

Sandurska Elżbieta, Szulc Aleksandra. A method of statistical analysis in the field of sports science when assumptions of parametric tests are not violated. *Journal of Education Health and Sport*. 2016;6(13):275-287. eISSN 2391-8306.

DOI <http://dx.doi.org/10.5281/zenodo.293762>

<http://ojs.ukw.edu.pl/index.php/johs/article/view/4278>

The journal has had 7 points in Ministry of Science and Higher Education parametric evaluation. Part B item 754 (09.12.2016).

754 Journal of Education, Health and Sport eISSN 2391-8306 7

© The Author (s) 2016;

This article is published with open access at Licensee Open Journal Systems of Kazimierz Wielki University in Bydgoszcz, Poland

Open Access. This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited. This is an open access article licensed under the terms of the Creative Commons Attribution Non Commercial License

(<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted, non commercial use, distribution and reproduction in any medium, provided the work is properly cited.

This is an open access article licensed under the terms of the Creative Commons Attribution Non Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted, non commercial use, distribution and reproduction in any medium, provided the work is properly cited.

The authors declare that there is no conflict of interests regarding the publication of this paper.

Received: 05.12.2016. Revised 20.12.2016. Accepted: 31.12.2016.

Metoda przeprowadzania analizy statystycznej w przypadku niespełnienia założeń testów parametrycznych, w kontekście badań z obszaru kultury fizycznej

A method of statistical analysis in the field of sports science when assumptions of parametric tests are not violated

Elżbieta Sandurska¹, Aleksandra Szulc²

¹Katedra Genetyki, Instytut Biologii Eksperymentalnej, Uniwersytet Kazimierza Wielkiego Chodkiewicza 30, 85-064 Bydgoszcz, Polska, ekoral@ukw.edu.pl

¹Department of Genetics, Institute of Experimental Biology, Kazimierz Wielki University, Chodkiewicza 30, 85-064 Bydgoszcz, Poland, ekoral@ukw.edu.pl

²Wydział Zarządzania i Ekonomii, Politechnika Gdańska, Gdańsk, Polska

²Faculty of Management and Economics, Gdansk University of Technology, Gdansk, Poland

Streszczenie

Wprowadzenie: Duża gama programów statystycznych pozwala na wykonywanie skomplikowanych analiz w prosty sposób, co powoduje, że autorzy artykułów chętnie wykonują obliczenia samodzielnie, bez zagłębiania się w szczegóły dotyczące wymogów i założeń stosowania odpowiednich testów. Takie powierzchowne podejście do analizy statystycznej prowadzić jednak może do błędnych wyników i niewłaściwych wniosków dotyczących badanego problemu.

Cel pracy: Omówienie problemów związanych z niespełnieniem założeń testów parametrycznych: testu t -Studenta oraz jednoczynnikowej analizy wariancji ANOVA, często używanych w badaniach biometrycznych z zakresu sportu, oraz wskazanie metod postępowania w przypadku braku ich spełnienia.

Skrócony opis stanu wiedzy: Test t -Studenta oraz ANOVA są testami parametrycznymi, w związku z tym wymagają założeń, m.in. normalności rozkładów oraz jednorodności wariancji w każdej grupie. W przypadku, gdy założenia nie są spełnione, pierwotna konstrukcja testu zostaje zachwiana, a wynik może być zafałszowany. Prostą metodą na ustabilizowanie wariancji oraz znormalizowanie danych jest transformacja ich przy pomocy przekształceń. W przypadku, gdy takie działania zawodzą, dobrą alternatywą są testy nieparametryczne, np. Manna-Whitneya, Kruskala-Wallisa czy kolejności par Wilcoxon.

Podsumowanie: Weryfikacja założeń testów parametrycznych pozwala na właściwy dobór narzędzi statystycznych, co jest podstawą prawidłowego przeprowadzenia analizy statystycznej. Stosując kilka prostych zasad, analiza prawidłowości pojawiających się w sporcie sprowadza się do pewnego, łatwo dającego się wdrożyć, schematu postępowania.

Słowa kluczowe: test t -Studenta, jednoczynnikowa analiza wariancji, przekształcenia matematyczne, statystyka, testy nieparametryczne

Abstract

Introduction: Application of statistical software typically does not require extensive statistical knowledge, allowing to easily perform even complex analyses. Consequently, test selection criteria and important assumptions may be easily overlooked or given insufficient consideration. In such cases, the results may likely lead to wrong conclusions.

Aim: To discuss issues related to assumption violations in the case of Student's t -test and one-way ANOVA, two parametric tests frequently used in the field of sports science, and to recommend solutions.

Description of the state of knowledge: Student's t -test and ANOVA are parametric tests, and therefore some of the assumptions that need to be satisfied include normal distribution of the data and homogeneity of variances in groups. If the assumptions are violated, the original design of the test is impaired, and the test may then be compromised giving spurious results. A simple method to normalize the data and to stabilize the variance is to use transformations. If such approach fails, a good alternative to consider is a nonparametric test, such as Mann-Whitney, the Kruskal-Wallis or Wilcoxon signed-rank tests.

Summary: Thorough verification of the parametric tests assumptions allows for correct selection of statistical tools, which is the basis of well-grounded statistical analysis. With

a few simple rules, testing patterns in the data characteristic for the study of sports science comes down to a straightforward procedure.

