PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 71(85) (2002), 79–89

KARAMATA'S CHARACTERIZATION THEOREM, FELLER, AND REGULAR VARIATION IN PROBABILITY THEORY

E. Seneta

Abstract. Karamata's Characterization Theorem provided the impetus for Feller's (1966) exposition of the theory of regularly varying functions within a probability theory context. We investigate the conditions under which this theorem holds, and indicate manifestations in the identification of the spectral functions of the stable laws. Regular variation of a distribution function occurred implicitly as a necessary and sufficient condition for convergence in the 1930's, in the probabilistic work of P. Lévy, Khinchin, and Feller; and more transparently in that of Gnedenko and of Doeblin. Explicit recognition of the relevance of the concept in probability was interrupted by World War 2. A final section of this paper traces the evolution of Feller's name and early mathematical career from his Balkan origins, with a view to illuminating his recognition of the relevance of regular variation and his connection with Karamata.

1. Introduction

This paper has as its focus the following version of Karamata's Characterization Theorem, which is stated without proof in Karamata [1933, p. 58]:

Theorem K. *If f is monotone on some interval $[A,\infty)$ and*

$$\lim_{x\to\infty}(f(x+t)-f(x))=\psi(t) \tag{1}$$

for $t_1,t_2\in R=(-\infty,\infty)$ such that t_1/t_2 is irrational and $\psi(t_1),\psi(t_2)$ are finite, then (1) *holds for all $t\in R$, and $\psi(t)=\gamma t$ for some real constant γ.*

A great impetus to the revival of interest in the theory of regularly varying functions came within a probabilistic context from the expectedly influential book of Feller (1966) in which Sections VIII.8 and VIII.9 (pp. 268–276), entitled respectively "Regular Variation" and "Asymptotic Properties of Regularly Varying Functions" begin with the statement:

2000 *Mathematics Subject Classification.* Primary 26A12; Secondary 60F05, 01A60, 60-03.
Key words and phrases. characterization; boundedness on finite intervals; regularly varying functions; limit theorems in probability theory; necessary and sufficient conditions; central limit problem; history and biography.

> The notion of regular variation (introduced by J. Karamata in 1930) ... finds an ever increasing number of applications in probability theory. The reason for this is partly explained in the next lemma, which is basic in spite of its simplicity.

The "next lemma" and its following Theorem 1 in Feller (1966) is in fact (essentially) Theorem K above. Feller's proof is a standard one using the multiplicative (Hamel) form of Cauchy's functional equation.

The only specific reference to Karamata's papers occurs at the beginning of VIII.9 (p. 272), enroute to the integral characterization (Representation Theorem) of a regularly varying function, where two successive footnotes state:

> [21] J. Karamata, *Sur un mode de croissance régulière*, Mathematica (Cluj), Vol. 4 (1930), pp. 38–53. Despite frequent references to this paper, no newer exposition seems to exist.
>
> [22] Although new, our proof of theorem 1 uses Karamata's ideas.

It was the first of these footnotes that led the present author to Karamata's original papers, and to the lucid writings of the Yugoslav School in general, since I had difficulty in understanding Feller's exposition. Indeed I established contact with Ranko Bojanić and received some material from Karamata himself from his home address in Geneva. The small book (Seneta, 1976) was the eventual result of a synthesis of the basics. (Its Preface gives more detail on its evolution).

It was intriguing to find that Feller's origins were also Balkan (like Karamata in 1902, he was born in Zagreb in 1906; and he spent his youth in this city) and that their lifetimes nearly coincided. We shall see that Feller continued Yugoslav academic contacts until 1939. Thus the questions arose: when did the concept of regularly varying functions make an explicit appearance within limit theorems of probability theory, and when did Feller become aware of Karamata's work? The exploration of these issues occupies the latter part of this paper.

