

Historia rozwoju statystycznych metod planowania i analizy doświadczeń rolniczych na świecie oraz w Polsce

A history of the development of statistical methods for designing and analyzing agricultural experiments in the world and in Poland

Wiesław Mądry✉, Dariusz Gozdowski

Katedra Biometrii, Wydział Rolnictwa i Biologii, Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

✉ e-mail: wieslaw_madry@sggw.pl

W pracy przedstawiono główne kierunki badań, ich chronologię oraz osiągnięcia w zakresie metod statystyki matematycznej w zastosowaniu do biometrii i doświadczalnictwa rolniczego, dokonywane od XVII wieku do czasów współczesnych. Uwzględniono dorobek uczonych na świecie oraz w Polsce. Podkreślono historyczne i współczesne znaczenie tych osiągnięć matematycznych i metodycznych dla rozwoju i postępu nauk empirycznych w ogóle, a zwłaszcza nauk rolniczych i biologicznych. Przedstawiono znaczenie zastosowania metod statystycznych w uznaniu empirycznych badań rolniczych, jako nauki rolnicze. Świadczenia i rozważania oraz autorskie osądy odkryć i wynalazków statystycznych na przestrzeni wieków i lat są udokumentowane i zilustrowane oryginalnym i publikacjami, a także realnymi dokonaniem pionierów statystyki, biometrii i doświadczalnictwa rolniczego oraz współczesnych uczonych w tych dziedzinach i specjalnościach.

Słowa kluczowe: metody statystyczne, metody wielowymiarowe, modele statystyczne, estymacja parametrów statystycznych, hipotezy, testowanie hipotez, układy doświadczalne, wnioskowanie statystyczne, statystycy, biometrycy

The paper presents the main directions, chronology and achievements in the field of mathematical statistical methods applied to biometrics and agricultural experimentation, conducted from the 17th century to modern times. The achievements of scientists globally and in Poland are taken into account. The historical and present importance of these mathematical and methodological findings for the development and progress of empirical sciences in general, and especially agricultural and biological sciences, is emphasized. The importance of using statistical methods in the recognition of empirical agricultural research as agricultural science is presented. Testimonies and considerations on statistical discoveries and inventions over the centuries are documented and illustrated by the original publications and real activities of the pioneers of statistics, biometrics and agricultural experimentation, as well as contemporary scientists in these fields.

Key words: statistical methods, multivariate methods, statistical models, estimation, hypothesis testing, experimental designs, statistical inference, statisticians, biometricians

Motto

Nauka jest gromadzeniem udokumentowanych faktów, tak jak dom jest tworzeniem konstrukcji z kamieni; ale, podobnie, jak stos kamieni nie jest domem, tak sam zbiór faktów bez ich filozoficznego widzenia nie musi być nauką [*Henri Poincaré (1854–1912)*].

Jesteśmy zobowiązani do badania zjawisk tym dokładniej, im trudniej przychodzi nam uznać je za istniejące (*Pierre Simon de Laplace (1749–1827)*, *Essai philosophique sur les probabilités, Paris 1814*).

1. Wprowadzenie

Źródłami wiedzy o zjawiskach rolniczych od zarania dziejów były (1) obserwacje zjawisk rzeczywistych, intuicyjne kojarzenie relacji i współistnienia zdarzeń oraz wnioskowanie o współzależnościach, przyczynowości i prognozowaniu (np. obserwacje zjawisk pogodowych i ich wpływu na rośliny) oraz (2) praktyka rolnicza, w której od tysiącleci rolnicy eksperymentowali z czynnikami przyrodniczymi i agrotechnicznymi. Rolnik, żyjący w bliskim kontakcie z przyrodą i produkcją rolniczą,

doskonalił swoje zdolności obserwacyjne i poznawcze. Jako odwieczny eksperymentator i kustosz ziemi użytkowanej rolniczo, rolnik dochodził do całkiem nowych i użytecznych odkryć. Do dzisiaj są one aktualne i fascynują swoją głębią i realizmem oraz pięknem ludowego sztuki. W Europie postęp dokonywany tą drogą zdobywania wiedzy rolniczej doprowadził do zwielokrotnienia produkcji roślinnej i zwierzęcej pomiędzy XVII, a XIX wiekiem. Jednakże, dopiero po udowodnieniu teorii o mineralnym odżywianiu się roślin przez chemika Justusa von Liebiga (1803–1873), opublikowanej w 1840 roku w dziele *Chemia organiczna i jej związek z rolnictwem i fizjologią* oraz na skutek rozwoju nauki i przemysłu w XIX wieku, zaczęły rozwijać się nauki rolnicze w Europie i Polsce. Głównym źródłem wiedzy rolniczej stawały się empiryczne badania naukowe, oparte na metodzie indukcyjno-dedukcyjnej. Empiryczne metody badań naukowych są zgodne ze współczesną (kształtującą się od XVII wieku), nowożytną filozofią nauki. Obejmują one sposoby intersubiektywnego

poznawania i tworzenia wiedzy, oparte na prawach logiki i rachunku prawdopodobieństwa. Intersubiektywność metody naukowej oznacza powtarzanie badań w takich samych lub podobnych warunkach, poprzez nowe obserwacje lub doświadczenia w celu weryfikowania wniosków, wyprowadzonych metodą indukcyjną „od szczegółu do ogółu” na podstawie badań w pierwszym etapie. Powtarzanie badań w celu naukowego uwiarygodnienia wniosków jest kluczowym wymogiem ustawicznego procesu autokorekcyjnego w metodzie naukowej. Słowami filozofa nauki Karla Poppera (1902–1994) „pojedyncze, nie powtórzone, badania empiryczne nie mają dla nauki wystarczającego znaczenia”. Luminarz statystyki matematycznej oraz biometrii i metod statystycznych w doświadczalnictwie, Ronald Aylmer Fisher (1890–1962) dodaje: „możemy uznać, że zjawisko jest udowodnione eksperymentalnie wówczas, gdy wiemy, jak przeprowadzić eksperyment, który rzadko zawiedzie w wykazaniu istotnych statystycznie faktów” (Fisher, 1935). R. Fisher ma tu na uwadze zarówno poprawny statystycznie i powtarzalny realnie eksperyment, jak i test statystyczny o dużej mocy (test zapewniający małe prawdopodobieństwo błędu II rodzaju) przy niskim poziomie istotności (test zapewniający małe prawdopodobieństwo błędu I rodzaju).

Od prawie tysiąca lat badania naukowe w fizyce, astronomii i chemii były wykonywane coraz częściej metodami empirycznymi. Jedno z pierwszych metodycznych podejść do empirii w nowoczesnym sensie jest widoczne w badaniach optycznych arabskiego matematyka, astronoma i fizyka Ibn al-Haythama (965–1040), wybitnego filozofa *złotej ery islamu*. Podstawą intelektualną naukowych metod empirycznych jest doktryna filozoficzna, nazywana empiryzmem. Głosi ona, że źródłem ludzkiego poznania są wyłącznie lub przede wszystkim bodźce zmysłowe docierające do naszego umysłu ze świata zewnętrznego, zaś wszelkie idee, teorie itp. są w stosunku do nich wtórne. Za twórcę nowoczesnego empiryzmu uznaje się brytyjskiego filozofa oraz poetę renesansu i baroku: Francisca Bacona (1561–1626). Zwolennikami empiryzmu byli genialni, uznawani za największych w dziejach świata, uczeni: Leonardo da Vinci (1452–1519) i Galileo Galilei — Galileusz (1564–1642). Galileusz (rówieśnik F. Bacona), będąc matematykiem oraz badaczem zjawisk fizycznych i pokrewnych, uważał, że sam eksperyment nie wystarcza w poznaniu naukowym. Krytykował pierwotne traktowanie empiryzmu. Sądził on, że samo nagromadzenie obserwacji o zdarzeniach nie stanowi nauki. Twierdził, że nauka ma za zadanie odkrywać powtarzalne współistnienia zdarzeń, czyli prawa przyrody, za pomocą rozumowania indukcyjnego, opartego na empirii. Współcześni badacze, także

w biologii i rolnictwie, w ten właśnie sposób postępują, traktując badania empiryczne i teoretyczne jako wzajemnie uzupełniające się w procesie wnioskowania o mechanizmach funkcjonowania świata.

Empiryczne metody badań rolniczych polegają przeważnie na przeprowadzaniu doświadczeń porównawczych, zwanych doświadczeniami czynnikowymi. Doświadczenie czynnikowe jest zbiorem działań wzbudzających określone reakcje zjawisk w pewnym materiale, pozostającym w tych samych warunkach, ale poddanym różnemu traktowaniu w postaci zmiennej przyczyny (czynnika). Doświadczenie jest naukowym sprawdzeniem tego, co się stanie z badanymi obiektami w konkretnych warunkach pod wpływem różnego traktowania. Podstawą filozoficzną doświadczeń czynnikowych jest tzw. kanon jedynej różnicy, sformułowany w 1843 roku przez wybitnego filozofa brytyjskiego, zwolennika empiryzmu, Johna Stuarta Milla (1806–1873). Kanon ten został opublikowany w jego pracy *System logiki dedukcyjnej i indukcyjnej*. Głosi on, że jeśli różne traktowanie materiału badawczego w tych samych warunkach powoduje różne reakcje obserwowanego zjawiska, to można wyprowadzić wniosek, że zastosowany czynnik ma wpływ na to zjawisko. W toku rozwoju badań doświadczalnych w rolnictwie i biologii stosowanej oraz towarzyszącej jemu ewolucji metodologicznej, około 150 lat temu wykształciło się pojęcie doświadczalnictwo (ang. experimentation). Ogólnie, doświadczalnictwo jest procesem naukowego odkrywania natury zjawisk, testowania hipotez lub demonstracji wybranych zjawisk, przeprowadzanym za pomocą metod doświadczalnych.

Doświadczalnictwo rolnicze (ang. *agricultural experimentation*) wyodrębniło się i ukształtowało jako pierwszy historycznie wzorzec doświadczalnictwa w ogóle. Jest ono obszerną, często interdyscyplinarną w dziedzinie nauk rolniczych, działalnością naukowo-badawczą i badawczo-wdrożeniową. Doświadczalnictwo rolnicze obejmuje: (a) wykonywanie optymalnie zaplanowanych czynnikowych ścisłych (kontrolowanych) doświadczeń laboratoryjnych, wazonowych i polowych (ang. *controlled factor experiments*), zakładanie doświadczeń w gospodarstwach rolnych (ang. *on-farm experiments*) lub prowadzenie badań obserwacyjno-pomiarowych (ang. *observational and measurement studies*), takich, jak ankiety i statystyka gospodarcza (ang. *surveys, censuses*), (b) wykorzystanie wyników z tych badań do wnioskowania o zjawiskach w celach naukowych lub wdrożeniowych. Każda dyscyplina nauk rolniczych, a nawet ich specjalności wypracowały swoje specyficzne doświadczalnictwo, z osobliwymi podstawami i metodami.

Termin doświadczalnictwo rolnicze ma też inne znaczenie. Określa on nazwę interdyscyplinarnej, metodyczno-statystycznej specjalności w dwóch dziedzinach naukowych, tj. nauk rolniczych i matematycznych, która ściśle towarzyszy eksperymentalnej działalności naukowo-badawczej (ang. *design of experiments, experimental design, design and analysis of experiments*). W tym znaczeniu doświadczalnictwo rolnicze zajmuje się (a) statystyczną metodyką planowania rolniczych ścisłych doświadczeń czynnikowych, doświadczeń w gospodarstwach rolnych i badań obserwacyjno-pomiarowych oraz (b) statystyczną analizą danych, interpretacją wyników i wnioskowaniem. Ta specjalność naukowa jest często nazywana statystycznymi metodami planowania i analizy doświadczeń rolniczych, statystyczną metodyką doświadczalnictwa rolniczego lub teorią doświadczenia rolniczego (Caliński, 2012). Początek ery tworzenia statystycznych metod planowania i analizy doświadczeń rolniczych przypada na koniec XIX wieku. Zbiega się on z początkiem przełomowego kształtowania nowego działu statystyki, którym jest statystyka matematyczna.

Statystyczna metodyka doświadczalnictwa rolniczego stała się wyodrębnionym i wyspecjalizowanym działem biometrii, która obejmuje szeroko pojęte zastosowania statystyki w naukach biologicznych i biologii stosowanej (głównie w naukach rolniczych i medycznych). Ta gałąź biometrii stanowi integralną składową całość metodologii rolniczych badań doświadczalnych (Oktaba, 2002; Caliński, 2012). Należy podkreślić, że dopiero zastosowanie statystyki matematycznej w doświadczalnictwie uczyniło niezaprzeczalnie z empirycznej metody doświadczalnej metodę naukową.

