Stochasticity, non-linearity and non-parametricity: statistical challenges in approximating socio-economic processes

Józef Pociecha[a]

Abstract. The aim of this article is to indicate the relevance of the challenges faced by statisticians in the analysis of socio-economic processes, which requires breaking many of the well-trodden paths of analysis. This leads to a preference for linear, parametric methods, supposedly allowing the discovery of cause-and-effect relationships. According to Kant's philosophy, 'a thing in itself' is not knowable and from this perspective, the main task of statistics is to approximate reality. In statistical studies, only stochastic approximation is made and determining the stochastic structure of processes is as important as studying their regression.

The presented review of the literature leads to the conclusion that stochasticity and non-linearity are the primary features of socio-economic processes and that nowadays their analysis is most effective when based on non-parametric methods. Thus, the paper presents a basic catalogue of methods used for studying non-linearity, the stochastic character and approximation of economic processes using non-parametric methods.

An additional aim of this paper is to emphasize the importance of fundamental statistical monographs to the development of statistical research methodology. The contribution of Polish publications to the advancement of contemporary knowledge is also discussed.

Keywords: approximation approach, stochastic approach, non-linear methods, non-parametric methods

JEL: C00, C10, C14

1. Introduction

The purpose of statistical research is to provide numerical characteristics of the reality surrounding us. When we conduct such research, it is also necessary to reflect on the philosophical and methodological foundations of this type of research. An important question thus arises whether it is possible to arrive at the truth by statistical means? Statisticians believe that it is possible (Pociecha, 2023). Generally speaking, statistics is a method of learning from experience and making decisions under conditions of uncertainty. In this view, the truth can be

[a] Krakow University of Economics, College of Economics, Finance and Law, Institute of Quantitative Methods in Social Sciences, Department of Statistics, ul. Rakowicka 27, 31-510 Kraków, Poland, e-mail: pociecha@uek.krakow.pl, ORCID: https://orcid.org/0000-0003-3140-481X.

understood as a judgment, with an acceptable probability (error), consistent with reality. Methods of statistical inference constitute one of the tools used for finding the truth in statistical terms.

Two outstanding mathematicians and statisticians of the 20th century, Richard von Mises and C. Radhakrishna Rao, emphasised that mathematical statistics is the universal way of arriving at the truth understood in probabilistic terms. In his book *Probability, Statistics and Truth* (von Mises, 1957), von Mises discusses the key issues of probability calculus and mathematical statistics in the light of the foundations of mathematics and the philosophy of science. He discusses the definition and essence of randomness and considers the definitions of probability, starting from the classical one, through the geometric, subjective concept of probability, to the theory of frequentist interpretation of probability. He then asks the question: What is statistics? The author then proceeds to offer comprehensive possible answers to it. Von Mises also pays attention to exploring the nature of causality. He perceives the transition from probability to knowledge of the real world as follows: 'Owing the original relation between the basic concepts and the observed primary phenomena, this theoretical structure permits us to draw conclusions concerning the world of reality. In order to allow a rationally justified application of this probability theory to reality, a quantitative probability concept must be defined in terms of potentially unlimited sequences of observations or experiments' (von Mises, 1957, pp. 7–8).

Rao, likewise, sees statistics as a tool for discovering the truth. In his book, *Statistics and Truth* (Rao, 1997), he uses examples from the natural sciences to show the scientific shift from determinism to probabilism. Quoting Max Born's thought, he writes: 'Today, the order has been reversed: chance has become the primary notion, mechanics an expression of its quantitative laws, and the overwhelming evidence of causality with all its attributes in the realm of ordinary experience is satisfactorily explained by the statistical laws of large numbers' (Rao, 1997, p. 20). Throughout the book, Rao emphasizes the view that statistics has become a tool for finding the truth in modern scientific research.

In addition to the views of eminent mathematicians and statisticians cited above, what inspired this article is a statement by Andrew Briggs, one of the pioneers of quantum technologies, professor at the University of Oxford. In an interview published in Poland, he states that 'life is nonlinear and stochastic' (Briggs et al., 2025, p. 46), which is developed in the book *Human Flourishing: Scientific insight and spiritual wisdom in uncertain times* (Briggs & Reiss, 2021). The authors emphasize that the statement 'life is nonlinear and stochastic' reflects the fact that life events and processes do not follow a straight, predictable path, but are influenced by chance and complex disproportionate relationships. Biological systems and individual human experiences exhibit non-linear dynamics, where small changes can lead to large, disproportionate effects. These systems and experiences are subject to stochasticity characterised by an inherent randomness and probability of events occurring over time.

Here, it should be noted that the full understanding of the reality around us is not possible. In philosophical terms, this was most fully expressed by Immanuel Kant, who formulated the concept of 'the thing in itself' as a reality that exists, but although inaccessible to human consciousness, its presence is manifested through phenomena that are the source of impressions. (Tatarkiewicz, 1978). We can observe these phenomena with greater or lesser accuracy, i.e. we are only able to approximate the reality under study. As previously mentioned, socio-economic processes are stochastic in nature, therefore, the stochastic approximation method is the most general statistical method for searching for the truth about the reality.

The aim of this article is to recall the philosophical and methodological foundations of conducting research and statistical analyses and, in this context, to explore the challenges faced by those conducting such type of research. The most important contemporary challenges facing statistics are as follows:

- The shift from discovering the regularity of the course of socio-economic processes to analysing the stochastic structure of the models constructed to represent them. This will allow a more precise determination of the uncertainty about their future course and thereby, the risk of developmental anomalies;
- Emphasizing that statistical or econometric modelling is only a form of approximation of the studied socio-economic reality, where the acceptable accuracy of the approximation is up to us;
- Assuming that non-linearity is an essential feature of the course of socio-economic processes. Their linear course is only a rare case, and not a reference point for conducting empirical research. Therefore, the basic methods of statistical analysis should be non-linear methods;
- Difficulty in modelling non-linear processes using parametric methods, as it requires defining the analytical form of the non-linear functions and determining the number of parameters involved, resulting in their computational complexity. Thus, non-parametric statistical modelling is more useful;
- The use of statistical learning methods for non-parametric process modelling.

The secondary aim of the work focuses on Polish statistical and economic thought and presents works that made a significant contribution to the world achievements in the field of statistical methods and were published in Polish. The aforementioned challenges will be discussed along with an indication of the recent research achievements in this area and suggestions for the application of the statistical learning paradigm in conducting socio-economic analyses.

2. Cause and chance

Causality is a philosophical category that describes the universal and necessary relationship between phenomena that take place in the objective reality, where one factor, which is the cause, leads to the occurrence of another factor, which is its effect. Causality describes the relationship between variables where a change in the cause leads to a change in the effect. The principle of causality is a fundamental philosophical and scientific concept proclaiming that every phenomenon has its cause and its effect, which describes the universal interdependence of phenomena in the surrounding reality. According to this principle, phenomena are related to each other in a causal relationship, where the cause precedes the effect as a set of conditions necessary and sufficient for its occurrence (Jadacki, 2020). Aristotle had formulated criteria essential to establish a cause-and-effect relationship (Aristotle's theory of four causes). These conditions are: logical temporal order, correlation, lack of alternative explanations (elimination of third factors) and a reliable causal mechanism (Nowakowski, 2017).

The philosophical concept which assumes that all events are inevitable and logical consequences of those preceding them is determinism. According to this idea, each state of the universe is shaped by its previous states and laws of nature, which excludes the influence of chance and free will. There are many types of determinism. The most important are: physical determinism, biological determinism and social determinism. According to physical determinism, all events and states of the universe are predetermined by previous causes and immutable laws of nature. Thus, any future is the only possible one when we know the initial state. According to biological determinism, human behaviours and traits are largely or entirely determined by genes and physiology. Social determinism assumes that an individual's behaviour and development are largely shaped by social factors such as the environment, culture, social class and interactions with other people.

