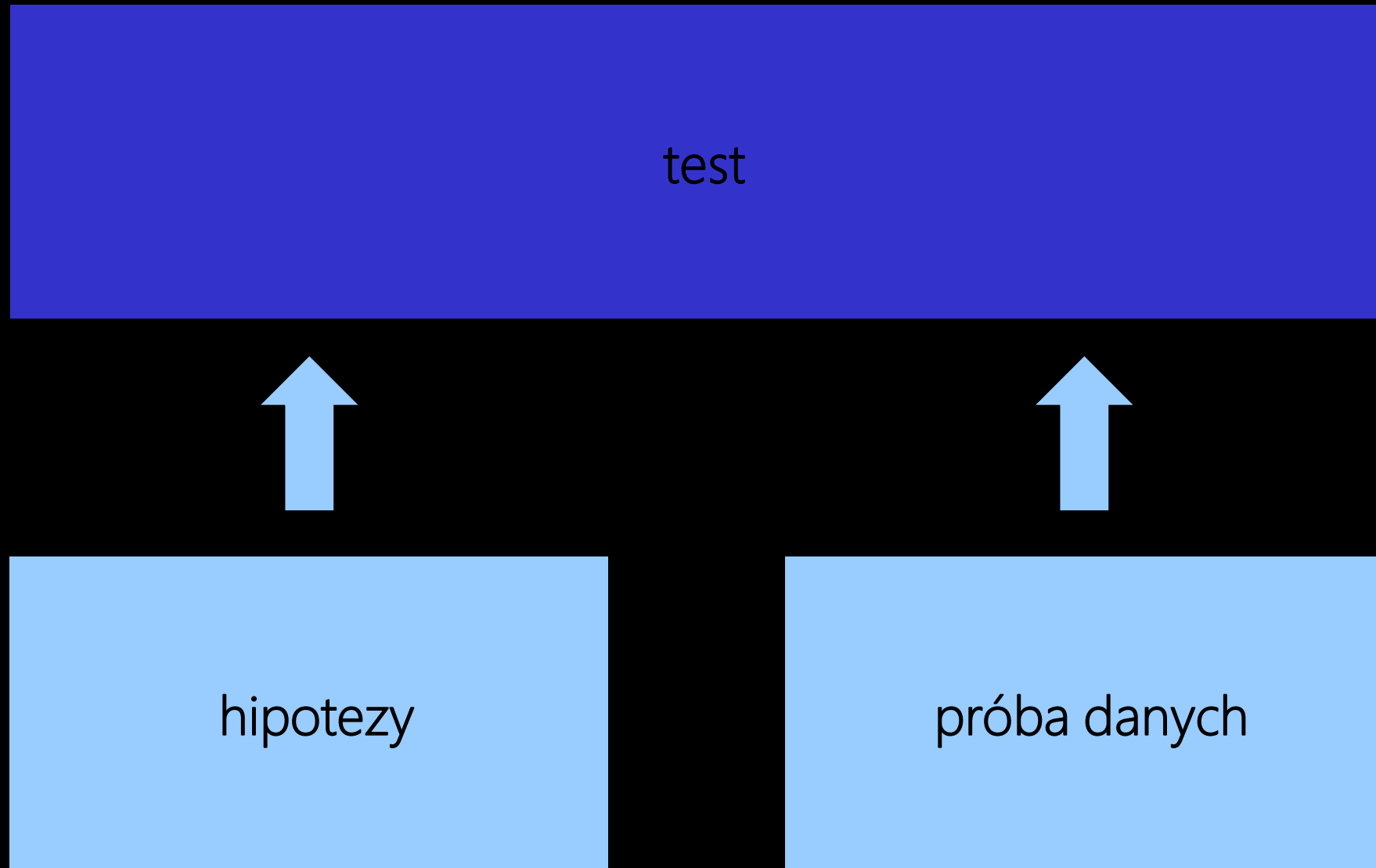


METODY STATYSTYCZNE W BIOLOGII

1. Wykład wstępny
2. Populacje i próby danych
3. Testowanie hipotez i estymacja parametrów
4. Planowanie eksperymentów biologicznych
5. Najczęściej wykorzystywane testy statystyczne I
6. Najczęściej wykorzystywane testy statystyczne II
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

