



## WSPOMAGANIE STATYSTYCZNEJ ANALIZY WYNIKÓW BADAŃ EMPIRYCZNYCH W STATISTICA 9

*Janusz Wątroba, StatSoft Polska Sp. z o.o.*

Badania empiryczne to proces wieloetapowy. Dla poprawnej ich realizacji badacz musi rozwiązać wiele problemów, w tym także takich, które nie dotyczą bezpośrednio dziedziny, w której są one prowadzone. Taka sytuacja może mieć miejsce w przypadku potrzeby zastosowania metod analizy danych. Dwie stosunkowo najczęściej zgłaszane trudności to: (1) dobór odpowiednich metod opracowania danych oraz (2) poprawna interpretacja wyników analiz. Waga problemu wynika z tego, że zarówno nieodpowiedni wybór metod, jak i błędna interpretacja wyników prowadzą zwykle do formułowania błędnych wniosków końcowych z badań.

W artykule omówiono wybrane zagadnienia związane z powyższymi trudnościami. Zwrócono uwagę na kryteria decydujące o wyborze stosownych technik analizy danych oraz na przykładach zilustrowano sposób interpretacji wyników. Przy okazji zaprezentowano również narzędzia programu *STATISTICA*, wspomagające kolejne etapy analizy danych.

### **Problem doboru statystycznych narzędzi do poprawnej weryfikacji hipotez badawczych**

W przypadku dobrze zaplanowanego badania wybór odpowiednich metod analizy statystycznej powinien być dokonywany wstępnie w fazie jego projektowania. Metody te są wybierane w oparciu o specyficzne cele oraz kluczowe hipotezy, które mają podlegać weryfikacji. Jednakże ostateczny wybór musi być przeprowadzony po wcześniejszym sprawdzeniu założeń wymaganych przez poszczególne metody. W związku z tym zazwyczaj jest bardzo ważne, aby rozpocząć analizę od eksploracyjnej analizy danych, która zwykle wymaga utworzenia wykresów i obliczenia statystyk opisowych dla oceny rozkładów zmiennych, źródeł zmienności, powiązań pomiędzy zmiennymi oraz niezależności obserwacji. Wstępne badanie dotyczy zazwyczaj rozkładu zmiennej zależnej, a także badania pod kątem odstających obserwacji. Jeśli założenia nie są spełnione, wówczas plan analizy powinien uwzględniać metody alternatywne. Dane o charakterze dychotomicznym lub skategoryzowanym powinny być analizowane przy pomocy metod przeznaczonych dla danych jakościowych. Podobnie metody „klasyczne” są generalnie odpowiednie dla danych, które w przybliżeniu podlegają rozkładowi normalnemu oraz charakteryzują się średnią lub dużą



liczebnością. W niektórych przypadkach, np. dla danych dotyczących zliczeń lub przeżyć z obserwacjami uciętymi, wymagane są specjalne techniki modelowania. Każdy model lub metoda statystyczna posiada swój odrębny zestaw założeń, które powinny być sprawdzane łącznie z dopasowaniem modelu. Wymaga to zazwyczaj badania odchyleń pomiędzy wartościami rzeczywistymi i przewidywanymi przez model (odchylenia te są nazywane resztami). Za pomocą wykresów często udaje się zidentyfikować trendy lub inne problemy z rozkładem reszt.

Przed rozpoczęciem analizy należy również wziąć pod uwagę, że większość metod statystycznych zakłada, że badane obserwacje są niezależne. Założenie to jest naruszone dla obserwacji, które zostały zebrane dla tych samych jednostek. Czasami sytuacja taka ma też miejsce w przypadku, gdy dane są w określony sposób „zgrupowane”, np. pacjenci leczeni przez tego samego lekarza. W takich przypadkach model statystyczny wykorzystywany do analizy danych powinien zawierać efekt lub efekty losowe.