Keywords: *t*-test, analysis of variance, mathematical transformations, statistics as topic, nonparametric test

Wprowadzenie

Statystyka stała się w dzisiejszych czasach dziedziną wiedzy obecną w różnorodnych aspektach codziennego życia. Stała się również podstawowym narzędziem badawczym wielu dziedzin i dyscyplin naukowych (Sobczyk, 1996). Jako element rachunku prawdopodobieństwa, używana jest z powodzeniem w naukach ścisłych, takich jak fizyka lub chemia. Jest niezbędna w naukach technicznych (np. kontrola jakości, wytrzymałość materiałów), w naukach ekonomicznych oraz w naukach przyrodniczych. Pozwala na łatwe podsumowanie i przedstawienie wyników badań empirycznych, co sprawia, że jest często wykorzystywana również przez badaczy zajmujących się kulturą fizyczną i sportem (Jankowicz-Szymańska et al., 2011; Wieczorek, 2011, Chmielewska et. al., 2014).

Wiele prac, poruszających zagadnienia z obszaru kultury fizycznej, dotyczy budowy i składu ludzkiego organizmu, siły mięśniowej wybranych grup mięśniowych, a także zdolności motorycznych (Jankowicz-Szymańska et al., 2011; Łopatka et al., 2012; Michnik et al., 2012; Wieczorek, 2011). Uzyskane wyniki pomiarowe wymagają odpowiedniej analizy i interpretacji statystycznej, która jest jednym z warunków opublikowania artykułu w recenzowanym czasopiśmie naukowym. Duża gama dostępnych programów statystycznych pozwala na wykonywanie nierzadko bardzo skomplikowanych analiz w bardzo prosty, niemal intuicyjny i szybki sposób, co powoduje, że autorzy artykułów, skuszeni prostotą używania takiego oprogramowania, chętnie wykonują obliczenia samodzielnie, bez zagłębiania się w szczegóły dotyczące wymogów i założeń stosowania odpowiednich testów. Jednakże sama znajomość obsługi programu statystycznego nie jest w pełni wystarczająca do pełnego zrozumienia wykonywanych analiz, jak również poprawnej interpretacji wyników. Powszechnym zjawiskiem jest posługiwanie się ogólnie znanymi testami statystycznymi, bez sprawdzenia ich założeń dotyczących rozkładu normalnego i jednorodności wariancji badanych cech, a co za tym idzie przydatności ich użycia. Takie postępowanie przyczynia się niejednokrotnie do publikacji błędnych wyników i wyciągania niewłaściwych wniosków dotyczących badanego problemu.

W przypadku, gdy założenia dotyczące rozkładu normalnego oraz jednorodności (homogeniczności) wariancji nie są spełnione, a następnie nie są uwzględnione w dalszych obliczeniach statystycznych, pierwotna konstrukcja testu zostaje zachwiana, a w konsekwencji jego wynik może być mocno zafałszowany (Stanisz, 2006). Brak spełnienia założeń testu oznacza konieczność podjęcia próby zmodyfikowania posiadanych danych (np. przez odrzucenie wartości skrajnie odstających), zmiany modelu (np. przez wyłączenie jednej z kilku grup z porównań) lub wprowadzenia drobnych przekształceń danych (pierwiastkowe, odwrotnościowe, itp.), które nie zaburzają wyniku porównań. W ostateczności można sięgnąć po jeden z mniej rygorystycznych pod względem założeń testów nieparametrycznych (Cleophas i Zwinderman, 2011).

Test *t*-Studenta oraz analiza wariancji ANOVA są testami parametrycznymi, co powoduje, że ich wykorzystanie zakłada spełnienie przez analizowane dane szeregu

założeń. Dzieje się tak dlatego, że testy parametryczne dotyczą wartości statystycznych populacji, takich jak np. średnia, w przeciwieństwie do nieparametrycznych, które dotyczą postaci funkcyjnej rozkładu (Sheskin, 2003). Najbardziej podstawowymi założeniami jest istnienie rozkładu normalnego zmiennej badanej w każdej z rozważanych grup oraz jednorodność wariancji tych rozkładów. Należy jednak przypomnieć, że testy parametryczne posiadają również inne założenia (Stanisz, 2007), często pomijane bądź traktowane jako spełnione, niejednokrotnie z powodu braku wiedzy badacza.

Bardzo ważnym założeniem jest możliwość stosowania testów parametrycznych tylko na danych zmierzonych na skali przedziałowej lub ilorazowej, tzn. takiej, która pozwala nam na operacje na danych jak np. dodawanie czy dzielenie (Walesiak, 1993). Niedopuszczalnym jest porównywanie ze sobą wyżej wymienionymi testami grup o charakterze porządkowym (np. pomiar zadowolenia klienta z wykonanej usługi, rangowany od 1 do 10) lub nominalnym, przypisującym obiekty do poszczególnych grup (np. kolor oczu, wykształcenie). Równie ważnym założeniem jak mierzalność danych jest losowość próby, czyli tzw. zasada randomizacji (Sheskin, 2003; Stanisz, 2007; Watała, 2005). Pamiętajmy, że dobrze dobrana próba powinna przedstawiać pewne odzwierciedlenie populacji, którą chcemy zbadać, np. wnioskując o kondycji fizycznej Polaków nie możemy do badań wybrać próby złożonej wyłącznie z osób uprawiających sport. Wyniki tak przeprowadzonej analizy statystycznej mogłyby być znacząco różne od faktycznej sprawności fizycznej przeciętnego Polaka. Dobrze dobrana grupa zapewnia spełnienie kolejnego założenia, jakim jest niezależność pomiarów (Ferguson i Takane, 1997), oznaczająca, że znajomość jednego pomiaru nie powinna dawać żadnej wskazówki o wartościach pozostałych pomiarów.