While the monotonicity of f in Theorem K is adequate for most probabilistic applications, even in our sequel, in fact it is not essential to the conclusion, and can be relaxed to its essence to give a more natural version of this Characterization Theorem. We proceed in this direction to a result in the spirit of Karamata (1933); and then show how to use the Characterization Theorem to identify the spectral functions for the stable probability laws in the Lévy Canonical Representation for the infinitely divisible laws. The final result of the identification in the probabilistic setting is of course well-known, but our Karamata-esque approach may be of interest.

2. Characterization Theory

LEMMA 1. *Let $h(x)$ be a real function defined and finite on $[A,\infty)$ for some A. Then a necessary and sufficient condition for*

$$h(x+1)-h(x)\to c \Longrightarrow h(x)/x\to c$$

as $x\to\infty$, where $-\infty<c<\infty$, is that $h(x)$ be bounded on each finite interval beyond some value x_0.

The sufficiency of the condition is a continuous variable version of Cauchy's theorem that the Cesáro limit of a sequence exists and is the same as the ordinary limit if this exists and is a finite number. A complete proof may be found in Seneta (1973; 1976, §1.7). The first of these references contains a discussion related to the present section.

THEOREM 1. *Let f be a function defined and finite on some interval $[A, \infty)$, and bounded on every finite subinterval of $[A, \infty)$. Further suppose that as $x \to \infty$*

$$\lim_{x\to\infty} (f(x+t) - f(x)) = \psi(t) \tag{2}$$

at $t = t_1, t_2 \in R$, with $\psi(t_1), \psi(t_2)$ finite. Then (2) holds for $t \in D$, where $D = \{t : t = qt_1 + rt_2,\ q, r \text{ integers}\}$ and $\psi(t) = \gamma t$, $t \in D$, where γ is independent of t.

PROOF. We first show that if (2) holds for any $t_1, t_2 \in R$ then for integers q, r

$$\lim_{x\to\infty} (f(x + qt_1 + rt_2) - f(x)) = q\psi(t_1) + r\psi(t_2). \tag{3}$$

Since

$$f(x + qt_1 + rt_2) - f(x) = f(x + qt_1 + rt_2) - f(x + rt_2) + f(x + rt_2) - f(x)$$

it is sufficient to show that for $j = 1, 2$

$$\lim_{y\to\infty} f(y + st_j) - f(y) = s\psi(t_j)$$

for integer s. This is clearly true if $s = 0$. If $s > 0$

$$\begin{aligned} f(y + st_j) - f(y) &= \sum_{i=1}^{s} (f(y + it_j) - f(y + (i-1)t_j)), \\ &\to s\psi(t_j) \end{aligned}$$

as $y \to \infty$. If $s < 0$, put $w = y + st_j$, so

$$\begin{aligned} f(y + st_j) - f(y) &= -(f(w - st_j) - f(w))\ , \\ &\to s\psi(t_j) \end{aligned}$$

since $-s$ is a positive integer, as $y \to \infty$, by the preceding. Thus(2) holds for $t \in D$, by(3).

Now in (2) put for $t \in D$ with $t > 0$, $v = x/t$, so that

$$\lim_{v\to\infty} f(t(v+1)) - f(vt) = \psi(t).$$

By Lemma 1, consequently

$$\lim_{v\to\infty} \frac{f(vt)}{v} = \psi(t)$$

so that

$$\lim_{x\to\infty} \frac{f(x)}{x} = \lim_{v\to\infty} \frac{f(vt)}{vt} = \frac{\psi(t)}{t}.$$

Thus for $t \in D$, $t > 0$,

$$\psi(t)/t = \gamma$$

is independent of t.

If $t \in D$, $t < 0$, in(2) put $w = -x/t$, so that

$$\lim_{w \to \infty} f(-t(w-1)) - f(-tw) = \psi(t).$$

But since $-t \in D$, we see from (2) that the limit is $-\psi(-t)$, and since $-t > 0$, by the preceding, $-\psi(-t) = \gamma t$. Thus for $t < 0$, also $\psi(t) = \gamma t$. Thus for all $t \in D$, $\psi(t) = \gamma t$. ☐

Corollary 1. *Under the conditions of Theorem* 1, *if* (2) *holds for all* $t \in [a, b]$, *a fixed closed subinterval of* R, *then* (2) *holds for all* $t \in R$, *and* $\psi(t) = \gamma t$, $t \in R$.

