NATIONAL ACADEMY OF SCIENCES

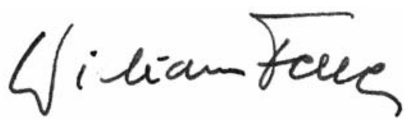
WILLIAM FELLER
1906—1970

A Biographical Memoir by
MURRAY ROSENBLATT

Any opinions expressed in this memoir are those of the author and do not necessarily reflect the views of the National Academy of Sciences.

Biographical Memoir

Photograph by Orren Jack Turner, Courtesy Princeton University Library.



WILLIAM FELLER

July 7, 1906–January 14, 1970

BY MURRAY ROSENBLATT

William Feller was one of the major figures in the development of interest and research in probability theory in the United States as well as internationally. He was born in Zagreb, Yugoslavia, on July 7, 1906, the son of Eugen Viktor Feller, a prosperous owner of a chemical factory, and Ida Perc. Feller was the youngest of eight brothers, one of twelve siblings. He was a student at the University of Zagreb (1923-1925) and received the equivalent of a master of science degree there. Feller then entered the University of Göttingen in 1925 and completed his doctorate with a thesis titled "Uber algebraisch rektifizierbare transzendente Kurven." His thesis advisor was Richard Courant. He left Göttingen in 1928 and took up a position as *privat dozent* at the University of Kiel in 1928. Feller left in 1933 after refusing to sign a Nazi oath. He spent a year in Copenhagen and then five years (1934-1939) in contact with Harald Cramér and Marcel Riesz in Sweden. On July 27, 1938, he married Clara Nielsen, a student of his in Kiel.

At the beginning of the 20th century the most incisive research in probability theory had been carried out in France and Russia. There was still a lack of effective basis for a mathematical theory of probability. There was a notion of a collective introduced by von Mises, defined as a sequence of observations with certain desirable asymptotic properties. An effective formalization of probability theory was given in a 1933 monograph of Kolmogorov that was based on measure theory.[1] An attempt to get a rigorous format for the von Mises collective led in the 1960s to an approach using algorithmic information theory by Kolmogorov and others. Feller's research made use of the measure theoretic foundations of probability theory as did most mathematical work.

Feller's first published paper in probability theory (1936) obtained necessary and sufficient conditions for the central limit theorem of probability theory. This was a culmination of earlier work of DeMoivre, Laplace, and Liapounov. The concern is with the asymptotic behavior of the partial sums $S_n=X_1+\ldots+X_n$ of a sequence of independent random variables $X_1,X_2,\ldots,X_n,\ldots$. Assuming that means $m_k=EX_k$ and variances $B_n^2=E(S_n^2)-(ES_n)^2$ exist, Feller showed that the normalized and centered sums $\sum_{j=1}^{n}(X_j-m_j)/B_n$ (with the summands $(X_j-m_j)/B_n$ uniformly asymptotically negligible) in distribution converge to the normal law

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du$$

as $n\to\infty$ if and only if the Lindeberg condition

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x|>\varepsilon B_n}x^2dF_k(x+m_k)\to 0$$

(F_k is the distribution function of X_k) as $n\to\infty$ is satisfied for each $\varepsilon>0$.

The actual fluctuating behavior of the sequence S_n was first properly formulated by Borel in 1909. After initial advances by Hausdorff, Hardy and Littlewood, and Khinchin, Kolmogorov obtained a law of the iterated logarithm, which states that under some strong boundedness conditions on the X_k that with probability one

$$\limsup_{n\to\infty} \frac{S_n - E(S_n)}{\sqrt{2B_n^2 \log\log B_n^2}} = 1$$

In a number of papers Feller through much research during his life extended and improved the law of the iterated logarithm.

A reviewing journal *Zeutralblatt für Mathematik* had been set up in 1931. The editor was Otto Neugebauer whose interests were in the history of mathematics and astronomy. Neugebauer resigned after the Nazi government's racist restrictions on reviewers were implemented. The American Mathematical Society with outside monetary support established a mathematics reviewing journal, *Mathematical Reviews*, with Neugebauer, Tamarkin, and Feller as effective editorial staff. The mathematical community is indebted to Feller for his help in setting up what became the primary mathematical reviewing journal in the world. Willy and Clara Feller moved to the United States in 1939 and Feller became an associate professor at Brown University.