Normalność rozkładu oraz jednorodność wariancji są kluczowymi założeniami, wymaganymi do przeprowadzenia testu t czy ANOVA. Innymi słowy testy te porównują wyniki uzyskanych badań w poszczególnych grupach, o ile mają one rozkłady normalne o podobnym rozrzucie danych. Mierząc parametry budowy i składu ciała oraz siły mięśniowej wybranych grup mięśniowych, jak również zdolności motoryczne dla próby małowciążeniowej, może się zdarzyć, że założenia te nie są spełnione. W takim przypadku nie musi to oznaczać rezygnacji z użycia testu parametrycznego. Prostą metodą na ustabilizowanie wariancji oraz znormalizowanie danych jest transformacja ich przy pomocy przekształceń (Keppel i Wickens, 2004). Najbardziej popularnym jest przekształcenie logarytmiczne, które pozwala na znormalizowanie danych dla prawie każdej zmiennej nieposiadającej rozkładu normalnego. Innymi alternatywnymi przekształceniami jest przekształcenie pierwiastkowe $Y = \sqrt{X}$, odwrotnościowe $Y = 1/X$, kwadratowe $Y = X^2$ czy cyklometryczne $Y = \arcsin \sqrt{X}$, gdzie X jest zmienną podlegającą przekształceniu. W przypadku, gdy kilka wartości jest bliskich zeru lub równych zeru proponowanymi przekształceniami są $Y = \ln(X + 1)$ (Bartlett, 1947), $Y = \sqrt{X + 0,5}$, $Y = 1/(X + 1)$.

Jeżeli wszelkie próby znormalizowania rozkładu czy ustabilizowania wariancji zawodzą nie oznacza to, że musimy całkowicie zrezygnować z porównań naszych prób. Bardzo dobrą alternatywą są tutaj testy nieparametryczne, takie jak test Manna-Whitneya, przy porównywaniu dwóch grup niezależnych, test Kruskala-Wallisa - jako odpowiednik jednoczynnikowej analizy wariancji oraz test kolejności par Wilcoxon, dla dwóch prób zależnych (Cleophas i Zwinderman, 2011). Ponadto test Kruskala-Wallisa, podobnie jak jednoczynnikowa analiza wariancji, pozwala na przeprowadzenie testów *post-hoc* (np. test Dunna), w przypadku odrzucenia hipotezy zerowej o równości średnich w rozważanych

grupach. Testy te z powodzeniem można przeprowadzić za pomocą programu STATISTICA, a analiza wyników jest analogiczna do interpretacji wyników testów parametrycznych. Dodatkowo nie wymagają żadnych założeń wstępnych, można je stosować dla danych porządkowych i jakościowych, a także w przypadku bardzo małych grup. Ze względu na fakt, że siła testów nieparametrycznych jest mniejsza niż siła testów parametrycznych, należy je stosować tylko wówczas, gdy nie możemy się posłużyć testem parametrycznym (Stanisz, 2006).

Celem niniejszej pracy jest omówienie problemów związanych z niespełnieniem założeń testów statystycznych wykorzystywanych często w badaniach biometrycznych z zakresu sportu, mianowicie testu *t*-Studenta oraz jednoczynnikowej analizy wariancji ANOVA (Stanisz, 2006; 2007). Na podstawie badań własnych autorów artykułu, zostały wykonane przykładowe analizy statystyczne wraz z omówieniem trudności wyłaniających się w trakcie prowadzonych analiz. Szczególną uwagę zwrócono na zasadność sprawdzenia założeń obu testów oraz wskazanie metod postępowania w przypadku braku ich spełnienia.

Materiały i metody

Do zobrazowania najczęściej występujących problemów, pojawiających się w trakcie analiz statystycznych użyto danych zebranych od 50 losowo wybranych uczniów 8 Liceum Ogólnokształcącego w Bydgoszczy w wieku 15 – 18 lat (39 dziewczyn i 11 chłopców). Dokonano pomiaru masy ciała, wysokości ciała, procentowej zawartości tkanki tłuszczowej (Tanita BC-418 MA) – (Szulc, 2013), siły ręki (Jamar Hydraulic Hand Dynamometer) oraz przeprowadzono próbę gibkości w staniu poprzez skłon tułowia w przód. W badaniu wykorzystano również platformę stykową (Fusion Sport-Smart Jump). Uczniowie wykonywali serię 5 wyskoków typu CMJ - countermovement jump, polegających na wyskoku pionowym z jednoczesnym wymachem rąk, a po 1 minutowej przerwie, wyskoków typu ACMJ – akimbo countermovement jump – z rękoma umieszczonymi na biodrach. Badanie dotyczyło czasu lotu (FT) oraz czasu kontaktu (CT) stóp z matą stykową. Do obliczeń wykorzystano trzy zbliżone i uśrednione wyniki. Na podstawie tych danych obliczono: wysokość wyskoku (H), wskaźnik siły reakcji liczony jako stosunek wysokości wyskoku do czasu kontaktu stóp z matą dynamometryczną (RSI_1) oraz FT/CT – stosunek czasu lotu do czasu kontaktu stóp z matą (RSI_2) – (Flanagan i Comyns, 2008, Flanagan, 2016). Badaną populację podzielono na 4 grupy zgodnie z ich masą ciała: poniżej 50 kg, od 50 do 60 kg, od 60 do 70 kg oraz powyżej 70 kg. Przy podziale celowo nie uwzględniono płci, gdyż dodatkowe wyodrębnienie kobiet i mężczyzn mogłoby wprowadzić grupy zbyt małoliczne, uniemożliwiające przeprowadzenie analizy wariancji. Zabieg ten uznano za dopuszczalny ze względu na poglądowy charakter niniejszej pracy.

