

## PLANOWANIE DOŚWIADCZEŃ I STATYSTYCZNA ANALIZA DANYCH EKSPERYMENTALNYCH NA PRZYKŁADZIE DOŚWIADCZENIA ROLNICZEGO

*Dariusz R. Mańkowski, Instytut Hodowli i Aklimatyzacji Roślin  
– Państwowy Instytut Badawczy w Radzikowie*

Eksperyment naukowy jest obecnie nieodzownym elementem wszelkich badań w naukach empirycznych (takich jak nauki przyrodnicze, medyczne, społeczne itd.). Oparcie pracy naukowej wyłącznie na wywodzie logicznym jest już niewystarczające. Wszelkie teorie i hipotezy muszą być sprawdzone empirycznie. Wymagane jest przeprowadzenie stosownego doświadczenia, którego rezultaty powinny być poparciem stawianych tez. Jednak doświadczenie takie powinno być założone w sposób poprawny, a dane uzyskane w jego wyniku powinny być poddane stosownej analizie statystycznej, której rezultaty pozwolą na wyciągnięcie poprawnych merytorycznie wniosków. I tak z procesu będącego elementem pracy naukowej powstał nowy dział statystyki matematycznej, zwany doświadczalnictwem.

### Doświadczalnictwo

Za prekursora i twórcę współczesnego doświadczalnictwa uważa się Ronalda Aylmera Fishera (1890–1962). Fisher w latach 1919–1933 był pracownikiem Stacji Doświadczalnej w Rothamsted w Wielkiej Brytanii (hrabstwo Hertfordshire). W swojej pracy zajmował się analizą danych pochodzących z doświadczeń prowadzonych w Rothamsted oraz naukowymi podstawami planowania eksperymentu. Za kluczowe dla powstania i rozwoju doświadczalnictwa uważa się jego prace: *Statistical Methods for Research Workers* (1925) oraz *The*

*Design of Experiments* (1935). W roku 1948 w Londynie Fisher powiedział, że *powstanie biometrii w XX wieku, podobnie jak geometrii w III wieku przed Chrystusem, wydaje się podkreślać jeden z wielkich krytycznych okresów w przyszłości myśli ludzkiej*. Pod terminem biometria należy tu rozumieć doświadczalnictwo w naukach przyrodniczych. To właśnie R. A. Fisher zaproponował pierwsze sformalizowane układy doświadczałne oraz analizę wariancji jako metodę statystycznej analizy danych pochodzących z tych doświadczeń. Zarówno zaproponowane przez niego układy doświadczałne, jak i metoda analizy danych są stosowane do dziś, co więcej – znalazły swoje zastosowanie nie tylko w naukach przyrodniczych, lecz również w innych naukach empirycznych.

Wspominając o Fisherze jako twórcy współczesnego doświadczalnictwa, nie można pominąć drugiego badacza, który również przyczynił się do rozwoju tego działu nauki. Mowa tu o Polaku: Jerzym Sławie-Neymanie. W latach 20. i na początku lat 30. XX wieku był on pracownikiem Instytutu Badań Rolniczych w Bydgoszczy, prowadził wykłady na SGGW w Warszawie, był kierownikiem Zakładu Biometrycznego Instytutu im. Nenckiego TNW w Warszawie, prowadził również wykłady na Uniwersytecie Jagiellońskim w Krakowie. Współorganizował Zakład Statystyki Matematycznej przy Wydziale Ogrodniczym SGGW w Warszawie. W roku 1934 objął stanowisko wykładowcy w University College w Londynie, a w roku 1939 został mianowany profesorem nadzwyczajnym Uniwersytetu Londyńskiego. Jerzy Sława-Neyman jest uważany za współtwórcę teorii testowania hipotez statystycznych oraz teorii estymacji przedziałowej.

Praca i osiągnięcia obydwu wspomnianych naukowców przyczyniły się do stworzenia ram doświadczalnictwa rolniczego (zwanego biometrią), a następnie do rozwoju doświadczalnictwa w innych dziedzinach nauk empirycznych.

### ***Randomizacja, replikacja, łączenie w bloki***

Przedmiotem badań przyrodniczych jest obserwacja oraz analiza zjawisk i procesów zachodzących w przyrodzie, a badanymi obiektami są najczęściej organizmy żywe. Fakt takiego doboru obiektów badań już na wstępie stwarza problemy z reguły nie spotykane w naukach

technicznych i ścisłych. Chodzi mianowicie o tzw. zmienność osobniczą. W przyrodzie dwa obiekty pochodzące z jednej populacji nigdy nie będą jednakowe. Nawet jeśli badania będziemy prowadzić na obiektach skrajnie spokrewnionych (np. na klonach), to zawsze będzie występować zmienność między tymi obiektami. Obserwowana zmienność może mieć różne podłoże – może być to zmienność genotypowa (okazuje się, że nawet w przypadku klonów występują różnice w ich genomie, np. tzw. zmienność soma-klonalna), może to też być zmienność spowodowana warunkami środowiska. Planując doświadczenie przyrodnicze, należy mieć wszystkie te źródła zmienności na uwadze.

Sposób na zminimalizowanie problemów wynikających ze zmienności obiektów żywych w doświadczeniach przyrodniczych zaproponował już Fisher. Zalecał on trzy reguły:

1. Randomizacja (*randomization*) – losowy dobór obiektów do badań oraz ich losowe rozmieszczenie w obrębie doświadczenia pozwalają na spełnienie warunku reprezentatywności próby oraz na zminimalizowanie niekontrolowanej zmienności powodowanej przez środowisko.
2. Replikacja (*replication*) – aby wynik był wiarygodny i obiektywnie potwierdzony, każda ocena powinna być wykonana w powtórzeniach. Dobór liczby powtórzeń jest zależny od rodzaju i charakteru badanego zjawiska oraz od zróżnicowania występującego wewnątrz badanej populacji.
3. Łączenie w bloki (*blocking*) – w przypadku występowania w obrębie doświadczenia tzw. zmienności systematycznej (niekontrolowanej zmienności wywoływanej najczęściej przez czynniki środowiskowe), w celu zminimalizowania wpływu tej zmienności na wynik doświadczenia zaleca się łączenie obiektów w bloki, w taki sposób by zmienność systematyczna wewnątrz bloku była minimalna, a zmienność systematyczna pomiędzy blokami była jak największa.