In the history of human thought, determinism was the original philosophical position, according to which all events in nature are predetermined. In antiquity, the most famous supporter of deterministic causality was Democritus, known primarily as the creator of the ancient theory of atomistic matter. Atomists tried to explain the phenomena of the surrounding world through a causal framework. They claimed that nothing happens without a reason, but everything happens for some reason and out of necessity. They understood the nature of causes in a material and mechanical way.

Determinism can take two forms, depending on the degree of determination: extreme and moderate (Pociecha, 2023). In the history of philosophy, the most extreme deterministic view was that of Laplace's 'mathematical demon'. As a spirit endowed with unlimited mathematical deduction, it could predict all future events knowing all the quantities that characterise the present state. In moderate determinism, on the other hand, certain restrictions are imposed on the functioning of cause-and-effect relationships. An example of such a restriction is the view that determinism is basically an intellectual construction, while the development of nature, which is spontaneous and creative, is conditioned by internal force and life momentum, inhibited by inert matter (Lemańska, 1998).

In philosophy, necessity is opposed to contingency. The word 'chance' refers to an unpredictable event or situation. In statistics, instead of the term chance, the term 'randomness' is used. It describes a situation in which we are not able to give a purpose, cause, order or predictable course of a given phenomenon. A random process, on the other hand, is a process whose results cannot be accurately predicted, but can only be described by its distribution. The mathematical equivalent of the concept of chance is a random event. In the calculus of probability, random event A is defined as a measurable subset of a set of elementary events Ω of a given random experience. Each subset contains single or any number of elements, called elementary events. The measurability requirement implies that possible events must form a sigma-body on Ω.

A philosophical position that recognises the objective existence of chance is indeterminism. It assumes that the same causes do not necessarily lead to the same effects. Thus, it makes predicting later phenomena on the basis of earlier ones impossible and negates the strict conditioning of all phenomena. It is based on the common observation that random events are common in the world around us. Science nowadays recognises physical, biological and social indeterminism.

Newton's classical physics, or classical mechanics, was an example of a deterministic understanding of the world. However, in as early as the second half of the 19th century, an Austrian physicist, Ludwig Boltzmann, provided the foundations of statistical mechanics, giving a statistical explanation of the second law of thermodynamics. Werner Heisenberg, a German physicist and philosopher of science, co-creator of the theory of quantum mechanics, winner of the Nobel Prize in physics, formulated the uncertainty principle, which states that there are pairs of quantities that cannot be measured with arbitrary accuracy at the same time. The act of measuring one quantity affects the system so that some of the information about the other quantity is lost. Heisenberg's uncertainty principle does not arise from the imperfection of measurement methods or instruments, but from the very nature of reality. Heisenberg proclaimed that in a world of the smallest material particles there is no causal conditioning; he pointed to the 'freedom of the will of the electron'. This position has shifted to the realm of philosophy, proving the existence of indeterminism in nature. Thus, the deterministic approach, which was the original philosophical understanding of the world, has been replaced by its indeterministic counterpart, which is now the dominant modern scientific view.

In biology, the theory of evolution is based on indeterministic behaviour, which explains the observed diversity of life through the accumulation of random mutations in the gene pool of populations that have not been in contact with each other. Influenced by various selection pressures, a different set of traits is selected from random mutations and, as a result, over time, populations become less and less similar to each other.

Social indeterminism is a philosophical and sociological view that denies the existence of strict, necessary laws governing social phenomena, which means that events in society are not strictly determined by the conditions that precede them. He rejects the idea that society is subject to the laws of determinism and instead emphasizes the role of chance, free will and other factors in shaping human behaviour and societies. Works exposing the indeterministic factor in the social sciences are puzzling. These include the work *Dynamics and Indeterminism in Developmental and Social Processes* by Fogel, Lyra and Valsinger (2014), in which the authors critically analyse the role of indeterminism in human psychological and behavioural development. They review the concepts of indeterminism and determinism in terms of dynamic systems thinking. They apply these general ideas to non-verbal communication systems, emphasizing the vagueness inherent in symbols and the creation of meanings in social systems. They also discuss indeterministic processes occurring within the individual, related to emotional, social and cognitive development.

Economic processes are also indeterministic in nature. Indeterminism in the economy assumes that economic events are not strictly determined by previous causes and their course is influenced by many unpredictable factors, including the free will to make decisions by the market participants. This means that even under identical initial conditions, different economic outcomes are possible. Intriguing are the works *The Concept of Indeterminism and Its Applications: Economics, Social Systems, Ethics, Artificial Intelligence, and Aesthetics*

(Katsenelinboigen, 1997) and *Indeterministic Economics* (Katsenelinboigen, 1992) by Aron Katsenelinboigen, a professor at Wharton School, University of Pennsylvania. In these studies, the author uses the concept of indeterminism to analyse various systems, including economics, social, ethics, artificial intelligence, and aesthetics systems. Katsenelinboigen argues that indeterminism is a fundamental aspect of complex systems, separate from uncertainty and requiring specialist understanding. He promotes his own 'theory of predisposition', which assumes that indeterminism can be understood as a tendency of a system to evolve in a certain way. An interesting book, written in Polish, on the role of indeterminism in economics is by Leszek J. Jasiński, titled *Analysis and Interpretation of Economic Research* (Jasiński, 2017), where the author focuses on the role of induction in economic research, the indeterministic nature of economic processes, the ways of overcoming uncertainty and the functioning paradigms in economics.

All the works cited here indicate the need to emphasize the stochastic character of socio-economic processes. When conducting statistical modelling of such processes, we should not only strive to discover the regularities of their course, but characterise the stochastic range of their variability as well. Equally important is to measure the uncertainty of economic processes resulting in uncertainty about future events or the outcomes of actions, inherently present in the economy. Uncertainty is a key element influencing decision-making and a prerequisite for determining risk. However, risk is a broader concept, encompassing situations in which we know the potential effects of our actions and the probability of such situations occurring.

In light of the statements above, the challenge facing statistics involves treating uncertainty and the related risk measurement as precisely as possible, while uncovering the patterns of past processes and determining their expected course in the future. From a mathematical point of view, this is a problem of optimising the stochastic approximation of the studied socio-economic reality.

3. Approximation as a method of achieving the cognitive goal of science

The main function of science is to acquire knowledge about the surrounding reality, and the cognitive goals and values indicated in the philosophy and methodology of science should serve this function. Witold Strawiński defines the goals of science as follows: 'The main social function of science is the cognitive function. Thus, the main goals and values pursued in scientific research are primarily cognitive goals and values. Other functions of science – educational, innovative, ‘emancipatory’ – are derived from its main function’ (Strawiński, 2011, p. 323).

The goal of scientific cognition is to obtain knowledge that is general, yet precise and simple. In terms of cognition, scientific knowledge is the best type of knowledge, which most adequately describes reality. Science owes its high cognitive status to the methods and language it uses. The use of scientific language and the application of generally accepted scientific rules determine whether or not a given statement is of a scientific nature. Consequently, the acquisition of knowledge about the reality that surrounds us is governed by the rules of scientific methodology, which is itself a metascience (Grobler, 2006). There are many typologies and classifications of scientific disciplines known in the literature (see, for example, Kamiński, 1981). The essence and paths of scientific cognition are presented in more detail, e.g. in the work by Bogdanienko (2018).

