Biography of Andrey Kolmogorov

Anastasiya Chuchkova

History of Mathematics

Stoyan Dimitrov

April 29th 2023

Intro (Summary)

Andrey Kolmogorov was one of the greatest mathematicians of the 20th century. He created over 500 publications and papers in his lifetime in topics including probability theory, turbulence, topology, set theory, classical mechanics, number theory, algorithmic information theory, and computational complexity. During his lifetime, he was greatly socially recognized and received numerous prizes, including the Wolf Prize (1980), Stalin Prize (1941), the Lenin Prize (1965), and a Lobachevsky Prize (1987) [1]. He was also one of the few Soviet mathematicians to be a part of a number of foreign academies and scientific societies like the USA National Society, Paris Academy of Sciences, German Academy of Sciences Leopoldina, and more [1]. This paper aims to give an overview of his early life, adult life, late adult life and discuss some of his mathematical discoveries, specifically in the field of Probability Theory.

Early life

Andrey Kolmogorov was born on April 25, 1903, in Tambov, Russia. However, he had no connection to the place and was born there by pure chance. His mother, Mariya Yakovlevna Kolmogorova, was "on a journey from the Crimea back to her home in Tunoshna near Yaroslavl" [2]. While stopping at a town on the way home, she started having contractions and gave birth to him. Sadly, Mariya died during childbirth. His father did not participate in his life, and little is known about him outside of him being an agriculturist and dying in a war in 1919 [2]. Kolmogorov ended up being raised by his aunt, Vera Yakovlevna Kolmogorova, who took him in from the moment his mother died. She is quoted to be "an independent woman who held high social ideals. She passed this over to her nephew, raising him in the sense of responsibility, independence of opinion, intolerance towards idleness and poorly performed tasks, and the desire to understand and not just to memorize." [1]. He thought of her his entire life as his mother and took care of her until she died in 1950 at the age of 87.

Andrey ended up being raised by Vera in Tunoshna, his mother's hometown. There he lived in his grandfather Yakov Stepanovich Kolmogorov's family estate. Yakov was a nobleman and wealthy landowner, which immensely helped Andrey have many opportunities in his early life, but also served as a burden, especially as the revolution neared.

During Andrey Kolmogorov's early life at his grandfather's estate, he was surrounded by love and care. From a very young age, his family's goal in regard to raising Andrey was to "cultivate [his] curiosity and an interest in books and nature." [3]. He attended a small school made by Vera and her friend, where he first developed his interest in mathematics. In the school, students, with the help of teachers, published a journal called "The Swallow of Spring" [3]. Andrey was the 'editor' of the mathematical section of the journal. Later in life, he remarked that his first "scientific publication" was in that journal. At the age of 5, he noticed a regularity in the sum of a series of odd numbers; $1 = 1^2$, $1 + 3 = 2^2$, $1 + 3 + 5 = 3^2$, $1 + 3 + 5 + 7 = 4^2$ and etc. [3]. He then published an article on his "discovery" in his school journal.

As Kolmogorov grew, his education and abilities grew exponentially. At the age of 7, in 1910, he and Vera moved to Moscow for him to attend one of the most progressive and advanced grammar schools ('gymnasiums') at the time [3]. This school was one of the only few co-educational schools of the time and had a lot of new experimental teaching methods that nurtured and gave special tailored attention to every child. Attending a school of that nature made Kolmogorov's range of interests extremely broad but comprehensive at the same time. In his early teen years, he was particularly interested in physics, chess, math, history, and biology. For example, at age 14, he already started studying higher mathematics [3].

However, his lifelong career as a mathematician was only confirmed much later in life, as he continued to explore a variety of subjects until the middle of his undergraduate degree.