Proof. The extension of (2) from $t \in [a, b]$ to all $t \in R$ is standard. For identifying ψ, the proof of the second part of Theorem 1 can be used by taking D as R. ☐

Corollary 2. *Theorem K.*

Proof. Boundedness on every finite subinterval of $[A, \infty)$ follows from monotonicity. The set D in Theorem 1 is dense in R by Kronecker's Theorem. (Hardy and Wright, Chapter XXIII, Theorem 438). $\psi(t)$ is monotone on D in the same sense as f is in R. Suppose f is nondecreasing. Let $t_0 \in R$. Then we can choose arbitrarily small $\epsilon_1, \epsilon_2 > 0$ such that $t_0 - \epsilon_1, t_0 + \epsilon_2 \in D$. Then

$$\begin{aligned}\gamma(t_0 - \epsilon_1) = \psi(t_0 - \epsilon_1) &\leqslant \liminf_{x \to \infty}(f(x + t_0) - f(x)) \\ &\leqslant \limsup_{x \to \infty}(f(x + t_0) - f(x)) \\ &\leqslant \psi(t_0 + \epsilon_2) = \gamma(t_0 + \epsilon_2)\end{aligned}$$

and letting $\epsilon_1, \epsilon_2 \to 0$ yields the result. ☐

Corollary 1 is a non-measure-theoretic version of Karamata's Characterization Theorem which accords with his general derivation (Karamata, 1933, p. 56) using in effect a continuous version of Cauchy's Theorem. Karamata does not state his assumptions, which led to this elegant method of proof being overlooked, and regarded as incorrect. The modern definition of a regularly varying function requires $f(x)$ to be defined and *measurable* on $[A, \infty)$ and to satisfy

$$f(x + t) - f(x) \to \psi(t)$$

as $x \to \infty$ for each t in a subset of positive measure of $(-\infty, \infty)$, where $\psi(t)$ is finite-valued for such t. Such functions f have the property that there exists a constant $B \geqslant A$ such that f is bounded on any finite subinterval of $[B, \infty)$ (Bojanić and Seneta, 1971), so that in fact this measurability requirement is more restrictive, and Corollary 1 may be used to define "weakly" regularly varying functions (Seneta, 1973; 1976, §1.7). Monotonicity implies both measurability and boundedness on finite intervals; but the convergence in Corollary 2 (Theorem K) is assumed only for a set of t which is countable, and thus of Lebesgue measure 0.

Theorem K can be re-expressed in more restricted form: If f is monotone on R, and for all $x \in R$

$$f(x + t) - f(x) = \psi(t) \tag{4}$$

at $t = t_1, t_2 \in R$ such that t_1/t_2 is irrational, and ψ is independent of x, then (4) holds for all $t \in R$, and $\psi(t) = \gamma t$ for all $t \in R$, with γ a constant independent of t. Moreover then

$$f(x) = f(0) + \gamma x, \ x \in R.$$

This is precisely the form that is needed to identify the form of the spectral functions for the class of stable probability laws in the Lévy Canonical Representation of their characteristic function which itself dates to about 1937. Indeed within this probabilistic context this form is in fact to be found in Khinchin (1938,pp. 95–97). An awareness of Karamata's work and its relevance to probability, however, did not come for at least 10 years more, with World War 2 intervening. The following result is a more general proposition along the lines of this restricted form, but which generalizes the monotonicity assumption. It is given in a "multiplicative" form more immediately relevant to probabilistic settings.