In 1931 a paper of Kolmogorov on analytic methods in probability theory discussed the differential equations satisfied by the transition probabilities of continuous time parameter Markov processes (random processes with the property that past and future of the process are conditionally independent given precise knowledge of the present).[2] The conditional probability that a system at time t at location x will at time $\tau > t$ be less than or equal to y is given by a function $F(t,x,\tau,y)$. It was known that in the case of a diffusion the function F would satisfy a Fokker-Planck or diffusion equation

$$\frac{\partial F}{\partial t} + a(t,x)\frac{\partial^2 F}{\partial x^2} + b(t,x)\frac{\partial F}{\partial t} = 0$$

in the backward variables t,x, where $a(t,x)$ and $b(t,x)$ represent the local fluctuation and local drift respectively at x at time t. Feller, under appropriate growth and smoothness conditions on $a(t,x)$ and $b(t,x)$ constructed a conditional distribution function F as a solution of the diffusion equation and demonstrated its uniqueness in (1937,1). Under additional conditions on the coefficients $a(t,x)$, $b(t,x)$ the conditional distribution function $F(t,x;\tau,y)$ is shown to have a density $f(t,x;\tau,y)=\frac{\partial}{\partial y}F(t,x;\tau,y)$ in y and the density satisfies the formal adjoint differential equation in τ,y.

$$-\frac{\partial f}{\partial \tau}-\frac{\partial^2}{\partial y^2}[a(\tau,y)f]+\frac{\partial}{\partial y}[b(\tau,y)f]=0$$

The case of purely discontinuous or jump Markov processes in continuous time was also examined and under appropriate conditions an integrodifferential equation is shown to be satisfied by the constructed transition probability function of the Markov process. The more complicated situation in which one has continuous excursions as well as jumps was also considered. These results were a strong extension of those of Kolmogorov.

The paper (1939) is an example of Feller's continued interest in applications, here in a biological context. He looks at the Lotka-Volterra equations, which are a deterministic predator-prey model as well as a deterministic model of population growth (and decrease). In setting up corresponding Markovian models it is noted that the mean response in the stochastic models may not agree precisely with the original deterministic models. A diffusion equation is shown to arise naturally if population is considered a continuous variable but an equation that is singular in having the coefficient $a(x)$ zero at a boundary point.

An amusing and insightful paper (1940,2) of Feller's related to some presumed statistical evidence for the existence of extrasensory perception (ESP). Experiments were carried out involving sequences of fair coin tosses. Track was kept of the successful versus unsuccessful guesses of the coin tossing up to time n, S_n. It was noted that there were often large excursions of S_n above zero as well as large excursions below zero. A large excursion above zero was interpreted as presence of ESP and a large excursion below zero as its departure. Feller incisively notes that the standard model of fair coin tossing accounts for such excursions without the extra introduction of special effects like ESP.

Feller moved to Cornell University in 1945 as professor of mathematics and remained there until 1950, when he left to go to Princeton University. While at Cornell he wrote the paper (1948) in which he derived the Kolmogorov-Smirnov limit theorems by methods of a simpler character than those used to derive these results originally. The limit theorems provide procedures for effective one and two sample statistical tests. An elegant paper (1949,1) written jointly with Paul Erdös and Harry Pollard used elementary methods to establish limit behavior of transition probabilities for countable state discrete time Markov chains under appropriate conditions (a result obtained earlier by Kolmogorov using arguments that were more elaborate). Feller's paper (1949,2) on recurrent events also appeared and extended the basic idea of an argument that often was used in the analysis of Markov chains and that was perhaps suggested by early work of Wolfgang Doeblin.

The year 1950 is marked by the publication of the first volume of Feller's *Introduction to Probability Theory and Its Applications*. The book gives an extended discussion of the nature of probability theory. To avoid measure theory the development is limited to sample spaces that are finite or countable. Results on fluctuation in coin tossing and random walks are dealt with. There is the usual discussion of conditional probability and stochastic independence. The binomial and Poisson distributions are introduced together with the classical deMoivre normal approximation to the binomial distribution. Definitions of random variables and expectation are then given. A law of large numbers and a law of iterated logarithm (for partial sums of 0-1 random variables) are given as limit laws in the context of finite sample size with increasing sample size so as to avoid a formal discussion using a background of measure theory. Branching processes and compound distributions are introduced using generating functions as a convenient tool. Recurrent events (as introduced by Feller) and renewal theory then follow and are considered in the context of Geiger counters or the servicing of machines. Stationary transition function countable state Markov chains are introduced as examples of dependent sequences. The Chapman-Kolmogorov equation is noted as a consequence of the Markov property (but not equivalent to it). The ergodic properties of Markov processes are then developed. An algebraic treatment is given for finite state Markov chains. Finally birth and death processes are introduced and considered as examples of countable state continuous time parameter processes. The book is remarkable with its extensive collection of interesting problems and its discussion of applications.

