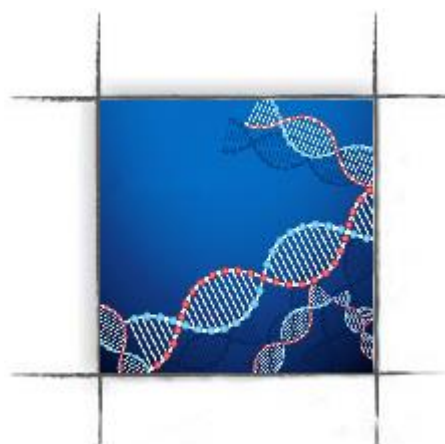


# StatSoft Polska



## *Zestaw Przyrodnika 1.0*

Podstawowe informacje o programie



## SPIS TREŚCI







<b>1. OPIS OGÓLNY</b>	<b>4</b>
1.1 Przegląd modułów programu	4
1.2 Zbiory danych	4
<b>2. PODSTAWOWE</b>	<b>5</b>
2.1 Statystyki podstawowe	5
2.2 Kreator regresji liniowej	6
2.2.1. Określenie rodzaju danych, wybór zmiennych	6
2.2.2. Transformacje zmiennych	6
2.2.3. Diagnostyka zmiennych	9
2.2.4. Współliniowość i liniowość	10
2.2.5. Interakcje	12
2.2.6. Budowa modelu wieloczynnikowego	13
2.2.7. Ocena wyników modelu wieloczynnikowego	14
2.3 Kreator testów	16
2.3.1. Możliwości programu	17
2.3.2. Testy dla pojedynczej zmiennej	17
2.3.3. Badanie istotności różnic	18
2.3.4. Dodatkowe możliwości programu	18
<b>3. UKŁADY CZYNNIKOWE ANOVA</b>	<b>20</b>
3.1. Doświadczenia jednoczynnikowe	21
3.1.1. Kwadrat łaciński	21
3.1.2. Bloki losowe	22
3.1.3. Układ całkowicie losowy	23
3.2. Doświadczenia dwuczynnikowe	24
3.2.1. Split-plot	24
3.2.2. Split-blok	25
3.2.3. Bloki losowe	26
3.2.4. Układ całkowicie losowy	27
3.3. Doświadczenia trójczynnikowe	28
3.3.1. Split-split-plot	28
3.3.2. Split-plot A-BC	30
3.3.3. Split-plot AB-C	31
3.3.4. Split-plot x split-blok	31
3.3.5. Split-plot-blok	33
3.3.6. Split-blok-plot	34
3.3.7. Bloki losowe	35
3.3.8. Układ całkowicie losowy	35
<b>4. ZAAWANSOWANE</b>	<b>36</b>
4.1 Doświadczenia kratowe	36
4.2 Serie doświadczeń	37
4.2.1. Układ AMMI	38
4.2.2. Serie doświadczeń split-plot	39

4.2.3. Serie bloków losowych	40
<b>4.3 Doświadczenia hodowlane</b>	<b>40</b>
4.3.1. Doświadczenia linia x tester	40
4.3.2. Diallel kwadratowy	41
4.3.3. Diallel trójkątny	41
4.3.4. Diallel kwadratowy bez przekątnej	42
4.3.5. Diallel trójkątny bez przekątnej	42
<b>4.4 Doświadczenia z wzorcem</b>	<b>43</b>
<b>4.5 AMOVA</b>	<b>45</b>
<b>5. WIELOWYMIAROWE</b>	<b>47</b>
5.1 Analiza aglomeracji	47
<b>6. LITERATURA</b>	<b>48</b>

## 1. Opis ogólny

### 1.1 Przegląd modułów programu

**Zestaw Przyrodnika** jest zbiorem modułów analitycznych i kreatorów służących do opracowania wyników badań i planowania doświadczeń w naukach przyrodniczych. Program składa się z trzech grup modułów:

				Doświadczenia kratowe Serie doświadczeń Doświadczenia hodowlane	Doświadczenia z wzorcem AMOVA		
Statystyki podstawowe	Układy ANOVA	Kreator regresji liniowej	Kreator testów			Analiza aglomeracji	O programie
Podstawowe				Zaawansowane		Wielowymiarowe	

- Podstawowe
  - Statystyki podstawowe
  - Układy ANOVA
  - Kreator regresji liniowej
  - Kreator testów
- Zaawansowane
  - Doświadczenia kratowe
  - Serie doświadczeń
  - Doświadczenia hodowlane
  - Doświadczenia z wzorcem
  - AMOVA
- Wielowymiarowe
  - Analiza aglomeracji

Grupa **Podstawowe** zawiera zestaw podstawowych i zazwyczaj najczęściej wykorzystywanych czterech modułów, pozwalających na obliczenie statystyk opisowych, przeprowadzenie testów istotności różnic oraz całościowej analizy regresji liniowej czy sprawdzenie normalności rozkładu. Są to czynności niezbędne w statystycznej analizie danych w każdej dziedzinie. Poza tym grupa ta zawiera moduł *Układy ANOVA* dający możliwość analizy i planowania 15 układów doświadczeń jedno-, dwu- i trójczynnikaowych swoistych dla nauk przyrodniczych, jak np. split-plot.

Grupa **Zaawansowane** składa się z modułów przeznaczonych do analiz będących w większości pewnymi wariantami analizy wariancji, stosowanych w naukach przyrodniczych. Wśród nich znajduje się *AMOVA* czyli analiza wariancji molekularnej, do danych genetycznych, oraz *Doświadczenia kratowe*, *Serie doświadczeń*, *Doświadczenia hodowlane* i *Doświadczenia z wzorcem* służące do analiz - i w niektórych przypadkach także planowania - upraw i hodowli.

Grupa **Wielowymiarowe** zawiera (w bieżącej wersji programu) pojedynczy moduł *Analiza aglomeracji*. Rozszerza on zakres stosowalności analizy aglomeracji względem podstawowego pakietu Statistica, jako, że zawiera osiem dodatkowych miar odległości wykorzystywanych w naukach przyrodniczych szczególnie do zerojedynkowych danych pochodzących z analiz genetycznych, m.in. odległość Ochiai czy Gowera.

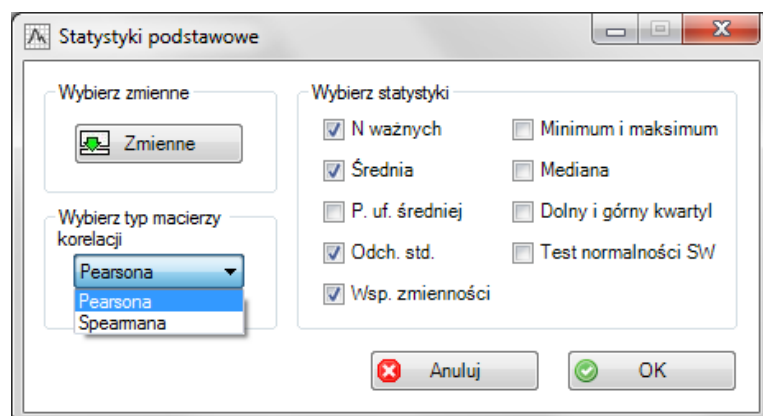
### 1.2 Zbiory danych

Wszystkie przykłady przedstawione w tej dokumentacji bazują na zestawie plików znajdujących się w katalogu *Zbiory danych*. Zawiera on przykładowe arkusze Statistica używane podczas wykonywania opisanych w kolejnych rozdziałach przykładów przedstawiających funkcjonalność odpowiednich modułów **Zestawu Przyrodnika**.

## 2. Podstawowe

### 2.1 Statystyki podstawowe

W module *Statystyki podstawowe* można dla zmiennych ilościowych obliczyć wybrane statystyki opisowe, macierze korelacji i sprawdzić normalność rozkładu, a dla zmiennych jakościowych można utworzyć tabele liczości.



Statystyki liczbowe, które możemy obliczyć dla wskazanego zestawu zmiennych ilościowych to:

- liczba ważnych przypadków (N ważnych)
- średnia
- 95% przedział ufności wartości średniej (zakłada normalność rozkładu)
- odchylenie standardowe
- współczynnik zmienności
- minimum i maksimum
- mediana
- dolny i górny kwartył
- wynik testu Shapiro-Wilka normalności rozkładu (jako wartość p)

Jeżeli wskażemy co najmniej dwie zmienne ilościowe, to oprócz wybranych statystyk opisowych, zostanie utworzona macierz współczynników korelacji liniowej (Pearsona) lub korelacji rang (Spearmana), co wybieramy przyciskiem *Wybierz typ macierzy korelacji*.

Dla wszystkich wskazanych zmiennych jakościowych zostaną utworzone tabele liczości.

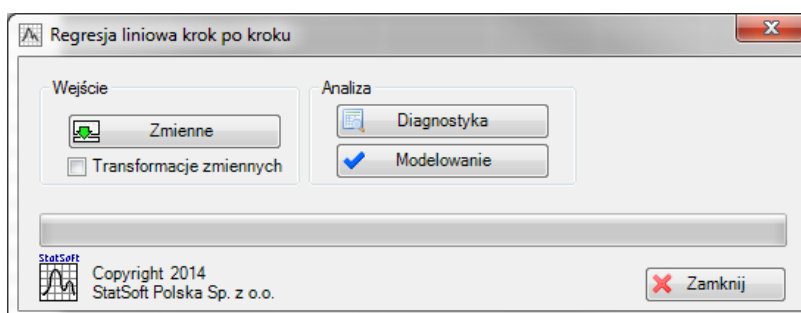
## 2.2 Kreator regresji liniowej

**Kreator regresji liniowej** jest narzędziem mającym na celu przeprowadzenie kompletnej analizy regresji liniowej krok po kroku, począwszy od transformacji zmiennych, poprzez narzędzia do diagnostyki i selekcji predyktorów, aż po narzędzia do oceny jakości modelu oraz oceny założeń związanych z budową modelu liniowego.

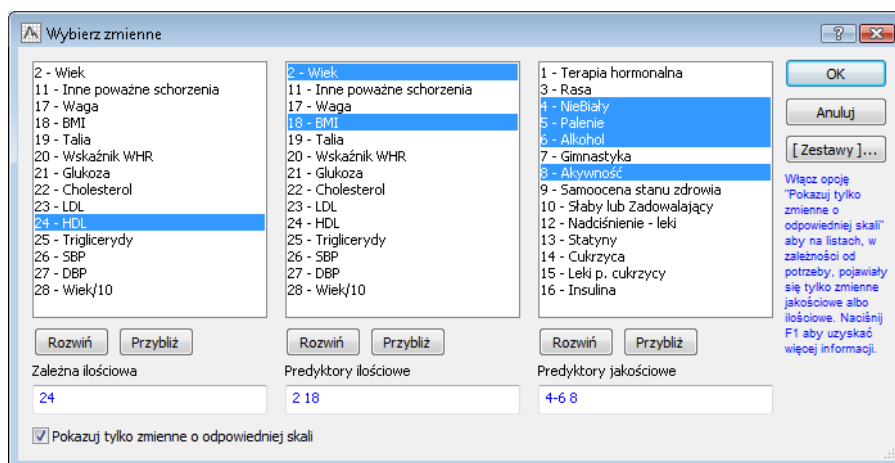
Możliwości *Kreatora* zostaną zaprezentowane na przykładzie pliku *HERS\_PL.sta* zawierającego dane z badań klinicznych HERS (*Heart and Estrogen/Progestin Study*), dotyczących badań nad skutecznością terapii hormonalnej w zapobieganiu nawracającym atakom serca oraz śmierci u 2763 kobiet po menopauzie ze zdiagnozowaną chorobą niedokrwienną serca.

### 2.2.1. Określenie rodzaju danych, wybór zmiennych

Aby rozpocząć analizę, z menu **Zestaw Plus** z grupy **Analizy** wybieramy opcję **Kreator regresji liniowej**, otwierając okno **Regresja liniowa krok po kroku**.



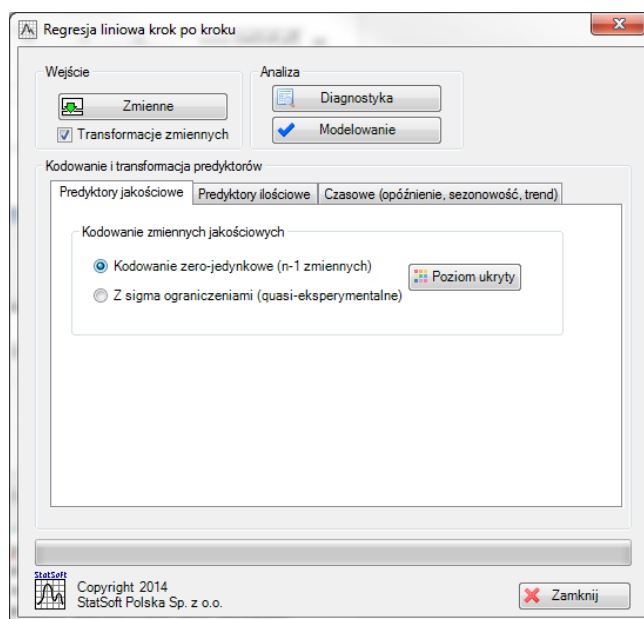
Pierwszą decyzją, jaką badacz musi podjąć, jest wybór zmiennych do analizy. W tym celu klikamy przycisk **Zmienne** i w oknie **Wybierz zmienne** wskazujemy zmienne do analizy. Na potrzeby naszego przykładu przyjmijmy, że naszym celem jest zbadanie wpływu wskaźnika BMI na HDL skorygowanego przez obecność innych zmiennych, takich jak: wiek, palenie, spożywanie alkoholu, rasa czy aktywność fizyczna. Poniżej okno ze zmiennymi wybranymi do naszej analizy.



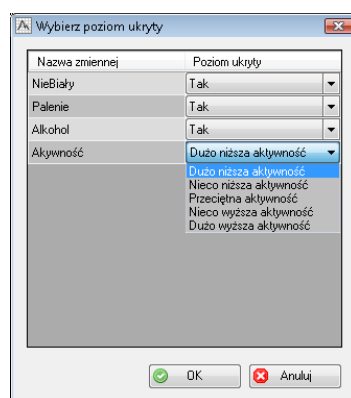
Po wybraniu zmiennych możemy przejść od razu do fazy modelowania, klikając przycisk **Modelowanie**, my jednak przed samym etapem modelowania wykonamy jeszcze dwa dodatkowe kroki związane z transformacją i diagnostyką zmiennych.

### 2.2.2. Transformacje zmiennych

Aby wykonać dodatkowe transformacje zmiennych, klikamy opcję **Transformacje zmiennych**, rozszerzając tym samym okno analizy o dodatkowe funkcjonalności widoczne poniżej.

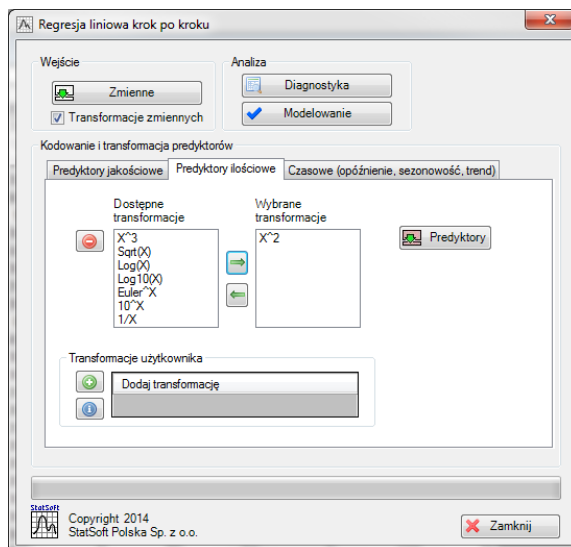



W obszarze **Kodowanie i transformacja predyktorów** na karcie **Predyktory jakościowe** możemy określić sposób kodowania zmiennych jakościowych użytych w modelu. Mamy do wyboru dwa schematy kodowania, spośród których wybieramy pierwszy – **Kodowanie zero-jedynkowe**. Istotnym elementem kodowania zmiennych jakościowych jest określenie poziomu odniesienia mającego wpływ na sposób interpretacji ocen współczynników regresji. Poziom ten specyfikujemy w oknie **Wybierz poziom ukryty**, które wyświetlone zostaje po kliknięciu przycisku **Poziom ukryty**.

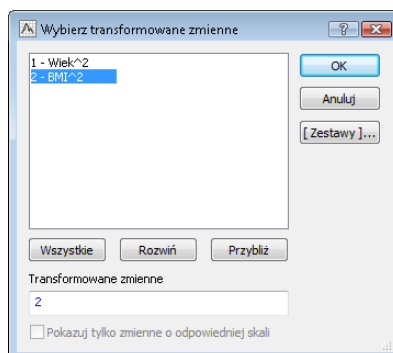


Przyjmijmy, że poziomem odniesienia dla zmiennej *Aktywność* będzie klasa *Dużo niższa aktywność*, natomiast dla pozostałych zmiennych będzie to klasa *Tak*.

Na karcie **Predyktory ilościowe** możemy definiować transformacje predyktorów ilościowych. Z listy dostępnych transformacji – domyślnie są to między innymi: kwadrat, logarytm czy pierwiastek – za pomocą strzałek umieszczamy wybrane transformacje na liście **Wybranych transformacji**. Następnie za pomocą przycisku **Predyktory** wskazujemy zmienne, jakie powinny zostać przekształcone.



W naszej analizie wybierzemy dla przykładu kwadrat predyktorów. Z listy **Dostępne transformacje** wybieramy opcję  $X^2$  i klikamy przycisk , przesuując tę transformację na listę **Wybranych transformacji**. Następnie klikamy przycisk **Predyktory**. W oknie wyboru zmiennych mamy możliwość wybrania interesujących nas zmiennych po transformacji.