THEOREM 2. *Suppose that $P(u)$ is defined and $\geqslant 0$ on $R^+ = (0, \infty)$, that $P(u) > 0$ and $P(u)$ and $1/P(u)$ are bounded on a closed finite subinterval $[c, d]$, $c < d$, of R^+, and that $P(u)$ satisfies*

$$\alpha P(u) = P(au), \ \beta P(u) = P(bu), \quad \forall u \in R^+ \tag{5}$$

for some two relatively prime integers $\alpha, \beta (\neq 1)$, and some $a, b > 0$. Then $P(u) = P(1)u^\gamma$, for some constant γ and for all $u = a^q b^r$, q, r integers, which is a set dense in R^+.

PROOF. Suppose $\ln a / \ln b = m/n$ for some integers m, n (i.e., $a^n = b^m$); then from (5), $\alpha^n P(u) = P(a^n u)$, $\beta^m P(u) = P(b^n u)$, $\forall u \in R^+$, so $\alpha^n P(u) = \beta^m P(u)$, and taking $u = u_0 \in [c, d]$ so that $P(u_0) > 0$, gives $\alpha^n = \beta^m$. This is a contradiction to α, β relatively prime. Hence the set $D = \{t : t = q \ln a + r \ln b, q, r \text{ integers}\}$ is dense in R (by Kronecker's Theorem). Hence the set $\{a^q b^r, \ q, r \text{ integers}\}$ is dense in R^+. But for any integers q, r,

$$P(a^q b^r u) = \alpha^q P(b^r u) = \alpha^q \beta^r P(u), \ u \in R^+ \ . \tag{6}$$

Now suppose for some $u_1 \in R^+$, $P(u_1) = 0$. Then $P(a^q b^r u_1) = 0$ for all integers q, r; but since the set $a^q b^r u_1$ is dense in R^+, for some q and r, $a^q b^r u_1 \in [c, d]$ and $P(u) > 0$ on $[c, d]$, giving a contradiction. Hence $P(u) > 0$ on R^+.

Now set $f(x) = \ln P(e^x)$, $x \in R$, so that(from(6))

$$f(t + x) - f(x) = q \ln \alpha + r \ln \beta, = \psi(t), t \in D, \tag{7}$$

independently of $x \in R$. Clearly, at $t = 0$ we may define $\psi(0) = 0$. Now choose $t_0 \in D$ such that $c < c + t_0 < d$. Then since $f(t_0 + x) = f(x) + \psi(t_0)$, $f(x)$, assumed bounded on $[c, d]$, is bounded on $[c + t_0, d + t_0]$, and continuing in this way, on $[c + nt_0, d + nt_0]$, integer $n \geqslant 0$, hence on $[c, d + nt_0]$. Thus f is bounded on any finite interval to the right of c, and similar reasoning gives it bounded on any finite interval to the left of d, hence it is bounded on any finite interval.

Now arguing as in the second part of the proof of Theorem 1, we obtain $\psi(t) = \gamma t$, $t \in D$, for some γ independent of t. Returning to (7) and putting $x = 0$ completes the proof, with $C = \exp\{f(0)\} = P(1)$. □

3. Identification of the Spectral Functions of the Stable Laws

Lukacs (1960, p. 90) calls the following result in probability theory the *Lévy Canonical Representation*: $\phi(t)$ is an infinitely divisible characteristic function if and only if it can be written in the form

$$\log \phi(t) = ita - \frac{\sigma^2}{2}t^2 + \int_{-\infty}^{-0} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) dM(u) + \int_{+0}^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2}\right) dN(u)$$

where $M(u)$, $N(u)$ and σ^2 satisfy the following four conditions:

(i) $M(u)$ and $N(u)$ are non-decreasing in the intervals $(-\infty, 0)$ and $(0, \infty)$ respectively.

(ii) $M(-\infty) = N(\infty) = 0$.

(iii) The integrals $\int_{-\epsilon}^{0} u^2 dM(u)$ and $\int_0^{\epsilon} u^2 dM(u)$ are finite for every $\epsilon > 0$.

(iv) The constant σ^2 is real and non-negative.

The representation is then unique.