Metody statystyczne planowania i analizy doświadczeń rolniczych zaczęły rozwijać się szybko, w ślad za postępem ilościowym i jakościowym doświadczalnej działalności badawczej w nauce i praktyce rolniczej oraz teorii statystyki matematycznej. Można bez wahania powiedzieć, że rewolucyjny rozwój teorii statystyki matematycznej, jako działu matematyki stosowanej, był stymulowany od przełomu XIX i XX wieku do połowy XX wieku, głównie przez rosnące potrzeby metodyki statystycznej, koniecznej do prowadzenia coraz bardziej zaawansowanych naukowo badań doświadczalnych w naukach rolniczych (chemii rolnej, hodowli i uprawie roślin) oraz naukach biologicznych. Później, aż do czasów współczesnych, dalszy nie słabnący postęp statystycznej metodyki doświadczeń był i jest sprzężony silnie z rozwojem zarówno agronomii, jak i biologii (fizjologii, genetyki), medycyny, ekologii, nauk środowiskowych oraz innych nauk wywodzących się z biologii. Także, rozwój teorii metod

statystycznych stymulował postęp w doświadczalnictwie rolniczym oraz w metodologii i zakresie jakościowym i ilościowym badań (Oktaba, 2002; Caliński, 2012).

Metody statystyczne planowania i analizy doświadczeń rolniczych, zgodnie z teorią statystyki matematycznej, obejmują dwa zagadnienia. Pierwszym jest odpowiedni układ doświadczalny. Druga problematyka dotyczy analizy statystycznej danych wraz z optymalnym wnioskowaniem o zjawisku w kategoriach parametrów zmiennych losowych w populacji generalnej. Dlatego metody statystyczne zapewniają możliwie wyczerpujące merytorycznie wnioskowanie indukcyjne o prawidłowościach badanych zjawisk na podstawie danych z przeprowadzonego doświadczenia. Tak uzyskiwane wnioski są powtarzalne, z założonym dużym prawdopodobieństwem, w kolejnych, takich samych lub podobnych eksperymentach. Dlatego, wnioski z dobrze przeprowadzonych i zinterpretowanych doświadczeń wzbogacają wiedzę naukową, zgodnie z filozofią nauki i zastosowaniem empirycznej metody naukowej.

Wnioskowanie statystyczne z danych doświadczalnych odbywa się na podstawie odpowiedniego modelu statystycznego tych danych. Model ten opisuje formalnie zależności przyczynowo-skutkowe lub/i relacje oraz współzależności różnych ekspresji badanego zjawiska. Wnioskowanie statystyczne polega na a) możliwie precyzyjnej ocenie (estymacji) parametrów modelu statystycznego, które charakteryzują prawidłowości badanych zjawisk oraz b) sprawdzaniu hipotez o tych parametrach.

W kolejnych rozdziałach będziemy odkrywać i kontemplować dokonania wielkich uczonych w Polsce i na świecie, których talent, wizja i pasja badawcza doprowadziły na przestrzeni dziejów do ukształtowania teorii statystyki matematycznej oraz biometrii i doświadczalnictwa. Ilustrują one wielkość matematyki i jej rolę w przybliżaniu naszego poznania i rozumienia praw materialnych do prawdy o rzeczywistości, koniecznych do zrównoważonego wykorzystania zasobów Ziemi. R. Fisher powiedział, że powstanie biometrii w XX wieku, podobnie jak geometrii w III wieku p.n.e., wydaje się podkreślać jeden z wielkich przełomowych okresów rozwoju myśli ludzkiej. Jest wielce znamienne, że wielcy matematycy, statystycy i biometrycy mieli zainteresowania przyrodnicze (fizyką, chemią, biologią, rolnictwem) i matematyczne oraz odebrali staranne wykształcenie w tych dziedzinach. Próba ukazania ich dzieła jest celem niniejszej pracy. Niektóre fragmenty tekstu tej pracy są zawarte w Księdze Pamiątkowej, przygotowywanej do wydania z okazji 100-lecia Doświadczalnictwa Rolniczego w Szkole Głównej Gospodarstwa Wiejskiego w Warszawie.

2. Rozwój statystyki i jej zastosowań w doświadczalnictwie do 1920 roku

To, co dzisiaj nazywamy statystyką, ma dość skomplikowaną historię (Oktaba, 2002; Ostasiewicz, 2012). W niniejszym rozdziale przybliżymy część tej historii, poczynając od czasów nowożytnych, kiedy kształtowała się statystyka jako sztuka gromadzenia i wykorzystywania danych ze spisów ludnościowych i gospodarczych, a także teoria rachunku prawdopodobieństwa, aż do końca XIX wieku i pierwszych dwóch dekad XX wieku (do początku pracy R. Fishera w Rothamsted w 1919 roku), kiedy to ukształtowały się fundamenty teorii statystyki matematycznej i jej pierwszych zastosowań w biologii, czyli podstawy biometrii i metod doświadczalnictwa.

Mimo że dobrze udokumentowana historia dyscypliny, określonej jako statystyka, jest niezbyt długa, to jednak jej korzenie sięgają państw antycznych, głównie: Egiptu, Chin, Babilonu, Grecji, Persji, Indii i Rzymu, w których posługiwano się metodami, dzisiaj określanymi jako statystyczne, tj. spisami ludności i inwentaryzacją administracyjną dóbr. Obecnie tego typu działalnością, tj. statystyką gospodarczą, w Polsce zajmuje się przede wszystkim Główny Urząd Statystyczny (GUS), który prowadzi swoją działalność od ponad 100 lat (od 1918 roku). Ta zasłużona dla gospodarki i statystyki instytucja organizuje, koordynuje i wykonuje, między innymi, powszechne spisy rolne. Działalność GUS była i jest silnie inspirująca dla rozwoju metod statystycznych, zwłaszcza metod doboru próby (Neyman, 1934; Kozak, 2004 a, b) oraz szacowania i prognozowania wyników zjawisk gospodarczych.

Statystyka nowożytna pojawiła się wraz z rozwojem filozofii nowożytnej, której rozwój zawdzięczamy głównie René Descartes'owi (1596–1650), uczonemu francuskiemu o spolszczonym nazwisku Kartezjusz. Także F. Bacon, jako jeden z głównych twórców podstaw empiryzmu, miał duży wpływ na rozwój statystyki. W jego czasach rozwój statystyki zaczął być inspirowany i stymulowany uznaniem empiryzmu, jako doktryny metodologicznej badań naukowych.

Ważnym wydarzeniem w procesie narodzin statystyki, jako dyscypliny naukowej było ukazanie się w 1662 roku książki Johna Graunta (1620–1674) *Naturalne i polityczne obserwacje nad biuletynami śmiertelności*, opartej na londyńskich biuletynach śmiertelności. Graunt był wybitnym prekursorem i przedstawicielem kierunku badań nad gospodarką i demografią, zwanym arytmetyką polityczną. Badania te były oparte na wnikliwej analizie ilościowej danych ze spisów i rejestracji administracyjnej. Jako pierwszy, Graunt wykrył, że wnikliwa analiza licznych danych, pochodzących z obserwacji,

prowadzi do ukazania prawidłowości rządzących zjawiskami masowymi. Określenie *Statistik*, wywodzące się od łacińskiego słowa *status*, czyli państwo, pojawiło się po raz pierwszy w języku niemieckim w 1749 roku. Było ono użyte w pracy Gottfrieda Achenwalla (1719–1772) w rozumieniu wiedzy o państwie. Aż do połowy XIX wieku, statystyka, jako sztuka analizy danych, rozwijała się w Europie zachodniej, głównie na potrzeby państwowostwa. Te potrzeby były aż do prawie końca XIX wieku, ważną siłą inspirującą postęp w statystyce. Od co najmniej połowy XIX wieku statystyką zaczęli interesować się też przyrodnicy, biolodzy, rolnicy i lekarze.

Pierwszym śladem zainteresowań rachunkiem prawdopodobieństwa i statystyką w Polsce są *Dyskursy*, autorstwa wybitnego uczonego Jana Śniadeckiego (1756–1830), a także jego rękopis z 1790 roku *Rachunek zdarzeń i przypadków losu* (Ostasiewicz, 2012). Dla propagowania i rozwoju myśli statystycznej w naszym kraju w zastosowaniu do gospodarki oraz administracji państwa polskiego pod zaborami w dobie Księstwa Warszawskiego i początku Królestwa Polskiego, wielce zasłużył się Stanisław Staszic (1755–1826), znakomity, wszechstronnie wykształcony we Francji (studiował w Collège de France i Instytucie Przyrodniczym w Paryżu) naukowiec, przyrodnik oraz działacz państwowy, naukowy, gospodarczy, edukacyjny i społeczny. Na początku XIX wieku Staszic opublikował pracę *Statystyka Polski*, gdzie starał się pokazać Napoleonowi obraz historyczny i gospodarczy kraju, w którym tworzone Księstwo Warszawskie. Staszic był jednym z najważniejszych inicjatorów założenia w roku 1816 Instytutu Agronomicznego w Marymoncie pod Warszawą, którego kontynuatorką tradycji i historii jest Szkoła Główna Gospodarstwa Wiejskiego w Warszawie.

Statystyka została formalnie uznana za naukę w początkach XIX wieku, kiedy to włączono ją jako sekcję do Brytyjskiego Towarzystwa Postępu Nauki i utworzono w roku 1834 Królewskie Towarzystwo Statystyczne. W drugiej połowie XIX wieku świadomość statystyczna w krajach Europy zachodniej była wysoka. Odbywały się międzynarodowe zjazdy statystyczne, znane były prace matematyków europejskich z rachunku prawdopodobieństwa. Ukazywały się też monografie. Ostasiewicz (2012) podaje, że w Polsce pierwsze książki na temat statystyki wydali w latach 1868–1870 Witold Załęski (1836–1908) oraz Zdzisław Korzybski (1834–1896). W. Załęski traktował metodę statystyczną jako jedną z metod poszukiwań naukowych. Uważał, że w świecie rzeczywistym zdarzenia zależą od przyczyn deterministycznych i losowych. Dlatego, głównym zadaniem statystyki jest usunięcie efektów losowych w analizie danych empirycznych i wykrycie prawa statystycznego,

które jest wyrazem prawa funkcjonowania świata rzeczywistego. Były to bardzo światłe poglądy o roli statystyki w empirycznym badaniu zjawisk. Dzisiaj statystyka spełnia tę samą rolę w naukowych badaniach empirycznych we wszystkich dziedzinach nauki. Około 1900 roku statystyka stała się już odrębną i dojrzałą dyscypliną naukową, która była wtedy na progu przełomowej ery wschodzenia i rozwoju teorii statystyki matematycznej, w której dotychczasowa statystyka została zamieniona w dział matematyki stosowanej, gwarantującej naukowy charakter empirycznej metody badań. Podstawą statystyki matematycznej jest teoria rachunku prawdopodobieństwa, tworzona zwykle z inspiracji empirycznej w celu opisu i analizy prawidłowości w zjawiskach masowych. Te zjawiska z natury mają charakter losowy lub deterministyczno-losowy, czyli nie są one w pełni kontrolowalne przez człowieka. Rozwój teorii probabilistycznej został zapoczątkowany w drugiej połowie XVII wieku przez francuskich matematyków: Blaise Pascala (1623–1662) i Pierre Fermata (1601–1665), którzy matematycznie uzasadnili prawidłowości występujące w grach losowych. Dalszy rozwój probabilistyki zawdzięczany jest m.in. takim matematykom, jak: Jakub. Bernoulli (1655–1705), który jako pierwszy uzasadnił i matematycznie określił tzw. prawo wielkich liczb, Abraham Moivre (1667–1754), który wprowadził pojęcie rozkładu prawdopodobieństwa oraz uogólnił prawo wielkich liczb Jakuba Bernoulliego (w pracy z 1733 roku), Daniel Bernoulli (1700–1782), który znacząco rozwinął probabilistyczną teorię błędów, Thomas Bayes (1702–1761) oraz Pierre Simon Laplace (1749–1827). Bayes był matematykiem najbardziej znanym ze sformułowania opublikowanego pośmiertnie twierdzenia Bayesa. Dotyczy ono prawdopodobieństwa warunkowego dwóch zdarzeń losowych. Na bazie tego twierdzenia została rozwinięta teoria wnioskowania bayesowskiego (statystyki bayesowskiej). Probabilistyczna teoria Bayesa znajduje obecnie liczne zastosowania, między innymi w metodach statystycznych, opartych na modelach liniowych i nieliniowych dla sklasyfikowanych danych doświadczalnych (da Silva i in., 2019). Laplace stworzył klasyczną definicję prawdopodobieństwa zdarzenia losowego oraz paradygmat klasycznego determinizmu. Teorie te przedstawił w monografiach *Théorie analytique des probabilités*, wydanej w 1812 roku oraz *Essai philosophique sur les probabilités*, wydanej w 1814 roku. Wymieniony paradygmat uznaje istnienie obiektywnych praw przyrody, które całkowicie determinują zjawiska rzeczywiste, więc człowiek, znając wszystkie prawa, którym podlegają zjawiska i mogąc je analizować, mógłby przewidzieć oraz częściowo kontrolować przebieg i rezultaty zjawisk. Pierre Simon

Laplace i Carl Friedrich Gauss (1777–1855) określili wzór funkcji gęstości prawdopodobieństw rozkładu normalnego zmiennej losowej, wzorując się na teorii błędów. Gauss, jako pierwszy, wyodrębnił składnik systema-tyczny i losowy w matematycznym opisie zjawisk masowych, w czym wyprzedził koncepcje regresyjne twórców statystyki matematycznej z końca XIX i pierwszych dwóch dekad XX wieku. Miał także znaczący udział, obok Adrien-Marie Legendre (1752–1833), w opracowaniu metody najmniejszych kwadratów, stanowiącej ważne kryterium w statystycznej estymacji parametrów zmiennych losowych w populacji generalnej. Zatem, ci dwaj wielcy matematycy: Laplace i Gauss wnieśli swój wczesny wkład w rozwój statystyki matematycznej, zanim opracowano jej teorie.