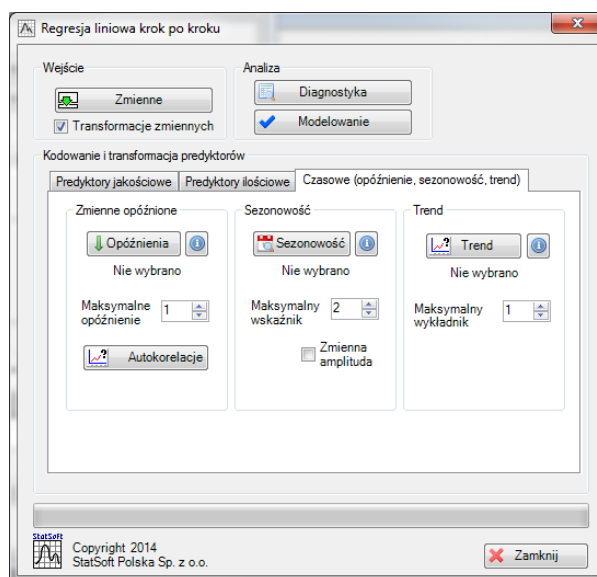


Wybieramy jedynie zmienną  $BMI^2$ , która zostanie dodana do listy dostępnych predyktorów.

Opcja **Transformacje użytkownika** umożliwia zdefiniowanie własnej transformacji na bazie funkcji dostępnych dla formuł arkusza. Zdefiniowana transformacja trafia na listę **Dostępnych transformacji** i jest pamiętana przez program po kolejnym uruchomieniu **Kreatora**.

W przypadku analizy danych czasowych użytkownik ma możliwość uzupełnienia zestawu predyktorów o opóźnienia, sezonowość oraz trend z odpowiednimi parametrami.

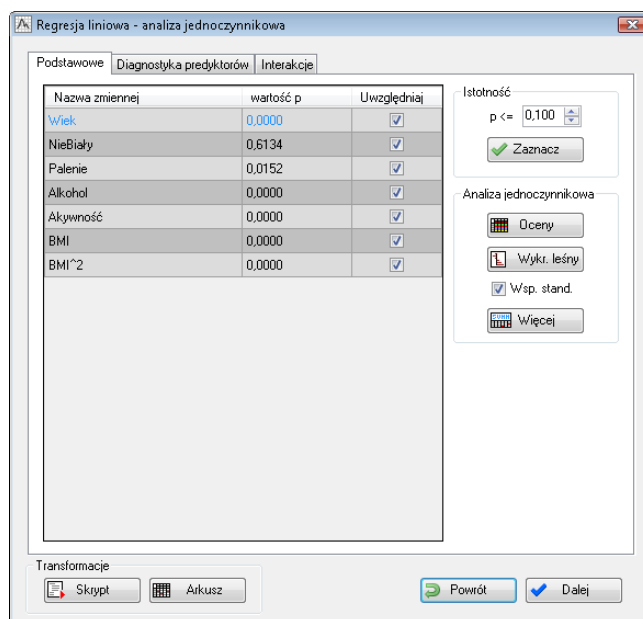




W naszym przykładzie, ze względu na charakter zbioru danych, nie skorzystamy z tej możliwości.

### 2.2.3. Diagnostyka zmiennych

Zanim przejdziemy do fazy modelowania, dokonamy jeszcze diagnostyki wybranych predyktorów. Klikamy przycisk **Diagnostyka** i przechodzimy do okna **Regresja liniowa – analiza jednoczynnikowa**.

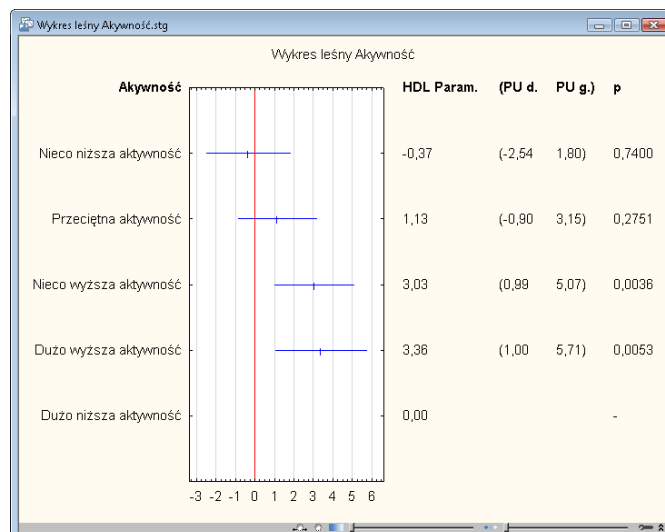


W tym oknie w obszarze **Analiza jednoczynnikowa** mamy możliwość przejrzania wyników prostej regresji dla każdego z wybranych predyktorów i oceny tym samym izolowanego wpływu poszczególnych zmiennych na zjawisko.

Kolumna **wartość p** w tabeli jest wartością poziomu prawdopodobieństwa testowego dla testu F, oceniającego istotność zmiennych zawartych w prostym modelu. Na tej podstawie możemy stwierdzić, że wartość oceny współczynnika regresji w modelu zawierającym zmienną *NieBiały* nie jest istotnie różna od zera, innymi słowy: średni poziom HDL w grupach określonych przez tę zmienną nie różni się istotnie.

W niektórych sytuacjach, zwłaszcza gdy chcemy budować model do celów predykcyjnych, tego typu kryterium może być przydatne na etapie wstępnej eliminacji zmiennych. W grupie **Istotność** możemy

określić graniczny poziom  $p$  i na tej podstawie automatycznie wykluczyć z dalszej analizy zmienne niepełniające wskazanego warunku. Działania tego typu musimy podejmować zawsze z należytą ostrożnością ponieważ zmienne nieistotne w modelu jednoczynnikowym mogą okazać się istotne po uwzględnieniu wpływu innych predyktorów. W naszym przypadku, gdy budujemy model opisowy, nie powinniśmy z góry wykluczać żadnej zmiennej. W modelu wieloczynnikowym jej korygujący wpływ na modelowane zagadnienie może okazać się istotny.



Za pomocą przycisku **Wykr. leśny** możemy przedstawić wartości ocen oraz przedziałów ufności współczynników regresji (lub standaryzowanych współczynników regresji) dla poszczególnych klas predyktorów jakościowych. Daje nam to szybki wgląd w siłę i charakter zależności pomiędzy zmienną zależną, a poszczególnymi klasami predyktora jakościowego. Powyżej przykładowy wykres leśny dla zmiennej **Aktywność**.

#### 2.2.4. Współliniowość i liniowość

Na karcie **Diagnostyka predyktorów** mamy możliwość oceny współliniowości wybranych predyktorów ilościowych oraz liniowości ich wpływu na zmienną zależną.

**Współliniowość**

Analiza czynnikowa

Min. ładunek: 0,70

Liczba czynników: 5

☐ Wybierz automatycznie

Liczba reprezentantów: 2

**Liniowość**

Dopasowanie

☒ Liniowe ☒ Gładkość

☒ Lowess 0,50

**Transformacje**

Współliniowość predyktorów jest okolicznością niekorzystnie wpływającą na proces budowy modelu. W skrajnych sytuacjach może wręcz uniemożliwić estymację parametrów regresji. Współliniowość ma również wpływ na precyzję oszacowań ocen współczynników regresji, zawyżając ich błędy standardowe i utrudniając wychwycenie efektów istotnie wpływających na badane zjawisko. Efektem współliniowości może być również odwrócenie znaków ocen współczynników regresji wielorakiej (w stosunku do znaków w regresji prostej), co może utrudnić bądź wręcz uniemożliwić poprawną interpretację modelu.

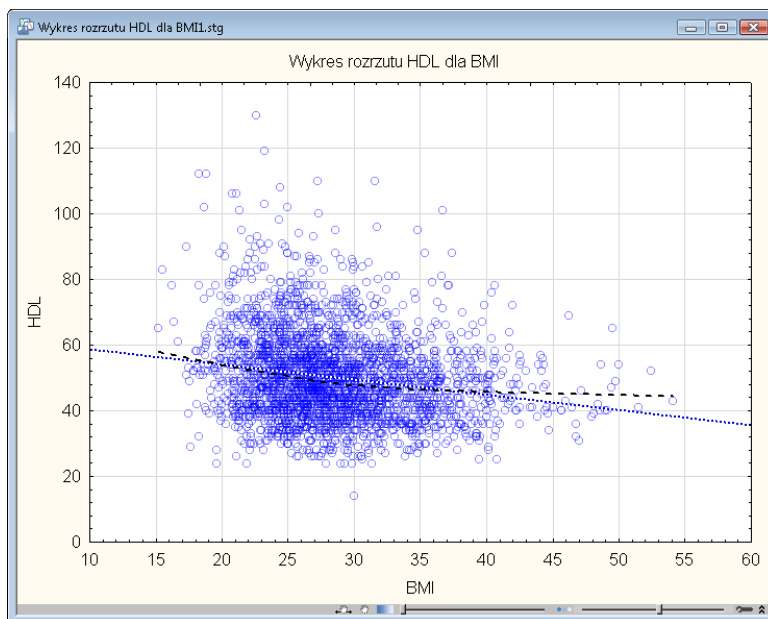
Współliniowość oceniamy za pomocą macierzy korelacji, na podstawie której możemy stwierdzić (co oczywiste) wysoki poziom korelacji pomiędzy zmiennymi *BMI* oraz *BMI*<sup>2</sup>.

Korelacje (HERS_PL.sta) N=2758					
Zmienna	Srednia	Odch. stand.	Wiek	BMI	BMI^2
Wiek	66,6523	6,8495	1,000000	-0,159448	-0,158965
BMI	28,5792	5,5178	-0,159448	1,000000	0,991144
BMI^2	847,2084	342,1263	-0,158965	0,991144	1,000000

W sytuacji, gdy liczba predyktorów ilościowych jest duża, wykonanie globalnej analizy korelacji mogłoby okazać się bardzo czasochłonne i nieefektywne, dlatego też użytkownik ma możliwość skorzystania z zaawansowanych metod identyfikacji nadmiernie skorelowanych zmiennych. Sposobem identyfikacji grup zmiennych wzajemnie ze sobą skorelowanych może być wykorzystanie np. analizy składowych głównych. Wykorzystując tę metodę, przeprowadzamy grupowanie cech, a następnie wybieramy reprezentanty każdej uzyskanej grupy, które tworzą zestaw danych wejściowych modelu. Metoda ta pozwala wychwycić skupiska cech wzajemnie ze sobą powiązanych i umożliwia automatyczną eliminację ze zbioru danych nadmiarowych zmiennych. Dostęp do opisanej powyżej metody jest możliwy za pomocą przycisku **Reprezentanty**.

W obszarze **Liniowość** za pomocą wykresów rozrzutu mamy możliwość oceny charakteru wpływu poszczególnych predyktorów na badane zjawisko. Przedstawione na wykresie dopasowanie LOWESS ułatwia ocenę rzeczywistego charakteru wpływu poszczególnych zmiennych na zjawisko oraz wychwycenie ewentualnych nieliniowości. Może to być podstawą do cofnięcia się do wcześniejszego etapu definiowania przekształceń predyktorów, tak by zlinearyzować ich wpływ na modelowane zjawisko.

Na podstawie uzyskanych wykresów możemy zauważyć, że ani BMI, ani BMI<sup>2</sup> nie spełniają w pełni założenia liniowości wpływu na zmienną zależną. Być może warto byłoby zastanowić się nad innym przekształceniem. Przed przystąpieniem do dalszych etapów analizy podejmujemy decyzję o eliminacji zmiennej BMI<sup>2</sup>. Na karcie **Podstawowe** anulujemy zaznaczenie opcji **Uwzględniaj** dla tej zmiennej.

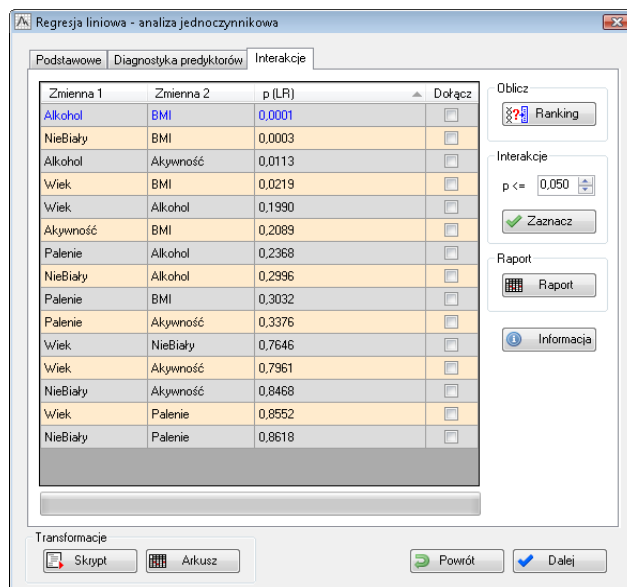


### 2.2.5. Interakcje

Budując modele za pomocą regresji liniowej, zakładamy, że poszczególne zmienne niezależne mogą mieć wpływ na modelowane zjawisko. Uwzględnianie nowych bądź eliminacja starych czynników powoduje zazwyczaj korektę siły wpływu pozostałych predyktorów. Zmiana wpływu poszczególnych czynników ryzyka na zjawisko nie jest jednak zależna od zmienności innych predyktorów. Możemy jednak wyobrazić sobie sytuację, w której siła wpływu zmiennej  $x_1$  na zjawisko zmienia się w sposób systematyczny wraz ze zmianą wartości zmiennej  $x_2$ . Tego typu wzajemny wpływ dwóch zmiennych na modelowane zjawisko nazywamy interakcją.

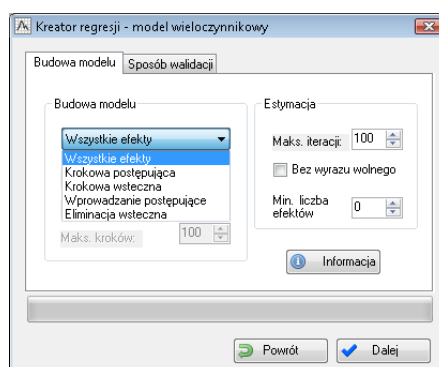
Aby przeprowadzić identyfikację interakcji, w oknie **Regresja liniowa – analiza jednoczynnikowa** przechodzimy na kartę **Interakcje**, a następnie klikamy przycisk **Ranking**, co spowoduje obliczenie wszystkich możliwych interakcji dwuczynnikowych dla zmiennych zaznaczonych na karcie **Podstawowe**. Po wykonanych obliczeniach na liście interakcji pojawi się zestaw par zmiennych tworzących interakcję oraz wartość prawdopodobieństwa testowego  $p$  informująca o istotności danej interakcji. Przykładowo na podstawie uzyskanych wyników możemy wysnuć przypuszczenie, że siła wpływu *BMI* na zmienną zależną zależy od wartości zmiennej *Alkohol* i warto byłoby taką interakcję uwzględnić w analizie<sup>1</sup>. Kliknięcie opcji **Dołącz** dla wybranej interakcji spowoduje uwzględnienie jej w modelu wieloczynnikowym.

<sup>1</sup> Technicznie interakcja jest po prostu dodatkową zmienną będącą iloczynem predyktorów, których dotyczy.



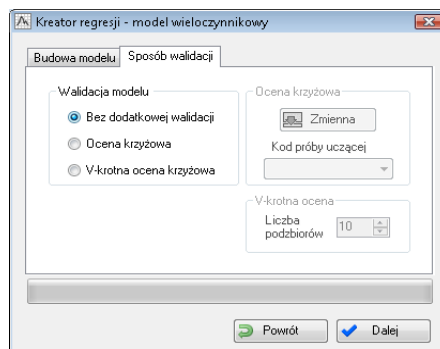
### 2.2.6. Budowa modelu wieloczynnikowego

Kliknięcie przycisku **Dalej** w oknie **Regresja liniowa – analiza jednoczynnikowa** spowoduje wyświetlenie okna **Kreator regresji – model wieloczynnikowy**, w którym możemy określić sposób budowy i walidacji modelu wieloczynnikowego.



Na karcie **Budowa modelu** określamy, czy model ma zostać zbudowany dla wszystkich zmiennych i interakcji, jakie zostały zaznaczone w poprzednim oknie (opcja **Wszystkie efekty**), czy też chcemy skorzystać ze schematów doboru zmiennych pozwalających na eliminację nieistotnych predyktorów z modelu. My wybierzemy opcję **Krokowa wsteczna**, aby ograniczyć liczbę zmiennych jedynie do istotnych (w kontekście modelu wieloczynnikowego) predyktorów.

Na karcie **Sposób walidacji** określamy sposób szacowania błędu predykcji modelu (ma to w szczególności znaczenie, jeżeli głównym celem naszego modelu jest predykcja). Obliczenie błędu predykcji na podstawie zbioru uczącego nie daje niestety obrazu zdolności predykcyjnych modelu. Błąd predykcji obliczony na tym zbiorze jest niedoszacowany w stosunku do jego rzeczywistej wartości. Rzeczywistą zdolność predykcyjną jesteśmy w stanie ocenić dopiero na podstawie zbioru, który nie brał udziału w procesie szacowania parametrów modelu. Podstawowym sposobem zbadania zdolności predykcyjnej na podstawie niezależnego zbioru danych jest podział zbioru danych na zbiór uczący, na którym szacujemy parametry modelu, oraz zbiór testowy, na którym oceniamy jego zdolność predykcyjną. Ten schemat walidacji dostępny jest za pomocą opcji **Ocena krzyżowa**.



Podział na próbę uczącą i testową nie zawsze jest dobrą strategią budowy i walidacji modelu. Zdarzają się sytuacje, w których ze względu na niewielką licznosc zbioru danych pozostawienie odlozonego zbioru niebiorącego udziału w procesie estymacji mogloby znacząco uszczuplić zestaw wzorców potrzebnych do oszacowania parametrów modelu.