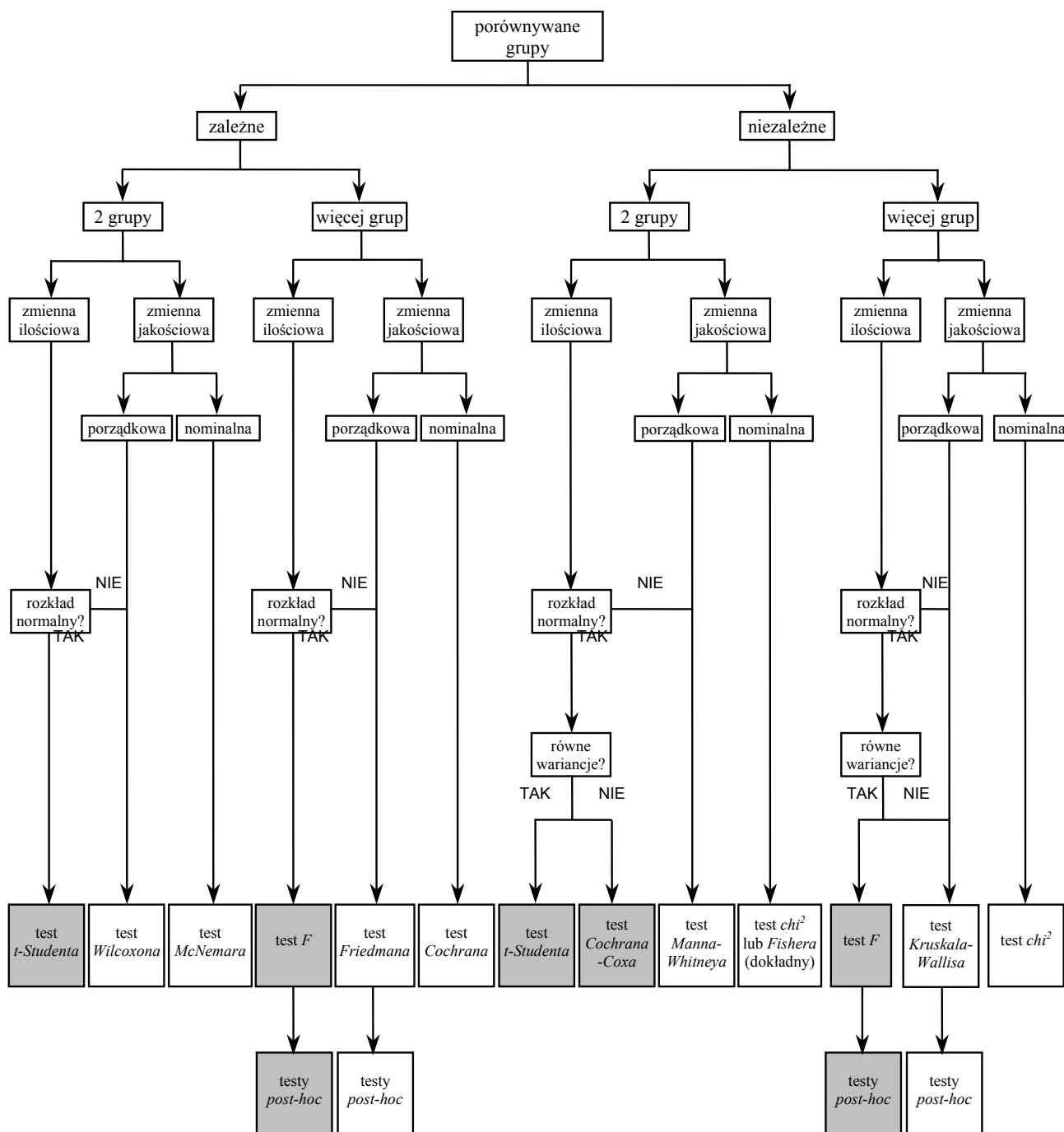
Trudno jest w ogólny sposób opisać analizę statystyczną, ponieważ każdy zbiór danych ma swoją specyfikę. Innym zagadnieniem, które często występuje zwłaszcza w przypadku badań medycznych, jest problem brakujących danych. Należy podkreślić, że sposób podejścia do takich danych zależy od tego, czy ich wystąpienie jest czysto losowe, czy też pojawiają się w sposób systematyczny. W tym drugim przypadku istnieje duże ryzyko wystąpienia zjawiska *obciążenia danych* (ang. *bias*).

Kluczowym zagadnieniem przy wyborze metod statystycznych, które mają być zastosowane, jest realizacja szczegółowych celów badawczych oraz sprawdzenie głównych hipotez. Po przełożeniu pytań i hipotez badawczych do postaci hipotez statystycznych trzeba zbadać inne związane z danymi warunki. W przypadku wielu badań występuje wiele zmiennych zależnych, stąd też analizę prowadzi się zazwyczaj odrębnie dla każdej z uwzględnionych zmiennych.

W przypadku gdy celem badawczym jest ocena zróżnicowania przeciętnego poziomu badanych cech w grupach badanych jednostek, do oceny statystycznej istotności różnic stosuje się tzw. *testy istotności*. Jest kilka kryteriów, które należy brać pod uwagę przy wyborze odpowiedniego testu. Pierwsze kryterium dotyczy schematu badań. Mogą być dwa rodzaje schematów badań: (1) zaplanowane w grupach zależnych oraz (2) zaplanowane w grupach niezależnych. Kolejne kryterium to liczba porównywanych grup (2 lub więcej niż 2). Następne kryteria dotyczą skali, na jakiej mierzono zmienną zależną (ilościowa, porządkowa lub nominalna), oraz założeń dotyczących normalności rozkładu badanej zmiennej i równości wariancji w obrębie porównywanych grup badanych jednostek.

Wymienione kryteria można przedstawić w postaci uproszczonego schematu, który w efektywny sposób wspomaga wybór odpowiedniego testu istotności.

Dla zilustrowania użyteczności schematu przedstawionego na rys. 1 zaprezentowane zostaną przykłady doboru testów istotności, analizy w środowisku *STATISTICA* oraz interpretacje wyników.



Rys. 1. Schemat pomocniczy do wyboru odpowiednich testów istotności różnic (testy parametryczne zaszarzone, testy nieparametryczne niezaszarzone).

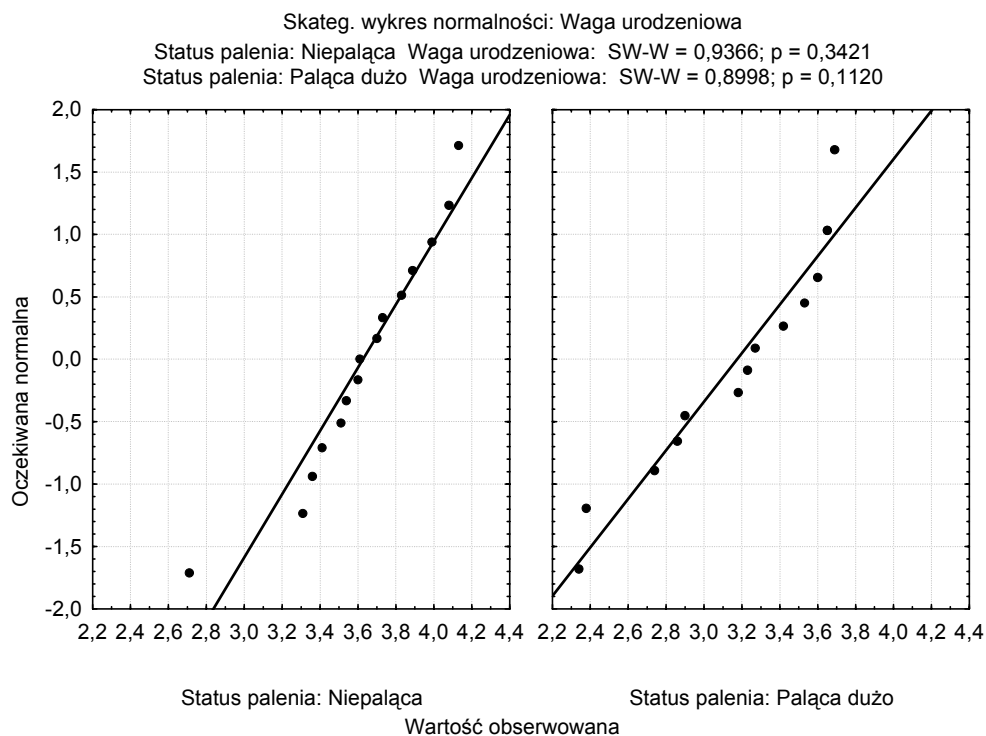
### *Czy palenie w czasie ciąży różnicuje wagę urodzeniową noworodków*