The cognitive goals pursued in scientific research depend on what area of knowledge a given scientific discipline represents. The sciences are generally divided into two basic classes: theoretical sciences and empirical sciences. Empirical sciences, based on the principle of a posterior inference, include the social sciences and economics (Sagan, 2016). It should be noted, however, that economics also uses elements of *a priori* inference, an example of which is mathematical economics. An exhaustive overview of the issues related to the philosophy and methodology of economic sciences was compiled by Burnewicz (2021). The cognitive function of economic sciences consists in identifying, describing and explaining economic phenomena and processes on a macro and micro scale, as well as in predicting their course (Kuciński, 2010). The position and specificity of the humanities and socio-economic sciences in the entire typology of sciences is emphasized in the book by Sosenko (2008).

The basic tool of quantitative economic research is the econometric model. Lawrence R. Klein, the winner of the 1980 Nobel Prize in Economics, defines the econometric model as follows: ‘A model is a schematic simplification, omitting irrelevant aspects in order to explain the inner workings, form, or construction of a more complex mechanism. Social systems are extremely complex, sometimes so complex that we are not able to fully explain all aspects of them at once. In such cases, we break the problem down into parts, but sometimes even that is not enough to fully understand it. That is why modelling is an important stage. The social model consists of simplistic assumptions, approximate but understandable dependencies, and a certain explanation of reality. It is not reality, but only a simplified image of reality that man is able to understand’ (Klein, 1982, p. 15).

However, Professor Zbigniew Czerwiński asks further questions and proposes some answers to them. ‘Is a model a description of facts, a stated regularity (of what degree of generality?), the formulation of a hypothesis or something else? It depends on how you understand the model. Regardless of what definition we take, it will not be harmful to econometrics to say that it has not been able to detect universal regularities (independent of the coordinates of time and space), accurate regularities (which work within the limits of measurement error), and non-trivial regularities (beyond common knowledge). Unfortunately, this is the difference between econometrics (and economics in general) from natural sciences such as physics, chemistry, biology, although many economists wanted to imitate these sciences. In explaining this state of affairs, one can refer to various circumstances: to the extraordinary complexity of economic phenomena; the variability of the environment in which these phenomena occur; the impossibility of experimentation (the study of the course of phenomena in artificially created conditions) to the extent that it is possible in the natural sciences; stochastic nature of economic (and more broadly: social) phenomena. These circumstances may explain the difference in the degree of development of the natural and social sciences, without violating the principle of the unity of science. However, the reasons for this difference can be found in the fundamental difference between the objects studied by both sciences. There is nature, there is society, and thus ultimately man, who has certain features that do not occur in inanimate nature and are incomparably less developed in plants and animals: memory (including the memory of the effects of previous actions), assessing (positively or negatively) the situation in which he finds himself, the ability to formulate goals of action and select means to achieve goals. Due to such circumstances, there is a thesis about the fundamental difference in tasks, cognitive capabilities and methods of proceeding in the social sciences and in the natural sciences’ (Czerwiński, 2002, pp. 443–444).

In light of the circumstances cited above, the question is to what extent can we achieve the cognitive goal of the socio-economic sciences. We must be content with the fact that our knowledge of socio-economic processes can only be a better or worse approximation of the surrounding socio-economic reality. Therefore, we must also formulate general criteria when to consider this approximation satisfactory. Such criteria are based on probabilistic grounds. Therefore, the most general method of cognition is the approximation method, because our cognition is always a simplification of the examined reality. The approximation of reality consists in simplifying it in such a way as to preserve all its essential features. It should be noted here that in its most general sense it is a philosophical concept, because we cannot fully know the reality around us and only approximate it.

The concept of approximation is commonly known as a mathematical term. It means replacing ‘true’ function *f*(x) by a simpler function *F*(x), defined in the same area, whose values depend on a certain number of parameters. An approximated function, also called an approximate, is replaced by a function called an approximant. The approximant brings us closer to an unknown primary function. Approximation is used in situations where we do not know or there is no analytical form of the approximate which would allow us to determine the value for any of its arguments; at the same time, the values of this unknown function are known for a certain set of its arguments. The approximation of an approximate by a selected approximant is always associated with the risk of approximation errors. The estimated magnitude of these errors is the criterion for choosing the best approximation method. If the set on which we measure the approximation error is discrete, then we speak of point approximation; however, if a set is specified on a range of real numbers, then we speak of integral approximation.

Generalised polynomials are most often used as approximant functions, which constitute the base function (*F*(x)), or otherwise called the approximation base. If these are additive linear functions, then this type of approximation is called linear approximation. Numerical methods, as a branch of applied mathematics, offer many ways of approximating an unknown or undefined approximate using appropriate computational techniques, formulated as approximation algorithms. These include, for example, rational approximation, which is the quotient of two generalised polynomials, sharing the same analytical form. These polynomials, however, differ in terms of their parameter values. The interpolating function takes the same values as the original function at the interpolation nodes.

This leads us to another type of approximation, i.e. interpolation, which consists in finding an interpolation function which takes the same values as the original function in the interpolation nodes and determines the approximate values of this function at points that are not nodes. Here, the simplest approach is to use polynomials as the base function. In practice, the Lagrange interpolation formulas, the Newtonian interpolation formulas or splined functions are most commonly used. A spline is an actual smooth function for which a family of subintervals exist in the domain of that function, such that the function is a polynomial on each of these intervals. In practice, splined functions of the third degree are often used (Fortuna et al., 1993).

An important approximation method is the mean-square approximation. It is used when the approximated function *f*(x) is known only on the discrete set of its *n* arguments. We assume that φ_j (x), $j=0, 1, ..., m$ is a system of basic functions. Then we look for a function φ(x) that approximates the given real function and provides its smoothing, with a mean-squared error given by the following general formula:

$$R = \sum_{i=0}^{n} [\varphi(x_i) - f(x_i)]^2 \quad (1)$$

Thus, we are looking for a function φ(x) that minimises the value of the error above. This leads to a known system of normal equations after substituting the values of the nodal points *xi* *(i = 0, 1, ..., n).* If we take polynomials as the base functions, then we obtain a polynomial approximation. However, as the degree of the polynomial increases, the calculations become increasingly more laborious, while the results obtained more uncertain. These difficulties can be eliminated by using orthogonal polynomials in the approximation process (Fortuna et al., 1993).

In the process of approximation, a situation tends to occur that the real function f(x) is a periodic function. Such a function must be approximated by trigonometric polynomials and this type of approximation is called trigonometric approximation. If *f*(x) is a continuous function with a period of 2π, the trigonometric polynomial takes the following form:

$$Q_n = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx), \quad (2)$$

where a_k, b_k are the trigonometric Fourier coefficients of function *f*(x) with respect to the orthogonal system of base functions (Fortuna et al., 1993).

Deterministic approximation methods are an important segment of numerical methods, widely used in practice. Their mathematical formalisation, detailed description and computational algorithms can be found e.g. in Kordecki and Selwat (2020). Approximation methods are also used to design deterministic self-learning systems (Wawrzyński, 2021).

4. Stochastic approximation as a cognitive method for studying reality

In the classical approach to approximation, the considered variables are not random variables. The actual approximated function is not known, but only its values in a finite number of points, usually called nodal points. However, if the considered variable is treated as a random variable, then its approximation is called stochastic approximation. A random variable is a function that assigns numbers to elementary events, i.e. a function that allows for a mapping that transfers probability studies from an inconvenient probabilistic space to a well-known Euclidean space. In stochastic approximation, the approximation error is also random.