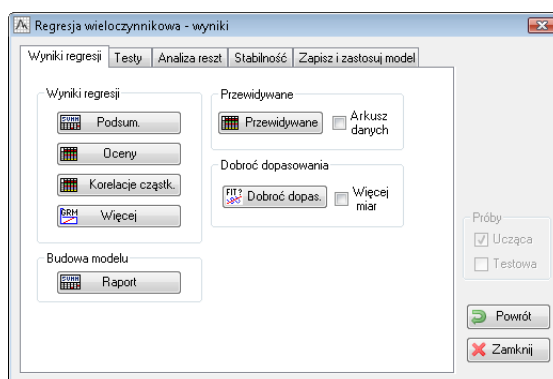
Alternatywą dla podziału na próbę uczącą i testową jest v-krotny sprawdzian krzyżowy, polegający na podziale zbioru danych na v podzbiorów, z których każdy pełni rolę zbioru testowego w jednej iteracji oceny. Globalna ocena błędu predykcji jest w tym przypadku uśrednieniem błędów pochodzących z poszczególnych zbiorów. Zalecana liczba podzbiorów najczęściej wynosi 10. Opisany powyżej schemat walidacji dostępny jest za pomocą opcji **v-krotna ocena krzyżowa**.

Nas interesuje przede wszystkim opis modelowanego zjawiska, a nie jakość predykcji, dlatego też pozostawiamy wybór opcji **Bez dodatkowej walidacji**. Aby wykonać i ocenić model wieloczynnikowy, w oknie **Kreator regresji – model wieloczynnikowy** klikamy przycisk **Dalej**.

### 2.2.7. Ocena wyników modelu wieloczynnikowego

Po zbudowaniu modelu wyświetlone zostaje okno **Regresja wieloczynnikowa – wyniki**, w którym mamy możliwość ocenić jakość dopasowania modelu oraz przetestować spełnienie założeń regresji za pomocą testów oraz metod graficznych.

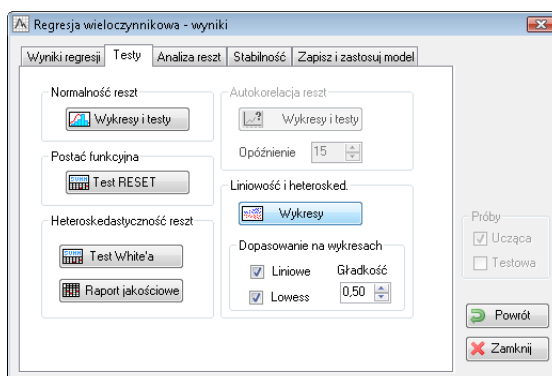
Karta **Wyniki regresji** umożliwia uzyskanie podstawowych wyników odnośnie zbudowanego modelu. Dostępne są oceny parametrów regresji (oraz standaryzowanych ocen parametrów regresji) wraz z oceną ich istotności oraz przedziałami ufności – przycisk **Oceny**. Użytkownik może również ocenić współzmiennosc analizowanych cech za pomocą szeregu miar (między innymi czynnika inflacji wariancji VIF), klikając przycisk **Korelacje cząstk.**



Jeśli wybraliśmy jedną z krokowych metod doboru parametrów modelu, przycisk **Raport** umożliwia wyświetlenie procesu doboru zmiennych do modelu.

W obszarze **Przewidywanie** możemy wygenerować arkusz z przewidywaniami na podstawie zbudowanego modelu regresji. Obszar **Dobroć dopas.** pozwala ocenić dobroć dopasowania modelu na podstawie szeregu miar dopasowania.

Na karcie **Testy** użytkownik może zbadać za pomocą testów spełnienie założeń odnośnie zbudowanego modelu. Dodatkowo możliwa jest mniej formalna ocena spełnienia tych założeń przy pomocy dedykowanych wykresów i raportów.

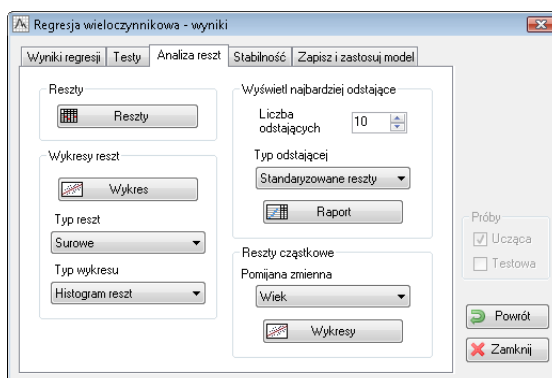


Obszar **Normalność reszt** pozwala obliczyć szereg testów normalności – między innymi dostępny jest test *Shapiro-Wilka* oraz test *Jarque-Bera*, raport uzupełniają: wykres normalności reszt oraz histogram reszt.

W programie zaimplementowano dodatkowo test postaci funkcyjnej *RESET Ramsaya*, test oceniający heteroskedastyczność reszt *White'a* oraz w przypadku modelowania danych czasowych test *Durbina-Watsona* wraz z przydatnymi wykresami.

Heteroskedastyczność reszt, a także liniowość wpływu poszczególnych predyktorów ilościowych można oceniać za pomocą wykresów reszty vs. predyktor z możliwością przedstawienia na wykresie dopasowania liniowego oraz LOWESS, szczególnie zalecanego podczas oceny tego typu wykresów. W przypadku predyktorów jakościowych możemy porównać wartości wariancji reszt w poszczególnych grupach i na tej podstawie ocenić skalę heteroskedastyczności.

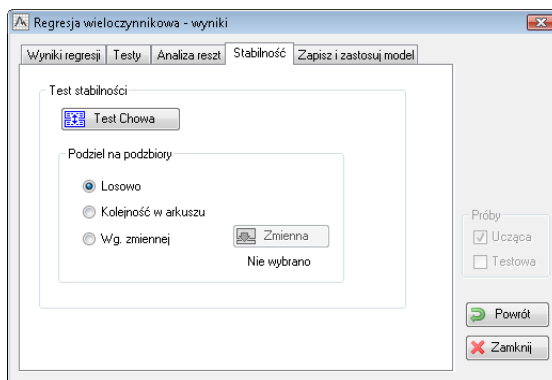
Karta **Analiza reszt** umożliwia wygenerowanie szeregu raportów i wykresów przedstawiających reszty oraz wskazać wartości odstające i wpływowe mające największy wpływ na wartości ocen parametrów regresji.



Użytkownik może także wyświetlić wskazaną liczbę przypadków najbardziej odstających, według wskazanego kryterium, oraz wygenerować wykres reszt częściowych i regresji częściowej.

Karta **Stabilność** pozwala na przeprowadzenie testu stabilności *Chowa*, który po podziale wejściowego zbioru danych na dwa podzbiory (w sposób losowy lub wskazany przez użytkownika), buduje na ich podstawie pomocnicze modele regresji, a ich oceny porównywane są z uzyskanymi w wyniku docelowej estymacji.





Opcja **Zapisz i zastosuj model** umożliwia zapis uzyskanych wyników w postaci kodu PMML, C oraz makra Visual Basic. W kodzie makra zawarte są nie tylko parametry modelu, ale również wszystkie ewentualne transformacje predyktorów.

Dodatkowo w sytuacji, gdy wybieramy dane czasowe, możliwe jest wykonanie prognozy oraz prognozy kroczącej (na podstawie wcześniejszych prognoz) dla analizowanego i nowego zbioru danych.

## 2.3 Kreator testów

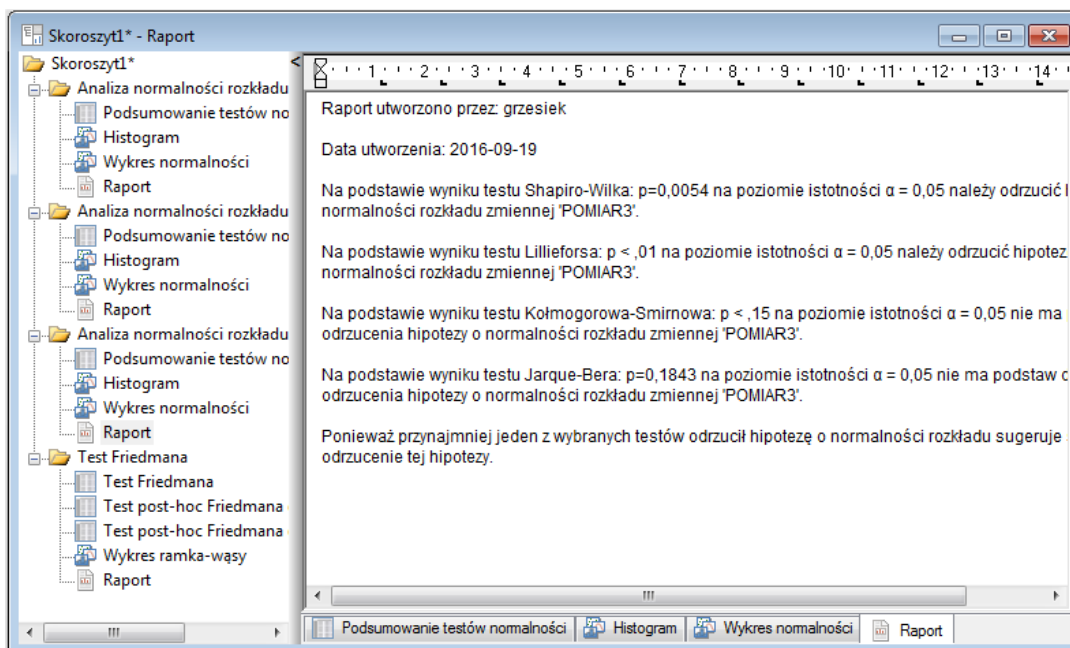
Testy istotności różnic dostępne w Zestawie Przyrodnika są udostępnione w ramach Kreatora testów statystycznych i dlatego też obejmują niemal wszystkie sytuacje spotykane w praktyce kiedy trzeba zbadać istotność różnic, nie tylko w obrębie nauk przyrodniczych. Możemy zatem ogólnie zbadać istotność różnic zmiennej ilościowej/porządkowej/jakościowej w zależności od czynnika grupującego oraz między zmiennymi wyrażającymi pomiary zależne. Dodatkową możliwością kreatora są testy pojedynczej zmiennej. Swoistą cechą kreatora jest automatyzacja przebiegu testu polegająca na samoczynnym sprawdzeniu założeń dot. m.in. rozkładu zmiennych ilościowych i wybraniu przez program odpowiedniego testu statystycznego. Co więcej, użytkownik ma możliwość wskazania niestandardowych kryteriów wyboru testów.

Korzystając z tego narzędzia badacz musi jedynie określić kwestie merytoryczne prowadzonej analizy, takie jak:

- *Jaką analizę chcemy przeprowadzić?*
- *Czy badane próby są zależne/niezależne od siebie?*
- *Ile grup analizujemy?*
- *Na jakiej skali mierzone są badane cechy?*

Wynikiem działania programu jest skoroszyt zawierający wyniki poszczególnych testów (dot. założeń jak i głównego pytania) razem z interpretacją oraz wykresy i dodatkowe analizy generowane standardowo przy danym rodzaju badań.





Z powyższego opisu można odnieść wrażenie, że kreator testów jest swoistą „czarną skrzynką” do której możemy „wrzucić” dane, powiedzieć co chcemy uzyskać, kliknąć 'Uruchom analizę' i uzyskać wynik. Tak rzeczywiście jest, ale nie do końca. Głównym zadaniem badacza podczas przeprowadzania analizy jest wskazanie aspektów merytorycznych i wybór zmiennych. W tym miejscu rzeczywiście można po prostu uruchomić analizę i uzyskać pełen raport ze wszystkimi możliwymi wynikami, jednakże możemy również każdą analizę dostosować do naszych potrzeb. Możemy wybrać które z możliwych testów chcemy zastosować, bądź też które wyniki nas interesują. Program automatycznie zapamiętuje wybrane opcje. Dodatkowym atutem kreatora testów jest możliwość wyeksportowania wyników do pliku MS Word.

W tym miejscu należałoby również zaznaczyć, że kreator testów ze względu na cel, który mu przyświeca, tj. pewna uniwersalność i prostota użytkowania, opiera się tylko na jednym z możliwych schematów testowania. Obsługuje on najczęściej występujące problemy analityczne. Bardziej złożone analizy czy też pewne dodatkowe opcje narzędzi tu użytych dostępne są w innych miejscach programu Statistica.

### 2.3.1. Możliwości programu

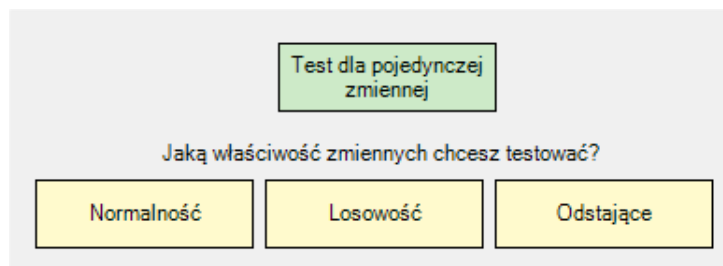
W obecnej wersji *Kreator testów statystycznych* umożliwia wykonanie dwóch rodzajów analiz:

- Testy dla pojedynczej zmiennej
- Badanie istotności różnic.

Wybranie jednej z opcji spowoduje wyświetlenie schematu w postaci drzewa, dzięki któremu badacz w prosty, intuicyjny sposób może określić jaki dokładnie typ analizy chce przeprowadzić.

### 2.3.2. Testy dla pojedynczej zmiennej

W celu przejścia do schematu dotyczącego analizy pojedynczej zmiennej wystarczy kliknąć odpowiednie pole rodzaju analiz. Na ekranie pojawi się poniższy schemat.



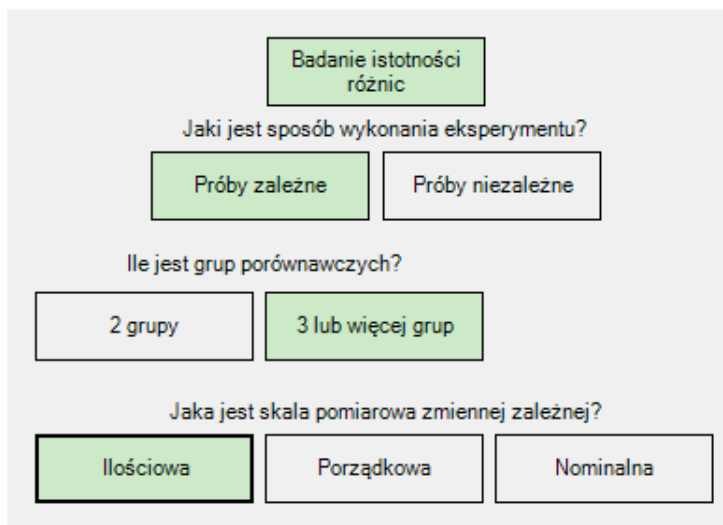
Jak można zauważyć, w chwili obecnej program oferuje trzy typy testów dla pojedynczej zmiennej:

- Normalność
- Losowość
- Obserwacje odstające.

W celu przeprowadzenia analizy wystarczy wybrać jeden z nich. Zostanie wtedy odblokowana możliwość wyboru zmiennych oraz wykonania analizy.

### 2.3.3. Badanie istotności różnic

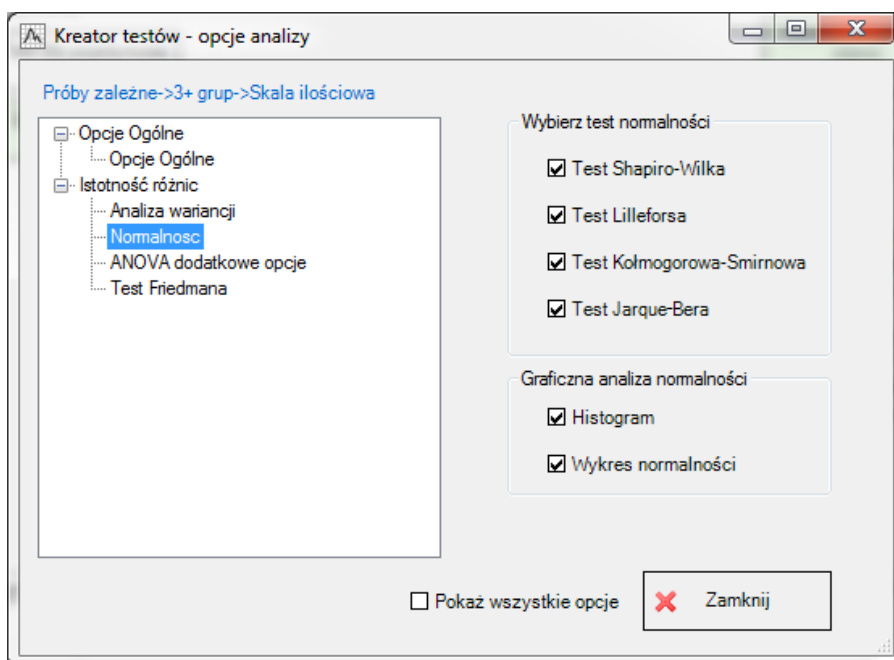
Po wybraniu tego rodzaju analizy, podobnie jak w przypadku analizy dla jednej zmiennej, pojawi się schemat ułatwiający dokładne określenie właściwej ścieżki postępowania. Kreator będzie prowadził użytkownika krok po kroku, zadając mu kolejne pytania pozwalające doprecyzować rodzaj analizowanego problemu. Przykładowy schemat decyzyjny zamieszczono poniżej.



Po udzieleniu odpowiedzi na powyższe pytania, użytkownik musi jedynie wybrać zmienne i uruchomić analizę. Program dokona sprawdzenia założeń związanych z daną klasą problemu, wybierze odpowiedni test i wygeneruje komplet wyników wraz z ich interpretacją.

### 2.3.4. Dodatkowe możliwości programu

Parametry programu są dostosowane do najbardziej typowych sytuacji. Bardziej doświadczeni analitycy mają jednakże możliwość określenia szeregu szczegółowych opcji dotyczących wyboru i konfiguracji testów.