Pierwszy z prezentowanych przykładów dotyczy oceny zróżnicowania przeciętnej wagi urodzeniowej pomiędzy noworodkami, których matki w czasie ciąży nie paliły lub paliły dużo papierosów. Dane zaczerpnięto z podręcznika Kirkwood i Sterne'a [2].



W tym przykładzie mamy do czynienia z dwiema niezależnymi grupami noworodków: (1) noworodki, których matki nie paliły w czasie ciąży oraz (2) noworodki, których matki paliły dużo w czasie ciąży. Waga urodzeniowa jest przykładem zmiennej wyrażonej na skali ilościowej (liczbowej). W takiej sytuacji należy w pierwszej kolejności sprawdzić, czy badana zmienna w porównywanych grupach podlega rozkładowi normalnemu. Jeśli zgromadzone wyniki sugerowałyby odrzucenie tego założenia, wówczas do oceny istotności różnicowania należało by zastosować nieparametryczny *test Manna-Whitneya*. W przypadku braku podstaw do odrzucenia założenia normalności należy dodatkowo sprawdzić równość wariancji w porównywanych grupach. Przy równych wariancjach do oceny istotności różnic należy zastosować *test t-Studenta* dla grup niezależnych. W przypadku nierównych wariancji odpowiednim testem jest *test Cochra-na-Coxa*.

Zgodnie z powyższymi uwagami, w pierwszej kolejności oceniono normalność rozkładu badanej zmiennej w porównywanych grupach noworodków. W tym celu utworzono wykresy normalności dla porównywanych grup oraz przeprowadzono test normalności Shapiro-Wilka. Wyniki przedstawia poniższy rysunek.



Rys. 2. Wykresy normalności z wynikami testu normalności Shapiro-Wilka.

Na podstawie wyników oceny normalności rozkładu widać, że do oceny istotności różnic należy użyć jednego spośród dwóch testów parametrycznych: *t-Studenta* lub *Cochrana-Coxa*. O wyborze decyduje ocena równości wariancji. Dokonano jej w oparciu o wyniki *testu Levene'a* i *testu Browna-Forsythe'a*. Wyniki zawiera tabela poniżej.



Zmienna	Testy t; Grupująca: Status palenia (Waga noworodków) Grupa 1: Niepaląca, Grupa 2 Paląca dużo					
	Levene'a F(1,df)	df Levene'a	p Levene'a	Brn-Fors F(1,df)	df Brn-Fors	p Brn-Fors
<b>Waga urodzeniowa</b>	<b>1,721061</b>	<b>27</b>	<b>0,200602</b>	<b>1,342601</b>	<b>27</b>	<b>0,256722</b>

Rys. 3. Wyniki testów do oceny równości wariancji.

Wyniki przeprowadzonych testów nie nakazują odrzucenia założenia równości wariancji. A zatem do oceny istotności różnicowania przeciętnej wagi urodzeniowej noworodków w porównywanych grupach można zastosować parametryczny *test t-Studenta* dla prób niezależnych.

Zmienna	Testy t; Grupująca: Status palenia (Waga noworodków) Grupa 1: Niepaląca, Grupa 2 Paląca dużo				
	Średnia Niepaląca	Średnia Paląca dużo	t	df	p
<b>Waga urodzeniowa</b>	<b>3,626800</b>	<b>3,174286</b>	<b>2,953730</b>	<b>27</b>	<b>0,006434</b>

Rys. 4. Wyniki *testu t-Studenta*.

Na podstawie zamieszczonych w tabeli wyników możemy stwierdzić, że oceniana różnica w przeciętnej wadze noworodków okazała się istotna statystycznie. Otrzymany wynik może sugerować, że palenie w czasie ciąży obniża wagę urodzeniową dziecka. Na podstawie przeprowadzonej analizy można ponadto podać punktową i przedziałową ocenę oczekiwanego efektu różnicowania.

Zmienna	Testy t; Grupująca: Status palenia (Waga noworodków) Grupa 1: Niepaląca, Grupa 2 Paląca dużo		
	Średnia 1 - Średnia 2	Ufność -95,000%	Ufność +95,000%
<b>Waga urodzeniowa</b>	<b>0,452514</b>	<b>0,138172</b>	<b>0,766857</b>

Rys. 5. Punktowa i przedziałowa ocena efektu różnicowania.

Jak widać, oczekiwana różnica wyniosła 0,45 kg. Przedział ufności dla różnicy (95%) obejmuje zakres (0,1382-0,7669). Na tej podstawie można przeprowadzić dalszą, kliniczną ocenę wyników przeprowadzonej analizy.