The *stable laws* are a class of the infinitely divisible laws for which, given any $b_1 > 0$, $b_2 > 0$, there exists a $b > 0$ such that (Lukacs, 1960, p. 99):

$$\sigma^2(b^2 - b_1^2 - b_2^2) = 0 \tag{8}$$

$$M(b^{-1}y) = M(b_1^{-1}y) + M(b_2^{-1}y) \text{ if } y < 0 \tag{9}$$

$$N(b^{-1}y) = N(b_1^{-1}y) + N(b_2^{-1}y) \text{ if } y > 0. \tag{10}$$

The purpose of this section is to identify the form of $N(y)$ for $y > 0$ and $M(y)$ for $y < 0$ using Theorem 2. Put $b_1 = b_2 = 1$ in (10), write $P(u) = -N(u)$, $u \in R^+$; then for all $u \in R^+$, there is an $A = 1/b > 0$ such that

$$P(Au) = 2P(u) \tag{11}$$

where $P(u) \geqslant 0$ and non-increasing. Thus (iterating (11)) $P(A^n u) = 2^n P(u)$, $u \in R^+$, each integer n, (positive or negative). Hence if there is a u_0 such that $P(u_0) > 0$, it follows that $0 < A < 1$, and (by monotonicity) that $P(u) > 0$, $u \in R^+$. Thus either $P(u) = 0$ for all u, or $P(u) > 0$ for all u and $P(u)$ and $1/P(u)$ are bounded on any interval $[c, d] \in R^+$. Applying (10) once more gives the existence of a $B > 0$ such that

$$P(Bu) = 3P(u), \ u \in R^+$$

and hence we may now apply Theorem 2 with $\alpha = 2$, $a = A$, $\beta = 3$, $b = B$ to obtain the result invoking the monotonicity of P, that $P(u) = P(1)u^\gamma$, $u \in R^+$, $\gamma < 0$ (which includes the case $P(u) = 0$, $u \in R^+$, which results if $P(1) = 0$). Thus for some constant $\gamma < 0$, $N(y) = N(1)y^\gamma$, $y > 0$ with $N(1) \leqslant 0$.

A similar argument can be applied to $P(u) = M(-u)$, $u \in R^+$, using the same A and B as with N, on account of (9) and (10), which results in the same $\gamma = \ln 2 / \ln A$ as with N giving finally $M(y) = M(-1)|y|^\gamma$ for $y < 0$, with $M(-1) \geqslant 0$.

The fact that $\gamma > -2$ follows from e.g., the finiteness of $\int_0^{\epsilon} u^2 dN(u)$ unless $N(1)$ is zero. Suppose now $N(1) \neq 0$ (so $A^\gamma = 2$, for $\gamma > -2$); supposing $\sigma^2 > 0$

(8) implies $A^{-2} = 2$, a contradiction. Similarly if $M(-1) \neq 0$. Thus $\sigma^2 > 0$ implies $N(1) = 0 = M(-1)$. If $\sigma^2 = 0$ not both $N(1)$ and $M(-1)$ can be zero, otherwise $\log \phi(t) = ita$. This corresponds to a distribution degenerate at a, a trivial case. This completes our identification.

Feller had a long-term preoccupation with representations of the Lévy type for infinitely divisible distributions, going back to a paper in Serbo-Croatian (Feller, 1938/1939). In Feller (1966) his development of infinitely divisible distributions in Section XVII.2 is substantially different from the more classic one of Lukacs (1960). Nevertheless, in his identification of the stable-law spectral measures in his version of a Canonical Representation he uses the regular variation ideas of his VIII.8.

In IX.8, Domains of Attraction, Feller (1966) uses regular variation explicitly in formulating a necessary and sufficient condition for a probability distribution function F to belong to the domain of attraction of some non-degenerate distribution (in fact the possible limit laws are just the stable laws). It is in the discovery of such necessary and sufficient conditions that regularly varying functions (although not recognized explicitly as such) make their appearance.

