

METODY STATYSTYCZNE W BIOLOGII

1. Wykład wstępny
2. Populacje i próby danych
3. Testowanie hipotez i estymacja parametrów
4. Planowanie eksperymentów biologicznych
- 5. Najczęściej wykorzystywane testy statystyczne**
6. Podsumowanie materiału, wspólna analiza przykładów, dyskusja
7. Regresja liniowa
8. Regresja nieliniowa
9. Określenie jakości dopasowania równania regresji liniowej i nieliniowej
10. Korelacja
11. Elementy statystycznego modelowania danych
12. Porównywanie modeli
13. Analiza wariancji
14. Analiza kowariancji
15. Podsumowanie materiału, wspólna analiza przykładów, dyskusja

TEST



HIPOTEZY



PRÓBA DANYCH

TEST T

TEST T - DEFINICJA

$$H_0: \mu_A = \mu_B$$

$$t = \frac{(\bar{x} - \bar{x}_A)^2 N_A + (\bar{x} - \bar{x}_B)^2 N_B}{\sum_{i=1}^{N_A} (x_{iA} - \bar{x}_A)^2 + \sum_{i=1}^{N_B} (x_{iB} - \bar{x}_B)^2} (N_A + N_B - 2) \sim t_{N_A + N_B - 2}$$

\bar{x} Średnia arytmetyczna obserwacji w obu próbach

\bar{x}_A / \bar{x}_B Średnia arytmetyczna próby A / B

x_{iA} / x_{iB} Wartość i -tej obserwacji w próbie A / B

N_A / N_B Liczebność próby A / B

TEST T - PRZYKŁAD

BMD	SEX
0.97	1
0.73	1
0.87	1
0.94	1
1.02	1
0.76	1
0.78	1
1.01	1
0.82	1
0.76	1
0.87	1
0.72	1
...	
0.91	2
1.02	2
0.87	2

1. Gęstość kości [g/cm²] 40 zdrowych osób dorosłych

2. Wartości znane dla mężczyzn i kobiet

H_0 : śr. gęstość kości kobiet jest taka jak mężczyzn

H_1 : śr. gęstość kości kobiet jest różna niż mężczyzn

$$H_0: \mu_K = \mu_M$$

$$H_1: \mu_K \neq \mu_M$$



$$\alpha_{\max} = 0.0500$$

$$t = 2.0566$$

$$\alpha_t = 0.0466$$



H_1

śr. gęstość kości u mężczyzn jest większa niż u kobiet

TEST T - ZAKRES STOSOWANIA

1. Porównanie średnich arytmetycznych

2. Dane ciągłe

3. Rozkład normalny

4. Podobne wariancje

5. Warianty testu:

- Pojedyncza próba $H_0: \mu = 1.0$
 - Dwie niezależne próby $H_0: \mu_A = \mu_B$
 - Dwie sparowane próby $H_0: \mu_A = \mu_B$
-

TEST T - DEFINICJA

$$H_0: \mu_A = \mu_B$$

$$t = \frac{\bar{D}\sqrt{N}}{\sqrt{\frac{\sum_{i=1}^N (D_i - \bar{D})^2}{N-1}}} \sim t_{N-1}, \quad \bar{D} = \frac{\sum_{i=1}^N (x_{iA} - x_{iB})}{N}$$

N

Liczba par obserwacji

x_{iA} / x_{iB}

Wartość i -tej obserwacji w próbie A / B

D_i

Różnica wartości obserwacji w i -tej parze

TEST T - PRZYKŁAD

Low CCT	High CCT
20.0	14.3
13.9	13.8
18.3	15.8
21.1	33.4
20.1	20.3
24.4	19.9
20.2	14.3
11.6	11.4
28.8	25.1
18.5	24.1

1. Badanie ciśnienia w gałce ocznej

2. Ciśnienie w 2 gałkach tej samej osoby

H_0 : ciśnienie nie zależy od grubości rogówki

H_1 : ciśnienie zależy od grubości rogówki

$$H_0: \mu_L = \mu_H$$

$$H_1: \mu_L \neq \mu_H$$



$$\alpha_{\max} = 0.0500$$

$$t = 0.2515$$

$$\alpha_t = 0.8071$$

H_0



ciśnienie w gałce ocznej nie zależy od grubości rogówki

TEST MANNA-WHITNEYA

TEST MANNA-WHITNEYA - DEFINICJA

$$H_0: \mu_A = \mu_B$$

$$U = \min \left(n_A n_B + \frac{n_B(n_B + 1)}{2} - \sum_{i=1}^{n_B} r_{Bi}, \quad n_B n_A + \frac{n_A(n_A + 1)}{2} - \sum_{i=1}^{n_A} r_{Ai} \right)$$