Analizy zostały wykonane przy użyciu programu STATISTICA 12 (StatSoft Polska). We wszystkich testach przyjęto poziom istotności równy 0.05. Przyjęto hipotezy zerowe, że płeć oraz masa ciała nie mają wpływu na osiągnięte rezultaty w przeprowadzonych testach sprawności fizycznej, jak również, że wyniki pomiędzy wyskokami typu CJM oraz ACJM nie różnią się między sobą.

Analizę statystyczną rozpoczęto od oszacowania podstawowych statystyk opisowych takich jak: średnia, mediana, odchylenie standardowe, górny i dolny kwartył łącznie oraz osobno dla grupy kobiet i mężczyzn, a także w czterech grupach wagowych. W dalszej kolejności postawiono szereg hipotez o równości średnich takich parametrów jak wysokość

wysokości, czasu kontaktu z matą kontaktową CT, czasu lotu FT, RSI_1 oraz RSI_2 w grupach dziewczyn i chłopców, tożsamych z hipotezami testu *t*-Studenta. Samego testu *t* nie przeprowadzono, gdyż jego wynik nie jest przedmiotem niniejszej rozprawy. Skupiono się na sprawdzeniu kilku założeń, a mianowicie, normalności rozkładu, który testowano metodą Shapiro-Wilka oraz jednorodności wariancji przy pomocy testu Levene'a. W przypadku niespełnionych założeń wykorzystano przekształcenie logarytmiczne danych w postaci $Y = \ln(X)$, gdzie *X* jest zmienną, która nie posiada rozkładu normalnego.

Wyżej opisane obliczenia powtórzono dla wszystkich wymienionych parametrów, sprawdzając tym razem hipotezę o równości średnich w czterech badanych grupach wagowych, tożsamą z hipotezą jednoczynnikowej analizy wariancji (ANOVA).

Podobne analizy wyznaczono dla założeń testu *t*-Studenta dla grup zależnych, który zakłada hipotezę o równości wartości średnich badanych parametrów w przypadku ćwiczenia wykonanego z zamachem rąk oraz bez. Testy *t* dla grup zależnych wymagają jedynie założenia o rozkładzie normalnym w każdej grupie, nie wymagają natomiast założenia o jednorodności wariancji (Stanisz, 1999). Procedura postępowania była analogiczna do powyższych.

Wyniki

Średnia masa ciała uczniów, bez względu na płeć, wyniosła $60,34 \pm 1,29$ kg, przy czym różnice pomiędzy dziewczynami i chłopcami były zauważalne i wyniosły odpowiednio: $57,29 \pm 1,21$ kg dla dziewczyn oraz $71,15 \pm 1,46$ kg dla chłopców. Podobne różnice zaobserwowano w pozostałych cechach, takich jak wysokość ciała: dla dziewczyn $166,06 \pm 0,9$ cm, dla chłopców $181,88 \pm 1,45$ cm, przy średniej ogólnej wynoszącej $169,29 \pm 1,19$ cm; oraz % zawartość tkanki tłuszczowej wynosząca odpowiednio $24,14 \pm 0,77$ % dla dziewczyn i $16,83 \pm 1,09$ % dla chłopców, przy średniej $22,65 \pm 0,78$ %. Różnice były wyraźnie widoczne również w sile rąk, która przyjmowała wartości zdecydowanie niższe dla dziewcząt: $279,5 \pm 7,08$ N dla ręki prawej oraz $262,67 \pm 7,6$ N dla ręki lewej, podczas gdy dla chłopców te wartości wynosiły odpowiednio $389,2 \pm 14,16$ N i $402,63 \pm 14,93$ N, kolejno dla ręki prawej oraz lewej. Pozostałe wartości statystyk opisowych zamieszczono w tabeli 1.

Przeprowadzone testy Shapiro-Wilka we wszystkich badanych grupach z wyszczególnieniem płci wykazały, że tylko w przypadku trzech zmiennych nie możemy odrzucić hipotezy o rozkładzie normalnym: czas lotu FT dla CMJ oraz ACMJ a także dla CT podczas CMJ. Wymienione trzy zmienne charakteryzowały się również jednorodnością wariancji w grupach dziewcząt i chłopców, a zatem założenia wstępne testu *t*-Studenta były spełnione i umożliwiały jego przeprowadzenie. W pozostałych przypadkach użyto przekształcenia logarytmicznego, które pozwoliło na znormalizowanie zmiennych (tabela 2).

Analiza tych samych zmiennych, tym razem podzielonych na 4 grupy w zależności od wartości masy ciała, pokazała podobne zależności. W przypadku 5 zmiennych: H, FT dla CMJ i ACMJ, a także CT (ACMJ) – została wykryta normalność w każdej badanej grupie. Użycie przekształcenia logarytmicznego pozwoliło na znormalizowanie zmiennych w pozostałych grupach za wyjątkiem zmiennej opisującej czas kontaktu z matą w CMJ (tabela 2).

Tabela 1. Statystyki opisowe podstawowych badanych cech uczniów, którzy wzięli udział w badaniu.