The concept of stochastic approximation was introduced in the early 1950s by American mathematicians, Herbert Robins and Sutton Monro (Robins & Monro, 1951). They presented an iterative method for finding the root of an unknown real function in cases where for each established value of an argument it is possible to obtain an estimate of the value of this function, but with a random error of a zero expected value. Inspired by this publication, American statisticians, Jack Kiefer and Jacob Wolfowitz (Kiefer & Wolfowitz, 1952), proposed an iterative method that would stochastically estimate the local maximum of a certain real function. Thus, two basic problems of stochastic approximation were formulated in the cited scientific articles along with a proposal of methods of solving them. Both the problems and the methods for their solving have been generalised to the multidimensional case, i.e. for finding the root of a function defined in k-dimensional Euclidean space and for searching for the local extrema of the real function of k-variables (Koronacki, 1989); these are known as RM (Robins-Monro) and KW (Kiefer-Wolfowitz) algorithms.

The stochastic approximation problems formulated above are the stochastic equivalents of deterministic problems involving the solution of systems of non-linear equations and optimisation in finite-dimensional spaces. However, what connects stochastic approximation with deterministic approximation is their iterative nature based on the successive approximations method. The RM and KW methods have gained great popularity, forming the theoretical basis for many stochastic approximation algorithms. Their formalisations and areas of application are comprehensively presented in a book by Koronacki (1989).

Soon after the idea of stochastic approximation in the field of regression research appeared in the world literature, it was transferred to Polish literature thanks to an important and yet somewhat forgotten book by Professor Zdzisław Hellwig, titled *Stochastic Approximation* (Hellwig, 1965). In the introduction to this work, the author writes that the subject of his interests is 'statistical and mathematical methods of research that can be applied in economics'. He continues stating that his paper focuses on 'multivariate random variables with particular emphasis on some aspects of the correlation and regression theory, especially important in economic applications. Some theoretical and practical shortcomings of the traditional way of studying and describing the relationship between statistical quantities prompted the author to attempt to frame the 'outline of the theory of stochastic approximation' presented in this paper. The main idea of the paper is the thesis that the regression function is nothing more than an approximation function, so in the theory of regression and correlation, the basic theorems and some calculation methods provided by the theory of approximation can be used' (Hellwig, 1965, p. 7).

In this work, the basic issue of approximation was formulated for the first time in Polish literature. Suppose we have two continuous variables X and Y, which are known to have some functional dependence, i.e.:

$$y = \psi (x), \tag{3}$$

where ψ (x) is an element of space C of continuous functions. The ψ(x) function is an approximate and it is unknown. In order to guess its form, an experiment is carried out through which *n* observations of the following type are obtained:

$$(x_1, y_1), (x_2, y_2), \dots (x_n, y_n). \tag{4}$$

They are called experimental points or random points, because it is assumed that variables X and Y are random variables. Having experimental points, one should guess the form of function ψ (x) or at least find such an *h* (x) function that could slightly differ from ψ (x) in a certain interval [a, b]. The *h* (x) function is called the approximant. For the first time in the Polish literature, the quoted book provided a mathematical formalisation of the basic problem of approximation and the formalisation of the idea of stochastic approximation in the regression range. Let us assume that function ψ (x) is continuous and differentiable over the entire interval [a, b]. Let us then consider a set of H_k functions $h(x, a_1, a_2, \dots a_k)$, where $a_1, a_2, \dots a_k$, are parameters that can assume a value of zero in certain cases. If function $h(x) \in Hk$, then at each point of interval [a, b], the following relation occurs:

$$\psi (x) - h (x) \mid < \Delta, \tag{5}$$

where Δ is a positive real number.

This approach to the basic issue of approximation (commonly referred to as the mathematical approach to approximation) disregards the debate as to which of the two competing functions h_1 (x) or h_2 (x) better approximates function ψ (x). They are both considered approximants and indistinguishable from the point of view of the approximation criterion (5).

The mathematical formalisation of the regression problem in terms of stochastic approximation is as follows: the starting point is a certain sequence of numbers given by formula (4), which are observations of a two-dimensional random variable (XY). Let a, b be two numbers, so chosen such that:

$$\int_{-\infty}^{\infty}\int_{a}^{b} f\ (xy)\ dxdy > 1 - \alpha \tag{6}$$

where α is a predetermined, small number the range [0,1].

Let R_k denote a set of functions ψ (x, α_1, , α_2 ... , α_k,) such that for a sufficiently large positive number Δ, the following relation is satisfied:

$$P\ [\ \mid Y - \psi\ (x)\ \mid < \Delta] = \int_{-\infty}^{\infty}\int_{a}^{b} f\ (xy)\ dxdy > 1 - \alpha\ . \tag{7}$$

Set R_k consists of approximated functions or regression. And the functions belonging to this set are approximates or regression functions in the population. Number Δ is called the tolerance or acceptable approximation error, while number 1−α is called the likelihood of the approximation.

On the other hand, if H_k denotes a set of functions h (x) with a number of parameters equal to k, such that h (x) ∈ H_k, then and only then the number of m points (x_i, y_i), chosen from sequence (4) and satisfying the relation:

$$y_i - h\ (x_i)\ \mid < \Delta \tag{8}$$

makes up for inequality:

$$\frac{m}{n} \geq 1 - \alpha, \tag{9}$$

where n is assumed to be large enough.

Set H_k is called the set of approximants, and functions h (x) ∈ H_k are called regression functions in the sample. Relation (9) implements that values y_i can differ only randomly from their approximated values h (x_i). Formula (7) can be expressed more conveniently in a slightly different form:

$$P[\psi\ (x) - \Delta < Y < \psi\ (x) + \Delta] \geq 1 - \alpha. \tag{10}$$

This formula resembles the well-known expression for confidence intervals.

Inequality

$$\psi\ (x) - \Delta < Y < \psi\ (x) + \Delta; x \in [a, b] \tag{11}$$

represents a certain area bounded by the lines $\psi\ (x) \mp \Delta$ and_x = a; x = b. This area has the property that the probability of a random event consisting in the fact that a random point (x,y) falling outside it is less than α. This area is called the tolerance area, and the complement of this area is the confidential area. The presented issue of stochastic approximation is not about determining the tolerance area itself, but about relating it to the concept of regression. In approximate terms, 'regression' does not mean a single approximation function, but a whole family of such functions.

The tolerance of an approximation can be a function of x. Then:

$$P\ [v_1(x) < Y < v_2(x)] \geq 1 - \alpha\ ; x \in [\mathrm{a, b}], \tag{12}$$

where:

$$v_1(x) \equiv \psi\ (x) - \Delta(x);\ v_2(x) \equiv \psi\ (x) - \Delta(x). \tag{13}$$

This means that the tolerance range can be different for different x. Relation (12) can be assigned an even more general form if the following condition is imposed on functions $v_1\ (x)$ and $v_2\ (x)$:

$$v_1\ (x) < v_2\ (x). \tag{14}$$

Then the normal space Q bounded by the curves $v_1(x)$ and $v_2(x)$ is called the tolerance area if:

$$P[(xy) \in Q] \geq 1 - \alpha. \tag{15}$$

Complementing the tolerance area is of critical importance.

To sum up, Hellwig (1965, p. 34) notes that an approach to the basic issue of approximation, where the search for the $h\ (x)$ function as the best approximant of $\psi(x)$ is abandoned, and the concept of the H_k set with indistinguishable elements, each of which can play the role of an approximant, is generally and theoretically correct because:

1. Each given function ψ(*x*) corresponds to an infinite H_k set of functions *h* (*x*), if the number of *k* parameters occurring in these functions can be arbitrarily large.
2. In the H_k set there is always an h_0 (*x*) function such that for any sequence (4),

$$\max_i \mid yi - h_0 (xi) \mid = 0. \tag{16}$$

It is enough to assume that $h_0\ (x)$ is a polynomial of degree *n* – 1, because for every sequence (4), it is possible to draw a polynomial of degree at most *n* – 1, which takes the values of y_1, y_2, ... y_n at points x_1, x_2, ... x_n. However, *n* must be predetermined and cannot depend on the length of the experimental data series.