Te trzy zasady stosuje się przy planowaniu doświadczeń do dnia dzisiejszego.

W dalszej części zostanie opisany sposób planowania doświadczenia z zakresu nauk rolniczych oraz analiza statystyczna uzyskanych z tego doświadczenia rezultatów.

## Planowanie doświadczenia

Załóżmy, że chcemy przeprowadzić doświadczenie, w którym zamierzamy w warunkach polowych jednocześnie przebadąć oddziaływanie dwóch czynników: deszczowania i nawożenia azotem na plonowanie pszenicy jarej.

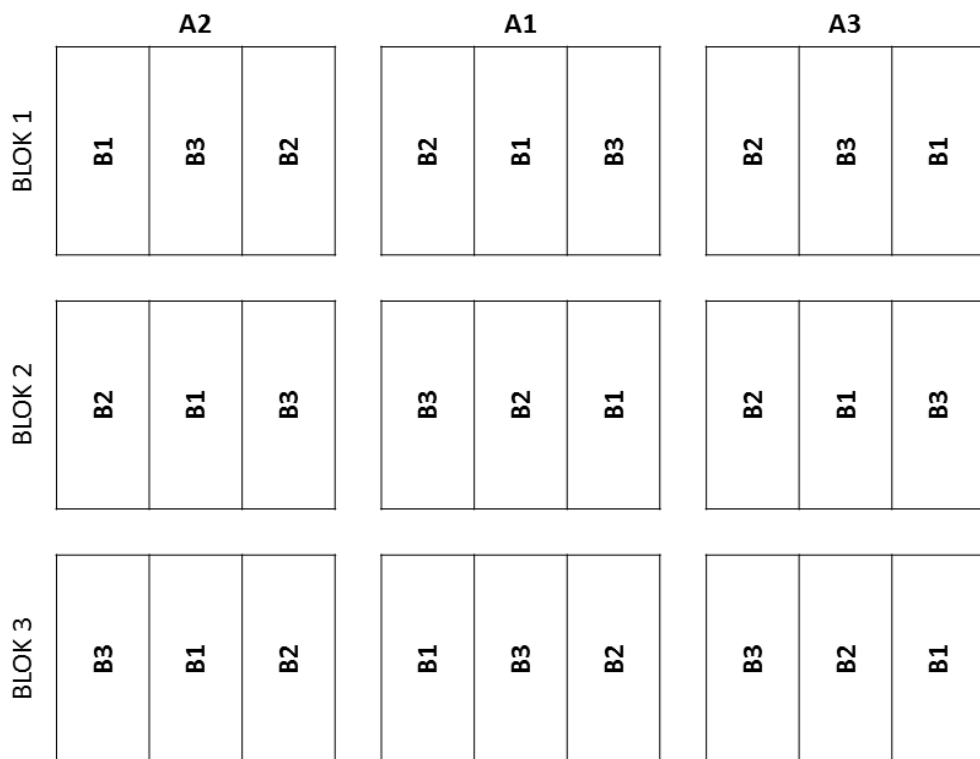
Pszenica jara jest zbożem wymagającym – wymaga dobrej jakości gleby, zasobnej w składniki odżywcze. Jest również bardzo wrażliwa na niedobór wody w okresie wegetacji. W latach suchych może występować potrzeba dodatkowego dostarczania roślinom na polu wody, w postaci np. deszczowania. W ramach planowanego eksperymentu będziemy chcieli zbadać, jaki w danym roku wpływ na osiągnięte plony pszenicy jarej ma fakt zastosowania deszczowania; dodatkowo sprawdzimy, jak na plonowanie wpływają rosnące dawki nawożenia azotowego oraz czy występuje interakcja pomiędzy nawożeniem azotowym a deszczowaniem.

Przystępując do planowania powyższego eksperymentu, zaczynamy od faktu, że ma to być doświadczenie polowe. Oznacza to, że rośliny wysiane na polu będą poddane działaniu niekontrolowanych czynników środowiskowych. W przypadku doświadczeń polowych zawsze możemy założyć, że w obrębie powierzchni doświadczalnej będzie występowała zmienność systematyczna wynikająca z warunków samego pola doświadczalnego. W większości przypadków doświadczenia zakłada się na polach znanych badaczowi. W takim przypadku jesteśmy w stanie wskazać ewentualny kierunek zmienności systematycznej na polu. Wówczas wiemy, że doświadczenie musi być zaplanowane w układzie blokowym (bloki powinny być rozmieszczone prostopadle dłuższym bokiem do kierunku zmienności systematycznej). Wybierzemy więc np. układ bloków losowych lub inny blokowy, a nie układ całkowicie losowy, który ma zastosowanie jedynie w przypadku, gdy mamy

pewność, że w obrębie doświadczenia nie występuje zmienność systematyczna (ma to miejsce np. w doświadczeniach laboratoryjnych lub w pełni kontrolowanych warunkach).

Kolejną sprawą jest kwestia rozmieszczenia badanych czynników i ich specyfiki. O ile czynnik nawożenia azotowego (załóżmy, że będziemy badać 7 poziomów tego czynnika: 0, 30, 60, 90, 120, 150, 180 kg N/ha) nie ma specjalnych wymagań technicznych co do jego zastosowania (najczęściej dawka nawożenia jest na poletkach rozsiewana ręcznie), to czynnik deszczowania (dwa poziomy: bez deszczowania, deszczowanie) jest bardziej problematyczny. Używane do deszczowania maszyny (tzw. deszczownie) mają określoną szerokość, a właściwie średnicę powierzchni roboczej, czyli powierzchni poddanej deszczowaniu. Sprawia to, że zastosowanie tego czynnika na małych poletkach jest technicznie niewykonalne, co więcej bezpośrednie sąsiedztwo poletek deszczowanych i niedeszczowanych może zatrzeć różnice w obserwowanych efektach działania czynnika poprzez podsiąk wody z poletek deszczowanych na niedeszczowane (tzw. efekt sąsiedzki). Najlepiej byłoby więc tak rozmieścić poletka deszczowane, by znajdowały się obok siebie oraz dodatkowo były oddzielone przestrzennie do poletek niedeszczowanych.