### ***Ocena skuteczności działania leku nasennego***

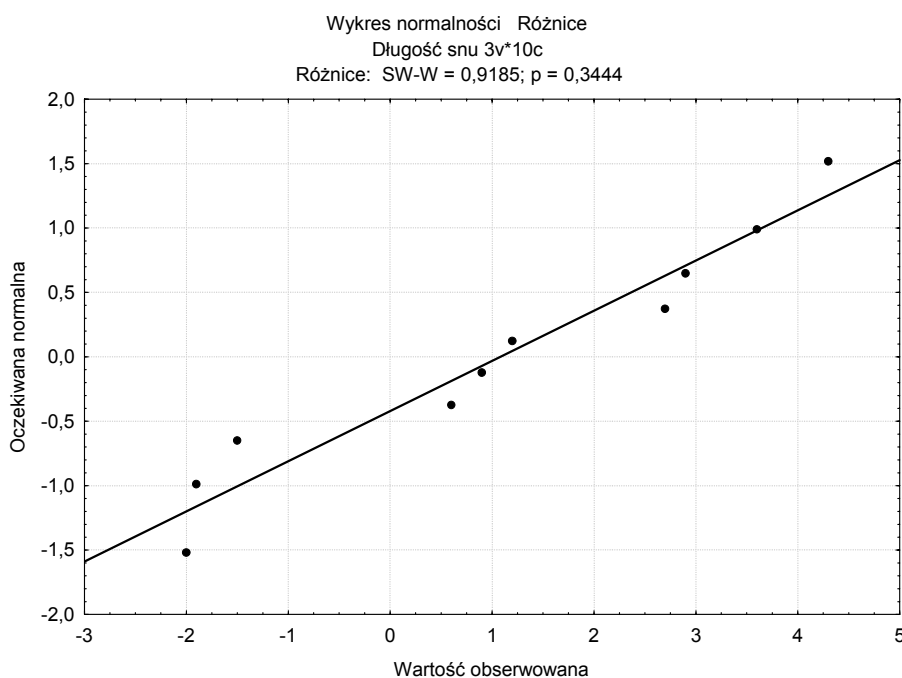
Dla oceny skuteczności działania pewnego leku nasennego przeprowadzono badania w losowo dobranej grupie 10 pacjentów, u których obserwowano problemy ze snem. U każdej z badanych osób obserwowano czas snu po podaniu leku nasennego i placebo (kolejność podawania środków też była ustalana losowo). Dane zaczerpnięto z podręcznika Kirkwood i Sterne'a [2].

Z opisu wynika, że badanie zostało zrealizowane w schemacie dla prób zależnych. Dane zostały zebrane w postaci dwóch serii pomiarów, po jednym pomiarze dla każdego z badanych pacjentów. Skuteczność leku w porównaniu do placebo oceniano czasem snu. Jest to zmienna typu ilościowego. W takiej sytuacji przed wyborem odpowiedniego testu należy



sprawdzić rozkład różnic pomiarów. Jeśli rozkład odbiega od normalnego, wówczas do oceny statystycznej istotności zróżnicowania przeciętnego czasu snu należy zastosować nieparametryczny *test Wilcoxona*. W przeciwnym przypadku odpowiednim testem jest parametryczny *test t-Studenta* dla prób zależnych.

Wyniki testowania normalności rozkładu różnic czasów trwania snu przedstawiono na wykresie normalności.



Rys. 6. Wykres normalności z wynikami testu normalności Shapiro-Wilka.

Wyniki wskazują na brak podstaw do odrzucenia założenia dotyczącego normalności rozkładu badanych różnic. Tak więc do oceny statystycznej istotności różnic przeciętnych czasów trwania snu przy stosowaniu ocenianego leku w stosunku do placebo należy zastosować parametryczny *test t-Studenta* dla prób zależnych.

Zmienna	Test T dla prób zależnych (Długość snu)									
	Zaznaczone różnice są istotne z $p < ,05000$									
	Średnia	Odch.st.	Ważnych	Różnica	Odch.st. Różnica	t	df	p	Ufność -95,000%	Ufność +95,000%
Lek nasenny	6,660000	0,879646								
Placebo	5,580000	1,520818	10	1,080000	2,308342	1,479529	9	0,173128	-0,571289	2,731289

Rys. 7. Wyniki *testu t-Studenta* dla prób zależnych.

Przedstawione w tabeli wyniki wskazują na brak podstaw do odrzucenia hipotezy o braku zróżnicowania przeciętnej długości snu przy podawaniu badanego leku w stosunku do placebo ( $p=0,1731$ ), mimo iż różnica pomiędzy odpowiednimi średnimi wyniosła 1,08 godz.