3. If *n* could be determined in advance, then one can look for such an h_0 (*x*) function in set H_k. which would meet the following condition:

$$\max_i \mid yi - h_0 (xi) \mid = inf\ \max_i \ \mid yi - h_0 (xi) \mid ;\ h\ (\mathrm{x}) \in H_\mathrm{k}. \tag{17}$$

Nevertheless, for computational reasons, the number of *k* parameters of the function *h* (*x*) ∈ H_k should be significantly lower than the number of data *n*. However, there are no objective criteria for determining the number *k*, so we are forced to choose it arbitrarily.

All methods of solving the basic problem of approximation are accompanied by various forms of arbitrariness; therefore, all these methods should be considered equivalent (Hellwig, 1965, p. 35). In this situation, the statistical interpretation of the stochastic approximation is as follows: we are looking for a family R_k of such *k*-parameter functions $\psi\ (x)$, which for the predetermined number α and Δ would satisfy the relation:

$$P[\ \mid \mathrm{Y} - \psi\ (x) \mid < \Delta \geq 1 - \alpha. \tag{18}$$

The general formalisation of the basic issue of approximation and the formalisation of stochastic approximation in a regression approach indicate that statistical regression is one of the forms of the approximation of real relations occurring in socio-economic processes. When approximating this reality, it is necessary to determine the tolerance that we accept, i.e. the acceptable discrepancy between the actual relationship between the explanatory variables we define and the variable we explain and its chosen approximant, in statistics called the permissible (average) error of estimation.

The above-mentioned general formulation of regression in the approximation approach, defining the confidence of approximation $1 - \alpha$ suggests that in regression studies, the methods of interval estimation should be used rather than point estimation, because in statistics, the degree of credibility is defined as the confidence factor.

Consequently, the approximation approach to regression indicates that the approximation function chosen on the basis of the sample is only one of many possible regression functions belonging to the set of approximants. What is more, the individual functions belonging to this set are indistinguishable in terms of the accepted level of tolerance. Thus, in econometric modelling, it is unjustified to label a given model 'the winner' simply because other models fall within an acceptable tolerance range. The approximation theory, followed by statistics, provides many forms of analytic functions for use as approximation models. As previously indicated, generalised polynomials, trigonometric polynomials or splines are the most commonly assumed. However, it should be noted that these models require the number of parameters to be determined, at least initially. The approximation approach to regression is thus formalised mainly in terms of parametric statistics.

Currently, computer science provides tools based on artificial intelligence in the form of machine learning, which is the equivalent of the deterministic approach to approximation, which is utilised by numerical and statistical learning methods, equivalent to stochastic approximation. Learning methods thus extend the range of statistical methods to both parametric and non-parametric approaches.

5. Non-linearity of socio-economic processes

Economists and statisticians have long observed that socio-economic processes do not evolve in a continuous or linear way, but are shaped by sudden changes in the economic and political situation, entailing disruptions that lead to changes in developmental directions. This is caused not only by macroeconomic factors, but also people's decisions and emotional behaviours which are not always rational. When analysing such processes, one cannot rely only on simple, linear models, but their complexity, dynamics and unpredictability need to be taken into account.

The non-linearity of socio-economic processes is caused by many factors. These include:

- The concurrence and interdependencies between different spheres of socio-economic life, politics and culture, which create dynamic and often unpredictable mechanisms of social behaviour;
- The presence of feedback effects between these spheres. The actions of one element of the system can affect others, which in turn affects the first element in a way that strengthens or weakens its influence;
- The risk of unpredictable, sudden changes occurring in the conditions of these processes, which is symbolised by the concept of a ‘black swan’. Sudden shifts such as financial crises, technological revolutions or pandemics can radically change the course of socio-economic processes in unexpected ways;
- The possibility of the ‘butterfly effect’: minor initial deviations can lead to significant, divergent effects in the long term;
- The stochastic nature of socio-economic processes. These processes are usually not deterministic in nature. Human decisions, free will and cultural factors play a key role here, introducing an element of unpredictability.

The literature on the non-linearity of the course of socio-economic processes is very extensive. The beginnings of non-linear thinking in economics can be traced back to Adam Smith’s theory, which assumes that growth has its limits caused by decreasing rates of return on investments. Formalised non-linear models of macroeconomic phenomena emerged with the works of several eminent 20th-century economists. These concerned the non-linear growth model presented by Nicholas Kaldor as the Kaldor-Hicks efficiency model, Richard Goodwin's macroeconomic model of economic growth and business cycles, and the classic non-linear relationship between basic macroeconomic processes, as is the Phillips curve (Orzeszko, 2016). Significant examples of the use of non-linear functions in microeconomic research are the Cobb-Douglas production function and the Tornquist consumption functions introduced into the economic literature in as early as the 1920s.

The non-linear nature of socio-economic processes has long been recognised in statistical and econometric research and has led to the successive introduction of non-linear macroeconomic models. It particularly concerns the study of the non-linearity of economic growth and, more broadly, socio-economic development. A current review of this type of work is included in Orlando et al. (2021). Specialist journals are also devoted to the study of non-linearity, an example of which is the *Journal of Nonlinear Sciences and Applications*. It publishes scientific articles on non-linear mathematical methods and their applications, mainly in the field of natural sciences. A current overview of the methods and their applications in the study of non-linear dynamics of economic processes can be found in Gardini et al. (2021).

The non-linearity of processes is revealed in time series analysis based on stochastic models. One of the monumental figures in this area of statistical research, along with G.E.P. Box and G.M. Jenkins, is C.W.J. Granger, winner of the 2003 Alfred Nobel Prize in Economics. He is the author of the statement that the world is almost certainly non-linear (Granger, 1989) and the author of the definition of causality (known as the Granger causality). It states that X causes Y if and only if including explanatory variable X in the model predicting explanatory variable Y increases the accuracy of the prediction. He is also the author of the concept of cointegration introduced in dynamic econometrics, initially in a linear version (Granger, 1981), and ten years later in a non-linear version (Granger & Hallman, 1991). Several generalisations of the concept of cointegration to cases of non-linear relationships have been proposed in the econometric literature based on Granger's proposal. The book by Fan & Yao (2005) still remains a fundamental work in the field of non-linear time series. One of the most recent monographs in this field is a work by Maitra (2025). This book is a comprehensive and accessible guide allowing for the understanding of advanced econometric methods and their use to analyse real data in the form of economic and financial time series, skilfully combining theory with practice. The work presents data filtering techniques, highlighting the role of the Kalman filter in improving model accuracy. Volatility modelling is also discussed, addressing common challenges in measuring and interpreting the variance of financial data. Moreover, hybrid approaches combining GARCH models with neural networks and dynamic volatility models for option pricing are presented, describing both the theoretical foundations and providing practical tools for their application. Finally, mode-switching models, including MSAR and STAR are discussed to show the non-linear behaviours and structural shifts of data in time series. The monograph presents a coherent framework for modelling the dynamic behaviour of financial time series, with a particular emphasis on volatility and structural change. It may be particularly useful for finance professionals and data scientists.

Many works on non-linearity have also been published in Polish. A monograph on non-linear processes and long-term relationships in economics by Bruzda (2007) is one of the most comprehensive studies in this field. Its scope includes the identification of stationary non-linear processes and verification of long-term relationships based on the Granger and Hallman method, and a parametric estimation and verification of non-linear long-term relationships. Moreover, it explores non-linear error correction models, threshold cointegration and it provides many examples of applications of non-linear cointegration analysis. Another extensive work devoted to the identification of non-linearities in financial and economic time series is (Orzeszko, 2016).