$$z = \frac{U - \frac{n_A n_B}{2}}{\sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}}} \sim N(0,1)$$

n_A / n_B

Liczba obserwacji w próbie A / B

r_{Ai} / r_{Bi}

Ranga i-tej obserwacji w próbie A / B

TEST MANNA-WHITNEYA - PRZYKŁAD

ŚREDNIE	WYSOKIE
5.5	6.0
6.0	7.0
5.0	7.5
7.0	6.0
5.5	7.5
6.0	8.0
7.0	11.0
6.0	9.0
8.0	8.0
7.0	11.0
6.0	8.0
7.0	8.0
6.0	7.0
8.0	7.0
6.0	7.0
7.0	9.0

1. Długość krewetki w zależności od zasolenia wody

H_0 : długość krewetki nie zależy do zasolenia

H_1 : długość krewetki zależy do zasolenia

$$H_0: \mu_S = \mu_W$$

$$H_1: \mu_S \neq \mu_W$$



$$\alpha_{\max} = 0.05$$

$$z = 3.09$$

$$\alpha_t = 0.002$$

$$U = 181$$

$$U_t = 46$$



H_1

długość krewetki zależy od zasolenia wody

TEST WILCOXONA

TEST WILCOXONA - DEFINICJA

$$H_0: \mu_A = \mu_B$$

$$W = \min \left(\sum_{i=1}^{n_1} r_i^-, \sum_{i=n_1}^{N-n_1} r_i^+ \right)$$

$$z = \frac{W - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}} \sim N(0,1)$$

N

Całkowita liczba par obserwacji

n_i

Liczba par o ujemnej różnicy wartości

r_i^- / r_i^+

Ranga i -tej obserwacji o ujemnej / dodatniej różnicy wartości

TEST WILCOXONA - PRZYKŁAD

NR OWCY	BEZ JAGN	Z JAGN
1	72.00	55.50
2	62.35	43.80
3	55.77	66.80
4	59.98	68.00
5	51.60	57.88
6	61.48	61.90
7	52.57	45.40
8	52.50	56.67
9	56.43	73.30
10	60.13	77.50
11	48.60	63.53
12	42.90	54.50
13	53.50	55.58
14	70.43	91.10
15	47.10	64.05
16	50.08	71.40

1. Czas pasienia się owcy z / bez jagnięcia

H_0 : czas pasienia nie zależy od obecności jagnięcia

H_1 : czas pasienia zależy od obecności jagnięcia

$$H_0: \mu_J = \mu_B \quad H_1: \mu_J \neq \mu_B$$



$$\alpha_{\max} = 0.0500$$

$$z = -2.02$$

$$\alpha_t = 0.0437$$

$$W = 29$$

$$W_t = 29$$

H_0

H_1

???

TEST F

TEST F - ZAKRES STOSOWANIA

1. Porównanie zmienności, różnice pomiędzy wieloma próbami danych

2. Dane ciągłe

3. Rozkład normalny

4. Warianty testu:

- Analiza wariancji

$$H_0 : \sigma_A^2 \leq \sigma_e^2$$

- Regresja

$$H_0 : \beta = 0$$

TEST F - DEFINICJA

$$H_0 : \sigma_A^2 \leq \sigma_e^2$$

$$F = \frac{\sum_{i=1}^{N_A} n_i (\bar{y}_i - \bar{y})^2}{\sum_{i=1}^{N_A} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2} \sim F_{N_A-1, N-N_A}$$

N Liczba obserwacji

N_A Liczba grup

n_i Liczba obserwacji w grupie i

\bar{y} Średnia wszystkich obserwacji

\bar{y}_i Średnia obserwacji w grupie i

y_{ij} Obserwacja j w grupie i

TEST F- PRZYKŁAD

A	B	C
3.06	3.41	2.92
2.60	3.23	2.88
2.55	3.93	3.25
2.42	3.74	2.64
2.35	3.18	3.28

1. Zawartość azotu w trzcinie (% suchej masy)

2. 3 lokalizacje (A, B, C)