Rozwiązaniem jest zastosowanie układu doświadczalnego *split-block* (układ pasów prostopadłych) - rys. 1. W układzie tym bloki ustawia się prostopadle dłuższym bokiem do kierunku zmienności, następnie w poprzek bloków wyznacza się pasy, którym przypisuje się poziomy czynnika A (czynnika wymagającego większych powierzchni do jego użycia) – jest to najczęściej przypisanie losowe, choć zdarzają się przypadki, że poziomy przypisuje się w sposób z góry ustalony; następnie w ramach mniejszych podjednostek (fragmentów bloków przypisanych do poszczególnych poziomów czynnika A) rozlosowuje się wszystkie poziomy czynnika B (mniej wymagającego).



Rys. 1. Schemat układu doświadczalnego *split-block*.

W naszym przypadku, po wydzieleniu na polu bloków (przyjmijmy, że doświadczenie założymy w czterech blokach) prostopadle do kierunku zmienności systematycznej, dzielimy powierzchnię pola na dwa pasy prostopadle do bloków. Następnie przypisujemy powstałym podjednostkom poziomy czynnika „deszczowanie” (po dwie sąsiadujące jednostki na poziom czynnika). Następnie w ramach każdej podjednostki rozlosowujemy wszystkie siedem poziomów czynnika „nawożenie N”. Schemat tak zaplanowanego doświadczenia przedstawiono na rys. 2.

	DESZCZOWANE	NIEDESZCZOWANE
BLOK I	120 0 60 150 90 180 30	120 60 90 30 0 180 150
BLOK II	60 30 180 120 150 90 0	90 120 180 0 150 30 60
BLOK III	0 120 60 0 180 30 90	30 150 120 60 90 0 180
BLOK IV	0 90 150 180 30 120 60	60 0 90 120 180 150 30

Rys. 2. Schemat planowanego doświadczenia.

Po zaplanowaniu doświadczenia i sporządzeniu planów tego eksperymentu można przystąpić do zakładania i przeprowadzenia doświadczenia. Należy przy tym pamiętać, by w trakcie trwania eksperymentu przeprowadzić wszelkie niezbędne pomiary cech z wykorzystaniem precyzyjnej techniki pomiarowej (tak by zminimalizować błąd pomiarowy) oraz by w miarę możliwości uzupełnić doświadczenie wnikliwymi obserwacjami, zarówno samego badanego zjawiska i procesu, jak i otaczającego środowiska, przede wszystkim w celu późniejszego wyjaśnienia uzyskanych wyników.

Gdy mamy już zebrane wyniki pomiarów obserwowanych cech (w postaci zestawu zmiennych), możemy przystąpić do opracowania uzyskanych danych liczbowych w sposób

pozwalający na otrzymanie z nich maksimum obiektywnych informacji, czyli do tak zwanej statystycznej analizy wyników.

## **Analiza wyników doświadczenia i wnioskowanie na podstawie wyników analiz statystycznych**

W przypadku analizy danych pochodzących z doświadczeń czynnikowych (a tak jest w naszym przypadku) najczęściej wykorzystywanym w analizie danych doświadczalnych testem statystycznym jest test F analizy wariancji, zwany popularnie analizą wariancji lub metodą ANOVA.

W klasycznym podejściu powinniśmy taką analizę rozpocząć od sprawdzenia założeń tego testu statystycznego, czyli sprawdzenia, czy zmienna zależna (w naszym przypadku plon) ma rozkład normalny oraz czy wariancje w ramach poszczególnych poziomów czynnika są równe. Jednak w praktyce najczęściej ten krok jest świadomie pomijany. Dzieje się tak dlatego, że zakłada się *a priori*, że założenia są spełnione. Jest w tym doza prawdy. W doświadczalnictwie rolniczym, przy założeniach dotyczących rozkładu cech, bardzo często odwołujemy się do Centralnego Twierdzenia Granicznego, czyli twierdzenia Linderberga-Leviego. Mówi ono, w swej ogólnej postaci, że rozkład sumy wartości zmiennych ilościowych o dowolnym rozkładzie asymptotycznie dąży do rozkładu normalnego (a plon to nic innego jak wyjątkowa postać sumy, czyli średnia – suma zbiorów podzielona przez powierzchnię). A zagłębiając się w interpretację matematyczną, można stwierdzić, że jeżeli na kształtowanie się danej zmiennej ilościowej ma wpływ nieskończona liczba innych losowych zmiennych, to rozkład tej zmiennej jest zgodny z rozkładem normalnym. To podejście rozwiązuje problem testowania zgodności rozkładów – na kształtowanie się plonu roślin rolniczych w warunkach polowych wpływa nieskończona liczba czynników w większości o charakterze losowym. Kwestia równości wariancji jest rozwiązana poprzez dobór obiektów doświadczalnych z jednej określonej populacji. W przypadku naszego eksperymentu, do wysiewu wykorzystano nasiona tej samej odmiany pszenicy jarej, pochodzące z jednej partii nasion uzyskanej z jednej plantacji nasiennej. Wszystkie odmiany



przed zarejestrowaniem w Polsce muszą przejść badania OWT (badania odrębności, wyrównania i trwałości), które potwierdzają między innymi wyrównanie materiału, czyli stabilną i stosunkowo niewielką wariancję cech użytkowych w całej populacji.

Przed przystąpieniem do właściwej analizy danych powinniśmy rozpisać model liniowy dla tej analizy. Dla bardziej złożonych układów doświadczalnych (a takim jest układ split-block) umożliwi to właściwą identyfikację wszystkich źródeł zmienności i wszystkich błędów doświadczalnych. Dla doświadczenia w układzie split-block model liniowy jest postaci:

$$y_{rijk} = \mu + \gamma_r + \alpha_i + \varepsilon_{ri \bullet k}^{(1)} + \beta_j + \varepsilon_{r \bullet jk}^{(2)} + \alpha\beta_{ij} + \varepsilon_{rijk}^{(3)}$$

gdzie:  $y_{rijk}$  – wartość  $k$ -tej zmiennej zależnej dla  $i$ -tego poziomu czynnika A,  $j$ -tego poziomu czynnika B w  $r$ -tym bloku;  $\mu$  – średnia ogólna;  $\gamma_r$  – efekt  $r$ -tego bloku;  $\alpha_i$  – efekt  $i$ -tego poziomu czynnika A;  $\varepsilon_{ri \bullet k}^{(1)}$  – błąd pierwszy,  $\varepsilon_{ri \bullet k}^{(1)} = \gamma\alpha_{ri}$ ;  $\beta_j$  – efekt  $j$ -tego poziomu czynnika B;  $\varepsilon_{r \bullet jk}^{(2)}$  – błąd drugi,  $\varepsilon_{r \bullet jk}^{(2)} = \gamma\beta_{rj}$ ;  $\alpha\beta_{ij}$  – efekt interakcji pomiędzy  $i$ -tym poziomem czynnika A oraz  $j$ -tym poziomem czynnika B;  $\varepsilon_{rijk}^{(3)}$  – błąd trzeci – błąd losowy.