Feller worked for eight years on the preparation of this volume. The volume was completed in the last year of Feller's tenure at Cornell University. The book is dedicated to Neugebauer. Gian-Carlo Rota remarked that the book was "one of the great events in mathematics of this century. Together with Weber's *Algebra* and Artin's *Geometric Algebra* this is the finest textbook in mathematics in this century. It is a delight to read and it will be immensely useful to scientists in all fields."[3]

The academic year 1949-1950 at Cornell University was a truly remarkable one in probability theory. The permanent and visiting faculty members were W. Feller, M. Kac, J. L. Doob, G. Elfving, G. Hunt, and K. L. Chung, and it was a most stimulating time for students.

I recall some impressions from my own days as a graduate student in the late 1940s at Cornell, where I took most of my courses in stochastics with Feller as a lecturer. Though I completed my doctorate with Mark Kac as adviser, I had an overwhelming impression of Will Feller as a man of supreme enthusiasm and occasional exaggeration that at times required some modification. An amusing example is given by a lecture where he introduced the three series theorem and turned to the student audience inquiring, “Isn’t it obvious?” Luckily we persuaded him to give a detailed presentation, and a series of two or more lectures on the theorem followed. He did give great insight in his lectures.

In 1950 he took a position at Princeton University as Eugene Higgins Professor of Mathematics. He held this position until his death on January 14, 1970, at 63 in the Memorial Hospital of New York.

The paper (1952,1) is an indication of Feller’s renewed interest in diffusion processes and their application in genetics. Stochastic processes as models in genetics and the theory of evolution are developed. Current methods at that time were due to R. A. Fisher and Sewall Wright for the most part. It is indicated how in a model of S. Wright the gene frequency $u(t,x)$ that satisfies a diffusion equation

$$u_t = \{\beta x(1-x)\,u\}_{xx} - \{[\gamma_2 - (\gamma_1+\gamma_2)\,x]\,u\}_x$$

is obtained by an appropriate limiting process from a bivariate discrete model. The equation has singular boundary points at 0 and 1. Here β, γ_1 and γ_2 are constants with the γ's denoting mutation rates.

In the 1950s Feller carried out his well-known research on one dimensional diffusion processes with stationary transition function $F(t,x;\tau,y)=P(\tau-t;x,y)$. He made use of appropriate modifications of the Hille-Yosida theory of semigroups. The stationary transition function $P(t)= P(t;x,y)$ generates a semigroup

$$P(t+s;x,y)= \int P(t;x,dz)P(s;z,y),\ \ t,s\geq 0$$

because of the Chapman-Kolmogorov equation. The transition function can be considered as an operator on appropriate spaces. Feller found it convenient to make the assumption that the transition function $P(t)$ acting on the space of continuous functions f

$$\{P(t)f\}(x)= \int P(t;x,dy)f(y)$$

takes continuous functions f into continuous functions $P(t)f$. Such transition functions are now usually called Feller transition functions. The derived operator L given by

$$\lim_{t\downarrow 0}(P(t)f-f)/t=Lf$$

is defined for a subclass of functions f and is in the case of the one dimensional diffusion the second order operator in x of the Fokker-Planck equation. A corresponding Markov diffusion process was determined by boundary conditions that might differ from those conventionally dealt with in the standard theory of differential equations. The boundary conditions could be regarded as a restraint on the class of functions on which L operated that in turn made the restrained L the infinitesimal generator of a properly defined transition function semigroup. He generalized the type of differential operator of diffusion theory. Every such operator (barring certain degenerate cases) could be written in the form $(d/d\mu)(d/d\sigma)$, with σ a scale function and μ an increasing function (or speed measure). In a natural way the general linear diffusion was related to a Wiener process (Norbert Wiener's model of Brownian motion) that was locally rescaled in space and speed. Aspects of this program were laid out in the papers (1954; 1959,1). Dynkin carried out research on a number of related problems.