1. Test F
2. Test Kruskala-Wallis
3. Test χ^2

- Zakres stosowania
- Definicja
- Przykład



Test F

Test F - zakres stosowania

1. Porównanie zmienności
 2. Dane ciągłe
 3. Rozkład normalny
 4. Warianty testu:
 - Analiza wariancji *one-way ANOVA* $H_0: \sigma_A^2 \leq \sigma_e^2$
 - Regresja
-

Test F - analiza wariancji

próba danych

1. Zawartość azotu w trzcinie (% suchej masy)
2. 3 lokalizacje (A, B, C), pomiar w 1996 r.
3. Flowermere, hrabstwo Cambridge

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



Test F - analiza wariancji

1. Określenie hipotez H_0 i H_1

H_0 : lokalizacje nie mają wpływu na zawartość azotu

H_1 : lokalizacje wpływają na zawartość azotu

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

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

2. Ustalenie poziomu istotności

$$\alpha_{MAX} = 0.05$$

3. Wybór i obliczenie wartości testu statystycznego:

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

Test F - analiza wariancji

3. Wybór i obliczenie wartości testu statystycznego

$$F = \frac{\frac{\sum_{i=1}^{N_g} n_i (\bar{y}_i - \bar{y})^2}{N_g - 1}}{\frac{\sum_{i=1}^{N_g} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2}{N - N_g}} = \frac{\frac{2.04}{3 - 1}}{\frac{1.03}{15 - 3}} = 11.95$$

4. Określenie rozkładu testu: $\sim F_{N_g-1, N-N_g}$

Excel 

5. Obliczenie wartości α_t : $\alpha_t = 0.00139$

R: `aov(surface~localisation,data=mays)`

6. Decyzja: $\alpha_t < \alpha_{\max}$

~~H_0~~ H_1

lokalizacje wpływają na zawartość azotu w suchej masie trzciny

Test F - zakres stosowania

1. Porównanie zmienności

2. Dane ciągłe

3. Rozkład normalny

4. Warianty testu:

- Analiza wariancji

- Regresja *regression*

$$H_0: \beta = 0$$

MASA CIAŁA	ZAW. TŁUSZCZU
89	28
88	27
66	24
59	23
93	29
73	25
82	29
77	25
100	30
67	23

1. Masa ciała [kg] i grubość tkanki tłuszczowej [mm]
2. 10 osób dorosłych, różnej płci, tej samej rasy



Test F - regresja

1. Określenie hipotez H_0 i H_1

H_0 : grubość tkanki tłuszczowej nie zależy od masy ciała

H_1 : grubość tkanki tłuszczowej zależy od masy ciała

$H_0: \beta_1 = 0$

$H_1: \beta_1 \neq 0$

2. Ustalenie poziomu istotności

$\alpha_{MAX} = 0.05$

3. Wybór i obliczenie wartości testu statystycznego:

$$F = \frac{\frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{N_\beta - 1}}{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N - N_\beta}}$$

3. Wybór i obliczenie wartości testu statystycznego

$$F = \frac{\frac{\sum_{i=1}^N (\hat{y}_i - \bar{y})^2}{N_{\beta} - 1}}{\frac{\sum_{i=1}^N (\hat{y}_i - y_i)^2}{N - N_{\beta}}} = \frac{\frac{54.32}{2 - 1}}{\frac{7.78}{10 - 2}} = 55.8$$

