

METODY STATYSTYCZNE W BIOLOGII

BIOLOGIA ST II ROK I (MGR)

Prowadzący: mgr **Łukasz Pawelec**, Zakład Antropologii, IBŚ, UPWr,

lukasz.pawelec@upwr.edu.pl, pokój 203 (obok sekretariatu, II piętro) , Zakład Antropologii,
Kozuchowska 5.

Konsultacje: Poniedziałek 11.30-13.00 (stacjonarnie/zdalnie) **po uprzednim umówieniu się mailowo!**

Jak będą wyglądały zajęcia?

Nr zajęć	tematy	data	Godziny (45')
1.	Ćwiczenia wstępne (eksport/tworzenie bazy danych, dodawanie /usuwanie rekordów, warunki selekcji, formuły, etykiety tekstowe, przekodowywanie zmiennych).	10.10	1
2.	Populacja a próba (sposoby pobierania próby, analiza licznosci próby i mocy testu)	17.10	1
3.	Estymacja parametrów (pojęcie błędu standardowego i przedziału ufności, graficzna prezentacja tych miar w postaci wykresów typu średnia-błędy oraz ramka-wąsy).	24.10	1

4.	Testowanie hipotez statystycznych I (podstawowe etapy testowania hipotez, testy dla porównania dwóch prób: t-Studenta, U Manna-Whitney' a, Wilcoxon)	7.11	1
5.	Testowanie hipotez statystycznych I - CIĄG DALSZY	14.11	1
6.	Testowanie hipotez statystycznych II (testy porównujące więcej niż dwie grupy: test Fishera, Kruskalla-Wallisa, ANOVA Friedmanna oraz właściwe testy post-hoc)	21.11	1
7.	Testowanie hipotez statystycznych II – CIĄG DALSZY	28.11	1
8.	Korelacje (korelacja liniowa [Pearsona] i nieliniowe [Spearmana, Kendalla, Gamma], korelacje cząstkowe).	5.12	1
9	Korelacje – CIĄG DALSZY	12.12	1
10	Regresja liniowa (równanie regresji liniowe, regresja wieloraka, regresja krokowa postępująca i wsteczna)	19.12	1
11	Regresja nieliniowa	9.01	1

12	Porównywanie cech nominalowych (test chi-2 jedno- i dwuzmiennowy, dokładny test Fishera, test McNemara)	16.01	1
13	Porównywanie cech nominalowych - CIĄG DALSZY	23.01	1
14	Powtórzenie wiadomości - ZADANIA	30.01	1
15	Zaliczenie	2.02 (czwartek)	1

Ćwiczenia 1

Ćwiczenia wstępne

Otwieramy pakiet Statistica 13.5

1 Otwórzmy program Statistica

2 Wprowadzanie danych do program

1) Import arkusza Excel

Plik→Otwórz *wybieramy Przykładowy_arkusz_EXCEL*→Import wszystkich arkuszy do skoroszytu/Import wybranego arkusza→Ok→OK

2) Utworzenie nowego arkusza Statistica

3 Otwieramy nowy arkusz

Plik→Nowy

4 Dodawanie zmiennych i przypadków

Zmienne → Dodaj

Przypadki → Dodaj

5 Nazywanie zmiennych

klikamy dwukrotnie w nagłówek kolumny i wpisujemy nazwę w polu „Nazwa”

- nazwijmy pierwszą zmienną (kolumnę) „Płeć”

6 Wprowadzanie danych/etykiety tekstowe

- Można je skopiować „ręcznie” np. z Excela lub wpisać na nowo

- zmienne mogą być liczbowe lub tekstowe (litery, wyrazy)

- zacznijmy wpisywać w kolejne pola kolumny płeć słowa „samiec” i „samica”

- program nada im etykiety tekstowe – Statistica to program matematyczny – nie rozumie słów tylko liczby!

7 Otwieramy plik „Wysiętek [1]”

- opis pliku

8 Wyliczanie wartości zmiennych na podstawie danych z innych zmiennych – Formuły matematyczne

- dodajmy dwie nowe zmienne (v11 i v12) i nazwijmy je: v11- Ubytek masy ciała, v12- BMI.

- wyliczmy wartości tych zmiennych za pomocą odpowiednich formuł:

$$v11: =v2-v10$$

$$v12: =v2/(v1/100)^2$$

i klikamy **OK**

- białe pole, gdzie wpisaliśmy przed chwilą formułę, może służyć jako NOTATNIK!

9 Selekcja wybranej części bazy danych – funkcja *select cases*

- aby je włączyć należy wcisnąć na klawiaturze **F8** lub **Fn+F8** (w zależności od typu klawiatury).

- mamy różne operatory logiczne i matematyczne

Arytmetyczne	+	Dodawanie
	-	Odejmowanie
	*	Mnożenie
	/	Dzielenie
	^ (lub **)	Potęgowanie
Logiczne	()	Nawiasy
	AND (lub &)	Koniunkcja
	OR (lub)	Alternatywa
	NOT (lub ~)	Negacja
Relacyjne	=	Równe
	<	Mniejsze
	>	Większe
	<=	Mniejsze lub równe
	>=	Większe lub równe
	<>	Różne

- można dowolnie włączać i wyłączać różne przypadki
- Aby widzieć, które przypadki zostały wybrane, najlepiej użyć specjalnego formatu podświetlania ważnych rekordów
- Spróbujcie wybrać osoby z SBP powyżej 135 – w której rundzie jest takich przypadków najwięcej i dlaczego
- wybierzcie osoby, których tętno w 2. Rundzie nie przekroczyło 100 ($v_6=2$ and $v_7<100$)
- wybierzcie te osoby, które mają:
 - a. wysokość ciała powyżej 185 cm **LUB** masę ciała powyżej 74 kg
 - b. wysokość ciała powyżej 185 cm **I** masę ciała powyżej 74 kg

Jaka jest różnica w liczbie ważnych przypadków?

10 Przekodowywanie danych, miary rozkładu

Zamiana zmiennych wyrażonych w różnych skalach

Wyróżniamy 3 główne rodzaje skal:

- a** nominalna (nominałowa, jakościowa)
- b** porządkowa (zmiennie dyskretne, skokowe)
- c** ilościowa (interwałowa, ilorazowa)

Kierunek przekształcania skal:

Ilościowa → porządkowa → nominalna

Nigdy w przeciwnym kierunku!!!

Przekodowywanie zmiennych

Aby przekształcić skale należy użyć funkcji **przekodowywania** zmiennych

Otwórzmy plik „Wysitek”

Spróbujmy przekodować zmienną BMI (v12) na 3 sposoby wg standardów WHO

$x \leq 25$ – norma

$25 < x \leq 30$ – nadwaga

$x > 30$ – otyłość

Tworzymy najpierw nową zmienną (13) – „BMI wg WHO”

Zapisz plik, jako **Wysitek [2]**