This section outlines issues related to the non-linearity of economic processes, particularly focusing on their financial aspects. It shows how economics and financial econometrics developed from non-linear models describing the functioning of economic systems, formulated around the middle of the 20th century by such eminent economists as N. Kaldor, J. Hicks, W. Leontief, R. Goodwin, P. Samuelson, to the latest achievements of non-linear financial econometrics. The published results of both theoretical and empirical research indicate enormous progress in this area of economics. Nonetheless, linear approximation of these processes still prevails. It is thus time to acknowledge that socio-economic processes are inherently non-linear and their linearity is either a special case of non-linearity or an oversimplification of them. A challenge faced by statistics is to improve the tools for detecting and describing the non-linearity of the studied processes.

6. Non-parametric modelling

Non-parametric modelling is a method that does not assume a specific analytical form for the population distribution. Instead, it relies on the shape of the data distribution in a sample or a training set, which allows the structure of the model to be determined based on the data. It is a more flexible method than parametric modelling, which requires a defined analytical form of the approximated function and a fixed number of parameters. Non-parametric modelling is an important segment of statistical methods. It is especially useful when the data do not follow a normal distribution or when it is unknown, and when the studied features are measured on weak, ordinal or nominal scales. Common applications of the non-parametric approach include the Mann-Whitney U test, the Kruskal-Wallis test or the chi-square test, which serve as an alternative to parametric tests.

The above-mentioned problems indicate that their statistical description and modelling in a parametric form is computationally complex and requires the adoption of many restrictive assumptions. An alternative is the non-parametric approach. The non-parametric regression model is the most widely used type. With the help of non-parametric regression, the relationship between the considered variables may be analysed without a predetermined analytical form of the regression function. The non-parametric regression model derives directly from the data. Estimating a non-parametric regression function involves selecting a function of a specific class that is flexible enough to fit well to the data set. For this reason, non-parametric regression models require a much larger number of observations than their parametric counterparts. We approximate not only the values of their parameters, but also the nature of the relationships between considered variables. Thus, as the dimensionality of the model increases, the number of observations necessary for a reliable estimation of the non-parametric regression function increases exponentially. This phenomenon, described as the curse of dimensionality, may be circumvented through semiparametric modelling (Pagan & Ullah, 2009) or the multimodelling approach (Gatnar, 2008).

Non-parametric regression methods are implemented within the statistical learning framework. In the general understanding of statistical learning, we consider variable Y (response variable), also understood as the explanatory (dependent) variable, and *k* of the explanatory variables (predictors) $X_1, X_2, \ldots, X_k$. We assume that between Y and $X = (X_1, X_2, \ldots, X_k)$ there is a certain relationship, which can generally be described as:

$$Y = f(X) + \xi, \tag{19}$$

where *f* is an unknown function that binds Y to X and ξ is a random component. The essence of statistical learning is to guess the actual function *f* by means of function *h*. It is one of the hypotheses concerning this unknown function *f*, which belongs to the hypothesis space *H* where *f* is located (Hastie et al., 2009).

There are two main reasons why we try to guess function *f*. The first one is of a practical nature and involves the prediction of Y based on the knowledge of X. The second one is of a cognitive nature and involves the inference of Y based on X. The task of prediction arises when a set of predictors X is available, but the corresponding values of response variable Y are not known. We then make a prediction of Y using the equation:

$$\hat{Y} = \hat{f}(X), \tag{20}$$

where $\hat{f}$ is one of the functions within the hypothetical space *H*. In this context, $\hat{f}$ is treated as a black box, where uncovering a specific character $\hat{f}$ is not the main objective, but rather achieving an as accurate as possible prediction of Y.

The main idea of statistical learning is to know the actual function *f*. We estimate it on the basis of a dataset, called a training dataset, which contains information about input and output data (x_{ij},y_i). In other words, we are looking for a function $\hat{f}$, for which Y ≈ $\hat{f}$(X) for any pair of observations from the (X,Y) set. For this purpose, we can use a parametric or non-parametric approach. Non-parametric statistical learning methods do not make clear assumptions about the analytical form of the functions concerning *f*. Instead, they seek a form of function *f* that fits the data from the training set as best as possible. The non-parametric approach may have a great advantage over the parametric approach. because by avoiding the adoption of a specific analytical form of function *f*, it can fit the empirical data more accurately. However, the non-parametric approach has the disadvantage that it does not reduce the number of the estimated parameters to only the relevant ones; thus, it requires a much larger teaching set (James et al., 2013).

Selecting the best $\hat{f}$ function that belongs to hypothetical space H should be based on a specific criterion of the matching quality of function $\hat{f}$ with the actual function *f*. From among the many methods of measuring the effectiveness of statistical learning at a specific $\hat{f}$, the mean squared estimation error (MSE) is the most commonly used:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{f}(x_i))^2, \tag{21}$$

where:

$\hat{f}(x_i)$ is the prediction of the actual *f* for the *i*-th observation.

If the estimation error is calculated for the data from the training set, then it is a measure of goodness of the fit for function f to empirical data y_i.

One of the simplest and most well-known non-parametric methods is the *k*-nearest neighbours method, which can be used for both classification and regression purposes. The *k*-nearest neighbours (KNN) regression for the *k*-considered number nearest neighbours and for the x_0 from the test set, first identifies *k* observations from the training set, the nearest x_0, denoted by n_0, and then it estimates $\hat{f}(x_0)$ using the average of all the values of y_i from the training set, belonging to n_0, i.e.:

$$\hat{f}(x_0) = \frac{1}{k}\sum_{x_i \in n_0} y_i, \tag{22}$$

where:

$\hat{f}(x_0)$ is the *k*-nearest neighbour estimator of the non-parametric regression function.

More precise non-parametric regression methods allow modelling non-linear relationships between data by using kernel estimators to transform data in empirical distribution classes. A kernel density estimator is designed to determine the distribution density of a random variable based on the obtained sample or training set. It is given by the formula:

$$\hat{f}_K(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(t_i) \tag{23}$$

where:

n – the sample size,

K_h – function of the kernel at smoothing parameter h,

t_i – standardised values of random variable X.

The Gaussian function is most often chosen as the kernel function; other functions are: the Epanecznikov kernel, uniform kernel, two-weight kernel, triangular kernel, radial kernel or product kernel. It is also possible to use a linear combination of kernels (Kulczycki, 2005).

A specific class of models in the decision-making theory are decision trees, which can also be used for regression or classification purposes. A regression tree is used to show hierarchically the regression of a continuous explanatory variable relative to its explanatory variables. On the other hand, if the explanatory variable is measured on a nominal or ordinal scale, then a classification tree is employed. More precise methods have been developed on the basis of decision trees such as bagging, boosting methods or random forests. Each of them involves creating multiple trees which are then combined to obtain a more precise prediction of the explanatory variable.

Non-hierarchical methods can also be used to study non-parametric regression. Among them, the Support Vectors Machines (SVM) method is widely used, both for classification and regression purposes. In the case of regression, it is called a support vector regression (SVR) (Vapnik, 2000). The key concept of SVR is the use of support vectors, i.e. hyperplane observations in multidimensional space. The selection of the best hyperplane is based on the method of an adaptive enrichment of the observation space and the search for a discriminatory hyperplane in such new spaces. Vapnik's idea is to solve an optimisation problem with a quadratic target function and linear constraints (square-linear optimisation) in Hilbert space (Koronacki & Ćwik, 2005).