W powyższym modelu liniowym widzimy wszystkie źródła zmienności, jakie musi zawierać nasza analiza. Dodatkowo widzimy, że w analizie występują trzy błędy doświadczalne: pierwszy będący efektem interakcyjnym bloków i czynnika A (dla testowania bloków i czynnika A), drugi będący efektem interakcyjnym bloków i czynnika B (dla testowania czynnika B) oraz trzeci – losowy (efektu interakcyjnego A×B).

Mając już określoną postać analizy, możemy przystąpić do jej wykonania. Wyniki analizy wariancji przedstawiono na rys. 3. W wyniku przeprowadzonej analizy możemy stwierdzić, że: występowały istotne różnice pomiędzy średnimi plonami na poletkach z deszczowaniem i bez deszczowania ( $F=732,7106$ ;  $p=0,000111$ ); występowały istotne różnice w obserwowanych wartościach średnich dla poszczególnych dawek nawożenia azotowego

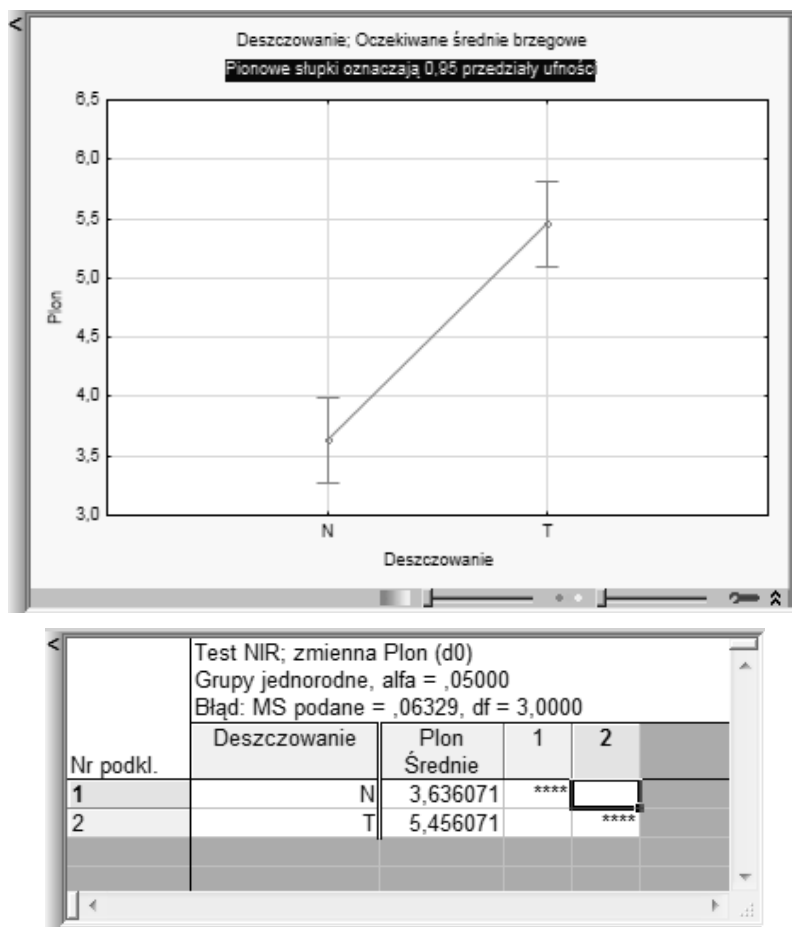
( $F=12,83249$ ;  $p=0,000012$ ) oraz występowała istotna interakcja pomiędzy nawożeniem azotowym a deszczowaniem ( $F=4,162$ ;  $p=0,008504$ ).

Jednowymiarowe testy istotności dla Plon (d0) Parametryzacja z sigma-ograniczeniami Dekompozycja efektywnych hipotez					
Efekt	SS	Stopnie swobody	MS	F	p
<b>Wyraz wolny</b>	1157,339	1	1157,339	1411,367	0,000000
Blok	0,08981	3	0,02994	0,4730	0,722815
Deszczowanie	46,37360	1	46,37360	732,7106	0,000111
Błąd I	0,18987	3	0,06329		
Nawożenie N	31,39189	6	5,231981	12,83249	0,000012
Błąd II	7,33884	18	0,407713		
Deszczowanie*Nawożenie N	20,479	6	3,413	4,162	0,008504
Błąd III	14,760	18	0,820		

Rys. 3. Wyniki analizy wariancji.

Występowanie istotnej statystyki testowej dla efektów głównych pozwala nam wnioskować jedynie o występowaniu różnic w obrębie średnich wartości zmiennej zależnej dla poziomów danego czynnika. Wiemy więc tylko, że przynajmniej jedna wartość średnie różni się istotnie od pozostałych. Nie wiemy jednak, które średnie się od siebie różnią (które poziomy czynnika powodują istotne zróżnicowanie średnich). Do identyfikacji tych różnic możemy wykorzystać testy porównań wielokrotnych (tzw. testy post-hoc). W naukach rolniczych najczęściej wykorzystywana jest w tym celu procedura porównań wielokrotnych Tukeya (Tukey HSD).

