierząt, z uwzględnieniem odpowiedzialnych za nie struktur.
 - 9) Rozmnażanie i rozwój. Zdający:
 - d) analizuje na podstawie schematu cykle rozwojowe zwierząt pasożytniczych [...].
- XVII. Ekologia.
3. Ekologia ekosystemu. Ochrona i gospodarka ekosystemami. Zdający:
 - 3) przedstawia adaptacje [...] pasożytów [...] do zdobywania pokarmu.

Zasady oceniania

1 pkt – za opisanie tych zmian w zachowaniu ślimaków, które polegają na częstszym przebywaniu w miejscach o dużym natężeniu oświetlenia.

0 pkt – za odpowiedź niespełniającą wymagań za 1 pkt albo za brak odpowiedzi.

Przykładowe rozwiązania

- Ślimaki nie chowają się aż tak bardzo w miejscach o małym natężeniu oświetlenia.
- Ślimaki częściej przebywają w miejscach odsoniętych.

Komentarz

We wstępie do zadania przedstawiono wyniki obserwacji w dwóch grupach ślimaków – zarażonych i niezarażonych *L. paradoxum*. Z wykresu możemy odczytać wartości minimalne i maksymalne oraz medianę natężenia światła w każdej z badanych grup. Przed wyciągnięciem wniosków warto zwrócić uwagę na to, że badane grupy nie są równoliczne. Grupa badawcza jest trzykrotnie mniejsza od grupy kontrolnej, a jednak w grupie badawczej rozstęp (różnica między wartością maksymalną i minimalną) jest prawie dwukrotnie większy niż w próbie kontrolnej. Oznacza to, że otoczenie, w jakim przebywają ślimaki zarażone, charakteryzuje się dużo większą zmiennością natężenia światła niż w przypadku ślimaków niezarażonych. Z czego to wynika? Wartości minimalne natężenia światła w dwóch grupach są takie same. Jednak mediana i wartość maksymalna są większe w grupie badawczej. O ile większa wartość maksymalna w grupie badawczej mogłaby być spowodowana pojedynczą wartością odstającą, to wyższa wartość mediany musi wynikać ze zmiany zachowania znacznej części osobników w próbie. Można wyciągnąć zatem wniosek, że zarażone ślimaki częściej przebywają w miejscach nasłonecznionych.