Przez drugą połowę XVII wieku, cały XVIII wiek oraz prawie cały XIX wiek teoria rachunku prawdopodobieństwa i statystyka budziły rosnące zainteresowanie wybitnych matematyków i fizyków oraz naukowców i administratorów, głównie w dziedzinie państwowostwa, arytmetyki politycznej i nauk przyrodniczych. Jednakże, dwie wymienione dyscypliny rozwijały się w znacznej izolacji, pierwsza miała charakter wybitnie matematyczny, choć była inspirowana empiryzmem, druga zaś miała wymiar empiryczny i aplikacyjny. Trwające dwa wieki prace arytmetyków politycznych, przyrodników, biologów oraz matematyków, rozwijających i coraz odważniej stosujących rachunek prawdopodobieństwa do wnioskowania z danych, doprowadziły w ostatniej ćwierci XIX wieku do znacznego zbliżenia ich celów badawczych oraz interdyscyplinarnej współpracy. Zaowocowała ona powstaniem nowego działu statystyki, jakim jest statystyka matematyczna. Prekursorami teorii statystyki matematycznej i jej pierwszych zastosowań w naukach biologicznych i rolniczych byli trzej Brytyjczycy, tj. Francis Galton (1822–1911), Karl Pearson (1857–1936) i William Sealy Gosset (1876–1937). Natomiast, największy wkład w rozwój teorii statystyki matematycznej i jej zastosowań wnieśli dwaj Brytyjczycy, tj. Ronald Fisher (1890–1962) i Egon Sharpe Pearson (1895–1980) oraz Polak Jerzy Sława-Neyman (1894–1981), publikujący od połowy 1920. lat pod nazwiskiem Jerzy Neyman (Statystycy Polscy, 2012).

Pierwsze idee, koncepcje i metody statystyki matematycznej, podobnie, jak wcześniejsza statystyka, zrodziły się z inspiracji i potrzeb praktycznych w biologicznych oraz przyrodniczo-technologicznych badaniach teoretycznych i empirycznych. Kategorie te dotyczyły populacji generalnej zdarzeń (jednostek) w zjawisku masowym oraz próby losowej i reprezentatywnej, teorii estymacji punktowej parametrów

statystycznych zmiennej losowej w populacji na podstawie dużej i małej próby, rozkładów prawdopodobieństw i częstości błędów i ich wariacji, statystycznej koncepcji korelacji i współczynnika korelacji prostej oraz liniowej regresji prostej. Do zdefiniowania i matematycznego ujęcia wielu z tych pierwszych, fundamentalnych koncepcji i technik statystyki matematycznej oraz ich zastosowania w teorii ewolucji, genetyce populacji cech skategoryzowanych i ilościowych, eugenice, medycynie, psychologii, a także w rolnictwie i technologii ogromnie przyczynili się Galton, K. Pearson oraz Gosset. Galton był pionierem zastosowania teorii rozkładu normalnego do dopasowania histogramów częstości na podstawie dużych prób, pracując nad tym zagadnieniami w latach 1870. oraz 1880. Na przełomie XIX i XX wieku, jeszcze przed powtórным odkryciem praw Mendla około 1900 roku, najpierw Galton, na gruncie ówczesnej wiedzy biologicznej, zaś później jego następca i kontynuator K. Pearson, z zastosowaniem teorii rachunku prawdopodobieństwa i statystyki matematycznej, stworzyli pierwszą szkołę biometrii oraz biometrycznej szkoły dziedziczenia i ewolucji.

K. Pearson w latach 1893–1904 bardzo rozwinął też techniki statystyczne na rzecz biometrii, szczególnie metody korelacji i regresji prostej oraz regresji wielokrotnej liniowej i nieliniowej, a także analizy składowych głównych. Wyniki jego pionierskich prac zostały opublikowane w 1901 roku (Pearson, 1901). Podkreślmy, że dwie pierwsze metody zostały opracowane pierwotnie, w latach 1880. w wersji graficznej i rachunkowej przez Galtona w trakcie rozwiązywania przez badacza różnych problemów w biologii: antropologii, genetyce ilościowej i eugenice. Były one szeroko promowane przez obu twórców, więc szybko znalazły zastosowanie w badaniach biologicznych, medycznych, antropologicznych, psychologicznych i innych, zaraz po ich publikacji w 1901 roku. K. Pearson był nie tylko prekursorem teorii estymacji parametrów, opracowując metodę momentów, ale teorii testowania hipotez statystycznych, wykorzystując rozkład chi-kwadrat i znaną w statystyce wartość p . Takie podejście K. Pearsona, podane w 1900 roku, choć ukierunkowało myślenie późniejszych twórców teorii testowania hipotez, Fishera, Neymana i E. Pearsona, to nie zostało ono przez nich przyjęte w pierwotnym kształcie (Johnson i Kotz, 1997). K. Pearson współzałożył w 1901 roku czasopismo *Biometrika*, które funkcjonuje do dziś i jest wciąż bardzo prestiżowe wśród statystyków i biometryków. Do około 1900 roku teoria estymacji (bez estymacji przedziałowej), oparta na dużej próbie, była już dobrze rozwinięta, głównie przez Galtona i K.

Pearsona. Jednakże, nie była ona wystarczająca dla Gosseta w prowadzonych przez niego pracach naukowo-wdrożeniowych. Jako dyplomowany chemik podjął on w 1899 roku prestiżową pracę analityka naukowego danych w browarze Guinnessa. Gosset miał do dyspozycji u swojego pracodawcy tylko małe zbiory danych (mało liczne próby). Zatem, opracował on własną, oryginalną teorię estymacji parametrów (dla przypadku średniej populacji i współczynnika korelacji prostej), opartej na małej próbie z wykorzystaniem probabilistycznych rozkładów dla błędów estymacji. W tym zakresie Gosset jest najbardziej znany ze swego pionierskiego odkrycia, opublikowanego w 1908 roku pod pseudonimem Student, jakim było opracowanie teorii rozkładu t (Student 1908). Rozkład t jest rodziną ciągłych rozkładów prawdopodobieństw, wywodzących się ze standaryzowanego rozkładu normalnego, który opisuje rozkład błędów estymacji średniej rozkładu normalnego w populacji wówczas, gdy liczebność próby jest mała, a odchylenie standardowe jest nieznane. W 1925 roku rozkład t został nazwany przez Fishera rozkładem Studenta, a później rozkładem t -Studenta, zaś oparty na nim podstawowy test statystyczny nazwano testem t -Studenta (Fisher, 1925).

Rozkład t -Studenta jest podstawą wielu innych późniejszych teorii i procedur statystycznych, takich, jak np. przedziały ufności dla średnich jednej i dwóch populacji, autorstwa Neymana (Neyman, 1937), a także analiza regresji i testy wielokrotnych porównań średnich obiektowych w doświadczalnictwie (Miller, 1981). Gosset był przyjacielem zarówno K. Pearsona (u którego odbył staż naukowy w University College of London w latach 1906/1907), jak i Fishera oraz E. Pearsona. Gosset ma także pionierski wkład do teorii eksperymentu. Planował on i nadzorował doświadczenia polowe z odmianami jęczmienia browarnego na użytek doskonalenia produkcji piwa w browarze Guinnessa, wkrótce po jego zatrudnieniu w 1899 roku. Fisher miał wielkie uznanie dla Gosseta za jego teorię wnioskowania z małych prób, niezwykłą intuicję badawczą, praktyczne nastawienie i prostotę podejścia w rozwijaniu statystyki matematycznej (Johnson i Kotz, 1997).

W pierwszych latach XX wieku, największy polski antropolog i biometryk tamtych czasów, Jan Czekanowski (1882–1965) stosował i promował metody korelacji i regresji wśród antropologów niemieckich, zyskując ich pełną akceptację i uznanie. Czekanowski, jako student Uniwersytetu w Zurychu (1902–1906) napisał artykuł o biometrii, a w roku 1913 opublikował pierwszy podręcznik biometrii w języku polskim *Zarys metod statystycznych w zastosowaniach do antropologii*. Dzieło ukazało się zaledwie

w dwa lata po pierwszym na świecie podręczniku statystyki matematycznej *An introduction to the theory of statistics*, autorstwa G. Yule'a. Podręcznik Czekanowskiego, zawierający m.in. wnioskowanie oparte na współczynniku korelacji oraz regresji prostej i wielokrotnej, odegrał wybitną rolę w rozpowszechnianiu biometrii wśród uczonych polskich przed I wojną światową i w okresie międzywojennym (Caliński, 2012; Statystycy Polscy, 2012). Statystyk Jan Czekanowski i geograf Eugeniusz Romer (1871–1954) odegrali znakomitą rolę ekspertów do spraw geograficznych, demograficznych i antropologicznych przy delegacji polskiej na konferencji pokojowej w Paryżu, odbywającej się w latach 1919–1920 i kończącej I wojnę światową. Pracowali oni tam od końca grudnia 1918 roku do końca października 1919 roku (Romer, 1989). Ich światłe argumenty merytoryczne, dotyczące praw Polski do odzyskania terytorium I Rzeczypospolitej w granicach z 1772 roku, miały duże znaczenie w negocjacjach pokojowych i pomogły przekonać zwycięskie mocarstwa w sprawie granic odrodzonej po zaborach Polski. Wśród Polaków, pionierską i ogromną rolę w zastosowaniach statystyki matematycznej w doświadczałnictwie rolniczym w końcu XIX wieku i pierwszych trzech dekadach XX wieku odegrał Edmund Załęski (1863–1932), z wykształcenia chemik, a z zainteresowań i dorobku także agrotechnik, hodowca roślin, doświadczałnik i statystyk, znakomita postać polskiej nauki rolniczej, prekursor teorii i zastosowań metod statystycznych w doświadczałnictwie i hodowli roślin w Polsce i na świecie. E. Załęski, pracował od 1888 roku w hodowli i nasiennictwie głównie buraka cukrowego i pszenicy, a od 1893 roku we własnej firmie hodowlanej (E. Załęski i S-ka), później w innych hodowlach. W tym czasie rozwijał swoją wiedzę matematyczną i statystyczną, wyniesioną z Politechniki w Rydze oraz konfrontował ją twórczo i wytrwale z praktyką doświadczałną w rozwijającej się hodowli roślin i ocenie odmian, doskonalił metodykę planowania i analizy doświadczeń. Doszedł on znacznie wcześniej przed Gossetem i Fisherem, do wspaniałych rozwiązań metodyczno-statystycznych i bardzo wysokiej precyzji doświadczeń rolniczych (nie osiągniętej jeszcze wtedy na zachodzie Europy), zwłaszcza w hodowli roślin. Znane jest zastosowanie metod E. Załęskiego w 1898 roku przez Aleksandra Janasza i Władysława Mayzela do doświadczeń zbiorowych, tzn. doświadczeń wielokrotnych. W 1907 roku E. Załęski wydał, jednocześnie w 5. językach, wyniki swoich osiągnięć w postaci publikacji *Instrukcja do urządzania doświadczeń porównawczych z różnemi odmianami buraków cukrowych*. Jest

ona uważana za pierwszy systematyczny wykład metodyki doświadczeń rolniczych z zastosowaniem rachunku prawdopodobieństwa i wczesnych metod statystycznych. Swoją pracą z 1907 roku E. Załęski wyprzedził co najmniej o rok pierwsze prace niemieckie i angielskie z tej dziedziny.