At 17, he published his first proper scientific report (with huge improvements from the one he published earlier at five). However, it got criticized by a leading historian and professor at the time, S.V. Bakhurshin, who told him not to publish it. Bakhurshin told him that he only found one proof, which is very little for a historian, and that he needed at least five [3]. Kolmogorov was disheartened at the time and later told this story as an anecdotal joke saying that Bakhurshin's comment might have influenced him to pursue mathematics instead of history as a career path since, in mathematics, you only need a singular proof.

Adult life-

As much as his childhood life was rich, easy, and even luxurious, he faced a lot of struggles in his late high school years right before entering university. Life in the post-Russian revolution of 1917 was very difficult for everyone, especially for the ex-nobleman/ex-aristocracy who were not used to living in harsh conditions. Because of this, Kolmogorov worked as a construction worker for the "Kazan-Ekaterinburg railway line whilst continuing to study on his own to complete high school exams" [4]. Even when he started attending the State University of Moscow in 1920, the first few years were harsh, with "lecture rooms cold and unheated in the winter of 1920/1921" [1]. At the time, students received grants and even some food that they had to ration for the entire year, but even that was very little and essentially the bare minimum to survive.

When enrolling at the University of Moscow, Kolmogorov also enrolled in Mendeleev Moscow Institute of Chemistry and Technology to study practical engineering as he was still unsure exactly what he wanted to dedicate his life to [3]. However, after a few courses, his interest in mathematics outweighed his doubts about the practicality of mathematics, and he dropped out of the engineering program to focus his whole attention on mathematics at the University of Moscow.

Between 1921-1922, only as an undergraduate, Kolmogorov independently researched and produced results that ended up being of international importance [2]. In June of 1922, he published a paper on the Fourier series that diverges almost everywhere [4]. His name started being known worldwide for the results contained in the paper as well as his young age and education level. In 1925 he finished his undergraduate degree in mathematics and started his graduate studies under Nikolai Luzin. Luzin was a well-known professor and researcher at Moscow University and is best known for his contributions to set theory analysis. He was so well known and respected that there was a popular nickname at the time for his equally well-known research group, Luzitania. Mathematics in Moscow in the 1920s is even classified by this term, Young Luzitania(1920-1923) and post-Luzitania(1923-1927) [1].

During his time at graduate school, Kolmogorov continued to be interested in a variety of topics in mathematics and ended up publishing 18 different papers and articles in various journals before receiving his Ph.D. One of these papers got published in Mathematicheskii Sbornik on the principle of the excluded middle in which "...he proved that [the] application of the law of the excluded middle in itself cannot lead to a contradiction" [1]. This was a key paper as this was the first systematic research in the world with regard to the field of intuitionistic logic, and he was the first person from the Soviet Union/Russia to publish a paper on mathematical logic that contained very significant results. This paper would then be used as a starting point for numerous other discoveries and works in mathematical logic [1].

Probability Theory

However, the results that are of most interest to me and that Kolmogorov is most widely known for are his works in Probability Theory. Kolmogorov first got interested in Probability Theory in 1924 [1]. However, he did not make his first publication in it until 1928.

In his first publication, he found the conditions necessary for the strong law of large numbers to hold, something that best mathematicians like Pafnuty Chebyshev and Andrey Markov (the older Markov brother) tried to figure out for decades prior [4]. In total, there are dozens of LLNs (Laws of Large Numbers). Generally, they can be categorized into two categories, weak laws and strong laws. The distinction between them is fairly simple," A LLN is called a Weak Law of Large Numbers (WLLN) if the sample mean converges in probability" [5]. On the other hand, a LLN is called a Strong Law of Large Numbers (SLLN) if the sample mean converges almost surely [5]. The reason for naming them 'weak' and 'strong' is because convergence in probability is oftentimes called weak convergence, and strong simply is the antonym of weak. Thus, to make the distinction, the LLN with sample mean that converges almost surely is called strong. Another reason is that SLLN comes with an assurance that something will happen vs. WLLN where there is an assurance that what we are after will happen with increasing probability [6]. Therefore SLLN, even intuitively, is 'stronger' than the weak law.