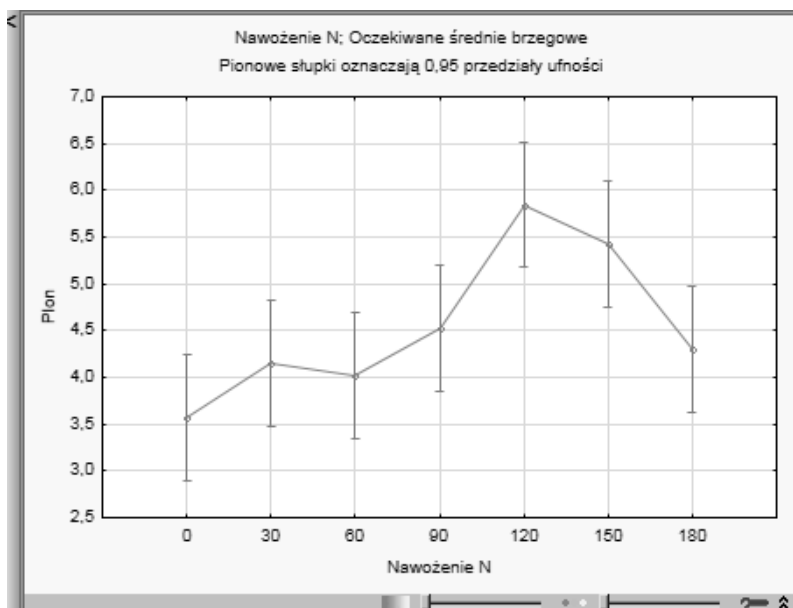
Dla czynnika „deszczowanie” stwierdzenie występowania istotnych różnic prowadzi do prostego wniosku, że poletka deszczowane różniły się istotnie średnim plonem od poletek niedeszczowanych. Wyniki procedury Tukeya dla tego czynnika oraz wykres średnich przedstawiono na rys. 4.



Rys. 4. Wykres wartości średnich i wyniki grupowania procedurą Tukeya dla czynnika „deszczowanie”.

Uzyskane wyniki pozwoliły na stwierdzenie, że deszczowanie pozwoliło na uzyskanie istotnie wyższych plonów pszenicy jarej, średnio o około 1,8 t/ha, niż dla poletek niedeszczowanych.

Dla czynnika „nawożenie N” wstępne wnioskowanie nie jest już takie oczywiste. Wynika to z większej liczby badanych poziomów tego czynnika. Na rys. 5 przedstawiono wyniki grupowania procedurą Tukeya oraz wykres średnich.



Test NIR; zmienna Plon (d0)  
Grupy jednorodne, alfa = ,05000  
Błąd: MS podane = ,40771, df = 18,000

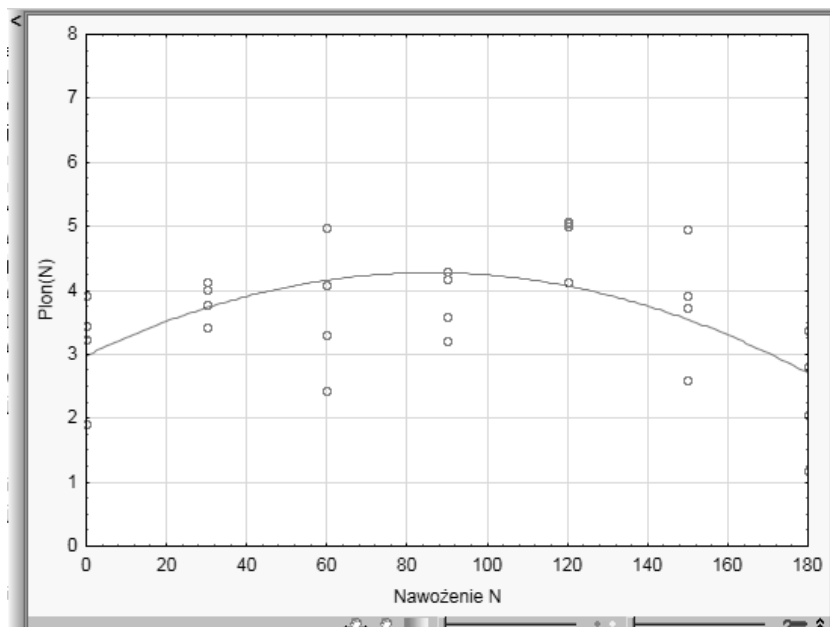
Nr podkl.	Nawożenie N	Plon Średnie	1	2	3
1	0	3,563750		****	
3	60	4,012500	****	****	
2	30	4,151250	****	****	
7	180	4,297500	****		
4	90	4,530000	****		
6	150	5,421250			****
5	120	5,846250			****

Rys. 5. Wykres wartości średnich i wyniki grupowania procedurą Tukeya dla czynnika „nawożenie N”.

Przeprowadzona analiza pozwoliła na wyodrębnienie 3 grup jednorodnych średnich plonów pszenicy jarej w ramach badanych poziomów nawożenia azotem. Jednak po uporządkowaniu poziomów nawożenia względem średnich plonów okazało się, że uzyskana kolejność z pozoru wydaje się nielogiczna. Jednak po wykonaniu wykresu średnich widać charakter zmienności średnich plonów pod względem wzrastających dawek nawożenia

azotem. Gdy w analizie wariancji badamy czynniki, których poziomy mają charakter ilościowy (a nie jakościowy, jak to było w przypadku „deszczowania”), taka sytuacja jest bardzo częsta. Niektórzy doświadczalnicy i statystycy zalecają nawet, by tego typu czynników nie poddawać analizie za pomocą procedur porównań wielokrotnych, a zamiast tego stosować analizę funkcji regresji dla wartości zmiennej zależnej względem zmieniających się poziomów czynnika. Jeśli spojrzymy na wykres średnich (rys. 5), to widzimy, że zmiany średnich plonów względem wzrastających dawek nawożenia azotem nie mają charakteru liniowego.