E. Załęski przedstawił tam także, opracowane przez siebie w ciągu prawie 20 lat pracy doświadczałnika, zasady stosowania wielu powtórzeń w doświadczeniach czynnikowych (nieznane jeszcze na Zachodzie) oraz koncepcję i zastosowanie metody wzorcowej (Statystycy Polscy, 2012). Metoda ta wykorzystuje wzorec (ang. *standard*), czyli ten sam obiekt doświadczałny, wysiewany w stałym odstępnie w różnych miejscach pola, do poprawienia obserwacji badanych obiektów ze względu na zmienność systematyczną żyzności glebowej. Polega ona na odtworzeniu za pomocą interpolacji liniowej hipotetycznych wartości danej cechy dla obiektów jednopowtórzeniowych, znajdujących się między systematycznie wysiewanymi wzorcami. Na poprawionych danych przeprowadzano analizy statystyczne, dostępne w tamtych latach, zaś od lat 1920., posługiwano się analizą wariancji. Edward Kostecki, bliski wieloletni współpracownik E. Załęskiego, wydał świadectwo o dziele swego mistrza z 1907 roku i jego znaczeniu w tamtym czasie, publikowane w *Gazecie Rolniczej* z 1933 roku (<http://dlibra.umcs.lublin.pl/dlibra/plain-content?id=6355>). Píše on „wyłuszczone w tej pracy zasady metodyczne, co ważniejsze, zasady te od razu stosowane przez Załęskiego i innych hodowców w praktyce, rzeczywiście o wiele wyprzedzały nie tylko oficjalnie obowiązujące przepisy, stosowane wówczas na Zachodzie, ale w wielu razach jeszcze dziś Zachód nie doszedł do ścisłości, osiągniętej metodami E. Załęskiego”. W 1908 roku na Międzynarodowym Zjeździe Rolniczym w Wiedniu, na wniosek E. Załęskiego podjęto rezolucję o stosowaniu metody najmniejszych kwadratów w doświadczałnictwie rolniczym do krytycznego wnioskowania. Późniejszy, dokonany po 1920 roku, imponujący wkład E. Załęskiego i jego licznych uczniów do teorii i praktyki doświadczałnictwa będzie opisany w rozdziale 3.

Alfred Hall (1864–1942), kierujący Stacją Doświadczałną w Rothamsted w latach 1902–1912, propagował i promował statystykę w licznych doświadczeniach rolniczych w tej sławnej rolniczej instytucji naukowej, dostrzegając wielką wartość danych zgromadzonych tamże od 1843 roku. Hall pracował na rzecz doskonalenia metodyki doświadczeń rolniczych w tych samych latach, co Załęski w Polsce, a Gosset w Anglii i Irlandii, którego polski uczyony cenił i współpracował z nim. W 1909 roku Hall opublikował

artykuł o znaczeniu błędu eksperymentalnego w doświadczeniach polowych (Hall, 1909). W tej pracy autor pisze „doświadczenia polowe, bez względu na badane czynniki, podlegają wielu źródłom błędów, głównie z powodu zmienności glebowej, zaś właściwa ocena zmienności błędu doświadczalnego decyduje o wiarygodności oceny efektów czynnikowych”. Dlatego Hall wykorzystał pewne podstawowe pojęcie statystyki (obliczanie średnich i ich błędów standardowych), aby określić, „jakie różnice średnich plonu dla dwóch obiektów doświadczalnych można uznać jako efekt różnego działania tych obiektów, a jakie różnice tych średnich wynikają ze zmienności błędu doświadczalnego”. Jak widać, rozumowanie Halla jest kolejnym, obok myśli Załęskiego i Gosseta, pierwowzorem nowej teorii testowania istotności statystycznej różnic średnich obiektowych, opracowanej w latach 1920. i 1930. zarówno przez Fishera, jak i wspólnie przez Neymana i E. Pearsona. Rok później Hall z agronomami podjęli próbę określenia liczby powtórzeń w doświadczeniach polowych w Rothamsted, które zapewniałyby stwierdzenie istotnych statystycznie różnic średnich obiektowych. Wyniki opublikowano w 1911 roku w *Journal of Agricultural Science*. To było prawdziwe prekursorstwo i zapowiedź tego, co miało nastąpić w Rothamsted 10 lat później. Kolejną zasługą Halla i współpracowników jest wdrożenie w 1910 roku i propagowanie na Zachodzie koncepcji doświadczalnego badania jednorodności pola. Takie badania polegają na wykonaniu doświadczeń beczynnikowych, nazywanych też doświadczeniami ślepych (ang. *uniformity trials, blank trials, dummy trials*), w których stosuje się wyrównaną uprawę jednej rośliny na polu doświadczalnym. Doświadczenia te służą ocenie zmienności przestrzennej żyzności glebowej w obrębie testowanego pola, na podstawie uzyskanych danych dla plonu lub innych zmiennych agronomicznych na najmniejszych poletkach, wydzielonych na tym polu. Dane te pozwalają także na wyznaczenie takiej wielkości i kształtu poletek oraz liczby poletek w bloku, które zapewniają możliwie minimalną ocenę wariancji błędu na testowanym polu doświadczalnym. Zatem, doświadczenia beczynnikowe przyczyniają się do zwiększenia precyzji doświadczeń i mocy testowania hipotez statystyczno-merytorycznych. Jest godne podkreślenia, że ideę doświadczenia beczynnikowego w swoich badaniach hodowlanych i odmianowych wprowadził i stosował też E. Załęski, już na początku XX wieku. Koncepcja doświadczeń beczynnikowych jest wciąż aktualna i stosowana w doświadczalnictwie polowym, chociaż stanowi ona duże wyzwanie praktyczne dla badacza. Analiza danych z doświadczeń beczynnikowych Halla

i współpracowników była ważną inspiracją dla Fishera przy tworzeniu blokowych układów doświadczalnych (Speed, 1992).

Należy podkreślić, że największy wpływ na rozwój statystycznej metodyki doświadczeń rolniczych na świecie od przełomu wieków do 1920 roku wywarli E. Załęski w Polsce oraz Gosset i Hall w Wielkiej Brytanii (Johnson i Kotz, 1997; Statystycy Polscy, 2012). Wszyscy oni bazowali na fundamentalnym dorobku F. Galtona i K. Pearsona w zakresie statystyki i biometrii.

W 1912 roku Hall odszedł z pracy w Stacji Doświadczalnej w Rothamsted. Zastosowanie metod statystycznych w tej instytucji naukowej miało zostać wznowione dopiero po I wojnie światowej. Stacja, pod kierownictwem w latach 1912–1943 chemika rolnego Edwarda Johna Russella (1872–1965), zatrudniła w jesieni 1919 roku matematyka oraz dobrze zapowiadającego się statystyka i genetyka, Ronalda A. Fishera. Wcześniej, w latach 1912–1918, opublikował on już kilka znaczących publikacji, dotyczących podstaw statystyki matematycznej i jej zastosowań w genetyce. Na uwagę zasługuje pionierska, fundamentalna dla genetyki ilościowej, praca z 1918 roku (Fisher, 1918). Zawiera ona matematyczne podstawy genetyki cech ilościowych i model liniowy: $P = G + E$, określający wartość fenotypową cechy ilościowej osobnika w populacji, jako sumę efektu genotypowego i środowiskowego. Równanie to było rezultatem dorobku genetyki pierwszych 15 lat jej rozwoju oraz myślenia matematycznego. Jest ono matematycznym modelem genetyczno-środowiskowego uwarunkowania zmienności fenotypowej organizmów żywych, będącym podstawą współczesnej genetyki ilościowej i jej zastosowań. Stacja Doświadczalna w Rothamsted oczekiwała od Fishera przeprowadzenia naukowej analizy dużych zbiorów danych z doświadczeń polowych, wykonywanych od 1843 roku, których nikt przed nim opracować nie potrafił, a on uczynił to za pomocą nowoczesnych metod statystycznych. Późniejsze rezultaty pracy Fishera oraz jego uczniów i następców, uzyskane w Rothamsted, przerosły najśmielsze oczekiwania. Stanowią one doniosły udział w rozwoju teorii statystyki matematycznej i jej zastosowań w doświadczalnictwie rolniczym. Te osiągnięcia będą przedstawione w rozdziale 3.

3. Osiągnięcia w metodyce doświadczalnictwa w Rothamsted, na świecie i w Polsce w latach 1920–1945

Znaczący rozwój nowoczesnej teorii i zastosowań statystyki matematycznej, poczynając od teorii estymacji, zaczął się dopiero od prac Gosseta, dotyczących małych prób statystycznych, a zwłaszcza od pracy z 1908 roku o rozkładzie *t*-Studenta. Rozwój tych dyscyplin został gwałtownie

przyspieszony od 1920 roku za sprawą dwóch największych statystyków, Fishera oraz Neymana, kiedy Fisher podjął swoją pracę w końcu 1919 roku w Stacji Doświadczalnej w Rothamsted, a Neyman po przyjeździe z Rosji do Polski w 1921 roku zaczął pracować najpierw w Państwowym Naukowym Instytucie Rolniczym w Bydgoszczy, a później w Szkole Głównej Gospodarstwa Wiejskiego w Warszawie (od 1923 roku) i Instytucie im. Nęcickiego w Warszawie (od 1928 roku). Stacja w Rothamsted była pierwszą w świecie instytucją naukową, w której Zakład Statystyki, założony przez Fishera w 1920 roku, był włączony do regularnej współpracy z naukowcami doświadczalnymi. Przyjęto także słuszne założenie, że rozwój metodyki statystycznej doświadczeń może następować tylko na skutek stymulacji przez rosnące potrzeby badaczy doświadczalników. Taka strategia organizacji pracy naukowej w tej sławnej instytucji okazała się owocna zarówno dla stacji, jak i świata, dając przez to wszystkim nam współczesnym jak najlepszy wzór do naśladownictwa. Zatem, jest zrozumiałe, że najważniejszym stymulatorem historycznego rozwoju nowoczesnej teorii statystyki w latach 1920. i 1930. było właśnie doświadczalnictwo rolnicze (Statystycy Polscy, 2012).

Imponujący wpływ na proces rozwoju statystyki matematycznej i jej zastosowań, jaki wywarli najwięksi statystycy: Fisher i Neyman, a także charakter i znaczenie ich dorobku przedstawił Erich L. Lehmann (1917–2009) w swojej obszernej monografii (Lehman, 2011). Ten autor cenionych prac i monografii w zakresie dziejów teorii statystyki matematycznej i jej zastosowań, pracował w Uniwersytecie Kalifornijskim w Berkeley, USA, tym samym, w którym pracował także Neyman od 1938 roku. Lehmann, uczeń i współpracownik Neymana, wyczerpująco dokumentuje i wyjaśnia dokonania Fishera i Neymana oraz ich wkład w opracowanie klasycznej teorii testowania hipotez statystycznych oraz planowania doświadczeń i badań obserwacyjnych lub quasi-doświadczeń. Ich wkład czasami się uzupełniał, czasami wiódł równolegle, a szczególnie na późniejszych etapach ich pracy często znajdował się w silnym sprzeczności. Ta cała historia i rezultat ich naukowych zmagania i intelektualnych postaw, jako ludzi o wielkich umysłach i sile ducha, jest nieprzebranym bogactwem, stanowiącym fundament, na którym stoi obecna statystyka i nauki empiryczne, które przez statystykę są naukami, a nie tylko rodzajem sztuki.

Fisher w swojej pionierskiej pracy z 1922 roku przedstawił fundamenty matematyczne wnioskowania statystycznego, skupiając się na ówczesnych swoich i innych dokonaniach, a dotyczyły one podstaw filozoficznych statystyki

matematycznej oraz estymacji parametrów populacyjnych (Fisher, 1922). Autor pisze, że „przedmiotem metod statystycznych jest redukcja danych”, polegająca na tym, aby „...cała informacja w danych była wiarygodnie przedstawiona za pomocą niewielu wielkości”. Chcąc zrealizować ten cel, uczyony przypomina (po Galtonie i K. Pearsonie) lub wprowadza swoje podstawowe pojęcia statystyczne (znane dzisiaj każdemu statystykowi i stosującemu statystykę). Należą do nich: hipotetyczna nieskończona populacja zdarzeń (ang. *hypothetical infinite population*), próba losowa (ang. *random sample*) reprezentująca tę populację, rozkład prawdopodobieństw populacji pod względem pewnej zmiennej losowej (ang. *law of distribution of this hypothetical population*) oraz parametry populacji (ang. *population parameters*). Te parametry dobrze opisują rozkład prawdopodobieństw zmiennej w populacji. Autor podaje trzy rodzaje problemów we wnioskowaniu statystycznym: (1) określenie matematycznego modelu populacji, czyli rozkładu zmiennej losowej, (2) wybór metod estymowania parametrów populacji, (3) określenie rozkładu prawdopodobieństw estymatorów. Definiuje też własności estymatorów, tj. zgodność (ang. *consistency*) i efektywność (ang. *efficiency*). Wprowadza metodę wiarygodności w estymowaniu parametrów populacyjnych i dowodzi, że estymatory parametrów rozkładu normalnego w populacji są dokładniejsze niż te uzyskane metodą momentów K. Pearsona.

Przełomowym etapem współistnienia rozwoju teorii statystyki matematycznej oraz metod planowania i analizy doświadczeń w latach 1920. było tworzenie podstaw matematycznych (teorii) testowania hipotez statystycznych, które są domniemaniami o wartościach zmiennej losowej w zjawisku masowym (populacji zdarzeń). Metody oparte na tych teoriach, stosowane w badaniach doświadczalnych, pozwalają na wiarygodne odróżnienie rzeczywistych skutków badanych przyczyn (czynników) od takich skutków, które pojawiły się w wyniku błędu losowego próbkowania lub jakiegokolwiek niekontrolowanej zmienności. Dwie, do dzisiaj uważane za klasyczne, chociaż oparte na różnych zasadach logiki probabilistycznej, teorie testowania hipotez statystycznych opracowali Fisher oraz wspólnie Neyman i E. Pearson. Pierwsza teoria, opracowywana przez Fishera od roku 1921 (Fisher, 1921), obejmuje testy istotności (ang. *tests of significance*). Druga zaś, stworzona kilka lat później, przez Neymana i E. Pearsona pod koniec lat 1920. i początku lat 1930. (Neyman, Pearson, 1928), uwzględnia testy hipotez statystycznych (ang. *tests of statistical hypotheses*). Obie, dojrzałe już teorie, zostały przedstawione nieco później (Neyman, Pearson 1933; Fisher, 1935a).