	Płeć	Średnia	Błąd stand.	Odch. stand.	Min - maks	Mediana	25 – 75%
masa ciała [kg]		60,34	1,29	9,11	47,7 – 77,4	56,85	52,2 – 67,3
	K	57,29	1,21	7,58	47,7 – 73,9	55,3	50,7 – 64
	M	71,15	1,46	4,83	64,4 – 77,4	70	66,5 – 76
wysokość ciała [cm]		169,29	1,2	8,39	155,4 – 189	168	163,5 – 175
	K	166,06	0,9	5,61	155,4 – 177	166,4	161 – 170,3
	M	181,88	1,44	4,58	175 – 189	182,75	177,5 – 185
% tkanki tłuszczowej		22,65	0,78	5,46	12 – 35,1	21,3	19 – 26,8
	K	24,14	0,78	4,86	17,3 – 35,1	23	19,8 – 28,4
	M	16,83	1,09	3,46	12 – 21,7	16,05	13,9 – 19,7
siła prawej ręki [N]		302,84	9,11	62,48	177,92 – 467,04	289,12	266,88 – 333,6
	K	279,5	7,08	43,08	177,92 – 378,08	266,88	266,88 – 311,36
	M	389,2	14,16	44,79	311,36 – 467,04	389,2	378,08 – 422,56
siła lewej ręki [N]		292,45	10,78	73,91	155,68 – 467,04	266,88	244,64 – 333,6
	K	262,67	7,6	46,25	155,68 – 355,84	266,88	222,4 – 289,12
	M	402,63	14,93	47,22	312,25 – 467,04	400,32	378,08 – 422,56
Gibkość		7,91	1,09	7,56	-12 – 22	7,45	2,15 – 14,05
	K	9,33	1,07	6,69	-7 – 22	9	5 – 14,4
	M	1,72	2,79	8,38	-12 – 16,4	1,7	-3,3 – 7

Analiza tych samych zmiennych, tym razem podzielonych na 4 grupy w zależności od wartości masy ciała, pokazała podobne zależności. W przypadku 5 zmiennych: H, FT dla CMJ i ACMJ, a także CT (ACMJ) – została wykryta normalność w każdej badanej grupie. Użycie przekształcenia logarytmicznego pozwoliło na znormalizowanie zmiennych w pozostałych grupach za wyjątkiem zmiennej opisującej czas kontaktu z matą w CMJ (tabela 2).

Przekształcenie logarytmiczne pomogło również zmienić rozkład badanych zmiennych celem przeprowadzenia testów *t*-Studenta dla grup zależnych. W porównaniach CMJ oraz ACMJ, tylko dla zmiennej FT test Shapiro-Wilka nie odrzucił hipotezy o normalności rozkładu. Ponownie jak w przypadku porównań wielokrotnych w grupach masy ciała, tak i tutaj zmienna przedstawiająca czas kontaktu stóp z matą, po przeprowadzeniu przekształcenia logarytmicznego, nadal naruszała założenie o normalności rozkładu (tabela 2).

Tabela 2. Wartości p w teście Shapiro-Wilka. Na czerwono zaznaczono wartości istotne statystycznie na poziomie istotności 0.05. H – wysokość skoku, CT – czas kontaktu stóp z matą dynamometryczną, RSI_1 – wskaźnik siły reaktywnej, FT – czas lotu, RSI_2 – stosunek czasu lotu do czasu kontaktu stóp z matą dynamometryczną, CJM - countermovement jump, ACJM – akimbo countermovement jump.

Zmienna	Bez przekształceń					Po przekształceniu $Y = \ln(X)$										
	Ogółem	Płeć		Grupy wagowe			Ogółem	Płeć		Grupy wagowe						
		K	M	40-50	50-60	60-70		70-80	K	M	40-50	50-60	60-70	70-80		
CJM	H	0,015	0,02	0,869	0,616	0,224	0,203	0,202	0,274	0,209	0,909					
	CT	<0,001	<0,001	0,566	0,356	0,239	<0,001	0,27	0,022	0,159	0,059	0,045	0,376	0,211	0,02	
	RSI_1	<0,001	<0,001	0,041	0,273	0,004	0,02	0,006	0,232	0,155	0,846	0,503	0,684	0,858	0,261	
	FT	0,093	0,093	0,895	0,628	0,501	0,351	0,28								
	RSI_2	<0,001	<0,001	0,007	0,124	0,003	0,031	0,002	0,083	0,095	0,39	0,238	0,363	0,571	0,111	
ACJM	H	0,01	0,02	0,572	0,848	0,191	0,251	0,977	0,18	0,223	0,733					
	CT	0,075	0,403	0,107	0,26	0,522	0,498	0,188	0,504							
	RSI_1	<0,001	0,002	0,873	0,346	0,043	0,564	0,188	0,296	0,402	0,922	0,749	0,756	0,646	0,91	
	FT	0,052	0,079	0,647	0,876	0,423	0,448	0,957								
	RSI_2	0,001	0,002	0,952	0,122	0,044	0,64	0,347	0,334	0,304	0,917	0,432	0,72	0,898	0,905	

Warto dodać, że mimo, iż nie znaleziono problemów ze spełnieniem założenia o jednorodności wariancji to zastosowane przekształcenia i tak ją stabilizowały we wszystkich grupach, w których zastosowano transformację. Świadczą o tym wyższe wartości p w teście Levene'a (tabela 3).