Podsumowanie regresji zmiennej zależnej: Plon (d0)						
R= ,39746319 R <sup>2</sup> = ,15797699 Popraw. R <sup>2</sup> = ,12620254						
F(2,53)=4,9718 p<,01050 Błąd std. estymacji: 1,3843						
N=56	b*	Bł. std. z b*	b	Bł. std. z b	t(53)	p
W. wolny			3,342202	0,427213	7,82328	0,000000
Nawożenie N	1,136662	0,454460	0,027804	0,011116	2,50113	0,015505
V3**2	-0,850011	0,454460	-0,000111	0,000059	-1,87037	0,066958

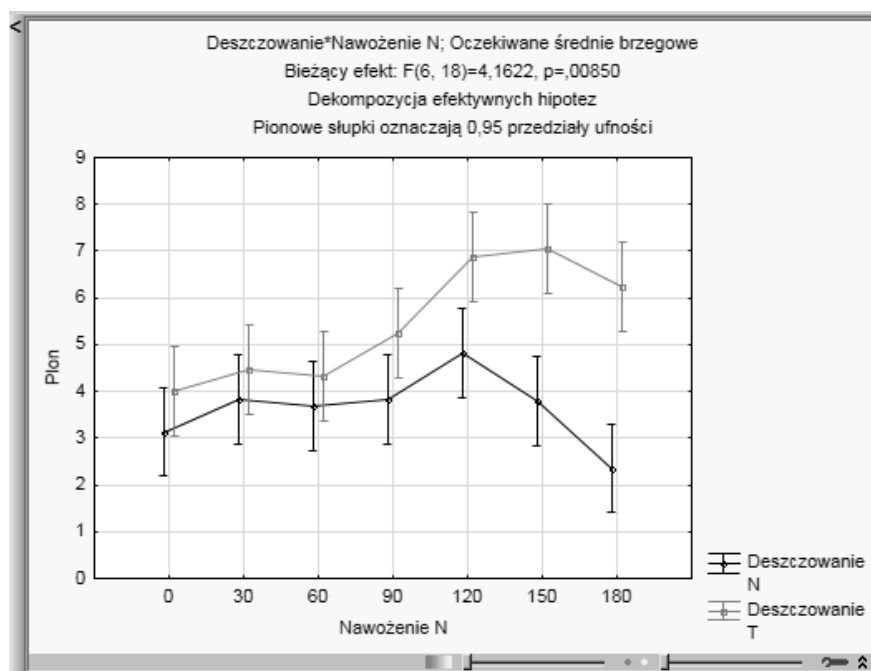


Rys. 6. Wyniki analizy regresji wielomianowej oraz wykres uzyskanej funkcji regresji.

Do analizy przebiegu tych zmian wykorzystamy więc model wielomianowy (tu wielomian kwadratowy) i analizę regresji nieliniowej z wykorzystaniem modeli linearyzowanych. Wyniki takiej analizy i wykres uzyskanej funkcji przedstawiono na rys. 6 (powyżej).

Uzyskane wyniki pozwalają na lepszy opis wpływu nawożenia azotowego na plonowanie pszenicy jarej.

Przeprowadzona analiza (rys. 3) wykazała występowanie istotnej interakcji pomiędzy deszczowaniem i nawożeniem azotowym. Przypomnijmy, że **interakcją** lub **współdziałaniem** dwóch czynników określa się zjawisko polegające na tym, że reakcja badanej cechy na poziomy jednego czynnika nie jest jednakowa dla wszystkich poziomów drugiego czynnika. Najszybszym narzędziem do analizy interakcji jest tak zwany wykres interakcji (rys. 7).



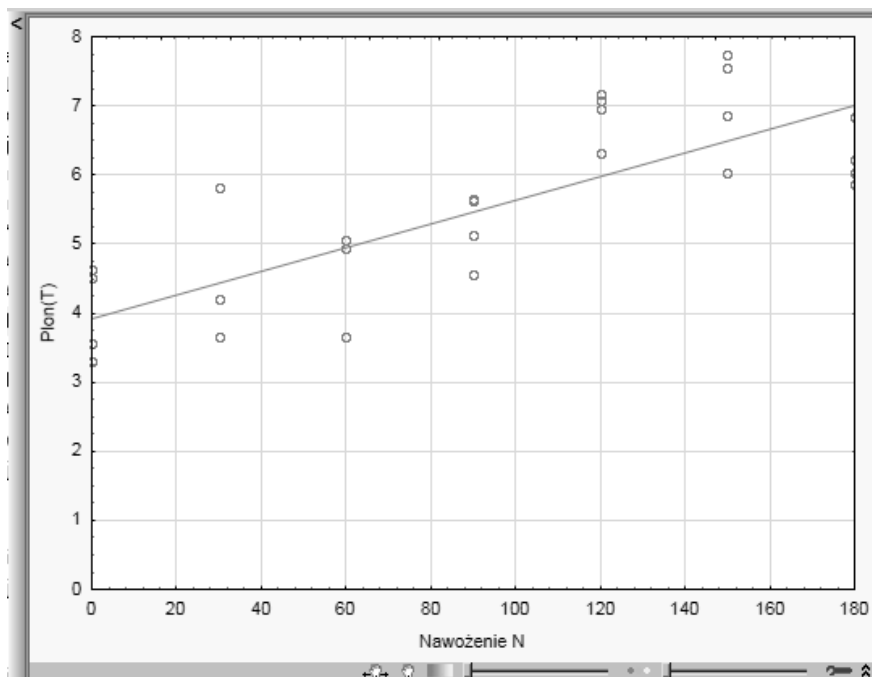
Rys. 7. Wykres interakcji.

Na sporządzonym wykresie interakcji widać, że zmiany w reakcji pszenicy ozimej na wzrastające dawki nawożenia pomiędzy poletkami deszczowanymi i niedeszczowanymi

widać wyraźnie przy dużych dawkach nawożenia azotowego (powyżej 120 kg N/ha). W przypadku analizy interakcji również możemy przeprowadzić analizę regresji względem czynnika ilościowego, ale musimy przeprowadzić oddzielną analizę dla każdego poziomu drugiego czynnika (jakościowego). W naszym przypadku powinniśmy więc wykonać dwie analizy regresji. Wyniki analiz przedstawiono na rys. 8.