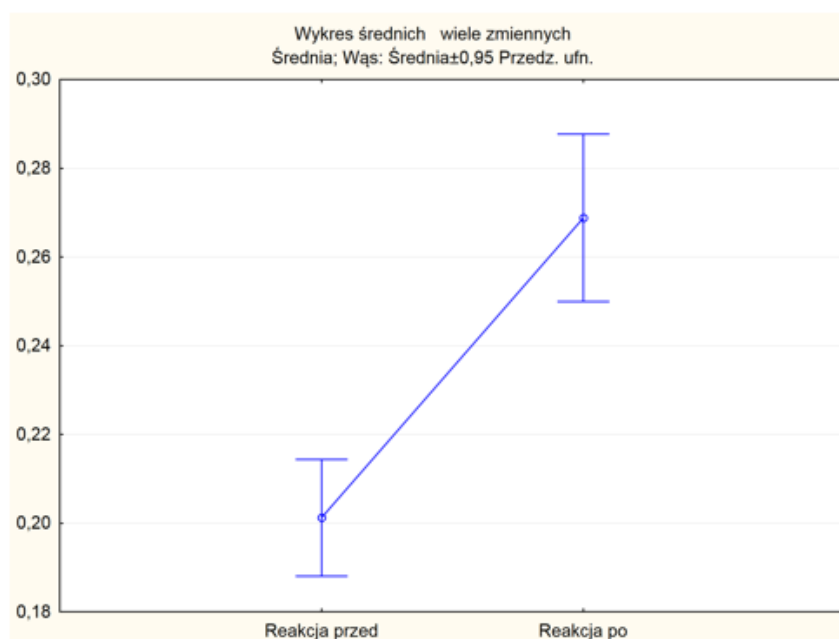


Ważnym atutem kreatora jest niewątpliwie możliwość zapisu raportu z analizy zarówno w formacie Statistica jak i MS Word. Poniżej fragment przykładowego raportu w formacie MS Word.

### 2.3. Statystyki opisowe

Zmienna	N ważnych	Średnia	Ufność - 95,000%	Ufność 95,000%	Odch.std
Reakcja przed	25,00	0,20	0,19	0,21	0,03
Reakcja po	25,00	0,27	0,25	0,29	0,05

### 2.4. Wykres interakcji



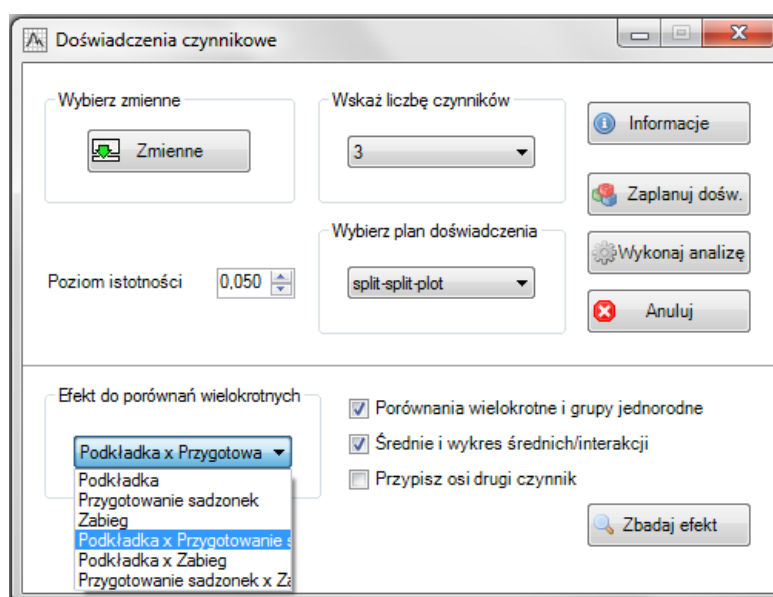
### 2.5. Interpretacja

Na podstawie wyniku testu t-Studenta dla prób zależnych:  $p < 0,0001$  na poziomie istotności  $\alpha = 0,05$  należy odrzucić hipotezę o równości średnich wartości zmiennej 'Reakcja przed' i zmiennej 'Reakcja po'.

### 3. Układy czynnikowe ANOVA

Dzięki modułowi *Doświadczenia czynnikowe* mamy możliwość analizy wariancji kilkunastu układów doświadczeń o 1, 2 lub 3 czynnikach (nie wliczamy do nich bloków), jak również zaplanowania takich doświadczeń.

Do analizy, po wybraniu układu doświadczenia, wystarczy już tylko wskazać zmienne oznaczające czynniki, zmienną blokową oraz zmienną zależną. Czynniki wskazujemy na liście zmiennych – w kolejności od nadrzędnego do podrzędnego, o ile ma miejsce taka hierarchia. Nie ma konieczności dokładnego określania układu i żmudnego testowania istotności różnych źródeł wariancji względem odrębnych członów typu składnik (błąd) losowy. Należy tylko użyć przycisku *Wykonaj analizę* i uzyskujemy arkusz z wynikami analizy wariancji. Po tym, okno rozwija się tak by mieć możliwość zbadania efektu - pojedynczego czynnika lub interakcji dwóch czynników, pod kątem porównań wielokrotnych, obejmujących nie tylko test post-hoc Tukeya lecz również podział na grupy jednorodne oraz wykresy średnich/interakcji.



W celu uzyskania planu doświadczenia, po wybraniu jego układu, trzeba tylko podać liczby poziomów czynników i bloków. Na przykład, do doświadczenia w układzie split-plot możemy zażądać następujących wartości: 4 poziomy czynnika I, 5 poziomów czynnika II oraz 3 bloki.

W większości układów czynniki tworzą pewną hierarchię i w opisie poszczególnych układów oznaczać je będziemy, zgodnie ze zwyczajem, idąc od góry do dołu hierarchii jako: A, B, C lub I, II, III. Na przykład, jeśli występuje hierarchia, to czynnik B jest podrzędny względem czynnika A, a czynnik C jest podrzędny względem obydwu pozostałych czynników.

### 3.1. Doświadczenia jednoczynnikowe

W doświadczeniach jednoczynnikowych badany jest wpływ pojedynczego czynnika na daną cechę ilościową. Do wyboru są trzy układy: kwadrat łaciński, bloki losowe i układ całkowicie losowy.

#### 3.1.1. Kwadrat łaciński

Układ ten stosujemy, gdy badanemu czynnikowi towarzyszą dwie zmienne zakłócające, które traktujemy jako zmienne blokowe, a liczba poziomów czynnika - równa  $a$  - jest taka sama jak którejkolwiek ze zmiennych blokowych. Wtenczas całe doświadczenie składa się z  $a^2$  przypadków, tak, że każdy poziom czynnika pojawia się dla pewnych  $a$  przypadków, zawsze łącznie obejmujących wszystkie  $a$  wartości dowolnej z obu zmiennych blokowych.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 1 i wybieramy plan doświadczenia *kwadrat łaciński* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy jedną liczbę - liczbę poziomów czynnika 1. Niech wynosi ona 6. Oznacza to, że planujemy doświadczenie 6 x 6 a więc o 6 poziomach badanego czynnika i o 36 przypadkach. Po zatwierdzeniu *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat* (6 zm., * 6 prz.)						
Analiza jednoczynnikowa, kwadrat łaciński						
# poziomów czynnika: 6						
	1	2	3	4	5	6
1	A6	A1	A5	A3	A4	A2
2	A2	A4	A1	A6	A5	A3
3	A4	A6	A3	A1	A2	A5
4	A5	A2	A6	A4	A3	A1
5	A1	A3	A2	A5	A6	A4
6	A3	A5	A4	A2	A1	A6

Schemat mówi np. że wartości nr 4 pierwszej zmiennej blokowej (odp. wierszom) oraz wartości nr 3 drugiej zmiennej blokowej (odp. kolumnom) przyporządkowano poziom nr 6 czynnika (ozn. A6).

#### Analiza



**Przykład** Otwieramy arkusz *kwadrat.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 1 i wybieramy plan doświadczenia *kwadrat łaciński*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych (*Zmienna zależna*, *Czynnik*, *Bloki*) odpowiednio zmienne nr 4, 3, 2 i 1, (dwie ostatnie

zaznaczamy na liście *Bloki*, kolejność nie ma znaczenia) po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku pojedynczy arkusz analizy wariancji.

### 3.1.2. Bloki losowe

W tym układzie występuje pewna liczba bloków, w których jednokrotnie pojawia się w kolejności losowej każdy z poziomów badanego czynnika. Łącznie więc wystąpi on tyle razy ile jest bloków.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 1 i wybieramy plan doświadczenia *bloki losowe* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy dwie liczby - liczbę poziomów czynnika 1 i liczbę bloków. Tutaj będzie to, odpowiednio, 7 i 5. Po zatwierdzeniu *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat\* (5 zm., \* 7 prz.)

Analiza jednoczynnikowa, bloki losowe  
# bloków: 5  
# poziomów czynnika: 7

	1 Blok 1	2 Blok 2	3 Blok 3	4 Blok 4	5 Blok 5
1	A5	A7	A7	A6	A4
2	A1	A1	A2	A1	A6
3	A7	A6	A3	A4	A3
4	A2	A4	A6	A5	A7
5	A6	A2	A5	A3	A1
6	A3	A3	A1	A2	A5
7	A4	A5	A4	A7	A2

Schemat mówi np. że w bloku 3 przypisujemy kolejno zgodnie z wyznaczonym kierunkiem obszaru doświadczalnego poziomy nr 7, 2, 3, 6, 5, 1, 4 czynnika.

#### Analiza



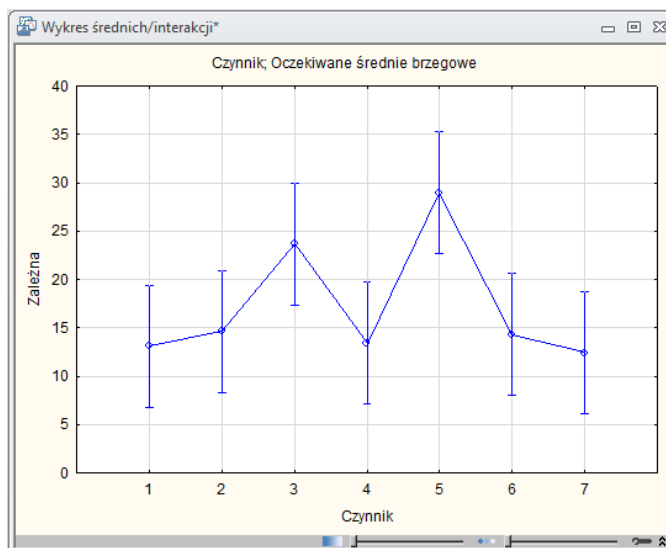
**Przykład** Otwieramy arkusz *jednoczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 1 i wybieramy plan doświadczenia *bloki losowe*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych *Zmienna zależna*, *Czynnik*, *Blok* zmienne nr 2, 1, 3 po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji.

Dane: Analiza jednoczynnikowa - bloki losowe - Wyniki\* (5 zm., \* 3 prz.)

Analiza jednoczynnikowa, bloki losowe  
Zmienna zależna: Zależna  
Precyzja doświadczenia: 34,8%

Efekt	SS	Stopnie swobody	MS	F	p
Blok	6,249	3	2,083	0,0578	0,981158
Czynnik	993,595	6	165,599	4,5983	0,005346
Błąd losowy	648,238	18	36,013		

Skoro czynnik okazał się istotny statystycznie ( $p = 0,005$ ), to powracamy do okna analizy, na którym teraz jest możliwość wykonania porównań wielokrotnych. Pozostajemy przy domyślnych ustawieniach i tylko naciskamy przycisk *Zbadaj efekt*. Wynikiem będą trzy arkusze - wartości  $p$  z testu post-hoc Tukeya, podział na grupy jednorodne, statystyki opisowe dla poziomów czynnika oraz wykres wartości średnich, który jest ukazany poniżej.



### 3.1.3. Układ całkowicie losowy

W układzie całkowicie losowym z jednym czynnikiem każdy jego poziom jest powtórzony w całym doświadczeniu tyle samo razy. Jeżeli powtórzenia potraktujemy jako bloki, to nie ma wymogu by każdy poziom czynnika występował w nich po jednym razie. Przydział tutaj odbywa się całkowicie losowo i w bloku mogą się pojawić nie wszystkie poziomy a więc niektóre z pozostałych wystąpią w nim więcej niż jeden raz.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 1 i wybieramy plan doświadczenia *całkowicie losowy* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy dwie liczby - liczbę poziomów czynnika 1 i liczbę powtórzeń. W przykładzie będzie to 5 i 6. Po zatwierdzeniu *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat\* (6 zm., \* 5 prz.)

Analiza jednoczynnikowa, ukł. całk. losowy  
# powtórzeń: 6  
# poziomów czynnika: 5

	1	2	3	4	5	6
1	A4	A1	A4	A5	A5	A5
2	A2	A5	A3	A5	A3	A4
3	A2	A4	A4	A3	A1	A5
4	A2	A1	A2	A3	A4	A3
5	A3	A1	A2	A1	A2	A1



Schemat mówi np. że w drugim powtórzeniu, które możemy utożsamić z pewnym obszarem, trzykrotnie pojawia się poziom nr 1 czynnika a poziomy nr 4 i nr 5 jednokrotnie. Kolejność przydzielania na obszarze taka sama jak w arkuszu.

### Analiza



**Przykład** Otwieramy arkusz *jednoczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 1 i wybieramy plan doświadczenia *całkowicie losowy*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych *Zmienna zależna*, *Czynnik zmienne nr 2* i 1 (zmiennej nr 3 nie wykorzystujemy) po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku pojedynczy arkusz analizy wariancji.

## 3.2. Doświadczenia dwuczynnikowe

W doświadczeniach dwuczynnikowych badany jest jednoczesny wpływ dwu czynników na daną cechę ilościową. Do wyboru są cztery układy: split-plot, split-blok, bloki losowe i układ całkowicie losowy. Niech czynnik A ma  $a$  poziomów a czynnik B  $b$  poziomów.

### 3.2.1. Split-plot

W układzie split-plot w każdym bloku cały obszar dzielony jest na  $a$  pól, do których losowo przypisujemy różne wartości czynnika A a każde z nich dzielimy na  $b$  poletek, do których losowo przypisujemy różne wartości czynnika B. O czynniku B mówimy wtedy, że jest *zagnieżdżony* względem czynnika A.

### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 2 i wybieramy plan doświadczenia *split-plot*, po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy trzy liczby - liczby poziomów czynnika 1 i czynnika 2 i liczbę bloków. Niech będą równe kolejno 3, 4 i 5. Po zatwierdzeniu *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat* (20 zm., * 6 prz.)													
Analiza dwuczynnikowa, split-plot													
# bloków: 5													
# poziomów czynnika I: 3													
# poziomów czynnika II: 4													
	1	2	3	4	5	6	7	8	9	10	11	12	13
	Blok 1				Blok 2				Blok 3				Blok 4
1	A3				A2				A3				A3
2	A3B1	A3B2	A3B4	A3B3	A2B4	A2B3	A2B2	A2B1	A3B3	A3B1	A3B4	A3B2	A3B4
3	A2				A3				A2				A2
4	A2B2	A2B1	A2B3	A2B4	A3B3	A3B2	A3B4	A3B1	A2B1	A2B2	A2B4	A2B3	A2B4
5	A1				A1				A1				A1
6	A1B2	A1B4	A1B1	A1B3	A1B1	A1B4	A1B2	A1B3	A1B3	A1B2	A1B4	A1B1	A1B1

Schemat mówi np. że w trzecim bloku, kolejnym trzem polom przypisujemy poziomy nr 2, 3, 1 pierwszego czynnika a czterem poletkom drugiego pola odpowiednio poziomy nr 3, 2, 4, 1 drugiego czynnika.

### Analiza





**Przykład** Otwieramy arkusz *dwuczynnikiowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 2 i wybieramy plan doświadczenia *split-plot*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych *Zmienna zależna*, *Czynniki*, *Blok* zmienne nr 4, 2 i 3, 1. Kolejność zmiennych na liście *Czynniki* ma znaczenie, wskazując 3 i 2 otrzymamy inne wyniki; wybrana kolejność oznacza, że czynnik *Zraszanie* traktujemy jako nadrzędny względem czynnika *Podłoże*. Po wybraniu zmiennych naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji.

Dane: Analiza dwuczynnika - split-plot - Wyniki\* (5 zm., 6 prz.)

Analiza dwuczynnika, split - plot  
Zmienna zależna: Plon  
Precyzja doświadczenia: 29,9%

Efekt	SS	Stopnie swobody	MS	F	p
Blok	33944	4	8486	1,0565	0,436811
Zraszanie	109438	2	54719	6,8123	0,018731
Błąd I (Blok * Zraszanie)	64259	8	8032		
Podłoże	31504	3	10501	1,5311	0,223079
Zraszanie * Podłoże	9624	6	1604	0,2339	0,962654
Błąd II	246910	36	6859		

W związku ze statystyczną istotnością czynnika *Zraszanie* powracamy do okna analizy, wskazujemy *Zraszanie* jako *Efekt do porównań wielokrotnych* i zatwierdzamy przyciskiem *Zbadaj efekt*. Utworzone zostaną wykres średnich oraz trzy arkusze, spośród których dwa, z wynikami testu post-hoc oraz z podziałem na grupy jednorodne, ukazano poniżej.

Dane: Porównania szczegółowe\* (4 zm., 3 prz.)

Test post-hoc Tukey  
Wartości p

Nr podkl.	Zras zani	{1}	{2}	{3}
1	G1	235,45	260,20	335,85
2	G2	0,624154	0,624154	0,001906
3	G3	0,001906	0,020275	0,020275

Dane: Grupy jednorodne\* (4 zm., 3 prz.)