4. Określenie rozkładu testu: $\sim F_{N_{\beta}-1, N-N_{\beta}}$

5. Obliczenie wartości α_t : $\alpha_t = 0.000071$

Excel:



```
R: lm(tluszcz~masa,data=human_weight)
```

6. Decyzja: $\alpha_t < \alpha_{\max}$

~~H_0~~

H_1

grubość tkanki tłuszczowej zależy od masy ciała

Test Kruskala-Wallis

Test Kruskala-Kallisa - zakres stosowania

1. Porównanie zmienności
2. Dane ciągłe lub porządkowe (rangi)
3. Dane nie pochodzą z rozkładu normalnego
4. Analiza wariancji

Test Kruskala-Wallis

próba danych

1. Wzrost dorosłych kobiet w USA
2. 3 przedziały wiekowe

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



Test Kruskala-Wallis

1. Określenie hipotez H_0 i H_1

H_0 : wzrost kobiet jest jednakowy w każdym p. wiekowym

H_1 : wzrost kobiet różni się w zależności od p. wiekowego

$$H_0: \sigma_{wiek}^2 \leq \sigma_e^2 \quad H_1: \sigma_{wiek}^2 > \sigma_e^2$$

2. Ustalenie poziomu istotności

$$\alpha_{MAX} = 0.05$$

3. Wybór i obliczenie wartości testu statystycznego:

N liczba obserwacji

N_A liczba grup

\bar{R}_i śr. ranking obserwacji
w grupie i

\bar{R} śr. ranking wszystkich obs.

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

Test Kruskala-Wallis

3. Wybór i obliczenie wartości testu statystycznego

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{N_A} n_i (\bar{R}_i - \bar{R})^2 = \frac{12}{15(15+1)} 5(7.2 - 8)^2 + 5(8 - 8)^2 + 5(8 - 8)^2 = 6.45$$

4. Określenie rozkładu testu: $\sim \chi^2_{3-1}$

5. Obliczenie wartości α_t : $\alpha_t = 0.0398$

Excel: przykład



6. Decyzja: $\alpha_t < \alpha_{\max}$



H_1

wzrost dorosłych kobiet różni się w poszczególnych przedziałach wiekowych

Test χ^2

TEST χ^2 - ZAKRES STOSOWANIA

1. Testowanie liczebności poszczególnych kategorii
2. Dane w formie liczebności \rightarrow liczba obserwacji w danej kategorii
~~%, średnia, prawdopodobieństwo~~
3. Dane podzielone na kategorie – nominalne, porządkowe ~~ciągłe~~
4. Nie należy stosować testu χ^2 gdy oczekiwana liczebność dla kategorii jest mała (<5)
5. Warianty testu:
 - Klasyfikacja jednoczynnikowa *One-way classification*
 - Klasyfikacja dwuczynnikowa

Test χ^2 – klasyfikacja jednoczynnikowa

próba danych

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. Klasyfikacja danych wg pojedynczego kryterium
2. Kolor kwiatów krokusa



TEST χ^2 – KLASYFIKACJA JEDNOCZYNNIKOWA

1. Określenie hipotez H_0 i H_1

H_0 : liczebności wystąpienia kolorów są jednakowe

H_1 : liczebności wystąpienia kolorów są różne

H_0 : $n_B = n_Z = n_C$

H_1 : $n_B \neq n_Z \neq n_C$

2. Ustalenie poziomu istotności

$$\alpha_{MAX} = 0.05$$

3. Wybór i obliczenie wartości testu statystycznego


$$\chi^2 = \sum_{i=1}^{N_K} \frac{(n_{obs}^i - n_{exp}^i)^2}{n_{exp}^i} = \frac{(n_{obs}^{bi} - n_{exp}^{bi})^2}{n_{exp}^{bi}} + \frac{(n_{obs}^{zó} - n_{exp}^{zó})^2}{n_{exp}^{zó}} + \frac{(n_{obs}^{cz} - n_{exp}^{cz})^2}{n_{exp}^{cz}}$$