H_0 : lokalizacje nie wpływają na zawartość azotu

H_1 : lokalizacje wpływają na zawartość azotu

$$H_0: \sigma_{\text{lok}}^2 \leq \sigma_e^2$$

$$H_1: \sigma_{\text{lok}}^2 > \sigma_e^2$$



$$\alpha_{\text{max}} = 0.0500$$

$$F = 11.95$$

$$\alpha_t = 0.00139$$

~~H_0~~

H_1

lokalizacje mają wpływ na zawartość azotu w trzcinie

TEST KRUSKALA-WALLISA

TEST KRUSKALA-WALLISA - DEFINICJA

$$H_0 : \sigma_A^2 \leq \sigma_e^2$$

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{N_A} n_i (\bar{R}_i - \bar{R})^2 \sim \chi_{N_A-1}^2$$

N Liczba obserwacji

N_A Liczba grup

\bar{R}_i Średni ranking obserwacji w grupie i

\bar{R} Średni ranking wszystkich obserwacji

TEST KRUSKALA-WALLISA- PRZYKŁAD

20-29	30-39	40-49
161.925	164.465	173.990
173.355	171.450	175.260
158.115	173.355	167.640
170.815	175.260	166.370
179.705	164.465	168.910

1. Wzrost dorosłych kobiet w USA

2. 3 przedziały wiekowe

H_0 : wzrost jest jednakowy w każdym p. wiekowym

H_1 : wzrost różni się w p. wiekowych

$H_0: \sigma_{wiek}^2 \leq \sigma_e^2$

$H_1: \sigma_{wiek}^2 > \sigma_e^2$



$$\alpha_{\max} = 0.0500$$

$$H = 6.45$$

$$\alpha_t = 0.0398$$

~~H_0~~

H_1

wiek wpływa wzrost dorosłych kobiet

TEST χ^2

TEST χ^2 - ZAKRES STOSOWANIA

1. Porównanie częstości
 2. Dane w formie częstości – ilość obserwacji w danej kategorii
 3. Minimalna liczba kategorii – 3
 4. Warianty testu: $H_0: N_A = N_B = N_C$
 - klasyfikacja jednoczynnikowa
 - klasyfikacja dwuczynnikowa
-

TEST χ^2 - DEFINICJA

$$H_0: N_A = \dots = N_K$$

$$\chi^2 = \sum_{i=1}^{N_K} \frac{(n_{obs}^i - n_{exp}^i)^2}{n_{exp}^i} = \frac{\left(n_{obs}^A - \frac{N}{N_K}\right)^2}{\frac{N}{N_K}} + \dots + \frac{\left(n_{obs}^K - \frac{N}{N_K}\right)^2}{\frac{N}{N_K}} \sim \chi_{N_K-1}^2$$

N Całkowita liczba obserwacji

N_K Liczba kategorii

n_{obs}^i Zaobserwowana liczba obserwacji w kategorii i

n_{exp}^i Oczekiwana liczba obserwacji w kategorii i

TEST χ^2 - PRZYKŁAD

KOLOR	CZĘSTOŚĆ
Biały	Żółty
Żółty	8
Żółty	Czerwony
Czerwony	5
Żółty	Biały
Żółty	4
Żółty	
Żółty	SUMA
Czerwony	17
Biały	
Czerwony	
Żółty	
Czerwony	
Czerwony	
Biały	
Żółty	
Biały	

1. Kolor kwiatów krokusa

H_0 : jednakowe częstości wystąpienia kolorów

H_1 : różne częstości wystąpienia kolorów

$$H_0: n_B = n_Z = n_C \quad H_1: n_B \neq n_Z \neq n_C$$



$$\alpha_{\max} = 0.05$$

$$t = 1.53$$

$$\alpha_t = 0.47$$

H_0



częstości wystąpienia poszczególnych kolorów krokusów są jednakowe

PORÓWNANIE DWU PRÓB



PRÓBY ZALEŻNE



**ROZKŁAD INNY NIŻ
NORMALNY**

TEST WILCOXONA

PORÓWNANIE DWU PRÓB



PRÓBY NIEZALEŻNE



**ROZKŁAD INNY NIŻ
NORMALNY**

TEST MANNA-WHITNEYA

PORÓWNANIE DWU PRÓB



PRÓBY ZALEŻNE



ROZKŁAD NORMALNY

TEST T

PORÓWNANIE WIELU PRÓB



ROZKŁAD NORMALNY

TEST F

PORÓWNANIE CZĘSTOŚCI

TEST χ^2

PORÓWNANIE WIELU PRÓB



**ROZKŁAD INNY NIŻ
NORMALNY**

TEST KRUSKALA-WALLISA

PORÓWNANIE DWU PRÓB



PRÓBY NIEZALEŻNE



ROZKŁAD NORMALNY

TEST T