Test post-hoc Tukey  
Grupy jednorodne

Nr podkl.	Zras zani	Plon Średnie	1	2
1	G1	235,4500	****	
2	G2	260,2000	****	
3	G3	335,8500		****

### 3.2.2. Split-blok

Układ split-blok jest zwany inaczej układem pasów prostopadłych. W każdym bloku cały obszar dzieli się na pasy w dwu prostopadłych kierunkach, w jednym na *a* pasów a w drugim na *b* pasów. Do obydwu podziałów przypisujemy losowo poziomy kolejnych czynników i wobec tego żaden z nich nie jest nadrzędny w stosunku do drugiego.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 2 i wybieramy plan doświadczenia *split-blok* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy trzy liczby - liczby poziomów czynnika 1 i czynnika 2 i liczbę bloków. Niech będą równe kolejno 6, 5 i 3. Po zatwierdzeniu *OK* utworzony zostanie arkusz ze

schematem doświadczenia, jak ten ukazany poniżej, a także arkusz do wprowadzania danych.

Dane: Schemat\* (15 zm. \* 6 prz.)

Analiza dwuczynnikowa, split-blok  
# bloków: 3  
# poziomów czynnika I: 6  
# poziomów czynnika II: 5

	1 Blok 1	2	3	4	5	6 Blok 2	7	8	9	10	11 Blok 3
1	A1B3	A1B4	A1B5	A1B1	A1B2	A4B3	A4B5	A4B1	A4B4	A4B2	A6B2
2	A4B3	A4B4	A4B5	A4B1	A4B2	A1B3	A1B5	A1B1	A1B4	A1B2	A3B2
3	A6B3	A6B4	A6B5	A6B1	A6B2	A6B3	A6B5	A6B1	A6B4	A6B2	A1B2
4	A5B3	A5B4	A5B5	A5B1	A5B2	A2B3	A2B5	A2B1	A2B4	A2B2	A2B2
5	A2B3	A2B4	A2B5	A2B1	A2B2	A5B3	A5B5	A5B1	A5B4	A5B2	A5B2
6	A3B3	A3B4	A3B5	A3B1	A3B2	A3B3	A3B5	A3B1	A3B4	A3B2	A4B2

Schemat mówi np. że w drugim bloku kolejnym pasom, na które podzielono cały obszar przypisujemy poziomy nr 4, 1, 6, 2, 5, 3 pierwszego czynnika a prostopadłym do nich pasom odpowiednio poziomy nr 3, 5, 1, 4, 2 drugiego czynnika.

### Analiza



**Przykład** Otwieramy arkusz *dwuczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 2 i wybieramy plan doświadczenia *split-blok*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych *Zmienna zależna*, *Czynniki*, *Blok* zmienne nr 4, 2 i 3, 1 po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji. Nie licząc efektu *Blok*, wyniki nie zależą od kolejności czynników.

Dane: Analiza dwuczynnikowa - split-blok - Wyniki\* (5 zm. \* 7 prz.)

Analiza dwuczynnikowa, split - blok  
Zmienna zależna: Plon  
Precyzja doświadczenia: 29,6%

Efekt	SS	Stopnie swobody	MS	F	p
<b>Blok</b>	33944	4	8486	1,0565	0,436811
Zraszanie	109438	2	54719	6,8123	0,018731
Błąd I (Blok * Zraszanie)	64259	8	8032		
Podłoże	31504	3	10501	1,4699	0,272161
Błąd II (Blok * Podłoże)	85732	12	7144		
Zraszanie * Podłoże	9624	6	1604	0,2388	0,959204
Błąd III	161178	24	6716		

### 3.2.3. Bloki losowe

W każdym z bloków tego układu cały obszar doświadczenia dzielimy na *ab* poletek - zazwyczaj dwoma prostopadłymi podziałami - i do każdego poletka przypisujemy losowo różne obiekty, tj. kombinacje obu czynników. Każdy obiekt pojawi się wtedy dokładnie raz w bloku.

### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 2 i wybieramy plan doświadczenia *bloki losowe* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy trzy liczby - liczby poziomów czynnika 1 i czynnika 2 i liczbę bloków. W przykładzie niech będzie to 4, 2 i 5. Po zatwierdzeniu przyciskiem *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej, a także arkusz do wprowadzania danych.

Dane: Schemat\* (5 zm. \* 8 prz.)

Analiza dwuczynnikowa, bloki losowe  
# bloków: 5  
# poziomów czynnika I: 4  
# poziomów czynnika II: 2

	1	2	3	4	5
	Blok 1	Blok 2	Blok 3	Blok 4	Blok 5
1	A4B2	A4B2	A2B1	A3B2	A1B1
2	A2B2	A1B1	A4B1	A1B1	A3B1
3	A3B2	A2B1	A1B1	A2B1	A3B2
4	A4B1	A1B2	A1B2	A3B1	A1B2
5	A1B1	A2B2	A3B1	A2B2	A2B1
6	A2B1	A3B2	A2B2	A1B2	A4B1
7	A3B1	A4B1	A4B2	A4B1	A2B2
8	A1B2	A3B1	A3B2	A4B2	A4B2

Schemat mówi np. że w drugim bloku, na szóstym poletku będziemy mieć kombinację poziomu nr 3 pierwszego czynnika z poziomem nr 2 drugiego czynnika (A3B2).

## Analiza



**Przykład** Otwieramy arkusz *dwuczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 2 i wybieramy plan doświadczenia *bloki losowe*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych *Zmienna zależna*, *Czynniki*, *Blok* zmienne nr 4, 2 i 3, 1, po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji. Żaden z wyników nie zależy od kolejności czynników.

Dane: Analiza dwuczynnikowa - bloki losowe - Wyniki\* (5 zm. \* 5 prz.)

Analiza dwuczynnikowa, bloki losowe  
Zmienna zależna: Plon  
Precyzja doświadczenia: 30,3%

Efekt	SS	Stopnie swobody	MS	F	p
<b>Blok</b>	33944	4	8486	1,1999	0,324333
Zraszanie	109438	2	54719	7,7374	0,001320
Podłoże	31504	3	10501	1,4849	0,231733
Zraszanie * Podłoże	9624	6	1604	0,2268	0,965850
Błąd losowy	311169	44	7072		

### 3.2.4. Układ całkowicie losowy

Tutaj każdy spośród  $ab$  obiektów, czyli kombinacji obu czynników, zostaje powtórzony tę samą liczbę razy w całym doświadczeniu przy czym prócz tego warunku przydział obiektów do poletek jest całkowicie losowy. Jeśli powtórzenia potraktujemy jako bloki, to dany obiekt może się nie pojawić w pewnym bloku a w innym wystąpić co najmniej dwukrotnie.

### 3.3. Doświadczenia trójczynnikiowe

W doświadczeniach trójczynnikiowych badany jest jednoczesny wpływ trzech czynników na daną cechę ilościową. Do wyboru jest osiem poniższych układów doświadczeń. Niech czynniki A, B i C mają odpowiednio  $a$ ,  $b$ ,  $c$  poziomów.

#### 3.3.1. Split-split-plot

W układzie split-split-plot w każdym bloku całą powierzchnię dzielimy na  $a$  pól, do których losowo przydzielamy różne poziomy czynnika A a każde z nich dzielimy na  $b$  poletek, do których losowo przydzielamy różne poziomy czynnika B, po czym dzielimy poletka na  $c$  małych poletek, do których losowo przydzielane są poziomy czynnika C. Mówimy wtedy, że czynnik B jest zagnieżdżony względem A i podobnie, że czynnik C jest zagnieżdżony względem czynnika B.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-split-plot* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy cztery liczby - liczby poziomów czynnika 1, czynnika 2 i czynnika 3 oraz liczbę bloków. Tutaj niech to będą 4, 3, 2 i 5. Po zatwierdzeniu OK utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat\* (30 zmn. \* 12 prz.)

Analiza trójczynnikiowa, split-split-plot  
 # bloków: 5  
 # poziomów czynnika I: 4  
 # poziomów czynnika II: 3  
 # poziomów czynnika III: 2

	1	2	3	4	5	6	7	8	9	10	11
	Blok 1						Blok 2				
1	A2						A3				
2	B1		B3		B2		B3		B2		B1
3	A2B1C2	A2B1C1	A2B3C2	A2B3C1	A2B2C1	A2B2C2	A3B3C2	A3B3C1	A3B2C2	A3B2C1	A3B1C2
4	A3						A2				
5	B3		B1		B2		B2		B1		B3
6	A3B3C1	A3B3C2	A3B1C1	A3B1C2	A3B2C2	A3B2C1	A2B2C1	A2B2C2	A2B1C1	A2B1C2	A2B3C1
7	A1						A4				
8	B2		B1		B3		B2		B3		B1
9	A1B2C1	A1B2C2	A1B1C1	A1B1C2	A1B3C2	A1B3C1	A4B2C1	A4B2C2	A4B3C2	A4B3C1	A4B1C1
10	A4						A1				
11	B2		B1		B3		B3		B2		B1
12	A4B2C2	A4B2C1	A4B1C1	A4B1C2	A4B3C1	A4B3C2	A1B3C2	A1B3C1	A1B2C1	A1B2C2	A1B1C1

Schemat mówi np. że w pierwszym bloku, przypisujemy poziomy nr 2, 3, 1, 4 pierwszego czynnika kolejnym polom a na drugim z tych pól przypisujemy poziomy nr 3, 1, 2 drugiego czynnika kolejnym poletkom zaś na trzecim poletku przypisujemy poziomy nr 2 i 1 trzeciego czynnika małym poletkom, na które podzielono to poletko (A3B2C2, A3B2C1).

## Analiza



**Przykład** Otwieramy arkusz *trojczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-split-plot*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na listach zmiennych *Zmienna zależna* i *Blok* zmienne nr 5 i 1, a na liście *Czynniki* zmienne nr 2, 3, 4 po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji:

Dane: Analiza trójczynnikowa - split-split-plot - Wyniki* (5 zm. * 11 prz.)					
Analiza trójczynnikowa, split - split - plot					
Zmienna zależna: Przyjęcie sadzonek					
Precyzja doświadczenia: 9,6%					
Efekt	SS	Stopnie swobody	MS	F	p
<b>Blok</b>	0,07063	5	0,01413	1,84	0,260548
Podkładka	0,00552	1	0,00552	0,72	0,435802
Błąd I (Blok * Podkładka)	0,03847	5	0,00769		
Przygotowanie sadzonek	1,79193	3	0,59731	16,04	0,000002
Podkładka * Przygotowanie sadzonek	2,66716	3	0,88905	23,87	0,000000
Błąd II	1,11728	30	0,03724		
Zabieg	0,00107	2	0,00054	0,12	0,886893
Podkładka * Zabieg	0,01844	2	0,00922	2,07	0,133354
Przygotowanie sadzonek * Zabieg	0,02128	6	0,00355	0,79	0,576783
Podkładka * Przygotowanie sadzonek * Zabieg	0,05940	6	0,00990	2,22	0,049575
Błąd III	0,35696	80	0,00446		

Ponownie, kolejność czynników na liście *Czynniki* ma znaczenie, ze względu na ich hierarchię. Gdybyśmy zamiast 2, 3, 4 wskazali kolejność 2, 4, 3, to jedyna interakcja trzeciego rzędu przestałaby być istotna statystycznie (na standardowym poziomie istotności  $\alpha = 0,05$ ).

Ze względu na istotność statystyczną pewnych efektów jesteśmy zainteresowani przeprowadzeniem porównań szczegółowych. Przyjrzyjmy się interakcji *Podkładka x Przygotowanie sadzonek*. Wracamy więc do okna analizy i wskazujemy *Podkładka x Przygotowanie sadzonek* jako *Efekt do porównań wielokrotnych*, po czym naciskamy przycisk *Zbadaj efekt*.

Doświadczenia czynnikowe

Wybierz zmienną: Zmienne

Wskaż liczbę czynników: 3

Wybierz plan doświadczenia: split-split-plot

Poziom istotności: 0,050

Informacje

Zaplanuj dośw.

Wykonaj analizę

Anuluj

Efekt do porównań wielokrotnych:

- Podkładka x Przygotowa
- Podkładka
- Przygotowanie sadzonek
- Zabieg
- Podkładka x Przygotowanie s
- Podkładka x Zabieg
- Przygotowanie sadzonek x Z

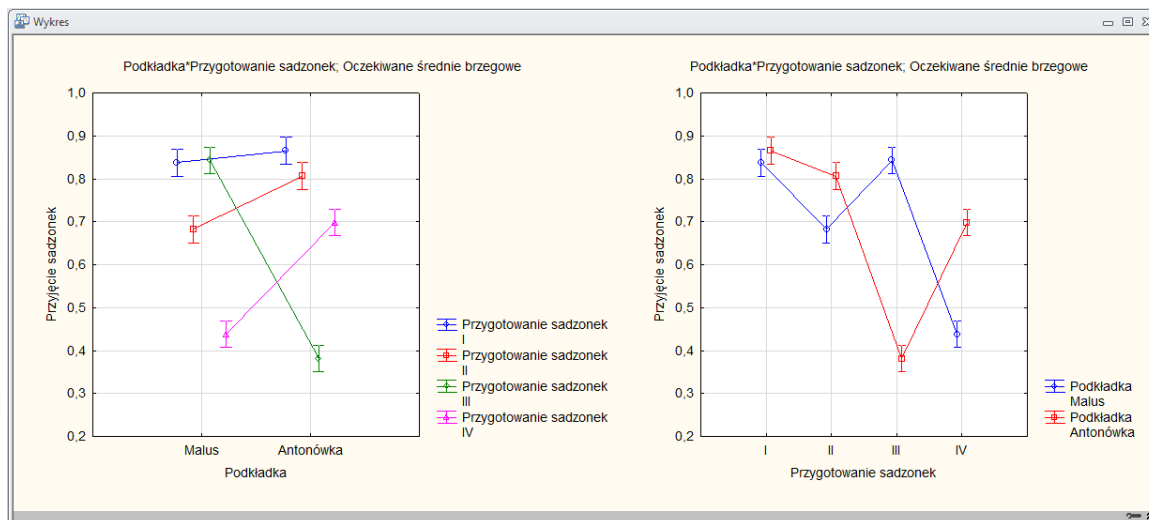
☒ Porównania wielokrotne i grupy jednorodne

☒ Średnie i wykres średnich/interakcji

☐ Przypisz osi drugi czynnik

Zbadaj efekt

Otrzymujemy zestaw wyników w postaci trzech arkuszy (wartości p, podział na grupy jednorodne, średnie brzegowe) i wykresu. Jeżeli badanym efektem jest interakcja dwu czynników, to czynne jest pole *Przypisz osi drugi czynnik*. Jeżeli w tym przypadku pozostawimy je domyślnie odznaczone, to na osi poziomej wykresu znajdzie się czynnik *Podkładka(L)*, w przeciwnym razie będzie to czynnik *Przygotowanie sadzonek(P)*.



### 3.3.2. Split-plot A-BC

Układ split-plot A-BC ma miejsce, gdy we wszystkich blokach każde z pól, do których losowo przydzielono poziomy czynnik A, dzielimy na *bc* poletek, do których następnie losowo przypisane zostają różne kombinacje czynników B i C.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-plot A-BC* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy cztery liczby - liczby poziomów czynnika 1, czynnika 2 i czynnika 3 oraz liczbę bloków. W przykładzie niech będą one równe 4, 2, 3 i 5. Po zatwierdzeniu OK utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej, a także arkusz do wprowadzania danych.

Dane: Schemat\* (30 zm., \* 8 prz.)

Analiza trójczynnikowa, split-plot A-BC  
 # bloków: 5  
 # poziomów czynnika I: 4  
 # poziomów czynnika II: 2  
 # poziomów czynnika III: 3

	1 Blok 1	2	3	4	5	6	7 Blok 2	8
1	A4						A1	
2	A4B2C3	A4B1C2	A4B1C3	A4B1C1	A4B2C1	A4B2C2	A1B2C3	A1B1C3
3	A1						A4	
4	A1B1C2	A1B1C3	A1B2C1	A1B2C3	A1B1C1	A1B2C2	A4B2C3	A4B2C2
5	A2						A2	
6	A2B2C3	A2B2C1	A2B1C2	A2B1C1	A2B2C2	A2B1C3	A2B2C3	A2B1C1
7	A3						A3	
8	A3B2C2	A3B1C1	A3B2C1	A3B1C3	A3B1C2	A3B2C3	A3B2C2	A3B1C2



Schemat mówi np. że w pierwszym bloku, przypisujemy poziomy nr 4, 1, 2, 3 pierwszego czynnika kolejnym polom a na trzecim z tych pól przypisujemy kolejnym jego poletkom kombinacje B2C3, B2C1, B1C2, B1C1, B2C2, B1C3 drugiego i trzeciego czynnika.

### 3.3.3. Split-plot AB-C

Ten układ powstaje, gdy we wszystkich blokach każde z  $ab$  pól z losowo przydzielonymi różnymi kombinacjami czynników A i B dzielimy na  $c$  poletek, do których losowo przypisuje się różne poziomy czynnika C.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-plot AB-C* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy cztery liczby - liczby poziomów czynnika 1, czynnika 2 i czynnika 3 oraz liczbę bloków. Niech wynoszą one 4, 2, 3 i 5. Po zatwierdzeniu OK utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat\* (15 zm., \* 16 prz.)