The most popular non-parametric models are neural networks. These are computational models inspired by the functioning of the human brain, made up of connected artificial neurons arranged in layers that process information and learn to recognise patterns from data. The application of networks to solve regression problems is common. The user of the neural model expects that the created network will be able to describe the relationships between variables, while assuming that the explained variable is continuous. Proponents of neural networks as a tool for conducting economic research point to their advantage over classical regression models, discrimination models and classical trend models. However, whether a neural network function provides the correct solution to the posed problem depends on two basic factors:

- the values of the weight coefficients of the neurons that make up the network;
- the structure (topology) of the network, which is determined by the number of layers, the number of neurons in the individual layers, the way neurons are connected and the adopted model of the neuron (the aggregation of input data, type of the activation function used) (Lula, 1999).

Many interesting books on non-parametric modelling have been published in Polish. These include works by (Bruzda (2007), Gatnar (2001) or a monograph combining problems related to non-linearity with a non-parametric approach by Orzeszko (2016). A representative example of many applications of non-parametric regression methods in socio-economic research can be found in the article by Trzęsiok (2013).

The world literature on non-parametric methods is extensive. A seminal study presenting the traditional approach to non-parametric methods is by Hollander et al. (2014). A comprehensive overview of non-parametric and semi-parametric models can be found in Härdle et al. (2004). Non-parametric modelling is also the subject of specialised scientific journals, among which the Journal of Non-parametric Statistics is well-known. It focuses on non-parametric statistics and related areas, including modelling, estimation, analysis and testing of statistical methods and algorithms. Regression and classification methods based on decision trees are central to the statistical learning paradigm. The basic monographs dedicated to these issues are by Hastie et al. (2009) and James et al. (2013).

The rapidly developing statistical methods include new branches of machine learning, such as deep learning and causal machine learning, which many recently published monographs are devoted to. One of the most interesting works in the field of deep learning is by Goodfellow et al. (2016), while in casual machine learning by Cunningham (2021).

7. Conclusions

The conducted extensive literature review along with the author's 50 years of research experience in the field of statistical analyses of socio-economic processes allow the formulation of the following remarks on the current challenges facing statistics:

- Statistical methods are tools for learning the truth about the reality that surrounds us. However, the truth discovered by statistical means must be understood in probabilistic terms. Thus, truth is a judgment, with an acceptable probability (error), consistent with reality;
- The surrounding reality is manifested through phenomena that are a source of impressions. We can observe these phenomena with greater or lesser accuracy and thus, only approximate the studied reality. Socio-economic processes are stochastic in nature, so the most general statistical method of searching for the truth about the reality is the stochastic approximation method;
- Throughout the history of human thought, the initial philosophical position was determinism, which assumed that all events in nature are predetermined. However, Werner Heisenberg demonstrated that in a world of the smallest material particles, causal conditioning does not exist, pointing to the 'free will of the electron'. The shift from physics to philosophy, proved the existence of indeterminism in nature. Thus, the deterministic approach, which was the original philosophical position regarding the understanding of the world, has been replaced by its indeterministic counterpart, which is the dominant modern scientific view. In biology, the theory of evolution is based on indeterministic behaviour. Social indeterminism is a philosophical and sociological view that denies the existence of strict, necessary laws governing social phenomena. It is based on the belief that events in society are not strictly determined by the conditions that precede them, emphasising the role of chance, free will and other factors in shaping human behaviour. Indeterminism in economics is based on the assumption that economic events are not strictly determined by previous causes and their course is influenced by many unpredictable factors, including the free will to make decisions by market participants. This means that even under identical initial conditions, different business results may be achieved;
- The challenge for statisticians is to balance uncertainty and the related risk measurement with the discovery of the course of past processes and forecasting their course in the future. From a mathematical point of view, it involves optimising the stochastic approximation of the studied socio-economic reality;
- Learning about socio-economic processes can only be a better or worse way to get closer to the socio-economic reality that surrounds us. Therefore, it is necessary to formulate general criteria for this approximation to be satisfactory. Such criteria are based on probabilistic foundations. Therefore, approximation remains the most general method of the cognition of reality, as human cognition is inherently its simplification;

- Soon after the emergence of stochastic approximation in the field of regression research in the world literature, it was introduced into the Polish academia by Zdzisław Hellwig in his book on stochastic approximation (Hellwig, 1965). The author deals with multidimensional random variables, particularly some aspects of correlation and the regression theory, which are especially important in economic applications. The main thesis is that the regression function is nothing more than an approximation function; therefore, the basic theorems and some computational methods provided by the theory of approximation can be used in the theory of regression and correlation;
- The problem of stochastic approximation is expressed by the following relation: $P[\ \mid Y - \psi\ (x)\ \mid < \Delta \geq 1 - \alpha$, where $\psi\ (x)$ is the approximant of an unknown function combining the components of a two-dimensional variable (X,Y), Δ is the tolerance or acceptable error of approximation, while 1− α is the credibility of the approximation. Statistical regression is one of the forms of the approximation of real relations occurring in socio-economic processes. When approximating this reality, it is necessary to determine tolerance Δ that we accept, i.e. the acceptable discrepancy between the actual relationship between the variables we define and its approximant we choose;
- The approximation approach to regression indicates that the approximation function chosen on the basis of the sample is only one of many possible regression functions belonging to the set of approximation functions. What is more, the individual functions belonging to this set are indistinguishable in terms of the accepted level of tolerance. In econometric modelling, stating that a given model 'wins' simply because others are within an acceptable tolerance range is unjustified;
- Socio-economic processes do not develop in a continuous and linear manner, but tend to be disrupted and shaped by sudden changes in the economic and political situation. This is due not only to macroeconomic factors, but also to people's decisions and emotional behaviours, which are not always rational. The analysis of such processes cannot rely solely on simple, linear models, but their complexity, dynamics and unpredictability need to be considered as well. The linearity of the course of socio-economic processes is not a rule, but a special case of their course;
- Process non-linearity is revealed in time series analysis through stochastic modelling. According to Clive Granger, the world is almost certainly non-linear. Despite significant advancements in financial economics and econometrics associated with the modelling of

non-linear socio-economic processes, improving the tools for detecting and describing the non-linearity of the studied processes remains a challenge for statisticians;

- Statistical description and parametric modelling of non-linear economic processes are computationally complex and require several restrictive assumptions. The non-parametric approach is a viable alternative; it does not require the assumption of a specific analytical form of the distribution in the population; instead, it relies on the shape of the data distribution in the sample or training set. It allows the structure of the model to be determined based on the data. Non-parametric modelling is thus an important segment of statistical methods, especially useful when the distribution of the data is non-normal or unknown, and when the studied characteristics are measured on weak, ordinal or nominal scales;
- Non-parametric methods, particularly non-parametric regression, represent a significant step towards the development of multivariate statistical methods based on artificial intelligence. Despite their increased flexibility, these methods are not free from limitations that affect model adequacy and interpretability, including their sensitivity to the quality of the data in the training sets. Measurement errors, data gaps or outliers can significantly distort the results of the model. The lack of restrictive assumptions can only be apparent, as they can be transferred to a lower, 'technical' level. Moreover, as the complexity of statistical learning models grows, problems with their substantive interpretation intensify;

In conclusion, the author maintains that the rapid development of statistical learning methods should be based on a philosophical foundation, containing axioms concerning the stochastic, non-linear and non-parametric nature of socio-economic processes, effectively modelled by stochastic approximation methods.

References

Bogdanienko, J. (2018). *Istota i problemy poznania naukowego*. CeDeWu.

Briggs, A., Łuczewski, M., & Puczkarski, P. (2025). Kwant bożego królestwa. *Tygodnik Powszechny*, (32), 45–47.

Briggs, A., & Reiss, M. J. (2021). *Human flourishing. Scientific insight and spiritual wisdom in uncertain times*. Oxford University Press.