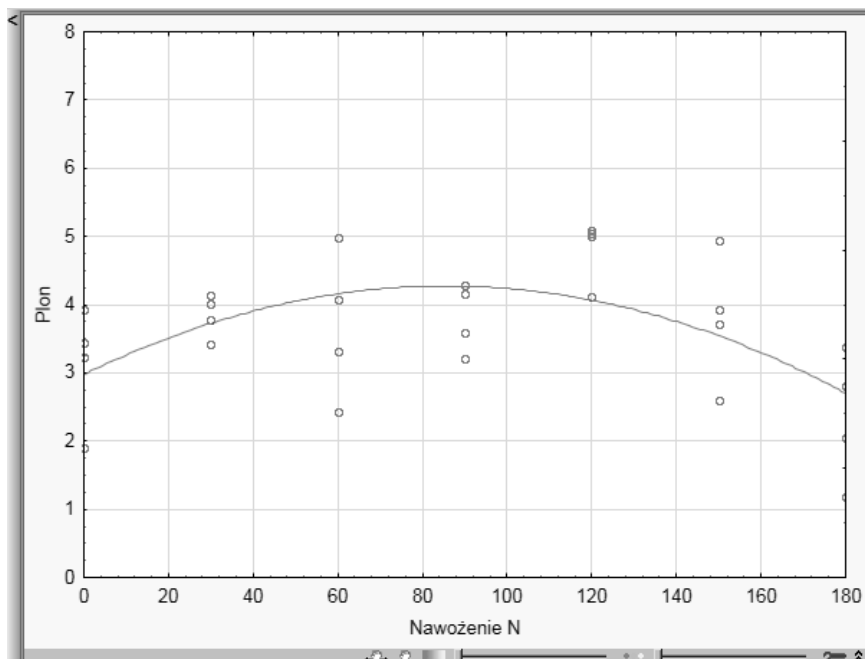
*deszczowane*

Podsumowanie regresji zmiennej zależnej: Plon (d0)						
R= ,78975776 R <sup>2</sup> = ,62371732 Popraw. R2= ,60924491						
F(1,26)=43,097 p<,00000 Błąd std. estymacji: ,83094						
Warunek uwzględniania: v2="T"						
N=28	b*	Bł. std. z b*	b	Bł. std. z b	t(26)	p
W. wolny			3,909732	0,283095	13,81069	0,000000
Nawożenie N	0,789758	0,120301	0,017182	0,002617	6,56483	0,000001



*niedeszczowane*

Podsumowanie regresji zmiennej zależnej: Plon(N) (Arkusz14)						
R= ,57356822 R^2= ,32898050 Popraw. R2= ,27529894						
F(2,25)=6,1284 p<,00683 Błąd std. estymacji: ,84402						
N=28	b*	Bł. std. z b*	b	Bł. std. z b	t(25)	p
W. wolny			2,978393	0,368360	8,08555	0,000000
Nawożenie N	1,86589	0,590703	0,030277	0,009585	3,15876	0,004110
V3**2	-2,04001	0,590703	-0,000177	0,000051	-3,45353	0,001983

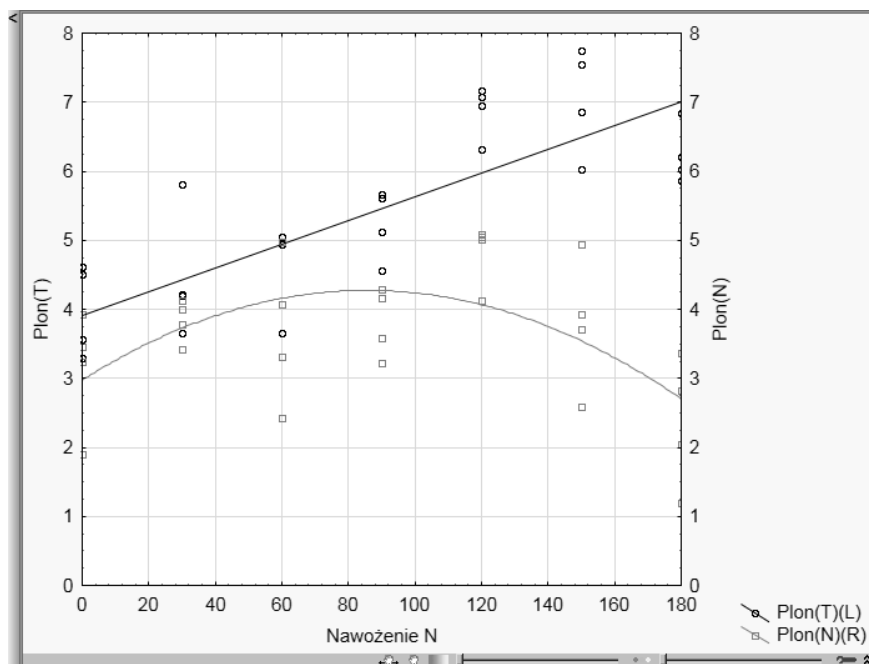


Rys. 8. Wyniki analiz regresji oraz wykres uzyskanych funkcji regresji.

W wyniku przeprowadzonych analiz uzyskano wielomianową funkcję regresji opisującą zmiany przeciętnego plonowania pszenicy jarej przy wzrastających dawkach nawożenia azotowego na polach bez deszczowania oraz liniową funkcję regresji opisującą zmiany przeciętnego plonowania pszenicy jarej przy wzrastających dawkach nawożenia azotowego na polach z deszczowaniem. Na podstawie tak różnych przebiegów reakcji można wywnioskować, że deszczowanie, z racji większego wymywania substancji odżywczej



zmniejszyło toksyczne działanie wysokich dawek nawożenia azotowego. Dodatkowo potwierdza to wykres, na którym przedstawiono wykresy obydwu funkcji (rys. 9).



Rys. 9. Wyniki analiz regresji oraz wykresy uzyskanych funkcji regresji.  
*Plon(T)* – plon na poletkach z deszczowaniem; *Plon(N)* – plon na poletkach bez deszczowania

## A co by było, gdyby ...

Przedstawiony przykład ma jedynie przedstawić, jak złożonym procesem jest planowanie eksperymentu, a później analiza danych pochodzących z tego eksperymentu. Omówiony układ doświadczalny jest układem klasycznym, dobrze opisanym w literaturze i podręcznikach do doświadczalnictwa. W praktyce często się jednak zdarza, że nie ma możliwości zastosowania układu klasycznego lub zastosowany układ wymaga istotnych modyfikacji, co z kolei wpływa na zmiany również po stronie statystycznej analizy wyników.