Test χ^2 – klasyfikacja jednoczynnikowa

3. Wybór i obliczenie wartości testu statystycznego

$$\begin{aligned}\chi^2 &= \sum_{i=1}^{N_K} \frac{(n_{obs} - n_{exp})^2}{n_{exp}} = \frac{(n_{obs}^{bi} - n_{exp}^{bi})^2}{n_{exp}^{bi}} + \frac{(n_{obs}^{zó} - n_{exp}^{zó})^2}{n_{exp}^{zó}} + \frac{(n_{obs}^{cz} - n_{exp}^{cz})^2}{n_{exp}^{cz}} \\ &= \frac{(4 - 5.7)^2}{5.7} + \frac{(8 - 5.7)^2}{5.7} + \frac{(5 - 5.7)^2}{5.7} = 1.53\end{aligned}$$

4. Określenie rozkładu testu: $\sim \chi_{K-1=2}^2$

5. Obliczenie wartości α_t : $\alpha_t = 0.47$ Excel: przykład 

6. Decyzja: $\alpha_t > \alpha_{max}$ H_0 ~~H_1~~

liczebności wystąpienia kolorów są jednakowe

Test χ^2 - zakres stosowania

1. Testowanie liczebności poszczególnych kategorii
2. Dane w formie liczebności \rightarrow liczba obserwacji w danej kategorii
~~%, średnia, prawdopodobieństwo~~
3. Dane podzielone na kategorie – nominalne, porządkowe ~~ciągłe~~
4. Nie należy stosować testu χ^2 gdy oczekiwana liczebność dla kategorii jest mała (<5)
5. Warianty testu:
 - Klasyfikacja jednoczynnikowa
 - Klasyfikacja dwuczynnikowa *Two-way classification*

Test χ^2 – klasyfikacja dwuczynnikowa

próba danych

1. Klasyfikacja danych wg dwu kryteriów
2. Liczebność słońi w Parku Narodowym Mikumi, Tanzania

		kategoria			
		samotny samiec	grupa samców	grupa rodzinna	grupa rodzinna + samiec
Pora	sucha	43	4	196	7
	deszczowa	92	17	195	8



Test χ^2 – klasyfikacja dwuczynnikowa

1. Określenie hipotez H_0 i H_1

H_0 : pora roku nie wpływa na zmiany liczebności słoni w poszczególnych kategoriach (oba kryteria są niezależne)

H_1 : pora roku wpływa na zmiany liczebności słoni w poszczególnych kategoriach (oba kryteria są zależne)

$$H_0: n_{Di} = n_{Si} \quad \text{dla } i=1...4 \quad H_1: n_{Di} \neq n_{Si}$$

2. Ustalenie poziomu istotności

$$\alpha_{MAX} = 0.05$$

3. Wybór i obliczenie wartości testu statystycznego

$$\chi^2 = \sum_{i=1}^{N_K} \sum_{j=1}^{N_M} \frac{(n_{obs}^{ij} - n_{exp}^{ij})^2}{n_{exp}^{ij}} = \frac{(n_{obs}^{D-1} - n_{exp}^{D-1})^2}{n_{exp}^{D-1}} + \dots + \frac{(n_{obs}^{S-4} - n_{exp}^{S-4})^2}{n_{exp}^{S-4}}$$

Test χ^2 – klasyfikacja dwuczynnikowa

$$\chi^2 = \sum_{i=1}^{N_K} \sum_{j=1}^{N_M} \frac{(n_{obs}^{ij} - n_{exp}^{ij})^2}{n_{exp}^{ij}}$$

liczebność klasy oczekiwana dla H_0

$$n_{exp}^{ij} = \frac{\sum_{i=1}^{N_K} n_{obs}^{i \cdot}}{N} \cdot \frac{\sum_{j=1}^{N_M} n_{obs}^{\cdot j}}{N} \cdot N = \frac{\sum_{i=1}^{N_K} n_{obs}^{i \cdot} \sum_{j=1}^{N_M} n_{obs}^{\cdot j}}{N}$$

	samo tny samie c	grupa samc ów	grupa rodzin na	grupa rodzin na + samie c	suma
sucha	43	4	196	7	250
deszczowa	92	17	195	8	312
suma	135	21	391	15	562

	samo tny samie c	grupa samc ów	grupa rodzin na	grupa rodzin na + samie c
sucha	60.5	9.34	173.93	6.67
deszczowa	74.95	11.66	217.07	8.33