Punktem wyjścia obu teorii testowania hipotez statystycznych jest założenie, że wnioskowanie statystyczne jest postępowaniem opartym na zasadzie naukowej indukcji, czyli wnioskowaniu „od szczegółu do ogółu lub, w języku statystycznym, od próby do populacji” (Okta, 2002; Lehmann, 2011). W podejściu Fishera formułuje się tylko jedną hipotezę — hipotezę zerową, H_0 , która odpowiada przyjętemu modelowi badawczemu. Wybierana jest statystyka testowa T o znanym rozkładzie prawdopodobieństwa, jeśli H_0 jest prawdziwa. Duża bezwzględna wartość T , obliczona na podstawie próby, a więc małe prawdopodobieństwo p ($p < \alpha$, gdzie α jest poziomem istotności, czyli prawdopodobieństwem błędu pierwszego rodzaju), dostarcza badaczowi dowodów przeciwko H_0 i upoważnia go do odrzucenia H_0 . Jeśli $p > \alpha$, hipoteza H_0 nie zostaje odrzucana, co stanowi poważny dalszy problem badawczy, ponieważ taki wniosek nie oznacza, że hipoteza H_0 jest prawie na pewno prawdziwa. W podejściu Neymana-Pearsona formułuje się dwie hipotezy, zarówno hipotezę zerową H_0 , jak i alternatywną H_1 . Sposób postępowania autorów podczas testowania hipotezy jest następujący: odrzucenie H_0 , jeśli $|T| \geq c$ i zaakceptowanie alternatywnej H_1 , albo przyjęcie H_0 , gdy $|T| < c$, gdzie c jest z góry ustaloną wartością krytyczną funkcji testowej, przy założonym prawdopodobieństwie błędu pierwszego rodzaju α i nieznanym prawdopodobieństwie błędu drugiego rodzaju β .

Obecnie, w praktyce zastosowań statystyki, teoria testowania hipotez Neymana-Pearsona z poziomem istotności α jest powszechnie uznawana jako norma w metodologii testowania hipotez statystycznych (Lehmann, 2011). Jednak metoda Fishera testowania istotności, gdzie wartość p jest tylko prawdopodobieństwem przeciwko prawdziwości hipotezy zerowej (nie zaś poziomem istotności α), zdominowała praktykę testowania. W obecnie stosowanym podejściu do testowania hipotez statystycznych zostały wykorzystane obydwie metody sprawdzania hipotez: Fishera oraz Neymana-Pearsona, tworząc swoistą, mieszańcową, procedurę użytkową (Greenland i in., 2016). Warto podkreślić, że w każdym komercyjnym pakiecie statystycznym wszędzie, gdzie testowane są hipotezy statystyczne, jest liczona wartość prawdopodobieństwa p . Patrząc uważnie na takie procedury statystyczne, widzimy, że ich stosowanie przez profesjonalnych badaczy nie musi być zagrożeniem dla wiarygodności wnioskowania statystycznego i naukowego — przecież badacz i tak nie jest w stanie określić w swoich badaniach prawdopodobieństwa błędu drugiego rodzaju β i mocy testu $1 - \beta$, chociaż może je wykorzystać jako źródło inspiracji w doskonaleniu teorii i praktyki metod

eksperymentalnych, zmierzających do redukcji wariancji błędów estymacji. Współcześnie w środowisku statystycznym i biometrycznym toczy się ożywiona dyskusja nad znaczeniem teorii Neymana, dotyczącej przedziałów ufności oraz teorii Fishera i Neymana-Pearsona, dotyczącej testowania hipotez, a także nad przydatnością naukową narzędzi wnioskowania statystycznego, opartych na tych teoriach (Hurlbert i Lombardi, 2009; Greenland i in., 2016).

Fisher, pracując bardzo twórczo i owocnie w okresie 14 lat w Stacji Doświadczalnej w Rothamsted w latach 1919–1933, zdominował historię rozwoju statystycznych metod planowania i analizy doświadczeń rolniczych między wojnami światowymi, 1918–1939. To były wspaniałe lata dokonywania przełomu i stawiania kroków milowych w doświadczalnictwie rolniczym. Jego rozwiązania metodyczne z tamtych czasów pozostają podstawą doświadczalnictwa rolniczego do obecnych dni (Speed, 1992). Fisher napisał dwie pionierskie monografie o statystycznej teorii doświadczalnictwa, wydane po raz pierwszy w roku 1925 (Fisher, 1925) i w roku 1935 (Fisher, 1935b) i mające wiele późniejszych wydań. Oryginalna terminologia metodyczno-statystyczna, zastosowana w tych monografiach i później utrwalona jako klasyczna, pochodzi z natury i metodyki doświadczeń rolniczych. W Rothamsted, poza opracowaniem teorii testowania hipotez statystycznych, Fisher wniósł wielki wkład do biometrii i doświadczalnictwa, tworząc podstawy estymacji parametrów za pomocą metody największej wiarygodności, procedurę analizy wariancji wraz z testem F (od nazwiska autora), idee wydzielania bloków i randomizacji w układach doświadczalnych, koncepcję nowych układów doświadczalnych: całkowicie losowego, losowanych bloków, split-plot i kwadratu łacińskiego oraz układów czynnikowych (ang. *factorial designs*). Nieograniczony dostęp do cennych retrospektywnych danych doświadczalnych w Stacji był bardzo pomocny w ocenie przydatności i skuteczności jego nowych, wymienionych wyżej, metod statystycznych. Stulecie rozpoczęcia pracy Fishera w Stacji w Rothamsted oraz wkład tamtejszych statystyków w rozwój nowoczesnych metod statystycznych zostały uczczone podczas międzynarodowej konferencji organizowanej pod auspicjami Biometric Society, która odbyła się w Rothamsted w lipcu 2019 roku.

W roku 1933 Fisher zakończył swoją chwalebłą pracę w Rothamsted i przeniósł się do University College w Londynie, aby znów zajmować się głównie genetyką statystyczną i eugeniką. Godnym następcą Fishera w Rothamsted był Frank Yates (1902–1994), zatrudniony przez niego w 1931 roku. Chociaż razem pracowali tylko przez dwa lata do 1933 roku, Fisher i Yates

pozostali bliskimi współpracownikami i przyjaciółmi przez kolejne 29 lat. Yates pracował w latach 1930. w kooperacji z Fisherem, nad koncepcjami nowych układów doświadczalnych, tj. kwadratu łacińskiego, układów o blokach niekompletnych (bardzo przydatnych w badaniach biologicznych i rolniczych), układów kratowych i układów czynnikowych. Yates'owi zawdzięczamy też teorię uwikłania w doświadczeniach czynnikowych (ang. *the theory of confounding in factorial experiments*) i ułamkowych układów czynnikowych (ang. *fractional replications*), układów split-plot, układów zrównoważonych i częściowo zrównoważonych o blokach niekompletnych (ang. *balanced and partially balanced incomplete block designs*) oraz układów kratowych i kwadratów quasi-łacińskich (ang. *lattice squares* oraz *quasi-Latin squares*). Bez tych wszystkich nowych układów doświadczalnych, doświadczalnictwo rolnicze oraz przyrodnicze i techniczne nie byłoby tak rozwinięte, jak jest już od prawie 80 lat. Yates zasłużył się też bardzo w szerokim zastosowaniu swoich nietrywialnych rachunkowo pomysłów, dokonując przełomu w komputeryzacji Stacji w Rothamsted i wyznaczając wzorce dla innych. Fisher i Yates w 1936 roku wydali obszernie tablice statystyczne, z cennym źródłem informacji niezbędnych dla użytkowników statystyki (Fisher, Yates, 1936).

Bliskim współpracownikiem Yates'a w Rothamsted był William Cochran (1909–1980), który pracował tam w latach 1934–1939. W ciągu tych sześciu lat pracy w Stacji Cochran stał się pionierem, wspólnie z Yates'em, w opracowywaniu metod analizy wyników z długookresowych doświadczeń rolniczych (ang. *long-term agricultural experiments*), głównie z doświadczeń plodozmianowych (Cochran, 1939).

W latach 1930. dokonał się także, zainicjowany przez Harolda Hotellinga (1895–1973), postęp w statystyce wielowymiarowej, później mocno rozwijanej oraz współcześnie szeroko wykorzystywanej w doświadczalnictwie. Ten prekursor statystyki wielowymiarowej opracował w 1931 roku uogólniony (wielowymiarowy) rozkład *t*-Studenta, nazywany od jego nazwiska rozkładem T^2 -Hotellinga, który jest spokrewniony z rozkładem *F*-Fishera (Hotelling, 1931). Rozkład T^2 jest podstawą testu T^2 -Hotellinga do testowania hipotez o braku różnic wielowymiarowych średnich populacji, przy założeniu rozkładu normalnego. Rozkład ten jest powiązany z odległością Mahalanobisa (Mahalanobis, 1930, 1936), opracowaną przez Prasanta Chandra Mahalanobisa (1893–1972). Ta miara jest ceniona i stosowana szeroko w biometrii, obok odległości Euklidesowej, zwłaszcza w metodach dyskryminacyjnych, analizie zmiennych kanonicznych (ang. *canonical variate analysis*) i analizie skupień (Caliński i in., 1985). W 1933 roku Hotelling opracował metodę analizy składowych

głównych (ang. *principal component analysis*, PCA) (Hotelling, 1933), której podstawy matematyczne już w 1901 roku przedstawił K. Pearson. Metoda PCA jest często stosowana do eksploracyjnej analizy danych w biometrii, głównie do poszukiwania tzw. czynników wspólnych wielu zmiennych w badanym zjawisku oraz wizualizacji wielowymiarowego podobieństwa obiektów, określonego za pomocą odległości euklidesowej i przybliżonego w małej liczbie (2–3) wymiarów (Johnson i Kotz, 1997). W metodzie PCA stosuje się rozkład macierzy kowariancji lub korelacji według wartości własnych lub rozkład macierzy obserwacji według wartości osobliwych. Metoda analizy składowych głównych stała się podstawą wielu klasycznych i najnowszych metod użytkowych, dostosowanych do różnych specyficznych zastosowań. Wśród nich, w doświadczalnictwie rolniczym najbardziej użyteczne są metody analizy czynnikowej (ang. *factor analysis/exploratory factor analysis*, EFA), analizy korelacji kanonicznych (ang. *canonical correlation analysis*), także pomysłu Hotellinga z 1935 roku, metody analizy skupień metodą *k*-średnich (ang. *k-means clustering*), a także metody oparte na statystycznych modelach addytywno-multiplikatywnych dla danych w klasyfikacji dwukierunkowej, takich, jak model AMMI (ang. *additive main effects and multiplicative interaction model*) i model GGE (ang. *genotype and genotype \times environment interaction model*). Te dwie metody zostały opracowane w końcu XX wieku i są stosowane w analizie danych z serii doświadczeń odmianowych, służącej pogłębionej ocenie i wizualizacji interakcji genotypowo-środowiskowej (Gauch i in., 2008).

W 1936 roku Fisher, już po pionierskich dokonaniach Hotellinga i Mahalanobisa, opublikował swoje kolejne przełomowe dzieło, mianowicie metodę liniowej analizy dyskryminacyjnej (Fisher 1936). Obecnie analizą dyskryminacyjną określane jest zbioru wielowymiarowych metod, prowadzących do znalezienia reguły klasyfikacyjnej na podstawie danych wielozmiennych dla obiektów o znanej przynależności do klas. Koncepcja analizy dyskryminacyjnej Fishera jest wykorzystywana w wielu współczesnych metodach analitycznych, m.in. w metodach z grupy *data mining* oraz *gene microarray data analysis*.