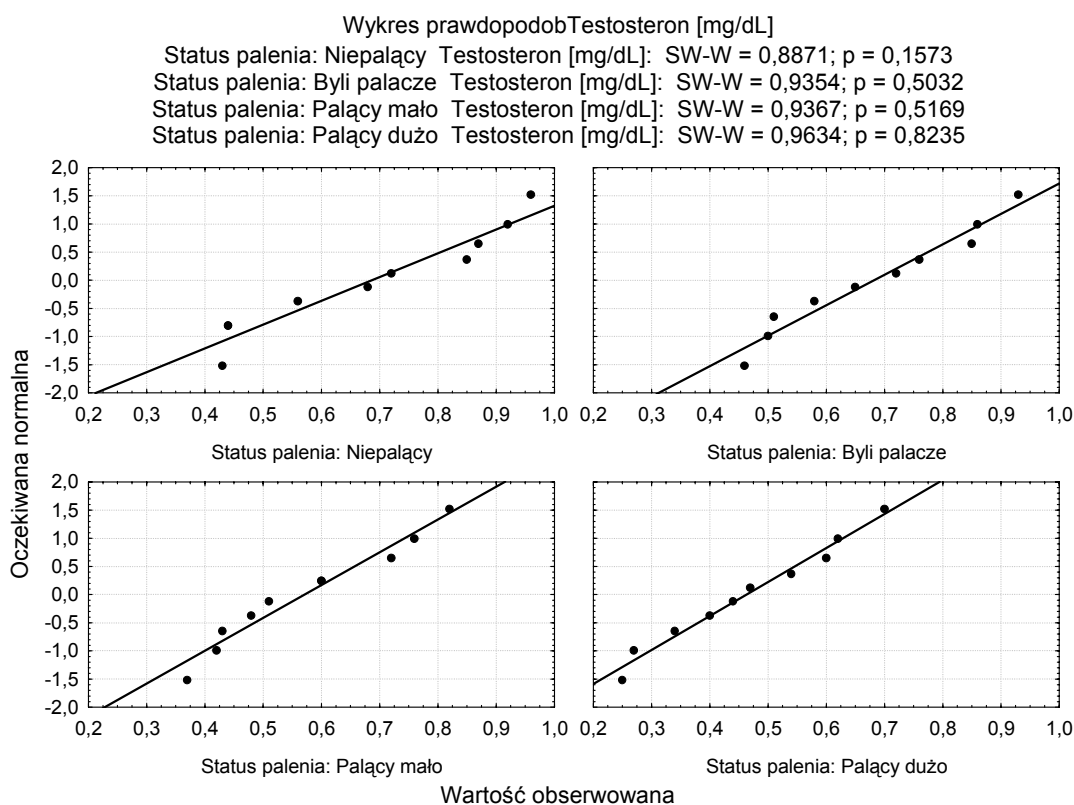


## Czy status palenia papierosów różnicuje przeciętny poziom testosteronu u mężczyzn

W celu odpowiedzi na to pytanie przeprowadzono badania wśród mężczyzn w wieku od 35 do 45 lat, którzy różnili się statusem palenia. Wzięto pod uwagę cztery grupy mężczyzn: niepalący (mężczyźni, którzy nigdy nie palili), byli palący (osoby, które przestały palić przynajmniej sześć miesięcy przed badaniami), palący mało (osoby, które paliły 10 papierosów dziennie lub mniej) i palący dużo (osoby, które paliły 30 lub więcej papierosów dziennie). Dane pochodzą z książki Le [5].

Z podanego wyżej opisu wynika, że mamy do czynienia z badaniem, w którym wystąpiły cztery niezależne grupy porównawcze. W takiej sytuacji do oceny zróżnicowania należy zastosować parametryczny *test F* lub nieparametryczny *test Kruskala-Wallisa*. O wyborze odpowiedniego testu decyduje spełnienie dwóch założeń: normalności rozkładu oraz równości wariancji.

Ocenę pierwszego z powyższych założeń przeprowadzono na podstawie wykresu normalności oraz *testu Shapiro-Wilka*.



Rys. 8. Wykresy normalności z wynikami testu normalności Shapiro-Wilka.

Dla oceny równości wariancji przeprowadzono dwa testy: *test Levene'a* oraz *test Browna-Forsythe'a*. Wyniki przedstawiono poniżej.



Test Levene'a jednorodności wariancji (Testosteron a palenie)								
Zaznaczone efekty są istotne z $p < ,05000$								
Zmienna	SS Efekt	df Efekt	MS Efekt	SS Błąd	df Błąd	MS Błąd	F	p
Testosteron [mg/dL]	0,017527	3	0,005842	0,238440	36	0,006623	0,882109	0,459521

Test jednorod. wariancji Browna-Forsythe'a (Testosteron a palenie)								
Zaznaczone efekty są istotne z $p < ,05000$								
Zmienna	SS Efekt	df Efekt	MS Efekt	SS Błąd	df Błąd	MS Błąd	F	p
Testosteron [mg/dL]	0,017527	3	0,005842	0,243420	36	0,006762	0,864062	0,468592

Rys. 9. Wyniki testów oceniających jednorodność wariancji.

Jak widać, obydwie brane pod uwagę przy wyborze testu istotności różnic założenia są spełnione (odpowiednie wartości prawdopodobieństw testowych znacznie przekraczają założony poziom istotności  $\alpha=0,05$ ). W takiej sytuacji do oceny statystycznej istotności różnic w przeciętnym poziomie testosteronu pomiędzy badanymi grupami mężczyzn należy zastosować *test F* (jednoczynnikowa analiza wariancji). Wyniki przedstawia poniższa tabela wyników analizy wariancji.

Analiza wariancji (Testosteron a palenie)								
Zaznaczone efekty są istotne z $p < ,05000$								
Zmienna	SS Efekt	df Efekt	MS Efekt	SS Błąd	df Błąd	MS Błąd	F	p
Testosteron [mg/dL]	0,339007	3	0,113002	1,071070	36	0,029752	3,798155	0,018294

Rys. 10. Tabela z wynikami jednoczynnikowej analizy wariancji.

Przedstawione wyniki sugerują, że hipotezę zakładającą równość średnich należy odrzucić. Poziom prawdopodobieństwa testowego  $p$  wyniósł  $0,018294$ . Jednocześnie otrzymane wyniki uprawniają do przeprowadzenia testów porównań typu *post-hoc*, które pozwalają badaczowi na stwierdzenie, które z badanych grup różnią się istotnie pod względem przeciętnego poziomu testosteronu. Do tego celu użyto dwóch testów (*test Newman-Keulsa* oraz *test Tukeya*), które w dużym stopniu gwarantują badaczowi utrzymanie błędu I rodzaju na przyjętym poziomie istotności przy przeprowadzaniu wielu jednoczesnych porównań [1, 4]. Wyniki przedstawiono poniżej.