Test χ^2 – klasyfikacja dwuczynnikowa

3. Wybór i obliczenie wartości testu statystycznego

$$\begin{aligned}\chi^2 &= \sum_{i=1}^{N_K} \sum_{j=1}^{N_M} \frac{(n_{obs}^{ij} - n_{exp}^{ij})^2}{n_{exp}^{ij}} = \\ &= \frac{(43 - 60.05)^2}{60.05} + \frac{(92 - 74.95)^2}{74.95} + \frac{(4 - 9.34)^2}{9.34} + \frac{(17 - 11.66)^2}{11.66} \\ &+ \frac{(196 - 173.93)^2}{173.93} + \frac{(195 - 217.07)^2}{217.07} + \frac{(7 - 6.67)^2}{6.67} + \frac{(8 - 8.33)^2}{8.33} = \\ &19.30\end{aligned}$$

4. Określenie rozkładu testu: $\sim \chi_{(K-1)(M-1)=3}^2$

5. Obliczenie wartości α_t : $\alpha_t = 0.0002$

Excel: przykład



Test χ^2 – klasyfikacja dwuczynnikowa

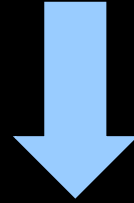
6. Decyzja: $\alpha_t < \alpha_{\max}$

~~H_0~~ H_1

liczebność poszczególnych grup słoni różni się w zależności od pory roku

wybór testu

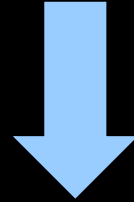
porównanie zmienności



rozkład normalny

test F

porównanie zmienności



rozkład inny niż normalny

test Kruskala-Wallisa

porównanie liczebności

test χ^2

Quiz – którego testu użyto ?

BMC Medical Genetics

Home [Articles](#)

Volume 21 Supplement 1

Selected Topics in "Systems Biology and Bioinformatics" - 2019: genetics

Research | [Open Access](#) | [Published: 22 October 2020](#)

Genetic polymorphisms of *PIP5K2A* and course of schizophrenia

[Evgeniya G. Poltavskaya](#) , [Olga Yu. Fedorenko](#), [Natalya M. Vyalova](#), [Elena G. Kornetova](#), [Nikolay A. Bokhan](#), [Anton J. M. Loonen](#) & [Svetlana A. Ivanova](#)

BMC Medical Genetics **21**, Article number: 171 (2020) | [Cite this article](#)

1615 Accesses | **1** Citations | [Metrics](#)

Quiz – którego testu użyto ?

Table 3 Association of *PIP5K2A* haplotypes with course of schizophrenia

From: [Genetic polymorphisms of *PIP5K2A* and course of schizophrenia](#)

SNPs	Haplotype	Frequency	Group 1, Group 2 frequencies	Chi square	P _{asym}	P _{Perm}
rs736203/rs943190	TT	0.571	0.565, 0.585	0.269	0.6043	0.9971
	CC	0.342	0.332, 0.366	0.836	0.3606	0.9437
	TC	0.046	0.053, 0.031	1.715	0.1904	0.7448
	CT	0.040	0.050, 0.018	4.463	0.0346	0.2097
rs10430590/rs1132816	AA	0.570	0.557, 0.608	1.397	0.2372	0.5328
	TA	0.211	0.211, 0.212	0.001	0.9750	1.0
	AG	0.127	0.128, 0.126	0.005	0.9451	0.9998
	TG	0.091	0.104, 0.054	3.990	0.0458	0.1130
rs943190/rs1132816	TA	0.497	0.501, 0.487	0.108	0.7430	0.9812
	CA	0.285	0.268, 0.334	2.827	0.0927	0.2276
	TG	0.117	0.117, 0.119	0.008	0.9309	0.9998
	CG	0.101	0.115, 0.061	4.276	0.0386	0.0946
rs11013052/rs1132816	CA	0.643	0.632, 0.675	1.075	0.2998	0.6804
	AA	0.139	0.137, 0.145	0.079	0.7791	0.9898
	AG	0.114	0.129, 0.073	4.119	0.0424	0.1423
	CG	0.104	0.103, 0.107	0.029	0.8652	0.9973
rs1132816/rs943194	AG	0.606	0.595, 0.637	1.004	0.3163	0.9144
	AC	0.175	0.173, 0.183	0.099	0.7535	0.9998
	GC	0.175	0.175, 0.103	5.158	0.0231	0.1419
	GG	0.063	0.058, 0.077	0.786	0.3754	0.9502