Feller also carried out analyses of countable state continuous time Markov chains with stationary transition probabilities. Here the transition probabilities are given by a matrix-valued function $P(t)=(p_{ij}(t))$ (i,j states of the process) with the semigroup property $P(t)P(s)=P(t+s)$, $t,s \geq 0$. If $P(t)$ is differentiable at zero with finite derivatives $q_{ij}=p'_{ij}(0)$, the equalities $\sum_j q_{ij}=0$ hold. The backward differential equation $QP(t)=P'(t)$ and the forward differential equation $P(t)Q=P'(t)$ are often referred to as the Kolmogorov differential equations ($Q=(q_{ij})$). In (1957,2) Feller described a general method of constructing transition probability functions $P(t)$ that satisfy the conditions on the q_{ij}'s. It is also shown how to construct transition functions $P(t)$ that satisfy both systems of differential equations. Research on topics of this type was also carried out by J. L. Doob, K. L. Chung, H. Reiter, and D. Kendall.

The second volume of Feller's *Introduction to Probability Theory and Its Applications* appeared in 1966. It was written so as to be independent of the first volume. Further, it was aimed to be of interest to a large audience ranging from a novice to an expert in the area. The book certainly succeeds, but it understandably could not be as popular as the first volume. The first few chapters deal with special distributions like the exponential, the uniform, and the normal. Chapter 4 introduces probability spaces and probability measures. Laws of large numbers, the Hausdorff moment problem, and the inversion formula for Laplace transforms follow in Chapter 7. The central limit theorem and ergodic theorems for Markov chains are obtained in Chapter 8. Infinitely divisible distributions follow in the next chapter. A host of additional topics follow in the remaining chapters of the book: Markov processes and semigroups, renewal theory, random walks on the real line, characteristic functions, expansions related to the central limit theory, the Berry-Esséen theorem on the error term in the central limit theorem, large deviations, and aspects of harmonic analysis. Many of the topics are dealt with in an elegant and succinct manner.

Feller was always interested in the problems of genetics. Toward the end of his life, as a permanent visiting professor at Rockefeller University, he had a close collaboration there with Professor Dobzhansky and colleagues. The paper (1966,2) is a result of this interaction and corrects an error in the theory of evolution due to assumption of constant population size in the case of a two-allele population.

The papers (1968, 1970) show Feller's continued research on problems related to the law of the iterated logarithm continued throughout his life.

Feller's investigations were greatly appreciated. He was elected to the National Academy of Sciences in 1960 and was a member of the American Philosophical Society, the American Academy of Arts and Sciences, and a foreign member of the Danish and Yugoslav academies of science. He was president of the Institute of Mathematical Statistics. His widow accepted the National Medal of Science on his behalf in 1970.

J. L. Doob, who was as influential as Feller in nurturing and developing the growth and interest in probability theory, remarked,

[A]part from his mathematics those who knew him personally will remember Feller most for his gusto, the pleasure with which he met life, the excitement with which he drew on his endless fund of anecdotes about life and its absurdities, particularly the absurdities involving mathematics and mathematicians. To listen to him deliver a mathematics lecture was a unique experience. . . In losing him, the world of mathematics has lost one of its strongest personalities as well as one of its strongest researchers.

A colleague of many years at Cornell University and Rockefeller University, also a remarkable researcher in probability theory, Mark Kac, said of Feller:

Feller was a man of enormous vitality. . . The intensity of his reactions was reflected in what his friends called the "Feller factor," an imprecisely defined number by which one had to scale down some of his pronouncements to get near the truth. . . But he was not stubborn and underneath the bluster, kind and generous. . . Much as he loved mathematics, his view of it was anything but parochial. . . I recall a conversation in which a colleague asked, . . . "What can the generals do that we mathematicians couldn't do better?" "Sleep during battle," said Feller and that was that. . . When he learned that his illness was terminal his courage and considerateness came poignantly to the fore. Having accepted the verdict himself he tried to make it easy for all of us to accept it too. He behaved so naturally and he took such interest in things around him that he made us almost forget from time to time that he was mortally ill. . . One of the most original, accomplished and colorful mathematicians of our time.