4. Regular Variation as Necessary and Sufficient

In his book, Khinchin (1938, p. 102, Theorem 4.5) gives as necessary and sufficient condition for F to be in the domain of the normal (Gaussian) law G the condition

$$\frac{x^2(1 - F(x) + F(-x))}{U(x)} \to 0 \tag{12}$$

as $x \to \infty$, where

$$U(x) = \int_{|y| \leqslant x} y^2 dF(y)$$

attributing the result equally, to independently written papers of Khinchin (1935), Lévy (1935), and Feller (1936). Khinchin gives the year of appearance as 1935 to all, an error perpetuated in the definitive book of Gnedenko and Kolmogorov of 1949 (an English translation appeared in 1954). Khinchin (1938, p. 101) states that at the time the domains of attraction of the other (i.e., non-Gaussian) stable laws were not known. The three papers were written in an intensely competitive milieu. In a follow-up paper, for example, Feller (1937, p. 30) cites a letter from Lévy that in spite of its later appearance, Lévy's paper was received substantially earlier than Feller's (October 1934 against May 1935) and had been presented to the Société Mathématique de France.

In fact—as Feller (1966, IX.8) later points out—(12) is equivalent to $U(x)$ being slowly varying at infinity in Karamata's sense. This is actually an extension to measures (of the density form e.g., Seneta, 1976, Theorem 2.1, p. 53) of one of Karamata's theorems, as Feller (1969) claims. The two papers (Feller, 1936, 1937) seem to be the first of Feller's writings on probabilistic topics. He was at the time already in Stockholm and in the company of Harald Cramér. The papers treat the Central Limit Theorem and concern themselves with not-necessarily identically distributed random summands. Regular variation (or slow variation) is not mentioned.

Lévy's (1935) derivation of a necessary and sufficient condition is also essentially "probabilistic". However Lévy (1935, p. 366), as preliminary, gives as a necessary and sufficient condition for the Weak Law of Large Numbers (WLLN) what appears to be a regular variation condition. Feller (1966, VII.7, p. 233) proves an elegant theorem that a WLLN holds for sums of independently and identically distributed random variables with distribution function F ($F(0) = 0$) if and only if

$$t(1 - F(t)) \Big/ \int_0^t x dF(x) \to 0$$

as $t \to \infty$. The proof is direct, but there is a footnote:

> [8]It will be seen in VIII.8 that [this condition] is equivalent to *regular variation of* $1 - F(x)$ *with exponent* -1. (The use of this would simplify the proofs).

A necessary and sufficient condition for F to be in the domains of attraction of the other (non-degenerate, non-normal) stable laws (i.e., of index α, $0 < \alpha < 2$) is credited in Gnedenko and Kolmogorov (1968) to what is a next generation of researchers, Gnedenko (1939) (following Khinchin) and Doeblin (1940) (following Lévy). We quote here Doeblin (1940, Théorème V, p. 81):

> La condition nécessaire et suffisante ... est ..., si $X \to \infty$
>
> $$F(-X) = h_1(X)X^{-\alpha},\ 1 - F(X) = h_2(X)X^{-\alpha}$$
>
> avec $h_1(kX)/h_1(X) \to 1$ quelque soit k, et que lim $h_1/(h_1 + h_2)$ existe.

The idea of a regularly and slowly varying function is very clear here, although there is still no mention of Karamata or of the explicit concept of a slowly varying function. We have not been able to see Gnedenko's paper. Doeblin's paper appeared in the first issue of *Studia Mathematica* to be published after Soviet troops had occupied (in late 1939 under the Nazi-Soviet pact) "Polish" Ukraine, and specifically the city of Lwów (L'viv, L'vov, Léopol, Lemberg) where the journal was published. Like all articles in that issue, it has a summary in Ukrainian. Presumably wartime conditions precluded for some time the article becoming available to non-Soviet readers. It is known that Feller once met Doeblin (who died during the war).