P_{asym} - asymptotic *P* value, P_{Perm} - empirical *P* value for 10,000 permutations. Group 1 – group of patients with continuous course of schizophrenia (*n* = 269), Group 2 – group of patients with episodic course of schizophrenia (*n* = 115). Bold text indicates *p*-value < 0.05

Quiz – którego testu użyć ?

Jakie hipotezy testowano ?

1. χ^2 Klasyfikacja pojedyncza

3. Test F regresja

4. Test F analiza wariancji

Quiz – którego testu użyć ?

J Rehabil Med 2018; 50: 914–919

ORIGINAL REPORT



Check for updates

EFFECT OF EXERCISE WITH A PELVIC REALIGNMENT DEVICE ON LOW-BACK AND PELVIC GIRDLE PAIN AFTER CHILDBIRTH: A RANDOMIZED CONTROL STUDY

Asuka SAKAMOTO, MS¹, Hiroshi NAKAGAWA, MD², Hitoshi NAKAGAWA, MD² and Kazuyoshi GAMADA, PhD¹
From the ¹Graduate School of Medical Technology and Health Welfare Sciences, Hiroshima International University, Hiroshima, and ²Obstetrics and Gynecology, Nakagawa Clinic, Hiroshima, Japan

- Badanie wpływu ćwiczeń po urodzeniu dziecka na ból miednicy
- 75 ciężarnych kobiet przydzielonych losowo do 3 grup: ćwiczenia z urządzeniem, ćwiczenia bez urządzenia, brak ćwiczeń
- Wyniki – klasyfikacja bólu po urodzeniu, brak rozkładu normalnego

Quiz – którego testu użyć ?

1. stepping



2. pelvic shifting



3. pelvic rotation with knee extension



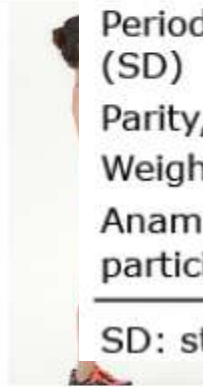
4. pelvic rotation with knee flexion



5. trunk flexion



trunk extension



Characteristics	Group R	Group E	Group C
Group size, <i>n</i>	25	25	25
Age, years, mean (SD)	32.4 (3.8)	31.4 (4.8)	30 (4.5)
Weight before pregnancy, kg, mean (SD)	51.5 (5.2)	52.4 (7.5)	50.4 (7.5)
Weight in third trimester, kg, mean (SD)	61.7 (4.6)	62.2 (8.0)	61.5 (7.9)
Period of pregnancy, weeks, mean (SD)	39.4 (0.9)	39.4 (0.8)	39.2 (1.2)
Parity, median (range)	1 (1-4)	1 (1-4)	2 (1-4)
Weight (new baby), g, mean (SD)	3,177 (292)	3,038 (322)	3,152 (428)
Anamnestic of LBP (number of participants, <i>n</i>)	5	6	3

SD: standard deviation; LBP: low-back pain.

Quiz – którego testu użyć ?

Jakie hipotezy testowano ?

1. χ^2 Klasyfikacja pojedyncza
2. χ^2 Klasyfikacja podwójna
3. Test F regresja

1. Test F
2. Test Kruskala-Wallis
3. Test χ^2