Henry McKean, a student of and co-researcher with Feller noted:

[H]is enthusiasm, his high standards, his indefatigable desire to make you understand "what's really going on." That was also his watchword when he lectured. He would get quite excited, his audience in his hand and come (almost) to the point. Then the hour would be over, and he would promise to tell us what's really going on next time. Only next time the subject would be not quite the same, and so a whole train of things was left hanging, somewhat in the manner of Tristram Shandy. But it didn't matter. We loved it and couldn't wait for the next (aborted) revelation. . . Back to Will himself. He was short, compact, with a mop of wooly gray hair, irrepressible. In conversation quick, always ready with an opinion (or two) addicted to exaggeration. If you knew the code, you applied the "Feller factor" (discount by 90%). . . I think of him often, hearing his voice, remembering him so full of fun.

THIS MEMORIAL IS BASED IN part on an obituary in The Annals of Mathematical Statistics 1970, vol. 41 and on accounts written by J. L. Doob and M. Kac in the *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability*. Helpful written remarks of H. P. McKean have also been used.

NOTES

1. A. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin: Springer, 1933.
2. A. Kolmogorov. über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Math. Ann.* 104(1931):415-458.
3. Back Cover Blurb. from W. Feller. *An Introduction to Probability Theory and Its Applications,* vol. 1, 3rd edition. New York: Wiley, 1968

SELECTED BIBLIOGRAPHY

1928

Über algebraisch rektifizierbare transzendente Kurven. *Math. Z.* 27:481-495.

1936

Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung. *Math. Z.* 40:521-559.

1937

[1] Zur Theorie der stochastischen Prozesse (Existenz und Eindeutigkeitssätze). *Math. Ann.* 113:113-160.

[2] Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung. II. *Math Z.* 42:301-312.

1939

Die Grundlagen der Volterraschen Theorie des Kampfes ums Dasein in wahrscheinlichkeitstheoretischer Behandlung. *Acta Biotheor.* A 5:11-40.

1940

[1] On the integro-differential equations of purely discontinuous Markov processes. *Trans. Am. Math. Soc.* 48:488-515.

[2] Statistical aspects of ESP. *J. Parapsychol.* 4(2):271-298.

1941

On the integral equation of renewal theory. *Ann. Math. Stat.* 12:243-267.

1943

The general form of the so-called law of the iterated logarithm. *Trans. Am. Math. Soc.* 54:373-402.

1945

The fundamental limit theorems in probability. *Bull. Am. Math. Soc.* 51:800-832.

1946

The law of the iterated logarithm for identically distributed random variables. *Ann. Math.* 47:631-638.

1948

On the Kolmogorov-Smirnov limit theorems for empirical distributions. *Ann. Math. Stat.* 19:177-189.

1949

[1] With P. Erdös and H. Pollard. A property of power series with positive coefficients. *Bull. Am. Math. Soc.* 55:201-204.

[2] Fluctuation theory of recurrent events. *Trans. Am. Math. Soc.* 67:98-119.

1950

An Introduction to Probability Theory and Its Applications, vol. 1. New York: Wiley.

1952

[1] Diffusion processes in genetics. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. J. Neyman. pp. 227-246: Berkeley and Los Angeles: University of California Press.

[2] Some recent trends in the mathematical theory of diffusion. *Proceedings of the International Congress of Mathematicians (2)*, pp. 322-339. Providence: American Mathematical Society.

1954

The general diffusion operator and positivity preserving semigroups in one dimension. *Ann. Math.* 60:417-436.

1957

[1] Generalized second order differential operators and their lateral conditions. *Illinois J. Math.* 1:459-504.

[2] On boundaries and lateral conditions for the Kolmogorov differential equations. *Ann. Math.* 65:527-570.

1959

[1] Differential operators with the positive maximum property. *Illinois J. Math.* 3:182-186.

[2] The birth and death process as diffusion process. *J. Math. Pure. Appl.* 38:301-345.

1966

[1] *An Introduction to Probability Theory and Its Applications*, vol. 2. New York: Wiley.

[2] On the influence of natural selection on population size. *Proc. Natl. Acad. Sci. U. S. A.* 55:733-738.

1968

An extension of the law of the iterated logarithm to variables without variance. *J. Math. Mech.* 18:343-356.

1970

On the oscillations of the sums of independent random variables. *Ann. Math.* 91:402-418.