Wielkie są także dokonania polskich prekursorów statystyków oraz naukowców doświadczalników (E. Załęski, J. Czekanowski, J. Neyman) w rozwijaniu metodyki statystycznej w doświadczalnictwie rolniczym i ich wkład do nauki światowej w tym zakresie. Spośród nich, największe zasługi wniósł Neyman, publikując prace z teorii doświadczalnictwa w latach 1923–1934. Jako pierwszy, przed Fisherem, zbudował adekwatny język probabilistyczny do opisu doświadczeń, w których posługiwano się

randomizacją. Dzięki temu możliwe stało się bardziej efektywne planowanie eksperymentów i — co ważniejsze — wiarygodne wnioskowanie. Swoje pierwsze przemyślenia na ten temat Neyman opublikował w języku polskim w roku 1923 (Spława-Neyman, 1923), gdy był zatrudniony w Państwowym Naukowym Instytucie Rolniczym w Bydgoszczy. Idea tej pracy dotarła do szerokiej społeczności statystycznej dopiero w roku 1990, dzięki jej tłumaczeniu przez D. M. Dąbrowską i T. P. Speeda dla czasopisma *Statistical Science*. Przełomowym nowatorstwem w tej pracy było przedstawienie idei eksperymentalnego układu całkowicie losowego oraz modelu probabilistycznego danych z takiego doświadczenia, jak również powiązanie planowania zrandomizowanych eksperymentów, z modelem probabilistycznym, pozwalającym na efektywne wnioskowanie statystyczne. Idea wspomnianego modelu jest podobna do tej, którą przedstawił Fisher w monografii z 1925 roku i znanej dzisiaj jako model liniowy analizy wariancji (model ANOVA). Można by więc uznać, że obaj luminarze teorii statystyki, biometrii i doświadczalnictwa, Neyman i Fisher, niezależnie od siebie są twórcami statystycznego modelowania i planowania ścisłych doświadczeń czynnikowych, nie tylko w rolnictwie. Podstawą takiego sądu jest przypuszczenie, a nawet założenie, że Fisher pracy Neymana w języku polskim nie czytał, a także wydanie w 1923 roku pierwszej pracy Fishera ze wstępnymi propozycjami analizy wariancji (ANOVA) w doświadczalnictwie (Fisher, Mackenzie, 1923) oraz wydanie wspomnianego wcześniej podręcznika Fishera w 1925 roku. W modelu Neymana pojawia się po raz pierwszy pojęcie prawdziwej wartości plonów, wiążące się z wartością oczekiwaną zmiennej losowej i prawdziwym efektem poziomu badanego czynnika. Z biegiem lat, na świecie powszechnie doceniono wkład Neymana w metodykę doświadczalnictwa rolniczego, uznając go, wspólnie z Fisherem, za ojca nie tylko teorii statystyki matematycznej, ale też statystycznej teorii eksperymentu. On sam (Neyman, 1979) wyjaśnia dominującą i inicjującą rolę Fishera w teorii eksperymentu, podając także swój udział w tworzeniu podstaw i roli randomizacji w planowaniu doświadczeń.

Należy wspomnieć, że także Gosset jest uważany, obok Fishera i Neymana, za pioniera metodyki statystycznej doświadczalnictwa rolniczego. Wykonywał i stale doskonalił statystyczne metody analizy danych z tych doświadczeń, zaś od 1920. lat współpracował efektywnie z Fisherem, który nazywał go „Faradayem statystyki”. Rezultaty ich współpracy były owocne dla obu uczonych. Na przykład, dyskusja statystyków w latach 1930., z dużym

udziałem Fishera, dotycząca znaczenia systematycznych i zrandomizowanych układów doświadczalnych, była rozpoczęta przez Gosseta w 1931 roku. Pomysły metodyczne w doświadczalnictwie polowym Gosseta, inspirowane też dyskusjami z Fisherem, były publikowane w latach 1923–1937 (Speed, 1992; Johnson i Kotz, 1997).

E. Załęski, który rozpoczął w końcu XIX wieku swoją pionierską działalność naukową w zakresie doświadczalnictwa rolniczego, swój ogromny i nowatorski dorobek zamknął w podręczniku *Metodyka doświadczeń rolniczych*, wydanym w 1927 roku, niespełna 2 lata po ukazaniu się w 1925 roku dwóch ważnych dzieł w Wielkiej Brytanii i Niemczech, tj. monografii *Statistical methods for research workers* autorstwa Fishera oraz *Der Feldversuch* autorstwa Theodora Roemera (1883–1951). E. Załęski w swym oryginalnym dziele, zupełnie nie nawiązującym do przełomowej monografii Fishera z 1925 roku, przedstawił zastosowanie w doświadczalnictwie rolniczym pierwszych procedur analizy statystycznej danych i wnioskowania, opracowanych na przełomie XIX i XX wieku przez przedstawicieli brytyjskiej wczesnej szkoły biometrii i doświadczalnictwa, tj. Galtona, K. Pearsona i Gosseta. Autor omówił tam zastosowanie teorii rozkładów prawdopodobieństw błędów doświadczalnych i estymacji parametrów, a także metodę korelacji i regresji liniowej. Przedstawił też logiczne i przyrodnicze przesłanki planowania wielopowtórzeniowych doświadczeń wazonowych, polowych oraz serii doświadczeń wielokrotnych (przez Autora nazywanych zbiorowymi) i wieloletnich, a także sposoby statystycznego analizowania danych. W planowaniu doświadczeń polowych naświetlił naturę i znaczenie zmienności fluktuacyjnej (nazwał ją doskonale losową) oraz zmienności systematycznej warunków doświadczalnych. Autor przedstawił także teorię i praktykę metod wzorcowych, praktykowanych we własnych badaniach polowych co najmniej od początku XX wieku, a także cenionych przez innych wspaniałych doświadczalników polskich o renomie światowej (m.in. Stefana Barbackiego oraz Józefa Przyborowskiego) i stosowanych aż do lat 1970. (Statystycy Polscy, 2012). Metoda wzorcowa, chociaż prosta w planowaniu doświadczeń i analizie danych oraz w tamtych czasach nowoczesna i bardzo pomocna w wiarygodnej ocenie roślinnych materiałów hodowlanych, przegrywała coraz bardziej konkurencję, najpierw z układami bloków niekompletnych Yates'a, poczynając od lat 1940., a potem, w latach 1980., z metodami analizy kowariancji danych z uwzględnieniem sąsiedztwa obiektów na polu doświadczalnym. Kolejna wartościowa praca E. Załęskiego *Regionalizacja, czyli dobór roślin uprawnych*,

wydana w 1929 roku jest oryginalną próbą oceny wyników z serii wielokrotnej i wieloletniej doświadczeń odmianowych, stanowiąc punkt wyjścia dla późniejszych prac z tego zakresu autorstwa polskich biometryków, tj. Jerzego Neymana, Stefana Barbackiego, Reginy Elandt, Tadeusza Calińskiego i innych.

E. Załęski i jego szkoła na Uniwersytecie Jagiellońskim mają też ogromne zasługi dla praktyki doświadczalnictwa rolniczego w Polsce i rekomendacji odmian oraz nawożenia roślin. Zakład Hodowli Roślin i Doświadczalnictwa na Wydziale Rolniczym Uniwersytetu Jagiellońskiego w Krakowie, kierowany przez E. Załęskiego w latach 1919–1932, a po jego śmierci przez ucznia i godnego następcę Józefa Przyborowskiego (1895–1939) w latach 1933–1939, współpracował z Sekcją Nasienną prężnie działającego Małopolskiego Towarzystwa Rolniczego, a zwłaszcza z Ministerstwem Rolnictwa i Reform Rolnych. Komisja Współpracy w Doświadczalnictwie tego ministerstwa, kierowana przez Przyborowskiego, koordynowała działalność badawczodoświadczalną nad ważnymi roślinami uprawnymi. Doświadczenia te były zakładane w skali ogólnokrajowej według najnowszych jednolitych, opracowanych oraz publikowanych przez Przyborowskiego, zasad statystycznego planowania doświadczeń pojedynczych i wielokrotnych oraz opracowywania danych (np. *Zasady organizacji i wykonywania doświadczeń odmianowych ze zbożami i ziemniakami*, Kraków, 1925). Zalecenia statystycznych metod planowania doświadczeń i analizy danych oraz uzyskane wyniki wraz z ich interpretacją w skali ogólnokrajowej były publikowane regularnie przez Przyborowskiego i Wileńskiego w latach 1930. w postaci zwartych monografii naukowo-wdrożeniowych (np. *Metoda przeprowadzania doświadczeń z zastosowaniem poletek wzorcowych*, Kraków, 1937; *Analiza zmienności wyników doświadczeń wielokrotnych*, Kraków, 1938) oraz w *Przeglądzie doświadczalnictwa rolniczego*. Są one dostępne m.in. w Bibliotece Wydziału Ekonomiczno-Rolniczego Uniwersytetu Rolniczego w Krakowie, będącego spadkobiercą i kontynuatorem dawnego Wydziału Rolniczego Uniwersytetu Jagiellońskiego w Krakowie. Przyborowski i Wileński z olbrzymim zaangażowaniem propagowali w ten koncepcyjny i praktyczny sposób zastosowanie nowoczesnej statystyki w doświadczalnictwie, głównie w zakresie hodowli i oceny odmian. Znaczące dokonania wybitnych przedwojennych doświadczalników teoretyków i praktyków: E. Załęskiego, Przyborowskiego, Wileńskiego oraz Barbackiego stworzyły odważny i śmiały pierwowzór powojennej oceny odmian i dzisiejszego, bardzo nowoczesnego w skali europejskiej, Porejestrowego Doświadczalnictwa Odmianowego, ale także były inspiracją

dla doświadczeń przedrejestrowych i rejestrowych. Wielce zasłużony w latach 1930. (i po II wojnie światowej) dla doświadczalnictwa polskiego był Stefan Barbacki (1902–1979), uczeń E. Załęskiego w Uniwersytecie Jagiellońskim w Krakowie. Pracował on przez 20 lat (1926–1945) w Państwowym Instytucie Naukowym Gospodarstwa Wiejskiego w Puławach. W 1935 roku Barbacki wydał znakomity podręcznik *Ogólna metodyka doświadczeń polowych w zarysie*. Ukazał się on w tym samym roku, co podręcznik *The design of experiments* autorstwa Fishera oraz *Complex experiments* autorstwa Yates'a. W tej monografii, Barbacki wprowadził w Polsce swoje oraz opracowane przez Fishera i Yates'a metody planowania doświadczeń pojedynczych i wielokrotnych oraz statystycznego analizowania wyników. Można dziś z całym przekonaniem powiedzieć, że o ile podręcznik E. Załęskiego z 1927 roku dawał podstawy współczesnej metodyki doświadczalnictwa, o tyle podręcznik Barbackiego z 1935 roku wprowadził do polskiego doświadczalnictwa rolniczego najnowsze osiągnięcia teoretyczne i metodyczne tamtych lat. W 1939 roku Barbacki opublikował monografię *Analiza zmienności w zagadnieniach doświadczalnictwa rolniczego*, której nakład uległ całkowitemu zniszczeniu we wrześniu 1939 roku. Oto fragmenty z ocalałych cudem szpał korektorskich tej monografii: „Metody statystyczne ułatwiają nam badanie doświadczalne, ale nie mogą stworzyć nic nowego. Jeśli zagadnienie w doświadczeniu jest źle postawione, to metody statystyczne tego nie zmieniają. Dadzą może odpowiedź ścisłą, ale nie na to, czego właściwie chcielibyśmy się dowiedzieć” (Caliński, 2012). Jako jedyny Polak, Barbacki wydał wspólną pracę z Fisherem (Barbacki, Fisher, 1936).

4. Postęp w statystycznej metodyce doświadczalnictwa po II wojnie światowej

Ważnym osiągnięciem w teorii doświadczalnictwa było opracowanie wielokrotnych porównań średnich obiektowych za pomocą jednoczesnych procedur statystycznych (testów istotności i przedziałów ufności), wykorzystywanych po wykonaniu analizy wariancji. Podstawą tych procedur była praca D. Newmana (Newman, 1939), w której autor wychodząc z idei Gosseta, sformułował i zilustrował pierwszy wielokrotny test rozstępu. Ogólne zasady wielokrotnych porównań średnich obiektowych zostały określone w ich obecnej formie w latach 1947–1955 przez trzech głównych badaczy, tj. D.B. Duncana, Henry'ego Scheffego (1907–1977) i Johna Tukeya (1915–2000). Nie było i wciąż nie ma pełnej zgody statystyków co do tego, które z procedur są najlepsze. Procedury wielokrotnych porównań możemy podzielić na 3 grupy: (1) analiza kontrastów (test

Scheffego), (2) procedury oparte na studentyzowanym rozstępie, umożliwiające grupowanie średnich (NIR, metoda Newmana-Keulsa, Tukeya i Duncana) oraz (3) wnioskowanie na podstawie przedziałów ufności (test Scheffego, Benferrioniego i Dunnetta) — (Miller, 1981).

Innym przełomowym wydarzeniem było wprowadzenie przez Sheffego w 1956 roku, liniowego modelu mieszanego (ang. *linear mixed model*) do Fisherowskiej analizy wariancji (Scheffe, 1959). Ta idea, choć już wcześniej rozpatrywana przez Yates'a w latach 1940. w teorii niekompletnych układów blokowych, w wydaniu Scheffego stała się krokiem milowym w rozwoju współczesnej teorii planowania i analizy prostych i złożonych doświadczeń z czynnikami stałymi (ang. *fixed factors*) i losowymi (ang. *random factors*), z powtarzającymi obserwacjami na tych samych jednostkach oraz w metodyce kompletnych i niekompletnych serii doświadczeń rolniczych.