Wystarczy na przykład, że opisanie wcześniej doświadczenie przeprowadzimy nie w jednej, ale w trzech lokalizacjach. Wówczas prowadząc łączną analizę wyników ze wszystkich

lokalizacji, nie wystarczy „dodać” czynnik lokalizacji. Problemem stają się bloki, gdyż na przykład pierwszy blok w pierwszej lokalizacji nie odpowiada pierwszemu blokowi w pozostałych lokalizacjach. Problem ten można rozwiązać, uwzględniając efekty zagnieżdżenia bloków w lokalizacjach. Zmieni się więc model liniowy analizy, a co za tym idzie cały jej późniejszy przebieg. W omawianym przykładzie taki model miałby postać:

$$y_{ijkl} = \mu + \zeta_l + \gamma_r(\zeta_l) + \varepsilon_{r\bullet\bullet kl}^{(1)} + \alpha_i + \varepsilon_{ri\bullet k\bullet}^{(2)} + \beta_j + \varepsilon_{r\bullet jk\bullet}^{(3)} + \alpha\beta_{ij} + \varepsilon_{ijkl}^{(4)}$$

gdzie:  $y_{ijkl}$  – wartość  $k$ -tej zmiennej zależnej dla  $i$ -tego poziomu czynnika A,  $j$ -tego poziomu czynnika B w  $r$ -tym bloku w  $l$ -tej lokalizacji;  $\mu$  – średnia ogólna;  $\zeta_l$  – efekt  $l$ -tej lokalizacji;  $\gamma_r(\zeta_l)$  – bloki zagnieżdżone w lokalizacjach;  $\varepsilon_{r\bullet\bullet kl}^{(1)}$  – błąd pierwszy,  $\varepsilon_{r\bullet\bullet kl}^{(1)} = \gamma_r(\zeta_l)$ ;  $\alpha_i$  – efekt  $i$ -tego poziomu czynnika A;  $\varepsilon_{ri\bullet k\bullet}^{(2)}$  – błąd pierwszy,  $\varepsilon_{ri\bullet k\bullet}^{(2)} = \alpha\gamma(\zeta)_{ril}$ ;  $\beta_j$  – efekt  $j$ -tego poziomu czynnika B;  $\varepsilon_{r\bullet jk\bullet}^{(3)}$  – błąd drugi,  $\varepsilon_{ri\bullet k\bullet}^{(3)} = \beta\gamma(\zeta)_{rjl}$ ;  $\alpha\beta_{ij}$  – efekt interakcji pomiędzy  $i$ -tym poziomem czynnika A oraz  $j$ -tym poziomem czynnika B;  $\varepsilon_{ijkl}^{(4)}$  – błąd trzeci – błąd losowy.

Jak widać, cała sytuacja się mocno komplikuje. Doświadczalnicy, pracujący na co dzień z rzeczywistymi eksperymentami, spotykają się z takimi sytuacjami dość często i wiedzą, jak sobie radzić z takimi problemami. Niestety często obserwuje się próby publikacji lub upowszechniania wyników, których analiza statystyczna jest przeprowadzona w sposób niewłaściwy lub samo doświadczenie zostało zaplanowane i założone w sposób uniemożliwiający poprawną analizę i wnioskowanie. Takie wyniki są dla nauki i rozwoju wiedzy bezwartościowe.

## Podsumowanie

Umiejętne zaplanowanie i przeprowadzenie eksperymentu wymaga uwzględnienia wielu zagadnień z pogranicza metodologii nauk empirycznych, biometrii, logiki i podstaw badawczego zjawiska lub procesu.

Najważniejszymi elementami, które powinny być brane pod uwagę przy planowaniu doświadczenia, a w szczególności doświadczenia rolniczego, są warunki, w których doświadczenie będzie przeprowadzone, staranne dobranie badanych czynników i ich poziomów, uwzględnienie jak największej liczby czynników towarzyszących danemu zjawisku bądź procesowi oraz dobór układu doświadczalnego, który wszystko to uwzględnia.

Zaplanowanie doświadczenia wiąże się bardzo ściśle z późniejszą analizą danych uzyskanych z tego eksperymentu. Błędy na etapie planowania mogą skutkować pogorszeniem jakości, bądź wręcz uniemożliwieniem przeprowadzenia właściwej analizy wyników. Z drugiej strony niewłaściwe podejście analityczne może doprowadzić do wyciągnięcia niewłaściwych lub wręcz błędnych wniosków.

Dlatego dopiero właściwe zaplanowanie eksperymentu i dobrze dobrana poprawna metoda analizy danych doświadczalnych mogą prowadzić do prawdziwych i wiarygodnych wniosków.

## Literatura

1. Box, G.E.P., Hunter, J.S., Hunter, W.G., 2005. Statistics for Experimenters — Design, Innovation, and Discovery. Second Edition. Wiley and Sons Inc., New Jersey, USA.
2. Cochran, W.G., Cox, G.M., 1992. Experimental design. John Wiley & Sons Inc., Hoboken, USA.
3. Elandt, R., 1964. Statystyka matematyczna w zastosowaniu do doświadczalnictwa rolniczego. PWN, Warszawa.
4. Fisher, R.A., 1925. Statistical Methods for Research Workers. Oliver & Boyd, Edinburgh and London, UK.
5. Fisher, R.A., 1935. The Design of Experiments. Oliver & Boyd, Edinburgh and London, UK.

6. Kala, R., 1996. Elementy wnioskowania parametrycznego dla przyrodników. Akademia Rolnicza w Poznaniu, Poznań.
7. Mądry, W., Mańkowski, D.R., Kaczmarek, Z., Krajewski, P., Studnicki, M., 2010. Metody statystyczne oparte na modelach liniowych w zastosowaniach do doświadczalnictwa, genetyki i hodowli roślin, Monografie i Rozprawy Naukowe IHAR-PIB. IHAR-PIB, Radzików.
8. Montgomery, D.C., 2005. Design and analysis of experiments. 6th edition. John Wiley & Sons Inc., Hoboken, NJ, USA.
9. Nawrocki, Z., 1967. Teoria i praktyka doświadczalnictwa rolniczego. PWRiL, Warszawa.
10. Oktała, W., 1966. Elementy statystyki matematycznej i metodyka doświadczalnictwa. PWN, Warszawa.
11. Oktała, W., 1971. Metody statystyki matematycznej w doświadczalnictwie. PWN, Warszawa.
12. Oktała, W., 2002. Historia teorii eksperymentu. Lubelskie Towarzystwo Naukowe, Lublin.
13. Trętowski, J., Wójcik, A.R., 1988. Metodyka doświadczeń rolniczych. WSRP, Siedlce.
14. Wójcik, A.R., Laudański, Z., 1989. Planowanie i wnioskowanie statystyczne w doświadczalnictwie. PWN, Warszawa.