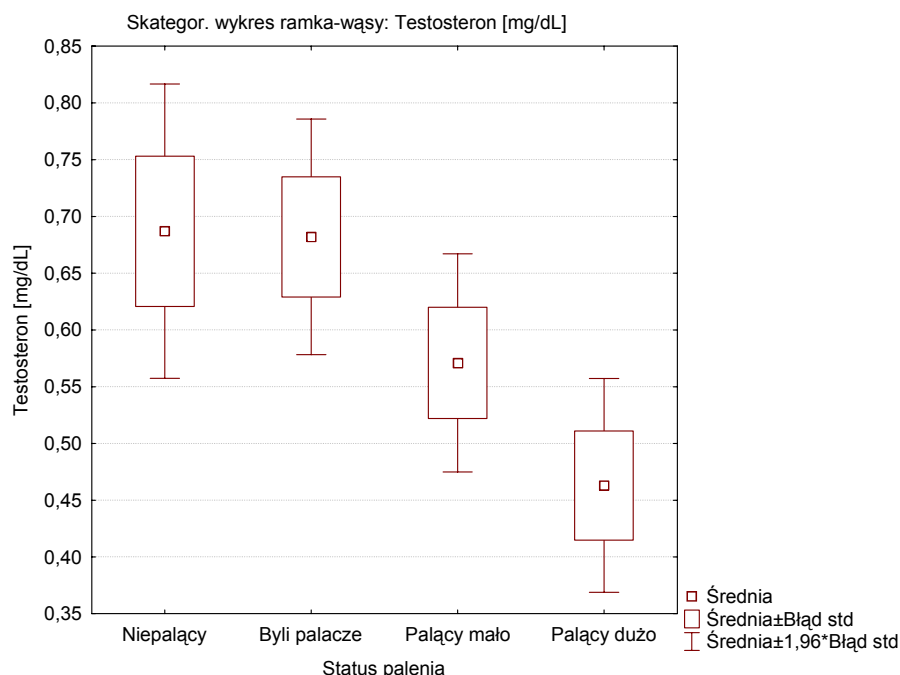
Status palenia	Test Newman-Keulsa; zm.: (Testosteron a)			
	{1}	{2}	{3}	{4}
Niepalący {1}	M=,68700	M=,68200	M=,57100	M=,46300
Byli palacze {2}	0,948759	0,948759	0,301245	0,030537
Palący mało {3}	0,301245	0,158917	0,158917	0,019843
Palący dużo {4}	0,030537	0,019843	0,170163	0,170163

Status palenia	Test RIR Tukeya; zmienna: Testosteron [mg/			
	{1}	{2}	{3}	{4}
Niepalący {1}	M=,68700	M=,68200	M=,57100	M=,46300
Byli palacze {2}	0,999912	0,999912	0,445923	0,030537
Palący mało {3}	0,445923	0,484154	0,484154	0,035666
Palący dużo {4}	0,030537	0,035666	0,507580	0,507580

Rys. 11. Wyniki testów porównań wielokrotnych.

Obydwa z przeprowadzonych testów dają w efekcie podobne wyniki. Tylko różnice pomiędzy niepalącymi i palącymi dużo oraz pomiędzy byłymi palaczami a palącymi dużo okazały się statystycznie istotne przy przyjęciu poziomu istotności  $\alpha=0,05$ . Pozostałe różnice były nieistotne. Zamieszczony poniżej wykres ramkowy ilustruje dodatkowo różnice pomiędzy badanymi grupami osób.



Rys. 12. Wykresy ramkowe w obrębie badanych grup osób.

Przedstawione wyniki analizy mogą posłużyć do przeprowadzenia dalszej, merytorycznej oceny zaobserwowanych różnic.

## Przykład modelowania nieliniowych związków pomiędzy zmiennymi

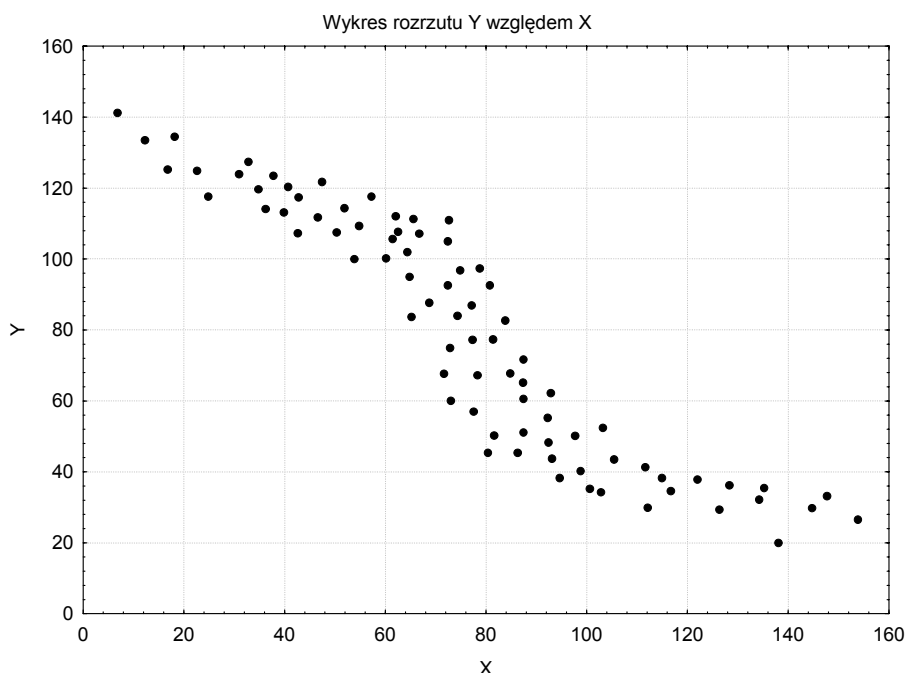
Zagadnienie oceny powiązań występujących pomiędzy zmiennymi opisującymi określone zjawisko lub proces jest często podejmowane w badaniach empirycznych. Zwykle towarzyszą temu dwa główne cele. Chodzi o lepsze poznanie mechanizmu generującego badane zjawisko oraz o ewentualne wykorzystanie modelu do przewidywania przebiegu zjawiska przy założeniu innych warunków.