Rozwijano twórczo i efektywnie teorię blokowych układów niekompletnych, zarówno na świecie, jak i w Polsce przez Tadeusza Calińskiego (1928-) oraz Jego uczniów w Poznańskiej Szkole Biometrii i współpracowników zagranicznych. Zaproponowano analizę wewnątrzblokową układu blokowego. Określono spójność oraz niespójność układów blokowych, ortogonalność oraz nieortogonalność układów, zrównoważenie ze względu na wariancję i efektywność, układy blokowe rozkładalne oraz afinicznie μ -rozkładalne (szeroko stosowane w doświadczeniach odmianowych), a także układy zrównoważonych bloków niekompletnych (BIB) i układy częściowo zrównoważonych bloków niekompletnych (PBIB) z bardzo licznymi modyfikacjami i przypadkami szczególnymi. Nastąpił także burzliwy rozwój teorii i zastosowań metod wielowymiarowych, głównie takich, jak wielozmienna analiza wariancji (MANOVA) i analiza zmiennych kanonicznych. Kluczowy udział w tym dziele miał Calyampudi Radhakrishna Rao (1920-), doktorant Fishera z roku 1948, Caliński (doktorant Barbackiego z roku 1961) i Jego uczniowie oraz John Gower (1930–2019), współpracownik Yates'a w Rothamsted.

Należy podkreślić wielkie zasługi wielu polskich doświadczalników-statystyków i rolników w rozwoju metodyki doświadczalnictwa rolniczego w skali krajowej i międzynarodowej po II wojnie światowej. Wielki wkład naukowy, dydaktyczny i organizacyjny w postęp statystycznej metodyki doświadczalnictwa rolniczego w tym okresie wnieśli głównie (w kolejności chronologicznej rozpoczęcia działalności zawodowej) Stefan Barbacki, Zygmunt Nawrocki, Regina Elandt, Wiktor Oktaba i Tadeusz Caliński (Statystycy Polscy, 2012).

W 1945 roku, Stefan Barbacki przeniósł się z Puław do Poznania na Wydział Rolniczo-Leśny

Uniwersytetu Poznańskiego (współtworząc i prowadząc Katedrę Doświadczalnictwa Rolniczego i Biometrii), przemianowany w 1951 roku w samodzielną Wyższą Szkołę Rolniczą. W poznańskim środowisku naukowym i rolniczym (hodowli roślin i oceny odmian) w Wielkopolsce oraz całym kraju promował i rozwijał swoje przedwojenne osiągnięcia i prze-myślenia w zakresie planowania eksperymentów oraz ich analizy statystycznej, wcielając także do praktyki idee Załęskiego. Wydana w 1951 roku jego książka *Doświadczenia kombinowane* jest niedoścignionym wzorem pisania w sposób prosty o rzeczach złożonych. Barbacki w 1961 roku współtworzył Zakład Genetyki Roślin PAN, który w roku 1979 został przekształcony, z wielkim udziałem Profesora, w Instytut Genetyki Roślin PAN. Był on także jednym z inicjatorów utworzenia w 1966 roku Centralnego Ośrodka Badania Odmian Roślin Uprawnych (COBORU), a następnie przez wiele lat był przewodniczącym jego Rady Naukowej (Statystycy Polscy, 2012). Barbacki zasłużył się ogromnie polskiej nauce, zarówno w okresie przedwojennym, jak i powojennym, jako znakomity naukowiec-doświadczalnik w zakresie metod statystycznych i praktyki badawczej oraz genetyki i hodowli roślin, a także jako wspaniały organizator i mentor nauki. Swoją wpływ odcisnął i niezatarty ślad zostawił zarówno przez dorobek własny, jak i swoich licznych wspaniałych uczniów. Jego uczniami są luminarze polskiej i światowej biometrii oraz doświadczalnictwa rolniczego: Regina Elandt, Tadeusz Caliński, Eugeniusz Bilski i wielu innych.

Zygmunt Nawrocki (1910–1978) kształcił się w czasie swych studiów na Uniwersytecie Jagiellońskim w Krakowie pod kierunkiem J. Przyborowskiego. Pracując w praktycznej hodowli roślin w czasie wojny i po jej zakończeniu, w 1950 roku uzyskał stopień doktora nauk rolniczych na Wydziale Rolnym Uniwersytetu im. Marii Curie-Skłodowskiej w Lublinie, na podstawie dysertacji *O metodzie dyskryminacji populacji hodowlanych, opartej na pomiarze wielu cech osobników do nich należących*. W 1951 rozpoczął pracę w Szkole Głównej Gospodarstwa Wiejskiego w Warszawie. Nawrocki ma ogromne zasługi dla rozwoju kształtującego się po wojnie środowiska biometryków i doświadczalników polskich w latach pięćdziesiątych, sześćdziesiątych i siedemdziesiątych XX wieku. Jest on wskrzesicielem po wojnie i kontynuatorem szkoły statystyczno-biometrycznej, stworzonej przed II wojną światową przez Jerzego Splawę-Neymana w Szkole Głównej Gospodarstwa Wiejskiego w Warszawie. Nawrocki jeszcze przed 1951 rokiem, jako jeden z pierwszych w Polsce, zastosował układ kratowy Yates'a w doświad-

zeniach polowych, a zwłaszcza hodowlanych, przy ocenie dużej liczby odmian. W latach 1960. był prekursorem odejścia w praktyce doświadczalnictwa rolniczego, zwłaszcza w hodowli roślin, od sztywnych schematów planowania doświadczeń. Wiele jego koncepcji (np. układ N, będący układem jednopowtórzeniowym dla dużej liczby obiektów hodowlanych oraz wielopowtórzeniowy dla odmiany wzorcowej) wprowadzono do praktyki doświadczalnictwa rolniczego w Polsce. Nawrocki zaadaptował i zastosował matematyczną teorię operatorów rzutowych zasłużonego biometryka holenderskiego Leo Corstena (1924–2013), do wykonywania analizy wariancji sklasyfikowanych danych nierównoważonych. Było to dużym krokiem naprzód w tamtych czasach, kiedy nie było jeszcze rozwiniętej teorii statystycznej modeli mieszanych oraz odpowiedniej technologii informatycznej (Statystycy Polscy, 2012). Nawrocki był od roku 1957 do swojej śmierci członkiem Rady Naukowej Instytutu Hodowli i Aklimatyzacji Roślin w Radzikowie oraz członkiem Rady Naukowej Instytutu Ziemiaka w Boninie. Uczniem (doktorantem) i współpracownikiem Zygmunta Nawrockiego był Zbigniew Laudański (1942–2017). Wniósł on bardzo duży wkład w dalszy rozwój metod Nawrockiego (pod względem teoretycznym przez wykorzystanie teorii operatorów rzutowych oraz informatycznym), stosowanych w hodowli roślin i naukach rolniczych.

Regina Elandt (1918–2011) ukształtowała swoje zainteresowania naukowe w poznańskiej szkole naukowej Barbackiego, uzyskując pod jego kierunkiem doktorat w 1955 roku na podstawie pracy *O pewnych testach interakcji w doświadczeniach wieloletnich i wielokrotnych. Zagadnienie rejonizacji*. Swoje liczne osiągnięcia naukowe nad metodami i zastosowaniami statystyki matematycznej w doświadczalnictwie rolniczym oraz genetyce i hodowli roślin zawarła m.in. w znakomitej monografii, wydanej w 1964 roku (Elandt, 1964). Jest to wielkie dzieło Autorki, znane i cenione przez Jej współczesnych, aż po dzień dzisiejszy, za nieprzemijającą nowoczesność i przystępność ujęcia ówczesnego dorobku biometrii i jej zastosowań. Inne osiągnięcia Elandt, wielce zasłużonej dla doświadczalnictwa w jej polskim okresie pracy (1946–1964), dotyczą zagadnień biometrycznych w zakresie genetyki człowieka oraz analizy przeżywalności i epidemiologii (Statystycy Polscy, 2012).

Szeroko znane, znaczące i oryginalne osiągnięcia naukowe Wiktora Oktaby (1920–2009) obejmują zróżnicowany zakres tematyczny i dotyczą następujących problemów: regresji, analizy wariancji jednej i wielu zmiennych, estymacji komponentów wariancyjnych, estymacji i weryfikacji hipotez dla jednozmiennych modeli

stałych niepełnego rzędu, teorii układów eksperymentalnych, wielozmiennej analizy ze szczególnym uwzględnieniem modeli Zyskinda-Martina, teorii brakujących obserwacji, algebry macierzy oraz historii statystyki. Oktaba stworzył Lubelską Szkołę Statystyki i Doświadczalnictwa. Wydał on wiele cennych podręczników akademickich (Statystycy Polscy, 2012). Był inicjatorem organizowania corocznych konferencji naukowych o nazwie Colloquium Biometryczne, od wielu lat, aż do chwili obecnej organizowane jako Międzynarodowe Colloquium Biometryczne. Drugim, wybitnym i zasłużonym statystykiem doświadczalnikiem w Lubelskiej Szkole był Tadeusz Przybysz (1929–2007). Do Jego najważniejszych osiągnięć naukowych należy rozwój metodyki bloków niekompletnych i doświadczeń płodozmianowych. Wprowadził uogólnienie metody Yates'a, metodę analizy i porównania płodozmianów z uwzględnieniem rośliny testowej. Sformułował kilka modeli dla doświadczeń płodozmianowych, takich jak hierarchiczny, kombinacji klasyfikacji krzyżowej z hierarchiczną oraz modelu wynikającego z połączenia układu z rozszczepionymi jednostkami z układem o blokach niekompletnych. Dla poszczególnych modeli Autor podał estymatory parametrów i właściwą postać analizy wariancji do testowania hipotez.

Tadeusz Caliński jest wśród Polaków najwybitniejszym po Jerzym Neymanie, statystykiem, biometrykiem i doświadczalnikiem, uznanym w Polsce i na świecie. Wniósł i stale wnosi dominujący wkład do statystyki, biometrii, w tym doświadczalnictwa rolniczego w skali światowej i krajowej. Rozwijając poznańską szkołę naukową Barbackiego, Caliński stworzył wspaniałą Poznańską Szkołę Biometrii. Dorobek Calińskiego i jego licznych znakomitych uczniów, który przyniósł im największą chwałę, mieści się w szerokim obszarze. Obejmuje on głównie: teorię niekompletnych układów blokowych, metody wielowymiarowe, jednoczesne procedury statystyczne, modele mieszane w zastosowaniu do oceny odmian w serii doświadczeń oraz analizę interakcji genotypowo-środowiskowej.

Znaczący wkład w rozwój metodyki statystycznej w doświadczalnictwie rolniczym w Polsce wnieśli: Leokadia Ubysz-Borucka (1919–1989), jej uczeń Jan Trętowski (1942–1993) oraz Ryszard Wójcik (1937–2003). Rozwinęli i znacząco wzbogacili oni (a także wspomniany wcześniej Z. Laudański) dorobek Nawrockiego, przyczyniając się do ukształtowania w Szkole Głównej Gospodarstwa Wiejskiego w Warszawie trzeciego w Polsce, po II wojnie światowej, ośrodka naukowo-dydaktycznego w zakresie biometrii i doświadczalnictwa rolniczego. Ich dorobek obejmuje głównie wielostronne doskonalenie i zastosowania układów eksperymentalnych

oraz metod jedno- i wielowymiarowych w doświadczeniach agrotechnicznych i hodowlanych oraz w ocenie odmian wielu gatunków roślin uprawnych na podstawie danych niezrównoważonych z serii doświadczeń. Ośrodek ten rozwija się prężnie, pracując głównie w obszarach zarówno zarysowanych przez wspomnianych jego założycieli, jak i w nowych obszarach nauk rolniczych, głównie, takich jak geostatystyka, rolnictwo precyzyjne, interpretacja i wykorzystanie danych satelitarnych i modelowanie matematyczne zjawisk.

Należy podkreślić także znaczący wkład Wrocławskiej Szkoły Statystycznej zainicjowanej przez Hugona Steinhausa (1887–1972), a później rozwijanej i kierowanej przez Juliana Perkala (1913–1965) do rozwoju biometrii, w skali nie tylko krajowej, ale też międzynarodowej. Dokonania Szkoły odbywały się głównie w Uniwersytecie Wrocławskim i Wyższej Szkole Rolniczej (dzisiaj Uniwersytet Przyrodniczy) we Wrocławiu. Ważnym nurtem badań wrocławskiej szkoły zastosowań matematyki była analiza wielowymiarowa, zaś jej najważniejsze osiągnięcia dotyczą nowej metody graficznej taksonomii, a zwłaszcza teorii i zastosowań (algorytmów) grafów. Diagram optymalnego grafu został nazwany dendrytem wrocławskim (Statystycy Polscy, 2012).

