Przegląd Statystyczny. Statistical Review, 2024, vol. 71, 4, 37–42 https://doi.org/10.59139/ps.2024.04.3

Topicality of the scientific achievements of Jerzy Spława-Neyman (1894–1981)[1]

Mirosław K. Krzyśko[a]

The year 2024 marks the 130th anniversary of Jerzy Spława-Neyman's birth. This is an opportunity to recall the life and achievements of this outstanding Polish statistician, one of the founders of modern mathematical statistics. His concepts and the procedures he initiated relating to interval estimation, statistical hypothesis testing, sampling methods and planning experiments have been used in research work in virtually every scientific discipline until today.

In the early 1930s, Spława-Neyman spotted a new form of the estimation problem, i.e. the problem of the estimation by 'an interval' or, more generally, 'by a set'. The inspiration to define confidence intervals came from a 1932 paper on the precision of the estimation of the regression coefficient by Wacław Pytkowski (Pytkowski, 1932), an assistant in the Department of Mathematical Statistics (headed by Spława-Neyman) at the Warsaw University of Life Sciences. The final version of the theory of confidence intervals was published in Spława-Neyman's two papers (1935, 1937a). The latter work was presented at a meeting of the Royal Statistical Society by Harold Jeffreys.

Interval estimation functions today as one of the standards in statistics. Although Spława-Neyman was not the only creator of this method, it is to him that we owe its modern form expressed in terms of frequency probability. The scholar's results in the field of confidence intervals have been continued – the examples include Efron's (1987) bootstrap method of confidence interval construction, Newcombe's (1998) confidence interval for the difference of two fractions, and, on the Polish ground, the results of the late Professor Ryszard Zieliński (2009, 2011) from the Mathematical Institute of the Polish Academy of Sciences and his son Professor Wojciech Zielinski (Jędrzejczak et al., 2021; Zieliński, 2017, 2022) from the Warsaw University of Life Sciences.

The second pillar of statistical inference, invented by Spława-Neyman with Egon Pearson, is the theory of statistical hypothesis testing. In 1928, they published their first joint paper, focused mainly on the test of the quotient of confidence. It appeared in two parts in the *Biometrika* journal (Neyman & Pearson, 1928a, 1928b). As a result

[1] The article is based on the paper delivered at the 42nd Conference Multivariate Statistical Analysis, held on 4th–6th November 2024 in Łódź.

[a] University of Kalisz, Inter-Faculty Department of Mathematics and Statistics, pl. Wojciecha Bogusławskiego 2, 62–800 Kalisz, Poland, e-mail: mkrzysko@amu.edu.pl, ORCID: https://orcid.org/0000-0001-8075-4432.

of further discussions with Pearson, Spława-Neyman formulated the testing problem in the language of optimisation, and in 1930 proved the basic lemma, today referred to as the Neyman-Pearson lemma. This is what Spława-Neyman said about the circumstances in which this discovery was made in his biographical note (Klonecki, 1970, pp. 21–23):

'I am able to point to a specific moment in which I understood how to formulate non-dogmatically the problem of the strongest test of a simple statistical hypothesis given a fixed alternative. Today this problem seems completely trivial, within the capacity of a novice doctoral student, but then, I must confess, it took Egon Pearson and me several years of work. The solution to this particular problem came to me one evening as I sat alone in the Statistical Laboratory of the University of Life Sciences in Warsaw, thinking persistently about something that could have been solved much earlier. The building was locked. Around 8 p.m. I heard voices calling out for me. It was my wife with some friends, who announced that it was time to go to the cinema. At first I felt irritated. Getting up from my desk to answer the call, I suddenly understood that for a given critical area and a given alternative hypothesis, the probability of making an error of the second kind could be calculated – it is represented by a specific integral. If so, the optimal critical area would be one that minimises this integral while at the same time the boundary condition would be related to the probability of the first-kind error. So we are dealing here with a specific problem of a calculus of variations, probably a very simple one. These thoughts came to me in a split second before I looked out the window to reassure my wife. They are still very clear in my memory, but I don't quite remember what film we were watching that evening. Probably one with Buster Keaton'.