Bruzda, J. (2007). *Procesy nieliniowe i zależności długookresowe w ekonomii. Analiza kointegracji nieliniowej*. Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Burnewicz, J. (2021). *Filozofia i metodologia nauk ekonomicznych*. Wydawnictwo Naukowe PWN.

Cunningham, S. (2021). *Causal Inference. The Mixtape.* Yale University Press.

Czerwiński, Z. (2002). *Moje zmagania z ekonomią*. Wydawnictwo Akademii Ekonomicznej w Poznaniu.

Fan, J., & Yao, Q. (2005). *Nonlinear Time Series. Nonparametric and Parametric Methods*. Springer.

Fogel, A., Lyra, M. C. D. P., & Valsiner, J. (Eds.). (2014). *Dynamics and Indeterminism in Developmental and Social Processes*. Psychology Press.

Fortuna, Z., Macukow, B., & Wąsowski J. (1993). *Metody numeryczne*. Wydawnictwa Naukowo-Techniczne.

Gardini, L., Lamantia, F., Szidarovszky, F., & Tramontana, F. (2021). Nonlinear dynamics in economic modeling. *Decisions in Economics and Finance*, *44*, 485–487. https://doi.org/10.1007/s10203-021-00353-8.

Gatnar, E. (2001). *Nieparametryczna metoda dyskryminacji i regresji*. Wydawnictwo Naukowe PWN.

Gatnar, E. (2008). *Podejście wielomodelowe w zagadnieniach dyskryminacji i regresji*. Wydawnictwo Naukowe PWN.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Granger, C. W. J. (1981). Some Properties of Time Series Data and Their Use in Econometric Model Specification. *Journal of Econometrics*, *16*(1), 121–130. https://doi.org/10.1016/0304-4076(81)90079-8.

Granger, C. W. J. (1989). *Forecasting in Business and Economics*. Accademic Press. https://doi.org/10.1016/C2013-0-10755-6.

Granger, C. W. J., & Hallman, J. (1991). Long Memory Series with Attractors. *Oxford Bulletin of Economics and Statistics*, *53*(1), 11–26. https://doi.org/10.1111/j.1468-0084.1991.mp53001002.x.

Grobler, A. (2006). Metodologia *nauk*. Aureus, Znak.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning. Data Mining, Inference, and Prediction* (2nd ed.). Springer.

Härdle, W., Müller, M., Sperlich, S., & Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Springer. https://doi.org/10.1007/978-3-642-17146-8.

Hellwig, Z. (1965). *Aproksymacja stochastyczna*. Państwowe Wydawnictwo Ekonomiczne.

Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Nonparmetric Statistical Methods*. Wiley. https://doi.org/10.1002/9781119196037.

Jadacki, J. (2020). Zagadka przyczynowości. *Przegląd Filozoficzny – Nowa Seria*, *29*(2), 17–51. https://doi.org/10.24425/pfns.2020.133136.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Application in R.* Springer.

Jasiński, L. J. (2017). *Analiza i interpretacja badań ekonomicznych*. Oficyna Wydawnicza Politechniki Warszawskiej.

Kamiński, S. (1981). *Pojęcie nauki i klasyfikacja nauk*. Towarzystwo Naukowe Katolickiego Uniwersytetu Lubelskiego.

Katsenelinboigen, A. (1992). *Indeterministic Economics*. Bloomsbury Publishing.

Katsenelinboigen, A. (1997). *The Concept of Indeterminism and its Applications. Economics, Social Systems, Ethics, Artificial Intelligence, and Aesthetics*. Bloomsbury Publishing.

Kiefer, J., & Wolfowitz, J. (1952). Stochastic Estimation of the Maximum of a Regression Function. *The Annals of Mathematical Statistics*, *23*(3), 462–466. https://doi.org/10.1214/aoms/1177729392.

Klein, L. R. (1982). *Wykłady z ekonometrii*. Państwowe Wydawnictwo Ekonomiczne.

Kordecki, W., & Selwat, K. (2020). *Metody numeryczne dla informatyków*. Helion.

Koronacki, J. (1989). *Aproksymacja stochastyczna. Metody optymalizacji w warunkach losowych*. Wydawnictwa Naukowo-Techniczne.

Koronacki, J., & Ćwik, J. (2005). *Statystyczne systemy uczące się*. Akademicka Oficyna Wydawnicza EXIT.

Kuciński, K. (Ed.). (2010). *Metodologia nauk ekonomicznych. Dylematy i wyzwania*. Difin.

Kulczycki, P. (2005). *Estymatory jądrowe w analizie systemowej*. Wydawnictwa Naukowo-Techniczne.

Lemańska, A. (1998). *Filozofia przyrody a nauki przyrodnicze*. Akademia Teologii Katolickiej.

Lula, P. (1999). *Jednokierunkowe sieci neuronowe w modelowaniu zjawisk ekonomicznych*. Wydawnictwo Akademii Ekonomicznej w Krakowie.

Maitra, S. (2025). *Non-Linearity in Econometric Modelling: Vol. 1. A Practical Approach*. Springer, Cham. https://doi.org/10.1007/978-3-032-06462-2.

von Mises, R. (1957). *Probability, Statistics and Truth*. John Allen & Unwin.

Nowakowski, A. (2017). Przyczynowa wizja świata. *Przegląd Filozoficzny – Nowa Seria*, *26*(1), 53–68. https://journals.pan.pl/Content/100487/PDF/P.Filoz.%201-17%203-A.Nowakowski.pdf?handler=pdf.

Orlando, G., Pisarchik, A. N., & Stoop, R. (Eds.). (2021). *Nonlinearities in Economics. An Interdisciplinary Approach to Economic Dynamics, Growth and Cycles*. Springer. https://doi.org/10.1007/978-3-030-70982-2.

Orzeszko, W. (2016). *Nieparametryczna identyfikacja nieliniowości w finansowych i ekonomicznych szeregach czasowych*. Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Pagan, A., & Ullah, A. (2009). *Nonparametric Econometrics*. Cambridge University Press. https://doi.org/10.1017/CBO9780511612503.

Pociecha, J. (2023). *Filozoficzne i metodologiczne podstawy współczesnych analiz statystycznych*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.

Rao, C. R. (1997). *Statistics and Truth. Putting Chance to Work*. World Scientific Publishing.

Robins, H., & Monro, S. (1951). A Stochastic Approximation Method. *The Annals of Mathematical Statistics*, *22*(3), 400–407. https://doi.org/10.1214/aoms/1177729586.

Sagan, A. (2016). *Metodologia badań ekonomicznych*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.

Sosenko, K. (2008). *Problemy filozofii i metodologii nauk dla ekonomistów*. Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.

Strawiński, W. (2011). Funkcja i cele nauki – zarys problematyki metodologicznej. *Zagadnienia Naukoznawstwa*, (3), 323–335. https://journals.pan.pl/dlibra/publication/108246/edition/93890/content.

Tatarkiewicz, W. (1978). *Historia filozofii* (Vol. 2). Wydawnictwo Naukowe PWN.

Trzęsiok, J. (2010). Wykorzystanie regresji nieparametrycznej do modelowania wielkości oszczędności gospodarstw domowych. *Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu*, (102), 99–108. https://www.ue.katowice.pl/fileadmin/_migrated/content_uploads/7_J.Trzesiok_Wykorzystanie_regresji_nieparametrycznej....pdf.

Vapnik, V. N. (2000). *The Nature of Statistical Learning Theory* (2nd ed.). Springer-Verlag. https://doi.org/10.1007/978-1-4757-3264-1.

Wawrzyński, P. (2021). *Uczące się systemy decyzyjne*. Oficyna Wydawnicza Politechniki Warszawskiej.