5. Najnowsze kierunki rozwoju statystycznej metodyki doświadczalnictwa

W ostatnich latach obserwowany jest znaczny rozwój metod statystycznych w doświadczalnictwie, które są bardziej wymagające obliczeniowo. Wynika to przede wszystkim z upowszechnienia komputerów i oprogramowania statystycznego (SAS, GENSTAT, Statistica, R, ASReml, IBM SPSS i XLStat). W takich okolicznościach, od przełomu XX i XXI wieku zaczęto szeroko rozwijać i stosować modele mieszane i metody do analizy sklasyfikowanych danych zrównoważonych, czyli kompletnych (ang. *balanced data*), a zwłaszcza danych niezrównoważonych, czyli niekompletnych (ang. *unbalanced data*). Dane takie są otrzymywane głównie w toku prowadzenia serii odmianowych doświadczeń wielokrotnych i wieloletnich w celu rzetelnej i wszechstronnej oceny wartości gospodarczej nowych odmian roślin uprawnych, wyrażonej w kategoriach ich stabilności i adaptacyjności pod względem ważnych cech rolniczych (Caliński i in., 2005; Smith i in., 2005; Van Eeuwijk i in., 2016; Studnicki i in., 2017). Rozwijane są także efektywnie zastosowania metod wielowymiarowych, które były już znacznie wcześniej opracowane w zakresie podstaw. Należą do nich metody analizy skupień i PCA oraz ich liczne modyfikacje dla zmiennych ilościowych i skategoryzowanych w zastosowaniu do klasyfikacji obiektów genetycznych (Crossa i Franco, 2004). Klasycznymi już metodami użytkowymi

w doświadczalnictwie rolniczym, wywodzącymi się z PCA są procedury statystyczne AMMI i GGE, oparte na stałych lub mieszanych dwóch modelach addytywno-multiplikatywnych, tj. na modelu AMMI oraz modelu GGE. Metody AMMI i GGE służą do oceny, wizualizacji i interpretacji rolniczej interakcji genotypowo-środowiskowej, bardzo ważnej w hodowli, ocenie i rekomendacji odmian roślin uprawnych, na podstawie danych zrównoważonych lub niezrównoważonych, które pochodzą z serii doświadczeń odmianowych (Smith i in., 2005; Gauch i in., 2008; Van Eeuwijk i in., 2016; da Silva i in., 2019).

W ostatnich latach coraz większą rolę odgrywają metody z zakresu geostatystyki, których celem jest analiza danych geograficznych o różnym zasięgu przestrzennym. W odniesieniu do badań rolniczych dotyczy to m.in. rolnictwa precyzyjnego, w tym zaawansowanych metod interpolacji różnych właściwości gleby. Nowym pojęciem z tego zakresu jest m.in. pedometria (ang. *pedometrics*), które powstało głównie z inicjatywy badaczy z Uniwersytetu w Sydney (Australia) oraz Wageningen (Holandia). Celem pedometrii jest zastosowanie różnych metod ilościowych, w szczególności geostatystycznych, mających na celu mapowanie właściwości gleby w różnej skali przestrzennej.

W związku z tym, że ciągle rośnie ilość dostępnych wyników z różnych doświadczeń rolniczych, to podobnie jak w innych dziedzinach nauki, coraz większe znaczenie odgrywają meta-analizy wyników z doświadczeń rolniczych. Mają tu zastosowanie różne metody statystyczne dla danych zagregowanych, w tym stałe i mieszane modele liniowe i nieliniowe dla danych niekompletnych. Coraz częściej w analizie danych z doświadczeń rolniczych stosowane są metody uczenia maszynowego (ang. *machine learning*), w tym modele wykorzystujące sztuczne sieci neuronowe (ang. *artificial neural networks, ANN*), algorytmy drzew decyzyjnych (ang. *decision trees*), maszyny wektorów nośnych (ang. *support vector machine, SVM*) oraz sieci Bayesowskie (ang. *Bayesian network*) do predykcji zmiennych (np. plonu) lub wystąpienia określonych zjawisk (np. porażenia roślin przez choroby). W tej kategorii mieszczą się także zaawansowane komputerowe modele symulacyjne roślin uprawnych (ang. *crop simulation models*), szczególnie przydatne do predykcji wegetacji i plonu roślin uprawnych w badaniach nad rolniczymi efektami zmian klimatu.

W związku z dynamicznym rozwojem biologii molekularnej powstała potrzeba opracowania metod statystycznych do analizy danych dotyczących ekspresji genów, ich ewolucji oraz struktury. Metodami statystycznymi z tego zakresu są między innymi metody mapowania

QTL (ang. *quantitative trait locus mapping*), mające na celu wyjaśnienie związków między fenotypem, a genotypem organizmów żywych. Mają tu zastosowanie między innymi takie metody, jak analiza wariancji, metody złożonego mapowania odległości (ang. *composite interval mapping, CIM*), jak również profilowanie ekspresji genów z wykorzystaniem mikromacierzy DNA. Jedną z metod wykorzystywanych do oceny podobieństwa genetycznego jest analiza wariancji molekularnej (ang. *analysis of molecular variance, AMOVA*). W filogenetyce molekularnej, czyli klasyfikacji genotypów na podstawie danych molekularnych (np. sekwencji DNA) są wykorzystywane metody próbkowania Monte Carlo łańcuchami Markowa (ang. *Markov Chain Monte Carlo, MCMC*).

Podziękowanie

Autorzy serdecznie dziękują z wdzięcznością Panu Profesorowi dr hab., dr h.c. multi Tadeuszowi Calińskiemu za przyjęcie trudu zapoznania się z pierwotną wersją tej pracy oraz dostrzeżenie słabości i życzliwe zaproponowanie wielce uzasadnionych zmian na rzecz wzbogacenia i udoskonalenia treści.

Autorzy kierują także szczerze podziękowania do dwóch anonimowych Recenzentów, których zasadnicze i szczegółowe uwagi pomogły nadać pracy obecny kształt, lepszy od tego, który charakteryzował manuskrypt przed recenzją.

Literatura

- Barbacki S., Fisher R. 1936. A test of the supposed precision of systematic arrangements. *Annals of Eugenics* 7: 189 — 193.
- Caliński T. 2012. Rozwój i osiągnięcia w biometrii polskiej. *Przegląd statystyczny, Numer specjalny I*: 47 — 52.
- Caliński T., Camussi A., Ottaviano E., Kaczmarek Z. 1985. Genetic distances based on quantitative traits. *Genetics*, 111:945 — 962.
- Caliński T., Czajka S., Kaczmarek Z., Krajewski P., Pilarczyk W. 2005. Analyzing multi-environment variety trials using randomization-derived mixed models. *Biometrics*, 61:448 — 455.
- Caliński T., Harabasz J. 1974. A dendrite method for cluster analysis. *Communications in Statistics* 3: 1 — 27.
- Cochran W. G. 1939. Long-term agricultural experiments. *J. Roy. Stat. Soc. (Suppl.)* 6: 104–148.
- Crossa J., Franco J. 2004. Statistical methods for classifying genotypes. *Euphytica* 137: 19 — 37.
- da Silva C. P., de Oliveira L. A., Nuvunga J. J., Pamplona A. K. A., Balestre M. 2019. Heterogeneity of variances in the Bayesian AMMI model for multi-environment trial studies. *Crop Sci.* 59: 2455 — 2472.
- Elandt R. 1964. *Statystyka matematyczna w zastosowaniu do doświadczalnictwa rolniczego*. PWN, Warszawa
- Fisher R. A. 1925. Applications of "Student's" distribution. *Metron* 5: 90–104.
- Fisher R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh* 52: 399 — 433.
- Fisher R. A. 1921. Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk. *The J. Agri. Sci.* 11: 107 — 135.
- Fisher R. A. 1922. On the mathematical foundations of theoretical statistics, *Philos. Trans. Roy. Soc., Lon. Ser. A*, 222: 309 — 368.
- Fisher R. A. 1925. *Statistical methods for research workers*, 1st ed. Oliver and Boyd, Edinburgh.
- Fisher R. A. 1935 a. The logic of inductive inference (with discussion). *J. Roy. Stat. Soc.* 98: 39 — 82.
- Fisher R. A. 1935 b. *The design of experiments*, 1st ed. Oliver and Boyd, Edinburgh.
- Fisher R. A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7: 179 — 188.
- Fisher R. A., Mackenzie W. A. 1923. Studies in crop variation. II. The manual response of different potato varieties. *J. Agric. Sci.* 13: 311 — 320.
- Fisher R. A., Yates F. 1936. *Statistical Tables for Biological, Agricultural and Medical Research*. Oliver and Boyd, Edinburgh.
- Gauch H.G., Jr., Piepho H.P., Annicchiarico P. 2008. Statistical analysis of yield trials by AMMI and GGE: Further considerations. *Crop Sci.* 48:866–889
- Greenland S., Senn S. J., Rothman K. J., Carlin J. B., Poole C., Goodman S. N., Altma D. G. 2016. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur. J. Epidemiol.* 31:337 — 350.
- Hall A. D. 1909. The experimental error in field trials. *J. Board Agr.* 16: 365 — 370.
- Hotelling H. 1931. The generalization of Student's ratio. *Ann. Math. Stat.* 2: 360 — 378.
- Hotelling H. 1933. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.*, 24: 417 — 441.
- Hurlbert S. H., Lombardi C. M. 2009. Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian. *Ann. Zool. Fennici* 46: 311 — 349.
- Johnson N. L., Kotz S. 1997. *Leading personalities in statistical sciences*. John Wiley & Sons, New York.
- Lehman E. L. 2011. *Fisher, Neyman and the creation of classical statistics*. Springer, New York.
- Kozak M. 2004 a. Alokacja próby między warstwy w przypadku cechy wielowymiarowej. *Wiadomości statystyczne* Nr 7: 13 — 21.
- Kozak M. 2004b. Efektywność schematów losowania w badaniach gospodarstw rolnych. *Wiadomości statystyczne* Nr 9:20 — 26
- Mahalanobis P. 1930. On tests and measures of group divergence I. Theoretical formulae. *J. and Proc. Asiat. Soc. of Bengal*, 26: 541 — 588.
- Mahalanobis P. 1936. On the generalized distance in statistics. *Proc. Nat. Inst. Sci.* 2:49 — 55.
- Miller R.G. 1981. *Simultaneous statistical inference*. Springer-Verlag, New York.
- Newman D. 1939. The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. *Biometrika* 31: 20 — 30.
- Neyman J. 1934. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. *J. Roy. Stat. Soc.*, 97:558 — 625
- Neyman J. 1937. Outline of a theory of statistical estimation based on the classical theory of probability. *Philos. Trans. Roy. Soc. London, Series A*. 236: 333 — 380.
- Neyman J. 1979. Narodziny statystyki matematycznej. *Wiadomości Matematyczne*, 22: 91 — 106.
- Neyman J., Pearson E. S. 1928. On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika* 20A, Pt. I: 175 — 240; Pt. II: 263 — 294.
- Neyman J., Pearson E. S. 1933. On the problem of the most efficient tests of statistical hypotheses. *Phil. Trans. Roy. Soc. London. Series A*, 231: 289 — 337.
- Oktaba W. 2002. *Historia teorii eksperymentu*. Lubelskie Towarzystwo Naukowe, Lublin.
- Ostasiewicz W. 2012. Rozwój myśli statystycznej w Polsce w XIX wieku. *Przegląd statystyczny, Numer specjalny I*: 34 — 46.

- Pearson K. 1901. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572).
- Scheffe H. 1959. *The analysis of variance*. John Wiley & Sons, Inc., N. Y.
- Smith A. B., Cullis B. R., Thompson R. 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. *J. Agric. Sci.* 143: 449 — 462
- Speed T. P. 1992. Introduction to Fisher (1926) — The arrangement of field experiments. W: *Breakthroughs in statistics*, Vol. II. (Ed by Kotz S., Johnson N. L.), Springer-Verlag, New York: 71 — 82.
- Splawa-Neyman J. 1923. Próba uzasadnienia zastosowań rachunku prawdopodobieństwa do doświadczeń polowych. *Roczn. Nauk Roln. i Leśnych*, 10: 1 — 51.
- Statystycy Polscy. 2012. (Redakcja Adamczewski W. i in.) GUS i Polskie Towarzystwo Statystyczne, Warszawa
- Student [W.S. Gosset]. 1908. The probable error of a mean. *Biometrika* 6: 1 — 25.
- Studnicki M., Paderewski J., Piepho H.P., Wójcik-Gront E. 2017. Prediction accuracy and consistency in cultivar ranking for factor-analytic linear mixed models for winter wheat multienvironmental trials *Crop Sci.* 57: 2506 — 2516.
- Van Eeuwijk F. A., Bustos-Korts D. V., Malosetti M. 2016. What should students in plant breeding know about the statistical aspects of genotype × environment interactions? *Crop Sci.* 56: 2119 — 2140.
- Yates F. 1935. Complex experiments. *J. Roy. Stat. Soc. (Suppl.)* 2: 181 — 247.
- Romer E. 1989. *Pamiętnik paryski (1918–1919)*. Wyd. Zakład Narodowy Imienia Ossolińskich, Wrocław.
- Załęski E. 1927. *Metodyka doświadczeń rolniczych*. Wydawnictwo Rozpraw Biologicznych, Lwów.