The results obtained by Spława-Neyman were incorporated into the paper he and Pearson wrote about uniformly strongest and uniformly best tests in a class of similar tests, presented in November 1932 by Pearson at a meeting of the Royal Statistical Society. After being accepted by the Society, the article was published in the *Philosophical Transactions of the Royal Society* in 1933 (Neyman & Pearson, 1933). This work has been of fundamental importance for the theory of hypothesis testing with a fixed sample size. Spława-Neyman and Pearson provided a template for mathematical statistics in general and for a general decision theory, later developed by Abraham Wald (1950). In 1992, the above-mentioned work was selected for the first volume (Volume I) of a series describing the most important achievements in the foundations of statistics in the 20th century, published by the German publishing house Springer.

The theories of statistical hypothesis testing have been widely accepted and are being further developed. On the Polish ground, examples are the smooth and adaptive tests developed by Professor Teresa Ledwina of the University of Wroclaw and her students, and by Professor Tadeusz Inglot of the Wroclaw University of Technology.

The results of the aforementioned studies were presented by Inglot in his 2021 book entitled *Lectures on Hypothesis Testing Theory* (Inglot, 2021).

In 1937, Spława-Neyman published a paper in a Scandinavian actuarial journal in which he gave an asymptotically-optimal solution to the problem of testing the consistency of an observation distribution with an assigned, completely known, continuous distribution (Neyman, 1937b). He also introduced sequences of local distributions – 'contiguous' distributions, which later became a standard tool of asymptotic statistics. This work was another milestone in the development of statistics.

In 1933, Spława-Neyman's monographic study, *The outline of the theory and practice of surveying the structure of the population by means of the representative method* (Neyman, 1933), was published by the Institute of Social Affairs, with which the scholar collaborated. This monograph pioneered optimal sampling methods – prior to Spława-Neyman's work, no such idea was known. It has since stimulated new research into the sampling theory. On the basis of the aforementioned monograph, a landmark paper entitled *On two different aspects of the representative method: the stratified sampling method and the purposive selection method* was written. It was presented at the Royal Statistical Society on 19th June 1934, and published in the *Journal of the Royal Statistical Society* that same year (Neyman, 1934). The paper contributed to the development of modern scientific sampling. In 1992, it was ranked among the greatest 20th-century achievements in statistical methodology and published in Springer's Volume II. Spława-Neyman's legacy has been continued, among other studies, in the works of Professor Janusz Wywiał (2020, 2023) of the University of Economics in Katowice and in the research of the latter's students.

Two papers on agricultural experimentation written by Spława-Neyman at the State Scientific and Agricultural Institute in Bydgoszcz (Neyman, 1923a, 1923b) mark the author's another important scientific achievement. In 1924, on the basis of the former paper, Spława-Neyman received a PhD in mathematics from Warsaw University. The eminent mathematician Wacław Sierpiński was his promoter. Some part of his doctoral thesis was translated into English in 1990 and published with comprehensive comments in the *Statistical Science* (Splawa-Neyman, 1990). Thanks to the latter paper (Neyman, 1923b), the scholar received his habilitation in 1928 from the University of Life Sciences in Warsaw.

In 1935, at a meeting of the Industrial and Agricultural Section of the Royal Statistical Society, Spława-Neyman presented a paper on orthogonal systems and randomised blocks, which he wrote together with his students Karolina Iwaszkiewicz and Stanislaw Kołodziejczyk (Neyman et al., 1935). The 70-page-long paper was printed in 1935 in the supplement to the *Journal of the Royal Statistical Society*. This publication, along with the work of Ronald Fisher, was important in the development of experience planning.

The research directions initiated by Spława-Neyman have been extensively developed in Poland. A good example here is the study carried out by the team centered around Professor Tadeusz Caliński (Caliński, 2007; Caliński et al., 2017; Caliński & Lejeune, 1998) of the University of Life Sciences in Poznań. Spława-Neyman's legacy has also been continued in the work of Professors Tomasz Górecki and Łukasz Smaga (2017, 2019) from the Adam Mickiewicz University in Poznań, who investigated the methods of variance analysis for functional data, i.e. continuous functions in Hilbert space. The subject of functional data has been comprehensively explored at the Adam Mickiewicz University in Poznań for many years. It is one of the main directions of development of modern mathematical statistics. Another example is the study of linear models conducted by Professor Roman Zmyślony (Fonseca et al., 2018; Zmyślony & Kozioł, 2021) of the University of Zielona Góra and his students and collaborators. As regards foreign continuators of Spława-Neyman's legacy, a number of significant scholars from many scientific centers from around the world could be mentioned, including a distinguished statistician Professor C. R. Rao (Baksalary et al., 1995; Baksalary et al., 1992).