The publication in Probability Theory that set the groundwork for Kolmogorov's later, very successful and revolutionary publications in probability theory was his publication titled 'A General Theory of Measure and the Calculus of Probabilities,' which was published in 1929. This publication details his draft for an axiom system for probability theory [1]. In the next 4 years, Kolmogorov worked on creating his first book, 'Foundations of the Calculus of Probabilities' which was published in German and titled 'Grundbegriffe der Wahrscheinlichkeitsrechnung'. Even though probability had been studied for centuries before that, this book set the formal, rigorous groundwork and established probability as a key scientific field [4][1]. In this book, Kolmogorov introduced the 5 axioms of probability shown below [7].

§ 1. Axioms[2]

Let E be a collection of elements $\xi, \eta, \zeta, \ldots$, which we shall call *elementary events*, and $\mathfrak{F}$ a set of subsets of E; the elements of the set $\mathfrak{F}$ will be called *random events.*

I. *$\mathfrak{F}$ is a field*[3] *of sets.*

II. *$\mathfrak{F}$ contains the set E.*

III. *To each set A in $\mathfrak{F}$ is assigned a non-negative real number $\mathsf{P}(A)$. This number $\mathsf{P}(A)$ is called the probability of the event A.*

IV. $\mathsf{P}(E)$ *equals* 1.

V. *If A and B have no element in common, then*

$$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$$

At first glance, this might appear quite strange, as every student who has studied statistics at any level knows that there are 3 axioms, not 5. But in reality, it's not strange at all, Axioms III-V are the 3 axioms we are all very familiar with, while axioms I-II establish the 'battle grounds' on which the study of probability will take place. Moreover, the reason why in modern teachings of probability theory, we only mention Axioms III-V is that they imply I and II. For example, given Axiom V, Axiom I must be true by default. That is because when we have A+B, we want P(A+B) to exist, and if there is no field, there might be some A and B for which P(A+B) does not exist [8]. This implies that there must be a field of sets. Moreover, given Axiom IV, P(E)=1, E must be in a field, which implies Axiom II [8]. As such, the 3 condensed theorems of probability are as follows.

Axiom 1 $P(A) \geq 0$, for all events $A \subseteq \Omega$.

Axiom 2 $P(\Omega) = 1$.

Axiom 3 If events A and B satisfy $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.

Axiom 1 simply states that for every event A, probability must be positive. This is sometimes called the axiom of nonnegativity. Axiom 2 states that the probability of an entire sample space, Ω, is 1, aka 100%. That is because sample space contains all possible outcomes of a random experiment. Thus the total probability must be 1, 100%. It is sometimes called the axiom of normalization. And the final Axiom, Axiom 3, basically states if A and B are disjoint events, then the probability of their union must be the summation of their individual probabilities.

This axiom extends past 2 events, A and B, and could be applied to 3 disjoint events, 4 disjoint events, 5 disjoint events, etc. So axiom 3 is often instead written as: If events $A_1, A_2, \ldots A_i$ are mutually disjoint then, $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$. This is sometimes called the axiom of finite additivity.

Like mentioned before, these axioms serve as the foundation for all of probability theory and can be used for various other theorems and laws in probability. To see that in action, I will show how to find the probability of a complement.

Derivation of the probability of a complement (Complement rule): $P(A^c) = 1 - P(A)$

By definition of a complement:

$$S = A \bigcup A^c \Rightarrow P(S) = P(A \bigcup A^c)$$

By the second axiom:

$$1 = P(S) = P(A \bigcup A^c)$$

Then, because A and A^c are disjoint, by the third axiom:

$$P(S) = P(A) + P(A^c)$$

Putting it all together:

$$1 = P(S) = P(A) + P(A^c)$$

Rearranging it we get:

$$P(A^c) = 1 - P(A)$$

Now we can use both the 3 Axioms and our complement rule to derive the other rules, like the probability of an empty set:

Derivation of the probability of an empty set: $P(\emptyset) = 0$
We know that the empty set, $\emptyset$, is the complement of the sample space, S:

$$\emptyset = S^c \Rightarrow P(\emptyset) = P(S^c)$$

Using the complement rule we derived above we know that:

$$P(\emptyset) = P(S^c) = 1 - P(S)$$

By the second axiom:

$$P(S) = 1$$
$$P(\emptyset) = P(S^c) = 1 - P(S) = 1 - 1$$
$$P(\emptyset) = P(S^c) = 1 - P(S) = 1 - 1 = 0$$
$$P(\emptyset) = 0$$

We can continue this process to derive all other theorems and rules with the help of the axioms and the rules we derived previously.

Late Adult Life-

Going back to the biography of Andrey Kolmogorov, there are numerous other accomplishments, theorems, laws, and general contributions in Mathematics and adjacent fields that he had made throughout his life. What was impressive about him is not just the breadth of his work, as he had created over 500 different books and articles between 1923 and 1977 [9]. But also the depth at which he studied every single subject.

Kolmogorov was not only interested in mathematics, though, and in his later years, he entered into the theory of pedagogy. He ended up being enormously influential in that field as well, writing textbooks, taking numerous teaching positions at Moscow State University, and being a member of the U.S.S.R. Academy of Pedagogical Sciences [9]. His teaching methods and style were very different from most other professors/teachers at the time, reminiscent of the way he himself was taught in his early life in a non-traditional school. He never explained anything, he wanted the students to arrive at the solutions through their own efforts and then come to him to discuss them [1]. He also stood out due to his respect for all of the students he had [1]. Needless to say, he was very loved by his students.

Sadly, he did not give the same love and respect to his Ph.D. mentor Nikolai Luzin. During the Great Purge in U.S.S.R., many high-ranking officers, business owners, artists, professors, and researchers deemed to be a ‘threat’ to U.S.S.R. and specifically Stalin, were, shot, exiled, sent to Gulag, and threatened. This included Luzin, who was accused of nepotism, plagiarism, and a fascist-type way of teaching in 1936 [10]. Many of Luzin’s previous and current students went on to testify against him, which included Kolmogorov. In the end, even though Luzin was convicted of the said crimes, his punishment was much milder than most other punishments during the Great Purge. Luzin was “neither expelled from the Academy nor arrested” [10], but instead, he simply lost all of his academic positions. There have been great debates about whether Kolmogorov, along with his peers, were forced or coerced into going against their mentor or if it was of their own free will.

In his late years, Kolmogorov developed Parkinson’s disease and went nearly blind. However, he continued to work at Moscow State University and be active in mathematics research until the day he died [11]. Andrey Kolmogorov died on October 20, 1987, in Moscow, Russia.

Bibliography:

[1] http://www.scholarpedia.org/article/Andrey_Nikolaevich_Kolmogorov
[2] https://mathshistory.st-andrews.ac.uk/Biographies/Kolmogorov/
[3]Chern, S. S., & Hirzebruch, F. (2001). *Wolf Prize in Mathematics*. World Scientific Publishing.
[4]https://valeman.medium.com/andrey-kolmogorov-one-of-the-greatest-mathematicians-of-the-xxst-century-4167ad02d10
[5] https://www.statlect.com/asymptotic-theory/law-of-large-numbers
[6]https://math.stackexchange.com/questions/2024255/what-is-the-difference-between-the-weak-and-strong-law-of-large-numbers
[7]Kolmogorov, A. (1956). *Foundations of the Theory of Probability*. CHELSEA PUBLISHING COMPANY NEW YOURK.
[8]Email conversation with Professor Dimitrov.
[9]https://www.britannica.com/biography/Andrey-Nikolayevich-Kolmogorov
[10]http://theor.jinr.ru/~kuzemsky/Luzinbio.html
[11]https://en.kuzminykh.org/2312-andrey-kolmogorov.html