Tabela 3. Wartości p w teście Levene'a dla porównań w grupach płci oraz masy ciała. H – wysokość skoku, CT – czas kontaktu stóp z matą dynamometryczną, RSI_1 – wskaźnik siły reaktywnej, RSI_2 – stosunek czasu lotu do czasu kontaktu stóp z matą dynamometryczną, CJM - countermovement jump, ACJM – akimbo countermovement jump

Zmienna	Płeć	Grupy wagowe			
		X	Y = ln(X)	X	Y = ln(X)
CJM	H	0,303	0,988		
	CT	0,398	0,872	0,749	0,928
	RSI_1	0,053	0,579	0,17	0,422
	RSI_2	0,207	0,962	0,301	0,804
ACJM	H	0,432	0,765		
	RSI_1	0,115	0,644	0,055	0,230
	RSI_2	0,206	0,549	0,4	0,687

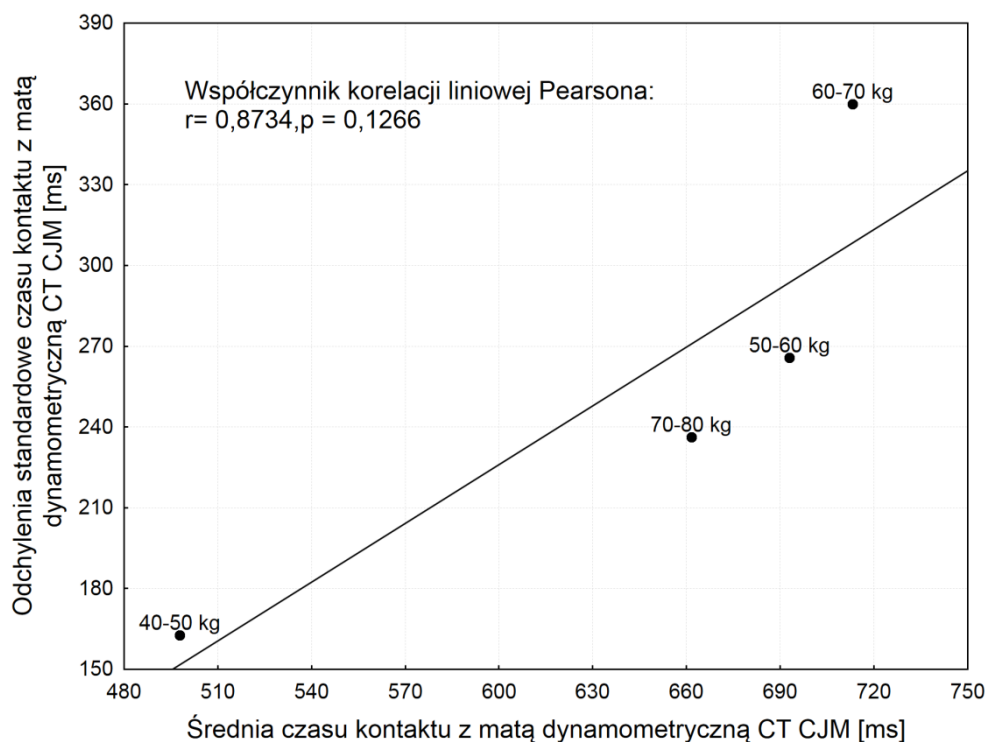
Podsumowanie

Znając zasady poprawnego doboru narzędzi statystycznych oraz sprawdzania i przeprowadzania testów, zweryfikowanie hipotez dotyczących porównania zdolności motorycznych uczniów w zależności od różnych czynników staje się nietrudnym zadaniem. Stosując kilka prostych zasad prawidłowej analizy statystycznej testowanie prawidłowości pojawiających się w sporcie sprowadza się do pewnego, łatwo dającego się wdrożyć, schematu postępowania (Szreder, 2010).

Znajomość alternatywnego sposobu postępowania w postaci wprowadzania transformacji umożliwia często wykorzystanie testów parametrycznych pomimo niespełnionych założeń wstępnych. Dzięki takiemu zabiegowi moc testu zostaje zachowana, a nawet zwiększona, będąc niejednokrotnie wyższą niż równoważnych testów nieparametrycznych, a tym samym prawdopodobieństwo popełnienia błędu drugiego rodzaju (nie odrzucenie fałszywej hipotezy zerowej) jest mniejsze (Watała, 2005).