Bardzo często przy modelowaniu stosuje się daleko idące uproszczenie rzeczywistości. Dzięki temu otrzymane modele są łatwiejsze w interpretacji. Jednakże zdarzają się również sytuacje, w których zjawiska są bardziej złożone i wówczas zachodzi potrzeba zastosowania narzędzi, które w adekwatny sposób pozwolą opisać przebieg modelowanych zjawisk [3].

Dla ilustracji budowy modeli dla zależności typu nieliniowego wygenerowano przykładowe dane. Zamieszczony poniżej wykres (zob. rys. 13) pokazuje rozmieszczenie danych w dwuwymiarowym układzie współrzędnych XY.

W pierwszej kolejności do danych dopasowano funkcję logistyczną o postaci:

$$f(x) = a - \frac{b}{1 + \exp(c \cdot (x - d))}$$



Rys. 13. Wykres rozrzutu zmiennych X i Y.

Poniżej zamieszczono tabelę zawierającą oceny parametrów modelu.

Model: $y=a-b/(1+\exp(c*(x-d)))$ (Reg_logist_dane_1)						
Zmn. zal. : Y						
Poziom ufności: 95.0% ( alfa=0.050)						
	Ocena	Błąd stand.	Wart. t df = 76	p	Doln. uf Granica	Górn. uf Granica
a	28,9202	3,771169	7,6688	0,000000	21,409	36,4312
b	-98,0809	6,151194	-15,9450	0,000000	-110,332	-85,8297
c	0,0730	0,010231	7,1314	0,000000	0,053	0,0933
d	77,0736	1,869261	41,2321	0,000000	73,351	80,7966

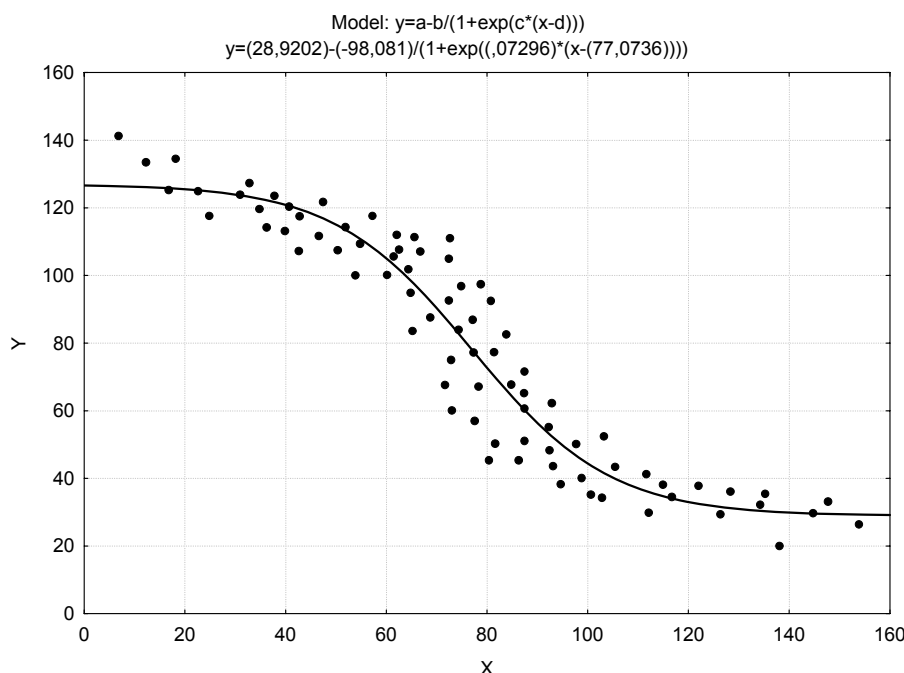
Rys. 14. Oceny parametrów dopasowanego modelu.

Dopasowany model można także przedstawić na wykresie (zob. rys. 15).

Biorąc pod uwagę charakter modelowanego zjawiska, funkcja ta ma jedną wadę. Można nią zamodelować jedynie sytuację, w której następuje zmiana z pewnego poziomu (stałego) mierzonej wielkości do innego poziomu, również o stałej wartości. W przypadku danych z opisywanego przykładu ich przebieg jest nieco inny. Widać dosyć wyraźnie, że zarówno przed zmianą poziomu obserwowanego zjawiska, jak i po tej zmianie występują raczej powolne zmiany poziomu modelowanego zjawiska niż stały poziom.

W związku z tym zaproponowano zmodyfikowaną funkcję logistyczną o postaci:

$$f(x) = p \cdot x + a - \frac{t \cdot x + b}{1 + \exp(c \cdot (x - d))}$$



Rys. 15. Oceny parametrów dopasowanego modelu.

Warto zauważyć, że w przypadku, gdy  $p = 0$  oraz  $t = 0$ , mamy sytuację, w której model pokrywa się z funkcją logistyczną. Krzywa opisuje wtedy przejście między dwoma stałymi poziomami (poziomem o wartościach:  $a$  i  $a - b$ ).