One more field of Spława-Neyman's scientific activity, relatively rarely recalled, is his development of the optimisation models for the American Army. The project, which started in February 1942, was carried out at the Berkely Statistical Laboratory with varying intensity until the end of the war.

Spława-Neyman saw almost unlimited applications of his methods and mathematical statistics in general in different areas of life and science, including medicine (the mechanism of the origin of cancer), astronomy (a model of the distribution of galaxies in interstellar matter), psychology (factor analysis), demography, or studies on the environmental pollution.

It is no exaggeration to say that Spława-Neyman was one of the most distinguished intellectuals of the 20th century. His scientific achievements paved the way for decades-long development of mathematical statistics, and are still relevant and creatively developed today.

References

Baksalary, J., Markiewicz, A., & Rao, C. R. (1995). Admissible linear estimation in the general Gauss-Markov model with respect to an arbitrary quadratic risk function. *Journal of Statistical Planning and Inference*, *44*(3), 341–347. https://doi.org/10.1016/0378-3758(94)00081-6.

Baksalary, J. K., Rao, C. R., & Markiewicz, A. (1992). A study of the influence of the 'natural restrictions' on estimation problems in the singular Gauss-Markov model. *Journal of Statistical Planning and Inference*, *31*(3), 335–351. https://doi.org/10.1016/0378-3758(92)90141-E.

Caliński, T. (2007). On some results of C. Radhakrishna Rao applicable to the analysis of multi-environment variety trials. *Journal of Statistical Theory and Practice*, *1*(3–4), 347–356. https://doi.org/10.1080/15598608.2007.10411845.

Caliński, T., Czajka, S., Kaczmarek, Z., Krajewski, P., Pilarczyk, W., Siatkowski, I., & Siatkowski, M. (2017). On a mixed model analysis of multi-environment variety trials: a reconsideration of the one-stage and two-stage models and analyses. *Statistical Papers*, *58*(2), 433–465. https://doi.org/10.1007/s00362-015-0706-y.

Caliński, T., & Lejeune, M. (1998). Dimensionality in Manova Tested by a Closed Testing Procedure. *Journal of Multivariate Analysis*, *65*(2), 181–194. https://doi.org/10.1006/jmva.1997.1722.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171–185. https://doi.org/10.2307/2289144.

Fonseca, M., Kozioł, A., & Zmyślony, R. (2018). Testing hypotheses of covariance structure in multivariate data. *Electronic Journal of Linear Algebra*, *33*(1), 53–62. https://journals.uwyo.edu/index.php/ela/article/view/1823/1823.

Górecki, T., & Smaga, Ł. (2017). Multivariate analysis of variance for functional data. *Journal of Applied Statistics*, *44*(12), 2172–2189. https://doi.org/10.1080/02664763.2016.1247791.

Górecki, T., & Smaga, Ł. (2019). fdANOVA: an R software package for analysis of variance for univariate and multivariate functional data. *Computational Statistics*, *34*(2), 571–597. https://doi.org/10.1007/s00180-018-0842-7.

Inglot, T. (2021). *Wykłady z teorii testowania hipotez*. Oficyna Wydawnicza GiS. https://www.gis.wroc.pl/pdf/tth_www.pdf.

Jędrzejczak, A., Pekasiewicz, D., & Zieliński, W. (2021). Confidence interval for quantile ratio of the Dagum distribution. *REVSTAT. Statistical Journal*, *19*(1), 87–97. https://doi.org/10.57805/revstat.v19i1.333.

Klonecki, W. (1970). [unpublished manuscript]. Colloquium Metodologiczne z Agrobiometrii.

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine*, *17*(8), 873–890. https://doi.org/10.1002/(SICI)1097-0258(19980430)17:8%3C873::AID-SIM779%3E3.0.CO;2-I.

Neyman, J. (1923a). Próba uzasadnienia zastosowań rachunku prawdopodobieństwa do doświadczeń polowych. *Roczniki Nauk Rolniczych i Leśnych*, *10*, 1–51.