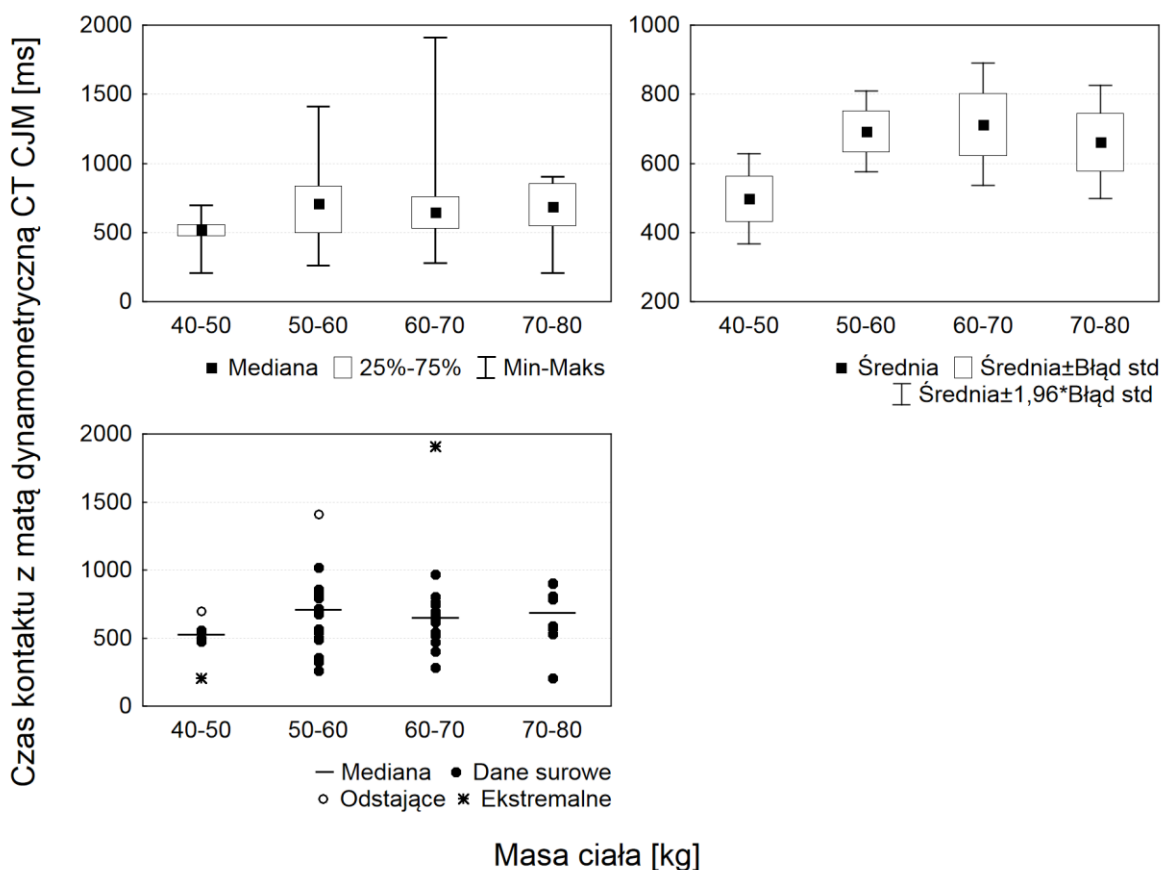
Zmienną, która nawet po przekształceniu logarytmicznym nie wykazywała rozkładu normalnego jest czas kontaktu stóp z matą kontaktową. Wielu statystyków uważa, że analiza wariancji jest w znacznym stopniu odporna na odchylenia od normalności (Lindman, 1974), zwłaszcza dla grup równolicznych, ale pod warunkiem, że średnie i odchylenia standardowe nie są ze sobą skorelowane. Chcąc odwołać się do tej prawidłowości powinniśmy to sprawdzić, np. na wykresie rozrzutu odchyłeń standardowych względem średnich (rycina 1). Jak widać w naszym przypadku punkty wyraźnie układają się według pewnej tendencji wzrostowej, a współczynnik korelacji liniowej Pearsona wyniósł 0,87 i był istotny statystycznie, a zatem w przypadku danych dotyczących wyskoków CMJ oraz ACMJ nie możemy oprzeć się na tej prawidłowości. Bardziej uzasadnionym w tym przypadku wydaje się być użycie testów nieparametrycznych. Dla porównania średniego czasu kontaktu stóp z matą w badanych grupach względem masy ciała można wykorzystać test Kruskala-Wallisa

(Feir-Walsh i Toothaker, 1974; Moder, 2010), natomiast do porównania CT w CMJ oraz ACMJ można skorzystać z testu kolejności par Wilcoxon (Stanisz, 2006).



Rycina 1. Zależność odchyłeń standardowych względem średnich czasu kontaktu stóp z matą dynamometryczną w wersji CJM.

Nie należy zapominać o tym, że prezentacja wyników w pracy naukowej jest równie ważna jak prawidłowo wykonana analiza statystyczna. Dla przykładu w przypadku używania testów nieparametrycznych istotną kwestią jest ich działanie na skali porządkowej, a zatem prezentując otrzymane wyniki w formie graficznej nie powinno tworzyć się wykresów ramka-wąsy z wartościami średnia – błąd standardowy – $1,96 \cdot \text{błąd standardowy}$, tylko użyć wartości mediana – 25/75 % - minimum/maksimum. Na rycinie 2 pokazano trzy rodzaje wykresu ramka-wąsy prezentujące te same dane, dotyczące czasu kontaktu stóp z matą dynamometryczną przy wyskoku CJM dla różnych grup wagowych. Pomimo bardziej reprezentatywnego wyglądu wykresu prezentującego średnie (ryc. 2b) do zobrazowania wyników testu Kruskala-Wallisa najodpowiedniejszy byłby wykres prezentujący mediany lub dane surowe (ryc. 2a, c) (McGill et al., 1978).



Rycina 2. Czas kontaktu stóp z matą dynamometryczną w wersji CJM z podziałem na klasy wagowe. w trzech wersjach wyglądu wykresu ramka-wąsy: (a) mediana – 25-75% – minimum/maksimum, (b) średnia – średnia \pm błąd standardowy – średnia \pm 1.96*błąd standardowy, (c) dane surowe wraz z wyszczególnionymi sytuacjami odstającymi i ekstremalnymi.

To co chcemy przekazać powinno być w pełni zrozumiałe dla czytelnika (Sokołowski, 2004), a sposób opisu statystycznego naszych rezultatów powinien być tak dobrany, aby uwidaczniać rzeczy najważniejsze i istotne dla udowodnienia wstępnych założeń opisywanych zjawisk, jednocześnie ukrywając informacje niewiele wnoszące do istoty sprawy.