Znaczenie poszczególnych parametrów w sytuacji ogólnej:

$d$  - punkt przegięcia,

$c$  - wpływa na nachylenie odcinka,

$p, a, t, b$  - odpowiadają za trendy przed przegięciem i po przegięciu,

mianowicie, gdy  $c > 0$ , dopasowana krzywa ma asymptotę w  $+\infty$  postaci  $px + a$ , a w  $-\infty$  postaci  $(p - t) \cdot x + a - b$ . W sytuacji gdy  $c < 0$ , na odwrót, dopasowana krzywa ma asymptotę w  $+\infty$  postaci  $(p - t) \cdot x + a - b$ , a w  $-\infty$  postaci  $p \cdot x + a$ .

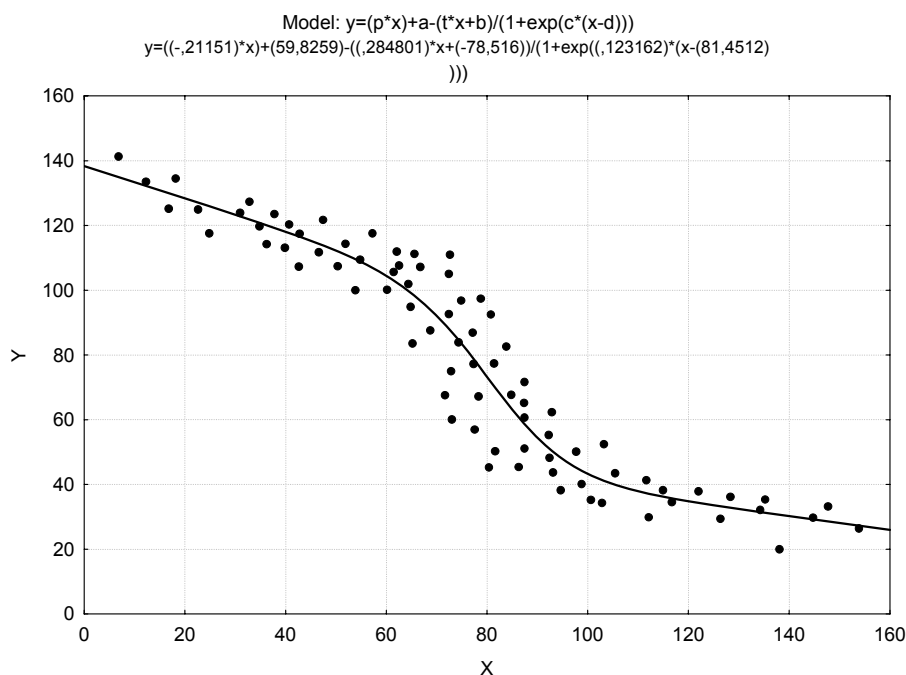
Tak zdefiniowany model dopasowano do danych. Wyniki modelowania przedstawiono w tabeli.

Model: $y=(p*x)+a-(t*x+b)/(1+\exp(c*(x-d)))$ (Reg_logist_dane_1)						
Zmn. zal. : Y						
Poziom ufności: 95.0% ( alfa=0.050)						
	Ocena	Błąd stand.	Wart. t df = 74	p	Doln. uf Granica	Górn. uf Granica
<b>p</b>	-0,2115	0,20692	-1,02221	0,310012	-0,624	0,2008
<b>a</b>	59,8259	27,02280	2,21390	0,029916	5,982	113,6700
<b>t</b>	0,2848	0,22678	1,25583	0,213129	-0,167	0,7367
<b>b</b>	-78,5156	26,11680	-3,00632	0,003610	-130,554	-26,4767
<b>c</b>	0,1232	0,04625	2,66269	0,009506	0,031	0,2153
<b>d</b>	81,4512	3,91197	20,82099	0,000000	73,656	89,2460

Rys. 16. Oceny parametrów dopasowanego modelu.



Tak jak w przypadku poprzedniego modelu wyniki modelowania przedstawiono również na wykresie.



Rys. 17. Oceny parametrów dopasowanego modelu.

Jak widać, model lepiej oddaje przebieg modelowanych danych. Dzięki temu badacz jest w stanie w sposób bardziej dokładny opisać modelowane zjawisko. Dopasowanie modelu lepiej odzwierciedlającego rzeczywiste dane umożliwia jego ewentualne wykorzystanie do przeprowadzania symulacji zjawiska w zależności od różnych warunków, w których ono przebiega.

## Podsumowanie

W pierwszej części przedstawiono kryteria decydujące o wyborze odpowiedniego testu istotności różnic. Przytoczono schemat ułatwiający wybór konkretnego testu parametrycznego lub nieparametrycznego. W przykładach pokazano sposób doboru testów oraz przebieg analizy w programie *STATISTICA*.

W drugiej części zaprezentowano dopasowanie do danych dwóch modeli. Pierwszy z nich to funkcja logistyczna, a drugi to zmodyfikowana funkcja logistyczna. Obydwie postaci funkcji mogą być wykorzystywane do opisu szerokiej klasy zjawisk przyrodniczych, ekonomicznych lub społecznych.



## Literatura

1. Cobb G.W., Introduction to Design and Analysis of Experiments, Springer, 1998.
2. Kirkwood B.R., Sterne J.A.C., *Essential Medical Statistics*, 2-nd ed., Blackwell Science, 2003.
3. Krzanowski W.J., An Introduction to Statistical Modelling, Arnold, 1998
4. Maxwell S.E., Designing Experiments and Analyzing Data. A Model Comparison Perspective, 2-nd ed., Lawrence Erlbaum Associates, 2004.
5. Le Ch.T., Introductory Biostatistics, Wiley-Interscience, 2003.