Neyman, J. (1923b). Próba rozwiązania niektórych zagadnień doświadczalnictwa rolniczego za pomocą statystyki matematycznej. *Miesięcznik Statystyczny*, *6*(9/12), 241–262.

Neyman, J. (1933). *Zarys teorji i praktyki badania struktury ludności metodą reprezentacyjną*. Instytut Spraw Społecznych.

Neyman, J. (1934). On the Two Different Aspects of the Representative Metod: the Metod of Stratified Sampling and the Metod of Purposive Selection. *Journal of the Royal Statistical Society*, *97*(4), 558–606. https://doi.org/10.2307/2342192.

Neyman, J. (1935). On the problem of confidence intervals. *Annals of Mathematical Statistics*, *6*(3), 111–116. https://doi.org/10.1214/aoms/1177732585.

Neyman, J. (1937a). Outline of a theory of statistical estimation based on the classical theory of probability. *Philosophical Transactions of the Royal Society A*, *236*, 333–380.

Neyman, J. (1937b). Smooth test for goodness of fit. *Skandinavisk Aktuarietidskrift*, *20*, 149–199.

Neyman, J., Iwaszkiewicz, K., & Kołodziejczyk, S. (1935). Statistical problems in agricultural experimentation (with discussion). *Supplement to the Journal of the Royal Statistical Society*. *Series B*, *2*(2), 107–180. https://doi.org/10.2307/2983637.

Neyman, J., & Pearson, E. S. (1928a). On the use and interpretation of certain test criteria for purpose of statistical inference. Part I. *Biometrika*, *20A*(1/2), 175–240. https://doi.org/10.1093/biomet/20A.1-2.175.

Neyman, J., & Pearson, E. S. (1928b). On the use and interpretation of certain test criteria for purpose of statistical inference. Part II. *Biometrika*, *20A*(3/4), 263–294. https://doi.org/10.1093/biomet/20A.3-4.263.

Neyman, J., & Pearson, E. S. (1933). On the Problem of the most Efficient Tests of Statistical Hypotheses. *Philosophical Transactions of the Royal Society A*, *231*, 289–337. https://doi.org/10.1098/rsta.1933.0009.

Pytkowski, W. (1932). *Zależność dochodów w małych gospodarstwach od ich powierzchni, nakładów i kapitału zainwestowanego w krowy*. Państwowy Instytut Naukowy Gospodarstwa Wiejskiego.

Splawa-Neyman, J. (1990). On the Application of Probability Theory to Agricultural Experiments. Essay on Principes. Section 9 (D. M. Dąbrowska & T. P. Speed, Trans.). *Statistical Science*, *5*(4), 465–472.

Wald, A. (1950). *Statistical Decision Functions*. Wiley.

Wywiał, J. L. (2020). Estimating the population mean using a continuous sampling design dependent on an auxiliary variable. *Statistics in Transition new series*, *21*(5), 1–16. https://doi.org/10.21307/stattrans-2020-052.

Wywiał, J. L. (2023). On the Maximum Likelihood Estimation of Population and Domain Means. *Journal of Statistical Theory and Practice*, *17*(40), 1–19. https://doi.org/10.1007/s42519-023-00337-4.

Zieliński, R. (2009). Przedział ufności dla frakcji. *Matematyka Stosowana*, *37*(10), 1–17. https://doi.org/10.14708/ma.v37i51/10.265.

Zieliński, R. (2011). *Statystyka matematyczna stosowana. Elementy*. Centrum Studiów Zaawansowanych Politechniki Warszawskiej.

Zieliński, W. (2017). The shortest Clopper-Pearson randomized confidence interval for binomial probability. *REVSTAT. Statistical Journal*, *15*(1), 141–153. https://doi.org/10.57805/revstat.v15i1.207.

Zieliński, W. (2022). The shortest confidence interval for Poisson mean. *Statistical Papers*, *63*(6), 2065–2072. https://doi.org/10.1007/s00362-022-01305-z.

Zmyślony, R., & Kozioł, A. (2021). Ratio F test for testing simultaneous hypotheses in models with blocked compound symmetric covariance structure. *Statistical Papers*, *62*(5), 2109–2118. https://doi.org/10.1007/s00362-020-01182-4.