Analiza trójczynnikowa, split-plot AB-C  
 # bloków: 5  
 # poziomów czynnika I: 4  
 # poziomów czynnika II: 2  
 # poziomów czynnika III: 3

	1 Blok 1	2	3	4 Blok 2	5	6	7 Blok 3	8	9	10 Blok 4
1	A1B2			A4B1			A3B1			A4B2
2	A1B2C1	A1B2C2	A1B2C3	A4B1C2	A4B1C1	A4B1C3	A3B1C1	A3B1C2	A3B1C3	A4B2C1
3	A2B2			A2B2			A4B1			A1B1
4	A2B2C3	A2B2C2	A2B2C1	A2B2C3	A2B2C2	A2B2C1	A4B1C3	A4B1C2	A4B1C1	A1B1C2
5	A1B1			A3B1			A2B2			A4B1
6	A1B1C3	A1B1C2	A1B1C1	A3B1C2	A3B1C1	A3B1C3	A2B2C2	A2B2C3	A2B2C1	A4B1C2
7	A3B2			A1B2			A3B2			A1B2
8	A3B2C2	A3B2C3	A3B2C1	A1B2C2	A1B2C1	A1B2C3	A3B2C3	A3B2C2	A3B2C1	A1B2C1
9	A3B1			A4B2			A2B1			A3B2
10	A3B1C1	A3B1C2	A3B1C3	A4B2C2	A4B2C3	A4B2C1	A2B1C2	A2B1C3	A2B1C1	A3B2C1
11	A2B1			A3B2			A1B2			A2B1
12	A2B1C3	A2B1C1	A2B1C2	A3B2C1	A3B2C3	A3B2C2	A1B2C1	A1B2C3	A1B2C2	A2B1C1
13	A4B1			A1B1			A4B2			A2B2
14	A4B1C2	A4B1C3	A4B1C1	A1B1C3	A1B1C1	A1B1C2	A4B2C3	A4B2C1	A4B2C2	A2B2C2
15	A4B2			A2B1			A1B1			A3B1
16	A4B2C3	A4B2C2	A4B2C1	A2B1C2	A2B1C3	A2B1C1	A1B1C3	A1B1C2	A1B1C1	A3B1C1

Schemat mówi np. że w trzecim bloku, kolejnym polom przypisujemy kombinacje A3B1, A4B1, A2B2, A3B2, A2B1, A1B2, A4B2, A1B1 dwóch pierwszych czynników a poletkom piątego z nich przypisujemy odpowiednio poziomy nr 2, 3, 1 trzeciego czynnika (A2B1C2, A2B1C3, A2B1C1).

### 3.3.4. Split-plot x split-blok

Układ ten powstaje, kiedy czynnik C jest zagnieżdżony w czynniku B, który wraz z czynnikiem A są w układzie split-blok. W każdym bloku cały obszar dzielony jest najpierw w układzie split-blok na  $ab$  poletek względem dwu pierwszych czynników. Następnie każdy pas o jednym poziomie czynnika B dzielony jest na  $c$  wąskich pasów, do których losowo przypisujemy różne poziomy trzeciego czynnika.

## Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-plot x split-blok* po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy cztery liczby - liczby poziomów czynnika 1, czynnika 2 i czynnika 3 oraz liczbę bloków. Niech będą one równe 4, 3, 2 i 5. Po zatwierdzeniu *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej, a także arkusz do wprowadzania danych.

Dane: Schemat\* (30 zm. \* 4 prz.)

Analiza trójczynnikowa, split-plot x split-blok  
 # bloków: 5  
 # poziomów czynnika I: 4  
 # poziomów czynnika II: 3  
 # poziomów czynnika III: 2

	1	2	3	4	5	6	7	8	9	10	11
	Blok 1						Blok 2				
1	A4B1C1	A4B1C2	A4B2C1	A4B2C2	A4B3C2	A4B3C1	A4B2C1	A4B2C2	A4B3C2	A4B3C1	A4B1C2
2	A1B1C1	A1B1C2	A1B2C1	A1B2C2	A1B3C2	A1B3C1	A2B2C1	A2B2C2	A2B3C2	A2B3C1	A2B1C2
3	A3B1C1	A3B1C2	A3B2C1	A3B2C2	A3B3C2	A3B3C1	A1B2C1	A1B2C2	A1B3C2	A1B3C1	A1B1C2
4	A2B1C1	A2B1C2	A2B2C1	A2B2C2	A2B3C2	A2B3C1	A3B2C1	A3B2C2	A3B3C2	A3B3C1	A3B1C2

Schemat mówi np. że w pierwszym bloku, pierwszym dwóm czynnikom w układzie split-blok przypisujemy odpowiednio poziomy nr 4, 1, 3, 2 w jednym kierunku (A4, A1, A3, A2) oraz nr 1, 2, 3 (B1, B2, B3) w prostym do niego kierunku. Następnie pasy o jednakowym poziomie drugiego czynnika dzielimy na dwa węższe pasy i przypisujemy im poziomy trzeciego czynnika, odpowiednio: nr 1 i 2 w pasie B1, nr 1 i 2 w pasie B2, nr 2 i 1 w pasie B3.

## Analiza



**Przykład** Otwieramy arkusz *trójczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-plot x split-blok*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na listach zmiennych *Zmienna zależna*, *Blok* zmienne nr 5 i 1 a na liście *Czynniki* zmienne nr 2, 3, 4 po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji. Zwróćmy uwagę, że jest tu aż pięć odrębnych źródeł błędu.

Dane: Analiza trójczynnikowa - split-plot x split-blok - Wyniki\* (5 zm. \* 13 prz.)

Analiza trójczynnikowa, split - plot x split-blok  
 Zmienna zależna: Przyjęcie sadzonek  
 Precyzja doświadczenia: 9,3%

Efekt	SS	Stopnie swobody	MS	F	p
<b>Blok</b>	0,07063	5	0,01413	1,84	0,260548
Podkładka	0,00552	1	0,00552	0,72	0,435802
Błąd I (Blok * Podkładka)	0,03847	5	0,00769		
Przygotowanie sadzonek	1,79193	3	0,59731	34,39	0,000001
Błąd II (Blok * Przygotowanie sadzonek)	0,26054	15	0,01737		
Podkładka * Przygotowanie sadzonek	2,66716	3	0,88905	15,57	0,000071
Błąd III (Blok * Podkładka * Przygotowanie sadzonek)	0,85674	15	0,05712		
Zabieg	0,00107	2	0,00054	0,11	0,893116
Przygotowanie sadzonek * Zabieg	0,02128	6	0,00355	0,75	0,613462
Błąd IV	0,18927	40	0,00473		
Podkładka * Zabieg	0,01844	2	0,00922	2,20	0,124114
Podkładka * Przygotowanie sadzonek * Zabieg	0,05940	6	0,00990	2,36	0,047880
Błąd V	0,16769	40	0,00419		



### 3.3.5. Split-plot-blok

Układ ten ma miejsce, kiedy cały obszar doświadczalny najpierw dzielimy na  $a$  pól z losowym przydziałem różnych poziomów czynnika A, a następnie na każdym z nich określamy układ split-blok pozostałych dwu czynników.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-plot-blok*, po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy cztery liczby - liczby poziomów czynnika 1, czynnika 2 i czynnika 3 oraz liczbę bloków. W przykładzie niech będą one równe 4, 3, 4 i 5. Po zatwierdzeniu OK utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej, a także arkusz do wprowadzania danych.

Dane: Schemat* (15 zm. * 20 prz.)								
Analiza trójczynnikowa, split-plot-blok								
# bloków: 5								
# poziomów czynnika I: 4								
# poziomów czynnika II: 3								
# poziomów czynnika III: 4								
	1	2	3	4	5	6	7	8
	Blok 1			Blok 2			Blok 3	
1	A1			A1			A4	
2	A1B3C3	A1B2C3	A1B1C3	A1B2C2	A1B3C2	A1B1C2	A4B3C1	A4B1C1
3	A1B3C1	A1B2C1	A1B1C1	A1B2C3	A1B3C3	A1B1C3	A4B3C4	A4B1C4
4	A1B3C4	A1B2C4	A1B1C4	A1B2C4	A1B3C4	A1B1C4	A4B3C3	A4B1C3
5	A1B3C2	A1B2C2	A1B1C2	A1B2C1	A1B3C1	A1B1C1	A4B3C2	A4B1C2
6	A4			A2			A3	
7	A4B1C4	A4B3C4	A4B2C4	A2B2C4	A2B3C4	A2B1C4	A3B3C2	A3B1C2
8	A4B1C3	A4B3C3	A4B2C3	A2B2C1	A2B3C1	A2B1C1	A3B3C1	A3B1C1
9	A4B1C1	A4B3C1	A4B2C1	A2B2C3	A2B3C3	A2B1C3	A3B3C4	A3B1C4
10	A4B1C2	A4B3C2	A4B2C2	A2B2C2	A2B3C2	A2B1C2	A3B3C3	A3B1C3
11	A3			A4			A2	
12	A3B1C4	A3B2C4	A3B3C4	A4B1C2	A4B2C2	A4B3C2	A2B2C3	A2B1C3
13	A3B1C1	A3B2C1	A3B3C1	A4B1C1	A4B2C1	A4B3C1	A2B2C2	A2B1C2
14	A3B1C2	A3B2C2	A3B3C2	A4B1C3	A4B2C3	A4B3C3	A2B2C4	A2B1C4
15	A3B1C3	A3B2C3	A3B3C3	A4B1C4	A4B2C4	A4B3C4	A2B2C1	A2B1C1
16	A2			A3			A1	
17	A2B1C2	A2B2C2	A2B3C2	A3B1C2	A3B3C2	A3B2C2	A1B3C2	A1B1C2
18	A2B1C4	A2B2C4	A2B3C4	A3B1C1	A3B3C1	A3B2C1	A1B3C4	A1B1C4
19	A2B1C1	A2B2C1	A2B3C1	A3B1C3	A3B3C3	A3B2C3	A1B3C1	A1B1C1
20	A2B1C3	A2B2C3	A2B3C3	A3B1C4	A3B3C4	A3B2C4	A1B3C3	A1B1C3

Schemat mówi np. że w drugim bloku, kolejnym polom przypisujemy poziomy nr 1, 2, 4, 3 pierwszego czynnika, a na każdym z nich wprowadzono kombinacje drugiego i trzeciego czynnika w układzie pasów prostokątnych. Na przykład, na czwartym z tych pól (A3), są poziomy nr 1, 3, 2 drugiego czynnika oraz poziomy nr 2, 1, 3, 4 trzeciego czynnika.

#### Analiza



**Przykład** Otwieramy arkusz *trojczynnikowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-plot-blok*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na listach zmiennych *Zmienna zależna*, *Blok* zmienne nr 5 i 1 a na liście *Czynniki* zmienne nr 4, 3, 2 po czym

naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji.

Dane: Analiza trójczynnikowa - split-plot-blok - Wyniki\* (5 zm., \* 12 prz.)

Analiza trójczynnikowa, split - plot - blok  
Zmienna zależna: Przyjęcie sadzonek  
Precyzja doświadczenia: 21,3%

Efekt	SS	Stopnie swobody	MS	F	p
<b>Blok</b>	0,07063	5	0,01413	1,923	0,177240
Zabieg	0,00107	2	0,00054	0,073	0,930073
Błąd I (Blok * Zabieg)	0,07346	10	0,00735		
Przygotowanie sadzonek	1,79193	3	0,59731	71,420	0,000000
Zabieg * Przygotowanie sadzonek	0,02128	6	0,00355	0,424	0,859109
Błąd II	0,37635	45	0,00836		
Podkładka	0,00552	1	0,00552	1,063	0,318833
Zabieg * Podkładka	0,01844	2	0,00922	1,777	0,202944
Błąd III	0,07782	15	0,00519		
Przygotowanie sadzonek * Podkładka	2,66716	3	0,88905	40,614	0,000000
Zabieg * Przygotowanie sadzonek * Podkładka	0,05940	6	0,00990	0,452	0,839561
Błąd IV	0,98507	45	0,02189		

### 3.3.6. Split-blok-plot

Gdy czynnik C jest zagnieżdżony względem kombinacji czynników A i B ułożonych w układzie split-blok, to stosujemy układ split-blok-plot. W każdym bloku cały obszar dzielony jest w układzie split-blok na  $ab$  poletek pod kątem dwu pierwszych czynników a następnie każde poletko dzielone jest na  $c$  małych poletek, do których losowo przypisujemy różne poziomy trzeciego czynnika.

#### Planowanie



**Przykład** W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-blok-plot*, po czym naciskamy przycisk *Zaplanuj doświadczenie*. Wpisujemy cztery liczby - liczby poziomów czynnika 1, czynnika 2 i czynnika 3 oraz liczbę bloków. Niech będą równe 4, 2, 3 i 4. Po zatwierdzeniu przyciskiem *OK* utworzony zostanie arkusz ze schematem doświadczenia, jak ten ukazany poniżej i oprócz niego także arkusz do wprowadzania danych.

Dane: Schemat\* (24 zm., \* 8 prz.)

Analiza trójczynnikowa, split-blok-plot  
# bloków: 4  
# poziomów czynnika I: 4  
# poziomów czynnika II: 2  
# poziomów czynnika III: 3

	1 Blok 1	2	3	4	5	6	7 Blok 2	8	9
1	A4B1			A4B2			A1B2		
2	A4B1C1	A4B1C2	A4B1C3	A4B2C1	A4B2C3	A4B2C2	A1B2C1	A1B2C3	A1B2C2
3	A1B1			A1B2			A2B2		
4	A1B1C2	A1B1C1	A1B1C3	A1B2C1	A1B2C2	A1B2C3	A2B2C2	A2B2C1	A2B2C3
5	A3B1			A3B2			A4B2		
6	A3B1C2	A3B1C1	A3B1C3	A3B2C1	A3B2C3	A3B2C2	A4B2C2	A4B2C1	A4B2C3
7	A2B1			A2B2			A3B2		
8	A2B1C3	A2B1C1	A2B1C2	A2B2C1	A2B2C2	A2B2C3	A3B2C1	A3B2C3	A3B2C2

Schemat mówi np. że w pierwszym bloku, poziomy odp. nr 4, 1, 3, 2 oraz nr 1, 2 dwu pierwszych czynników zostały rozmieszczone w układzie pasów prostopadłych a każde z 8 powstałych poletek podzielono na 3 małe poletka, do których przypisano losowo poziomy trzeciego czynnika. Przykładowo, małym poletkom A3B1 przypisano kolejno poziomy nr 2, 1, 3 trzeciego czynnika.

## Analiza



**Przykład** Otwieramy arkusz *trojczynnikiowe.sta*. W module *Układy ANOVA* wskazujemy liczbę czynników równą 3 i wybieramy plan doświadczenia *split-blok-plot*. Następnie po kliknięciu przycisku wyboru zmiennych wskazujemy na listach zmiennych *Zmienna zależna*, *Blok* zmienne nr 5 i 1 a na liście *Czynniki* zmienne nr 3, 4, 2 po czym naciskamy przycisk *Wykonaj analizę*. Dostajemy w wyniku poniższy arkusz analizy wariancji.

Dane: Analiza trójczynnikowa - split-blok-plot - Wyniki* (5 zmn. * 12 prz.)					
Analiza trójczynnikowa, split-blok-plot Zmienna zależna: Przyjęcie sadzonek Precyzja doświadczenia 19,2%					
Efekt	SS	Stopnie swobody	MS	F	p
<b>Blok</b>	0,07063	5	0,01413	0,813	0,558404
Przygotowanie sadzonek	1,79193	3	0,59731	34,389	0,000001
Błąd I (Blok * Przygotowanie sadzonek)	0,26054	15	0,01737		
Zabieg	0,00107	2	0,00054	0,073	0,930073
Błąd II (Blok * Zabieg)	0,07346	10	0,00735		
Przygotowanie sadzonek * Zabieg	0,02128	6	0,00355	0,919	0,495420
Błąd III (Blok * Przygotowanie sadzonek * Zabieg)	0,11581	30	0,00386		
Podkładka	0,00552	1	0,00552	0,311	0,578913
Przygotowanie sadzonek * Podkładka	2,66716	3	0,88905	50,187	0,000000
Zabieg * Podkładka	0,01844	2	0,00922	0,520	0,596903
Przygotowanie sadzonek * Zabieg * Podkładka	0,05940	6	0,00990	0,559	0,761223
Błąd IV	1,06290	60	0,01771		

### 3.3.7. Bloki losowe

W doświadczeniu trójczynnikowym o układzie bloków losowych wykonywany jest losowy przydział wszystkich obiektów do abc małych poletek, na które dzielony jest w każdym bloku obszar, tak, że każdy obiekt wystąpi jednokrotnie i w związku z tym w całym doświadczeniu pojawi się tyle razy ile jest bloków.

### 3.3.8. Układ całkowicie losowy

Tutaj każdy spośród abc obiektów, czyli kombinacji poziomów czynników, zostaje powtórzony tę samą liczbę razy w całym doświadczeniu przy czym prócz tego warunku przydział obiektów do małych poletek jest całkowicie losowy. Jeśli powtórzenia potraktujemy jako bloki, to dany obiekt może się nie pojawić w pewnym bloku a w innym wystąpić co najmniej dwukrotnie.

## 4. Zaawansowane

### 4.1 Doświadczenia kratowe

Moduł *Doświadczenia kratowe* pozwala zarówno na analizę jak i planowanie doświadczeń kratowych. Bada się w nich pojedynczy czynnik a obiekty (zabiegi), będące jego poziomami, są pogrupowane w równoliczne bloki, które z kolei są pogrupowane w kraty. Grupowanie jest wykonywane w ten sposób, że w każdej kratce każdy obiekt pojawia się dokładnie jeden raz. Układ ten stosuje się gdy badanych obiektów jest dosyć dużo, znacznie więcej niż może wejść do pojedynczego bloku.

#### Analiza

Do analizy wyników doświadczenia kratowego wskazujemy cztery zmienne: zależną oraz oznaczające: kratę, blok, obiekt. Wynikiem jest tabela analizy wariancji o efektach: kraty, blok zagnieżdżony w kratce, obiekt(zabieg) i na końcu z błędem losowym.



**Przykład** W pliku *kratowe.sta* znajdują się dane o liczbie rozwiniętych sadzonek w zależności od uprzednio zastosowanego jednego z 20 zabiegów sadowniczych. Zabiegi pogrupowano w bloki długości 4, w związku z czym w każdej kratce znalazło się 5 bloków. Wchodzimy do modułu *Doświadczenia kratowe* na zakładkę *Analiza*, wybieramy poziom istotności po czym naciskamy przycisk wyboru zmiennych.