References

1. Bartlett MS. The use of transformations. *Biometrics*. 1947; 3(1), 39-52.
2. Boneau CA. The effects of violations of assumptions underlying the t test. *Psychological bulletin*. 1960; 57(1), 49.
3. Chmielewska K., Chmielewski J., Szulc A. Statistical prediction methods in public health. *Acta Universitatis Matthiae Belii Physical Education and Sport*, 2014, Vol. 6 (2), 15-26
4. Cleophas TJ. Zwinderman AH. Non-parametric tests. Cleophas TJ. Zwinderman AH. *Statistical Analysis of Clinical Data on a Pocket Calculator*. Springer Netherlands, 2011
5. Fay MP, Proschan MA. Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Statistics surveys*. 2010; 4, 1.

6. Feir-Walsh BJ, Toothaker LE. An empirical comparison of the ANOVA F-test, normal scores test and Kruskal-Wallis test under violation of assumptions. *Educational and Psychological Measurement*. 1974; 34(4), 789-799.
7. Ferguson GA, Takane Y. *Analiza statystyczna w psychologii i pedagogice*. PWN. Warszawa; 1997.
8. Flanagan E. The Reactive Strength Index Revisited. (<https://www.trainwithpush.com/blog/reactive-strength-index-revisited>. 26.12. 2016).
9. Flanagan, E.P. and Comyns, T.M., 2008. The use of contact time and the reactive strength index to optimize fast stretch-shortening cycle training. *Strength & Conditioning Journal*, 30(5), 32-38.
10. Good PI, Hardin JW. *Common errors in statistics (and how to avoid them)*. John Wiley & Sons; 2012.
11. Jankowicz-Szymańska A, Wojtanowski W, Chronowski M, Ridan T. Porównanie motoryczności dzieci pełnosprawnych, niesłyszących i niepełnosprawnych intelektualnie w stopniu lekkim. *Instytut Rozwoju Służb Społecznych*. 2011.
12. Jankowska D, Majka A. Zastosowanie alternatywnych metod rozstrzygnięcia o istnieniu różnic pomiędzy średnimi w badaniach rolniczych. *Annales Universitatis Mariae Curie-Skłodowska. Sectio E, Agricultura*. 2009; 64(3), 50-57.
13. Keppel G, Wickens TD. *Design and Analysis: A researcher's handbook*. Upper Saddle River, NJ: Pearson Prentice Hall. 2004.
14. Lindman HR. *Analysis of variance in complex experimental designs*. WH Freeman & Co. 1974.
15. Lindman HR. *Analysis of variance in experimental design*. Springer, New York. 1991.
16. Łopatka M, Guzik-Kopyto A, Michnik R, Rycerski W. Badania antropometryczne kończyny górnej oraz pomiar siły ścisku dłoni i kciuka. *Aktualne Problemy Biomechaniki*. 2012.
17. McGill R, Tukey JW, Larsen WA. Variations of box plots. *The American Statistician*, 1978, 32.1: 12-16.
18. Michnik R, Jurkojć J, Czaplą K. Biomechaniczna ocena zdolności siłowych siatkarek. *Modelowanie Inżynierskie*. 2012; 13(44), 217-222.
19. Moder K. Alternatives to F-test in one way ANOVA in case of heterogeneity of variances (a simulation study). *Psychological Test and Assessment Modeling*. 2010; 52(4), 343-353.
20. Reinhart A. *Statistics Done Wrong: The Woefully Complete Guide*. No Starch Press. 2015.
21. Scariano SM, Davenport JM. The effects of violations of independence assumptions in the one-way ANOVA. *The American Statistician*. 1987; 41(2), 123-129.
22. Sheskin DJ. *Handbook of parametric and nonparametric statistical procedures*. crc Press, 2003.
23. Sobczyk M. *Statystyka*. Wydawnictwo Naukowe PWN. 1996.
24. Sokołowski A. O niewłaściwym stosowaniu metod statystycznych. *Materiały szkoleniowe. StatSoft Polska [online]*, 2004.
25. Stanisław A. Podstawy statystyki dla prowadzących badania naukowe. Odcinek 10: Testy istotności różnic dla prób zależnych. *Medycyna Praktyczna*. 1999, 7-8: 207-209.
26. Stanisław A. *Przystępny kurs statystyki: z zastosowaniem STATISTICA PL na przykładach z medycyny. Statystyki podstawowe*. StatSoft, 2006.

27. Stanisław A. Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny: Modele liniowe i nieliniowe. StatSoft, 2007.
28. StatSoft, Inc. (2014). STATISTICA (data analysis software system), version 12. www.statsoft.com.
29. Szreder M. O weryfikacji i falsyfikacji hipotez. Przegląd Statystyczny, 2010, 57.2-3: 82-88.
30. Szulc A. Biofizyczne i środowiskowe podstawy analizy składu ciała metodą BIA. Journal of Health Sciences - 2013, Vol. 3 (13), 410-427
31. Walesiak M. Strategie postępowania w badaniach statystycznych w przypadku zbioru zmiennych mierzonych na skalach różnego typu. 1993.
32. Watała C, Różalski M, Boncler M, Kaźmierczak P. Badania i publikacje w naukach biomedycznych: tom 1. Planowanie i prowadzenie badań, tom 2. Przygotowywanie publikacji; alfa-Medica Press, Bielsko-Biała. 2011.
33. Watała C, Zakład Zaburzeń Krzepnięcia Krwi KDŁ. Jak skutecznie wykorzystywać metody statystyczne w planowaniu i przeprowadzaniu eksperymentu naukowego? Statystyka i data mining w badaniach naukowych, StatSoft, Warszawa. 2005.
34. Wieczorek M. Rozwój fizyczny i sprawność fizyczna dzieci głuchych jako pozytywny miernik ich zdrowia fizycznego. Prob. Hig. Epidemiol. 2011; 92(2), 304-310.