Na kolejnych listach zmiennych: *Zmienna zależna*, *Kraty*, *Obiekt* i *Zmienna z blokami*, wskazujemy odpowiednie zmienne (*Liczba sadzonek*, *Kraty*, *Zabieg* i *Blok*), po czym klikamy *OK*.

Dane: Analiza doświadczenia kratowego. Zmienna zależna: Liczba sadzonek...

Efekt	SS	Stopnie swobody	MS	F	p
Krata	433,6	2	216,8	0,6379	0,536486
Blok(Krata)	5861,5	12	488,5	1,4371	0,211790
<b>Zabieg</b>	<b>69522,6</b>	<b>19</b>	<b>3659,1</b>	<b>10,7658</b>	<b>0,000000</b>
Błąd losowy	8836,9	26	339,9		

Do skrótytu wyników trafia powyższy arkusz o postaci tabeli analizy wariancji. Widzimy, że rodzaj zabiegu istotnie wpływa na liczbę sadzonek a ani przydzielony blok ani krata, mająca charakter powtórzenia, nie okazały się istotne, co jest zresztą pożądane. W przypadku statystycznej istotności zabiegu, do skrótytu trafia również arkusz z podziałem rodzajów zabiegu na grupy jednorodne i z wartościami średnimi zmiennej zależnej.

## Planowanie

Aby zaplanować doświadczenie kratowe, należy tylko podać liczbę krat równą  $k$ , liczbę bloków w kracie równą  $b$  i liczbę obiektów w bloku równą  $m$ . Wówczas wszystkich różnych obiektów będzie  $bm$  i zostaną one rozmieszczone w kratkach w taki sposób żeby w jak największym stopniu spełnić warunek zrównoważenia (z idealnym zrównoważeniem mamy do czynienia, kiedy każda para obiektów wspólnie występuje w jednakowej liczbie bloków). Odpowiednie liczby należy wpisać w polach zakładki *Planowanie*.

Przykładowo, mając 21 obiektów i chcąc je umieszczać po 7 w bloku, potrzeba 3 bloków na kratę i jeżeli są 4 kraty, to należy wpisać kolejne te wartości w polach *Liczba obiektów w bloku*, *Liczba bloków w kracie*, *Liczba krat*, jak przedstawiono poniżej, a następnie kliknąć OK. Wynikiem jest arkusz o postaci gotowej do wprowadzania wartości zmiennej zależnej.

## 4.2 Serie doświadczeń

Moduł *Serie doświadczeń* służy do analizy doświadczenia o zadanym układzie, które jest powtarzane. Często oznacza to powtarzanie w różnych lokalizacjach lub latach - stąd nazwa „serie”. Dostępne są trzy układy: *AMMI*, *Seria split-plotów*, *Seria bloków losowych*.

Układ doświadczenia	Układ serii doświadczeń	Uwagi
dwuczynnikowe z interakcją	trójczynnikowy AMMI	'seria' - czynnik losowy
split-plot	split-split-plot*	'seria' - czynnik stały, 'blok' - czynnik losowy
bloki losowe	hierarchiczno-krzyżowy	

\*z drobną zmianą - włączeniem bloku do błędu I

#### 4.2.1. Układ AMMI

AMMI to szczególnego rodzaju doświadczenie dwuczynnikowe, w którym interakcja czynników rozkładana jest na składowe za pomocą analizy składowych głównych. Zmienna zależna domyślnie oznacza plon a czynniki odmianę(genotyp) i miejsce (lokalizację). Postać wyników dotyczących powyższych wielkości sprawia, że analizę AMMI (ang. *additive main effects and multiplicative interaction*) można uważać za rodzaj rozszerzonej analizy wariancji, gdzie istotność składowych członów interakcji badanych czynników jest oceniana metodą Corneliusa.

Do skróty z wynikami trafiają arkusz postaci jak w zwykłej analizie wariancji, dwa wykresy, zwane biplotami oraz drugi arkusz, z miarami adaptacji odmian. Podawane są następujące trzy miary: miara nadrzędności plonowania, miara niezawodności przewagi plonowania i miara Kanga stabilności. Na biplotach widnieją znaczniki dwóch kształtów, jednego dla odmian a drugiego dla miejsc. Ich współrzędne wynikają ze składowych interakcji obu tych czynników i dlatego wykresy te mogą być uznane miarodajnymi, w przypadku biplotu pierwszego rodzaju jeśli interakcja w ogóle jest istotna a w przypadku biplotu drugiego rodzaju jeśli dwie pierwsze składowe są istotne. Możemy z nich odczytywać to czy dana odmiana i dane miejsce sprzyjają sobie, w sensie, że plon jest obfitszy niż wynikałoby to z sumy osobnych ich wpływów, czy ma miejsce coś przeciwnego.



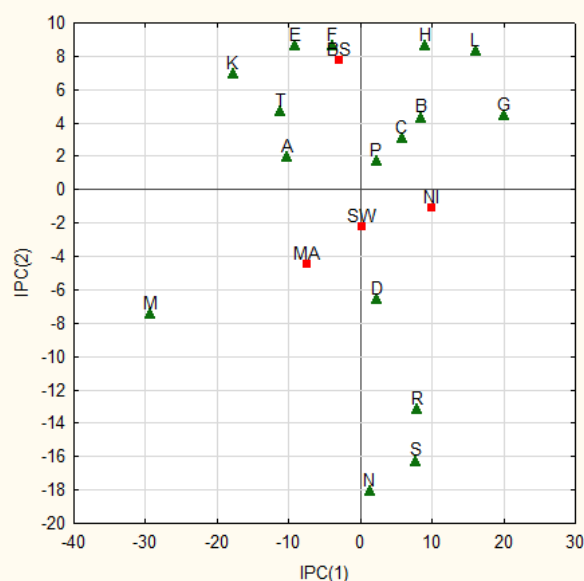
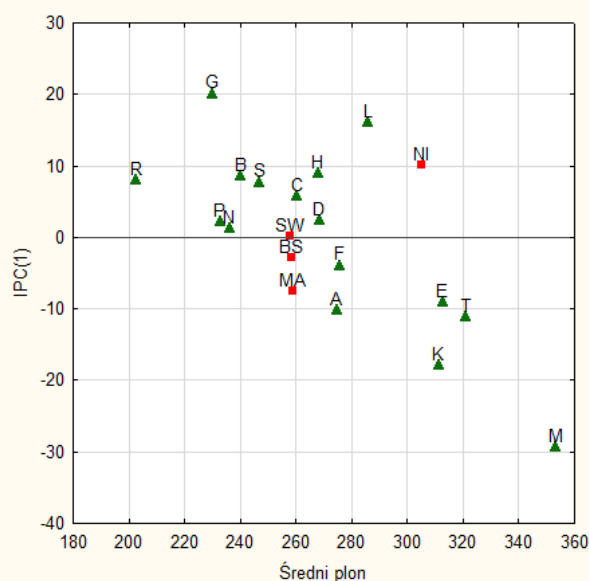
**Przykład** W pliku *ammi.sta* znajdują się dane o wielkości plonu w zależności od odmiany zboża i miejsca uprawy. Do doświadczenia weszło 16 odmian badanych w 4 miejscach, wszystko w 3 powtórzeniach. Zauważmy, że występują braki danych - nie jest to ograniczeniem w analizie. Wchodzimy do modułu *Serie doświadczeń* i wybieramy układ *AMMI*, (możemy też wybrać poziom istotności) po czym naciskamy przycisk wyboru zmiennych. Na kolejnych czterech listach zmiennych: *Zależna*, *Odmiana/Genotyp*, *Miejsce/Środowisko*, *Powtórzenie* wskazujemy odp. zmienne nr 4, 1, 2, 3 i zatwierdzamy *OK*.

Najpierw patrzymy na wyniki w pierwszym arkuszu (drugi arkusz zawiera po prostu wartości liczbowe trzech rodzajów wskaźników adaptacji odmian).

Dane: AMMI - analiza wariancji* (7 zmn. * 9 prz.)							
Źródło zmienności	Liczba stopni swobody	Suma kwadratów	Średni kwadrat	Odsetek zmienności	Względem interakcji	Statystyka F	Wartość p
Ogół	189	2932147	15514	100,0			
Miejsce	3	77758	25919	2,7		4,479	0,039969
Powtórzenie(Miejsce)	8	46298	5787	1,6		0,816	0,590271
Odmiana	15	282514	18834	9,6		2,654	0,001694
Odmiana x Miejsce	45	1688266	37517	57,6		5,287	0,000000
Składowa nr 1	17	1237434	72790	42,2	73,3	5,287	0,000000
Składowa nr 2	15	327957	21864	11,2	19,4	2,269	0,001243
Składowa nr 3	13	122792	9446	4,2	7,3	1,332	0,204156
Błąd losowy	118	837312	7096	28,6			

Obydwa badane czynniki, tj. miejsce i odmiana okazały się istotne, jak również ich interakcja, które istotne są dwie pierwsze składowe. Kierujemy teraz uwagę na wykresy oparte o składowe główne – tzw. *biploty*.





Jak już wiadomo, pierwszy biplot (lewy) możemy uważać za miarodajny i go w ogóle odczytywać gdy interakcja jest istotna a drugi biplot (prawy) - gdy co najmniej dwie pierwsze składowe są istotne. W naszym przypadku właśnie tak jest. Na obu odmiany mają zielone znaczniki a miejsca czerwone. Pierwszy wykres mówi, że odmiana i miejsce mające znaczniki po tej samej stronie czarnej poziomej kreski ( $IPC(1) = 0$ ) sprzyjają sobie, np. odmiana 'B' z miejscem 'NI'. Z drugiego wykresu możemy wyczytać, że np. odmiana 'E' lub też odmiana 'F' jest dobra w miejscu 'BS' jako, że ich znaczniki leżą w tym samym kierunku od przecięcia się czarnych kresek. To samo miejsce nie sprzyja natomiast odmianom 'R', 'D' czy 'S' ze względu na przeciwieństwo kierunków.

#### 4.2.2. Serie doświadczeń split-plot

Seria doświadczeń split-plot przypomina doświadczenie w układzie split-split-plot. Jediną różnicą tutaj jest włączenie bloku do błędu I. Zostanie to bardziej naświetlone w poniższym przykładzie.



**Przykład** Plik *seria split plot.sta* zawiera dane o odsetku przyjęcia się sadzonek w zależności od dwu czynników, bloku i roku, który traktujemy jako czynnik serii. Wchodzimy do modułu *Serie doświadczeń* i wybieramy układ *Seria split-plotów*, (możemy też wybrać poziom istotności) po czym naciskamy przycisk wyboru zmiennych. Na kolejnych czterech listach zmiennych: *Zależna*, *Seria*, *Blok*, *Czynniki* wskazujemy odp. zmienne nr 5, 4, 1, (2 i 3) i zatwierdzamy OK. Poniżej wynikowy arkusz.

Dane: Analiza wariancji\* (5 zmnn. \* 10 prz.)

Efekt	Zmienna zależna: Przyjęcie sadzonek				
	SS	Stopnie swobody	MS	F	p
Rok	0,00130	2	0,00065	0,057	0,945262
Błąd I	0,17269	15	0,01151		
Podkładka	0,00123	1	0,00123	0,213	0,651100
Rok * Podkładka	0,00629	2	0,00314	0,544	0,591360
Błąd II	0,08663	15	0,00578		
Przygotowanie sadzonek	1,73650	3	0,57883	50,054	0,000000
Rok * Przygotowanie sadzonek	0,01126	6	0,00188	0,162	0,985987
Podkładka * Przygotowanie sadzonek	2,60459	3	0,86820	75,076	0,000000
Rok * Podkładka * Przygotowanie sadzonek	0,00367	6	0,00061	0,053	0,999374
Błąd III	1,04078	90	0,01156		

Dostalibyśmy niemal te same wyniki traktując to doświadczenie jako split-split-plot o czynnikach (4 + 2 + 3), czyli z czynnikiem serii jako pierwszym i rzecz jasna z tą samą zmienną blokową. Jedyna różnica dotyczy istotności czynnika serii, ponieważ jest ona testowana względem błędu I, do którego włącza się tutaj sumę kwadratów odp. blokom, podczas gdy w dośw. split-split-plot blok jest osobnym efektem.

#### 4.2.3. Serie bloków losowych

Bloki losowe oznaczają tutaj doświadczenie z pojedynczym czynnikiem oraz zmienną blokową. Ich seria nosi nazwę *układu hierarchiczno-krzyżowego*.

Aby wykonać analizę, należy w module *Serie doświadczeń* wybrać układ *Seria bloków losowych* i wskazać zmienne na kolejnych czterech listach: *Zależna*, *Seria*, *Blok*, *Obiekt* i zatwierdzić przyciskiem *OK*. Wspomniany pojedynczy czynnik zaznaczamy na liście *Obiekt*. Warto też nadmienić, że w efektach wynikowej tabeli analizy wariancji blok zostaje zagnieżdżony w czynniku serii.

### 4.3 Doświadczenia hodowlane

Moduł *Doświadczenia hodowlane* daje możliwość analizy pięciu różnych układów doświadczeń z krzyżowaniem form matecznych i form ojcowskich: linia x tester (top-cross) i cztery rodzaje tzw. układów diallelicznych. Dla każdego z nich skoroszyt z wynikami zawiera cztery arkusze: tabelę analizy wariancji, tabele z obliczonymi wartościami GCA i wartościami SCA oraz ze współczynnikami odziedziczalności w wąskim i szerokim sensie. Oprócz wskazania zmiennych oznaczających formy i zmiennej zależnej możemy ale nie musimy wskazać zmienną blokową oraz wybrać model stały lub losowy.

#### 4.3.1. Doświadczenia linia x tester

W doświadczeniach tego rodzaju *linia* odpowiada formie matecznej a *tester* ojcowskiej. W przeciwieństwie do układów diallelicznych oba zestawy form nie muszą być jednakowe, a nawet mogą być zupełnie odrębne i nierównoliczne. W analizie wariancji ocenianymi efektami są linia, tester i ich interakcja.





**Przykład** Otwieramy plik *top cross.sta*, uruchamiamy moduł *Doświadczenia hodowlane* i wskazujemy układ badania *Top-cross*. Pozostawiamy domyślnie ustawione efekty stałe. Za pomocą przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych, tj. *Zmienna zależna*, *Linia*, *Tester*, *Zmienna z blokami* zmienne nr 1, 2, 3, 4 i zatwierdzamy przyciskiem *OK*. Poniżej widoczny jest arkusz z wynikami analizy wariancji. Zauważmy, że otrzymalibyśmy inne wyniki gdybyśmy nie wskazali zmiennej blokowej i/lub wybrali efekty losowe zamiast stałych.

Dane: Analiza wariancji\* (5 zmn. \* 5 prz.)

Doświadczenie linia x tester  
Efekty stałe

Źródło zmienności	SS	Stopnie swobody	MS	F	p
Linia	8156	4	2039	0,609	0,657290
Tester	8067	4	2017	0,602	0,662022
Linia * Tester	79931	16	4996	1,492	0,118727
Blok	26593	4	6648		
Błąd losowy	321530	96	3349		

#### 4.3.2. Diallel kwadratowy

Diallel kwadratowy oznacza pełen układ  $p^2$  kombinacji krzyżowania każdej pary form matecznych i ojcowskich, gdzie  $p$  to liczba form. Ocenianymi w analizie wariancji efektami są forma(odmiana), krzyżowanie proste i krzyżowanie odwrotne.



**Przykład** Otwieramy plik *dialkw.sta*, uruchamiamy moduł *Doświadczenia hodowlane* i wskazujemy układ badania *Diallel kwadratowy*. Pozostawiamy domyślnie ustawione efekty stałe. Za pomocą przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych, tj. *Zmienna zależna*, *Forma mateczyna*, *Forma ojcowska*, *Zmienna z blokami* zmienne nr 1, 2, 3, 4 i zatwierdzamy *OK*. Poniżej pokazano arkusz z wynikami analizy wariancji. Otrzymalibyśmy inne wyniki gdybyśmy nie wskazali zmiennej blokowej i/lub wybrali efekty losowe zamiast stałych.

Dane: Analiza wariancji\* (5 zmn. \* 5 prz.)

Diallel kwadratowy  
Efekty stałe

Źródło zmienności	1 SS	2 Stopnie swobody	3 MS	4 F	5 p
Odmiana	4516	4	1129	0,337	0,852335
Krzyżowanie	27870	10	2787	0,832	0,598873
Krzyżowanie odwrotne	63767	10	6377	1,904	0,053725
Blok	26593	4	6648		
Błąd losowy	321530	96	3349		

#### 4.3.3. Diallel trójkątny

W diallelu trójkątnym rozpatruje się  $p(p+1)/2$  kombinacji - potomstwo krzyżowania prostego oraz potomstwo badanych form, gdzie  $p$  to liczba form. Ocenianymi w analizie wariancji efektami są forma(odmiana) i krzyżowanie proste.



**Przykład** Otwieramy plik *dialtr.sta*, uruchamiamy moduł *Doświadczenia hodowlane* i wskazujemy układ badania *Diallel trójkątny*. Pozostawiamy domyślnie ustawione efekty stałe. Za pomocą przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych, tj. *Zmienna zależna*, *Forma matczyzna*, *Forma ojcowska*, *Zmienna z blokami* zmienne nr 1, 2, 3, 4 i zatwierdzamy OK. Poniżej przedstawiono arkusz z wynikami analizy wariancji. Wyniki różniłyby się od poniższych gdybyśmy nie wskazali zmiennej blokowej i/lub wybrali efekty losowe zamiast stałych.

Dane: Analiza wariancji* (5 zm. * 4 prz.)					
Źródło zmienności	Diallel trójkątny				
	Efekty stałe				
	1 SS	2 Stopnie swobody	3 MS	4 F	5 p
Odmiana	14103	4	3526	1,017	0,406516
Krzyżowanie	40747	10	4075	1,175	0,326712
Blok	26710	4	6678		
Błąd losowy	194137	56	3467		

#### 4.3.4. Diallel kwadratowy bez przekątnej

Diallel kwadratowy bez przekątnej to układ bez uwzględniania potomstwa badanych form - wykonywane jest tylko krzyżowanie proste i odwrotne. Łączna liczba kombinacji wynosi zatem  $p(p - 1)$ , gdzie  $p$  to liczba form. W analizie wariancji ocenianymi efektami są, tak samo jak w pełnym diallelu kwadratowym: forma(odmiana), krzyżowanie proste i krzyżowanie odwrotne.



**Przykład** Otwieramy plik *dialkwbp.sta*, uruchamiamy moduł *Doświadczenia hodowlane* i wskazujemy układ badania *Diallel kwadratowy bez przekątnej*. Pozostawiamy domyślnie ustawione efekty stałe. Za pomocą przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych, tj. *Zmienna zależna*, *Linia*, *Tester*, *Zmienna z blokami* zmienne nr 1, 2, 3, 4 i zatwierdzamy OK. Poniżej widać arkusz z wynikami analizy wariancji. Otrzymalibyśmy inne wyniki gdybyśmy nie wskazali zmiennej blokowej i/lub wybrali efekty losowe zamiast stałych.

Dane: Analiza wariancji* (5 zm. * 5 prz.)					
Źródło zmienności	Diallel kwadratowy bez przekątnej				
	Efekty stałe				
	1 SS	2 Stopnie swobody	3 MS	4 F	5 p
Odmiana	3120	4	780	0,241	0,914079
Krzyżowanie	14621	5	2924	0,904	0,482763
Krzyżowanie odwrotne	63767	10	6377	1,972	0,048141
Blok	34305	4	8576		
Błąd losowy	245710	76	3233		

#### 4.3.5. Diallel trójkątny bez przekątnej

Diallel trójkątny bez przekątnej obejmuje jedynie  $p(p - 1)/2$  krzyżowań prostych, gdzie  $p$  to liczba form. Nie uwzględnia się w nim krzyżowań odwrotnych ani potomstwa badanych form

rodzicielskich. W analizie wariancji ocenia się te same co w zwykłym diallelu trójkątnym dwa efekty: formę i krzyżowanie proste.



**Przykład** Otwieramy plik *dialtrbp.sta*, uruchamiamy moduł *Doświadczenia hodowlane* i wskazujemy układ badania *Diallel trójkątny bez przekątnej*. Pozostawiamy domyślnie ustawione efekty stałe. Za pomocą przycisku wyboru zmiennych wskazujemy na kolejnych listach zmiennych, tj. *Zmienna zależna*, *Linia*, *Tester* oraz *Zmienna z blokami*, zmienne nr 1, 2, 3, 4 i zatwierdzamy *OK*. Poniżej widnieje arkusz z wynikami analizy wariancji. Wyniki różniłyby się od poniższych gdybyśmy nie wskazali zmiennej blokowej i/lub wybrali efekty losowe zamiast stałych.

Dane: Analiza wariancji* (5 zmnn. * 4 prz.)					
Źródło zmienności	Diallel trójkątny bez przekątnej				
	Efekty stałe				
	1 SS	2 Stopnie swobody	3 MS	4 F	5 p
Odmiana	18653	4	4663	1,500	0,222737
Krzyżowanie	21408	5	4282	1,377	0,255828
Blok	40821	4	10205		
Błąd losowy	111918	36	3109		

#### 4.4 Doświadczenia z wzorcem

Moduł *Doświadczenia z wzorcem* pozwala na analizę i planowanie tzw. doświadczeń jednopowtórzeniowych z wzorcem. W tym układzie w każdym bloku występują obiekty wzorcowe i niewzorcowe. Każdy obiekt niewzorcowy występuje tylko w jednym z bloków i to jednokrotnie, zaś każdy obiekt wzorcowy występuje we wszystkich blokach, po tyle samo razy, z tym, że krotność ta nie musi być jednakowa u wszystkich obiektów wzorcowych.

##### Analiza

W wynikach analizy doświadczenia jednopowtórzeniowego z wzorcem są trzy testowane efekty: obiekty wzorcowe, obiekty niewzorcowe i kontrast.



**Przykład** W pliku *wzorzec.sta* znajdują się wartości zmiennej zależnej w kolejnych blokach dla obiektów wzorcowych ('P1', 'P2', 'P3') oraz dla obiektów niewzorcowych (o nazwach zaczynających się od 'B'). Obiekt wzorcowy 'P1' występuje w każdym bloku dwukrotnie, 'P2' czterokrotnie a 'P3' jednokrotnie. Nie ma zmiennej oznaczającej czy obiekt jest wzorcowy czy niewzorcowy, ponieważ nigdy nie jest potrzebna - obiekty wzorcowe to te, które występują więcej niż raz. Wchodzimy do modułu *Doświadczenia z wzorcem* na zakładkę *Analiza* po czym naciskamy przycisk wyboru zmiennych. Na kolejnych trzech listach zmiennych - *Zmienna zależna*, *Obiekt* i *Zmienna z blokami* - wskazujemy odp. zmienne nr 3, 2, 1 i zatwierdzamy *OK*. Otrzymujemy następujące wyniki.

Dane: Analiza doświadczenia jednopowtórzeniowego ze wzorcem. Zmienna zależna: Zależna\* (5 zm. \* 5 prz.)

Efekt	1 SS	2 Stopnie swobody	3 MS	4 F	5 p
<b>Blok</b>	4,198	5	0,840		
Obiekty niewzorcowe	20,597	17	1,212	1,01	0,469
Obiekty wzorcowe	14,455	2	7,228	6,04	0,006
Kontrast	74,809	1	74,809	62,56	0,000
Błąd losowy	40,660	34	1,196		

Wartości  $p$  w skrajnej prawej kolumnie mówią (zakładamy standardowy poziom istotności  $\alpha = 0,05$ ), że obiekty niewzorcowe nie różnią się istotnie od siebie we wpływie na poziom zmiennej zależnej podczas gdy obiekty wzorcowe różnią się. Dodatkowo, kontrast okazał się istotny statystycznie, co oznacza, że obiekty niewzorcowe różnią się od obiektów wzorcowych.

## Planowanie

Aby zaplanować doświadczenie jednopowtórzeniowe ze wzorcem podajemy liczbę bloków równą  $b$ , liczbę obiektów niewzorcowych w bloku równą  $m$ , liczbę obiektów wzorcowych w bloku równą  $n$  oraz następnie  $n$  liczb – ozn.  $k_1, \dots, k_n$  – oznaczające krotności występowania w bloku obiektów wzorcowych. Wówczas wszystkich obiektów niewzorcowych będzie  $bm$  i będą one wszystkie różne, natomiast  $i$ -ty obiekt wzorcowy wystąpi w całym doświadczeniu łącznie  $bk_i$  razy.



**Przykład** W celu zaplanowania doświadczenia wchodzimy do modułu *Doświadczenia z wzorcem* na zakładkę *Planowanie* i wpisujemy odpowiednie liczby w polach *Liczba bloków*, *Liczba obiektów niewzorcowych w bloku* i *Liczba obiektów wzorcowych w bloku*, tutaj(L) są to wartości 4, 7 i 4, z czego m.in. wynika, że łącznie jest 28 obiektów niewzorcowych. Następnie naciskamy przycisk *Krotności ob. wzorcowych* i w oknie, które się otworzy(P) wpisujemy w prawej kolumnie krotności kolejnych obiektów, przykładowo obiekt wzorcowy nr 2 pojawi się w każdym bloku pięciokrotnie.

Doświadczenia z wzorcem

Analiza Planowanie

Liczba bloków 4

Liczba obiektów niewzorcowych w bloku 7

Liczba obiektów wzorcowych w bloku 4

Krotności ob. wzorcowych

Anuluj OK

Krotności obiektów wzorcowych

Obiekt wzorcowy	Krotność
1	2
2	5
3	1
4	3

Dla każdego obiektu wzorcowego wprowadź jego krotność

OK Anuluj

Po potwierdzeniu OK utworzony zostanie skoroszyt z dwoma arkuszami, ukazanym niżej schematem przydziału obiektu do bloków oraz arkuszem do wprowadzania danych.

Obiekty niewzorcowe są oznaczone kolejnymi liczbami całkowitymi a obiekty wzorcowe dodatkowo literą 'P'.

Dane: Schemat\* (4 zm., \* 18 prz.)

Doświadczenie jednopowtórzeniowe ze wzorcem - plan

# bloków: 4  
# obiektów niewzorcowych: 7  
# obiektów wzorcowych: 4

	1	2	3	4
	Blok 1	Blok 2	Blok 3	Blok 4
1	24	20	8	19
2	18	10	12	6
3	14	2	27	17
4	3	1	13	25
5	5	11	23	22
6	15	9	7	26
7	28	4	21	16
8	P1	P1	P1	P1
9	P1	P1	P1	P1
10	P2	P2	P2	P2
11	P2	P2	P2	P2
12	P2	P2	P2	P2
13	P2	P2	P2	P2
14	P2	P2	P2	P2
15	P3	P3	P3	P3
16	P4	P4	P4	P4
17	P4	P4	P4	P4
18	P4	P4	P4	P4

## 4.5 AMOVA

Rozwinięcie angielskiego skrótu AMOVA to *analiza wariancji molekularnej* (*Analysis of MOlecular VAriance*). Analiza ta ma zastosowanie do danych genetycznych. Wejściowym zbiorem danych jest arkusz posiadający zmienną wyznaczającą podział wszystkich przypadków na grupy, dodatkowo może mieć, ale niekoniecznie, zmienną wyznaczającą podział grup na podgrupy. Wszystkie pozostałe zmienne oznaczają pojedyncze markery genetyczne, zazwyczaj są to zmienne zerojedynkowe.

Po wskazaniu zmiennej grupującej oraz być może zmiennej wskazującej podgrupy, jak i ustaleniu liczby permutacji do testów, otrzymujemy wynik w postaci tabeli przypominającej tę jak przy zwykłej analizie wariancji. Oprócz sumy kwadratów, średniego kwadratu i liczby stopni swobody, dla każdego źródła zmienności widnieje jeszcze odsetek zmienności, statystyka  $\Phi$  oraz istotność, do której nie podaje się dokładnej wartości p ze względu na zastosowanie testów permutacyjnych (obliczenia na podstawie Excoffier i in.. (1992)), lecz same zakresy. Istotność oznacza istotnie różne od zera wartości statystyk  $\Phi$ , co się przekłada na różnicowanie markerów przez przynależność do różnych grup/podgrup. Ze względu na nieznaną postać rozkładu z próby, do testowania istotności zastosowane zostały testy permutacyjne o domyślnej liczbie permutacji = 5000. Ze względu na charakter tej metody, czas obliczeń może być rzędu minut jak również nie podaje się dokładnych wartości p lecz symbole, które odpowiadają poniższym zakresom:

Symbol	Zakres
---	$p \geq 0,05$
*	$0,02 \leq p < 0,05$
**	$0,01 \leq p < 0,02$
***	$0,001 \leq p < 0,01$
****	$p < 0,001$

Odsetek zmienności jest obliczany poprzez przyrównanie średnich kwadratów do ich wartości oczekiwanych i dlatego może przyjmować wartości ujemne. Zakłada się, że z grupą, podgrupą i pojedynczym przypadkiem związane są niezależne od siebie komponenty wariancyjne.

Statystykę  $\Phi$  interpretuje się jako pewnego rodzaju współczynnik korelacji. Zakładając, że w danych jest też zmienna 'podgrupa', dla 'pojedynczego osobnika' jest to wsp. korelacji losowych przypadków w obrębie tej samej podgrupy względem przypadków wziętych z całej populacji. Natomiast dla 'podgrupy' (zagnieżdżonej w 'grupie') jest to względem przypadków wziętych z grupy. Dla 'grupy' jest to wsp. korelacji losowych przypadków w obrębie tej samej grupy względem całej populacji. Jeżeli podział jest tylko na 'grupy', to będzie tylko jedna statystyka  $\Phi$ , zamiast trzech, ale interpretuje się ją wówczas tak samo jak przy dwustopniowym podziale.



**Przykład** Plik *amova.sta* zawiera zerojedynkowe dane 178 markerów genetycznych zebranych od osobników z populacji podzielonej na grupy, składające się z kolei z dalszych podgrup. Po uruchomieniu modułu AMOVA pozostawiamy domyślną liczbę permutacji do testów i przechodzimy do wyboru zmiennych, wskazując zmienne 1 i 2 odpowiednio na listach *Podział nadrzędny* i *Podział podrzędny* i klikamy OK. Po chwili zostanie utworzony następujący arkusz z wynikami.

Dane: AMOVA - wyniki* (6 zmn. * 4 prz.)						
Źródło zmienności	1 % zmienności	2 # stop.swob.	3 Suma kwadratów	4 Średni kwadrat	5 Statystyka $\Phi$	6 Istotność
(Podgrupa)	83,9	18	540,4500	30,0250	0,1609	***
Podgrupa(Grupa)	3,0	8	258,1958	32,2745	0,0349	---
Grupa	13,1	1	97,9970	97,9970	0,1306	**
Ogół		27	896,6429	33,2090		

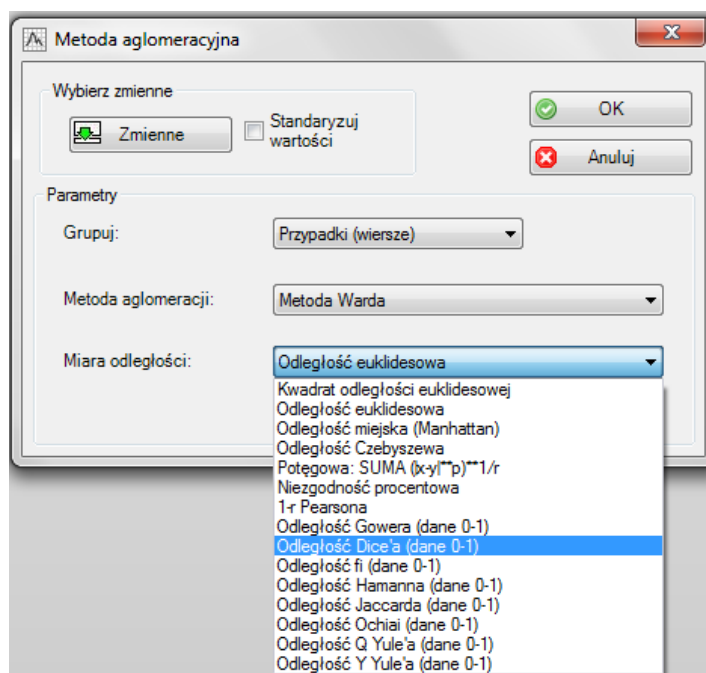
Wyniki te mówią, że osobniki z tej samej podgrupy są istotnie skorelowane ze sobą ( $p < 0,01$ ) w odniesieniu do całej populacji lecz nie ( $p > 0,05$ ) w odniesieniu do ich grupy jak i to, że osobniki w obrębie grupy są istotnie skorelowane ze sobą ( $p < 0,02$ ) w odniesieniu do całej populacji.



## 5. Wielowymiarowe

### 5.1 Analiza aglomeracji

Analiza skupień opiera się na miarach podobieństwa i odległościach między badanymi obiektami. W Zestawie Przyrodnika jej metody udostępniono w module *Analiza aglomeracji*.



Analizę wykonuje się i interpretuje w analogiczny sposób co w podstawowym module *Analiza skupień* pakietu Statistica. Względem niego znajduje się tutaj więcej rodzajów miar odległości. Dodano je z myślą o wykorzystaniu do danych zerojedynkowych z analiz wielocechowych w genetyce. Poniżej zestawiono dodane miary podając wzory odpowiadających im miar zgodności: jeśli miara zgodności wynosi  $r$ , to miara odległości wynosi  $1-r$ . Do określenia ich wartości dla pary obiektów  $X$  i  $Y$  trzeba wyznaczyć liczności  $a, b, c, d$  odpowiednio par  $(1, 1), (1, 0), (0, 1)$  i  $(0, 0)$  w przypisanych im obserwacjach.

Nazwa miary	Wzór
Gowera	$(a + d)/(a + b + c + d)$
Jaccarda	$a/(a + b + c)$
Dice'a	$2a/(2a + b + c)$
Hamanna	$(a + d - b - c)/(a + b + c + d)$
Y Yule'a	$((ad)^{1/2} - (bc)^{1/2})/((ad)^{1/2} + (bc)^{1/2})$
Q Yule'a	$(ad - bc)/(ad + bc)$
Ochiai	$a/((a + b)(a + c))^{1/2}$
Fi ( $\Phi$ )	$(ad - bc)/((a + b)(a + c)(b + d)(c + d))^{1/2}$



## 6. Literatura

1. Ceranka B., Chudzik H. (1977), *Doświadczenia jednopowtórzeniowe z wzorcami*, Siódme Colloquium Metodologiczne z Agro-Biometrii, PAN.
2. Excoffier L., Smouse P.E., Quattro J.M. (1992), *Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data*, Genetics 131.
3. Iwańska M. (2010), *Przydatność różnych miar statystycznych do oceny stopnia szerokiej adaptacji odmian pszenicy ozimej* (praca doktorska), Wydział Rolnictwa i Biologii SGGW w Warszawie.
4. Mądry W., Mańkowski D., Krajewski P., Kaczmarek Z., Studnicki M. (2010), *Metody statystyczne oparte na modelach liniowych w zastosowaniach do doświadczeń rolnictwa, genetyki i hodowli roślin*, Monografie i rozprawy naukowe 34/2010, IHAR.
5. Mańkowski D., Laudański Z., Janaszek M. (2011), *Przydatność wybranych miar podobieństwa dla danych binarnych do analiz wielocechowych w badaniach molekularnych*, Biuletyn IHAR nr 262.