It seems that the explicit connection between regular variation and probability limit theorems was made after 1943. In a well-known paper on extreme values in French, in an American journal, the Karamata connection is still absent even though Gnedenko (1943, Théorème 4, p. 439) gives as a necessary and sufficient condition:

$$\lim_{x \to \infty} \frac{1 - F(kx)}{1 - F(x)} = k^{-\alpha}$$

for every $k > 0$. This is again, like Doeblin, only a short step to a connection with Karamata's general theory.

The paper of Karamata (1930) is actually listed in the references to Gnedenko and Kolmogorov's book in its 1954 English translation (and so was likely listed in the 1949 Russian version) but there appears to be no mention within the book. It is likely that it was the English version of Gnedenko and Kolmogorov's book which led to a recognition of the relevance of Karamata's theory by Feller and others. Karamata's Tauberian Theorem, already well-known from 1930, which also plays an important part in probability theory, and also involves regular variation of F, may have contributed to this recognition.

5. William Feller

Despite Feller's Balkan origins, it would thus seem that he made the probabilistic connection with Karamata's theory relatively late in his career. He contributed (Feller, 1969) to the volume of *L'Enseignement Mathématique* ("Dédié à la mémoire de Jovan Karamata") but gives in his article no clue about any closer contact. It is therefore appropriate to use this present Karamata commemoration to set out a few little-known details concerning Feller's pre-American period, and some parallels with Karamata.

Karamata was born February 1, 1902 in Zagreb, and died in Geneva on August 14, 1967. Feller was born in Zagreb on July 7, 1906, and died January 14, 1970 in New York. At the time of their birth and up to the end of World War I, Zagreb as a city of the Austro-Hungarian Empire was called Agram. Feller's family was German-speaking. Feller's Catholic mother, in anticipation of a difficult birth determined to give her offspring the name of the saint on whose day in the (German) Catholic saint's list the birth occurred. The saint's name was Willibald. The family decided to use the Croatian Catholic list, which gave the name Vilim, which Feller kept until coming to Germany. He attended the University in Zagreb from 1923 to 1925, and from 1925 to 1928 he was at the University of Göttingen where in 1926 (at the age of 20) he earned his doctorate. Here he became acquainted with David Hilbert and Richard Courant who encouraged him to a mathematical career. In 1928 he went as Privat-dozent to the University of Kiel, but left in 1933 after refusing to sign a Nazi oath. Of interest to historians of probability are two joint papers with Erhard Tornier in 1933 (not on probabilistic topics). During a subsequent years in Copenhagen, he came to know Harald and Nils Bohr, and during five years (1934–1939) at the University of Stockholm, he had as senior colleagues Marcel Riesz and Harald Cramér. It is during this last period, presumably under the influence of Cramér, that his probabilistic papers (such as Feller 1936, 1937) begin to appear. These two papers mention both Cramér and Riesz.

In Germany, since "Vilim" sounded strange, he changed his name to the very common German name "Willy", and it is under the name Willy Feller that his German language papers (Feller 1936, 1937, 1939) appear. Actually Feller (1939) is a German abbreviation of the fuller paper (Feller, 1938/1939) written in Serbo-Croatian, and dated Stockholm, September 1937, which has underneath the title the annotation: Napisao član dopisnik Vilim (W.) Feller [By Corresponding Member Vilim (W.) Feller – E.S.]. Articles from the *Radovi* of the Yugoslav Academy in Zagreb (actually *Jugoslavenska Akademija Znanosti i Umjetnosti. Radovi.*), generally written in Serbo-Croatian, would be republished in abbreviated form in the *Bulletin International* in a more international language (French, German, English, Italian or Latin). There has consequently been confusion of various kinds over language and pagination in citing the two versions (Feller 1938/1939, 1939). It is notable, however, that in 1939 Feller was still in touch with language and country of his youth, and indeed as we have seen, was a Corresponding Member of the Yugoslav Academy of Zagreb. Karamata had been elected as Corresponding Member of this Academy in 1933. Feller's list of publications shows an earlier item [1934] in the *Bulletin International.*

The Feller (1938/1939, 1939) work is especially interesting since it begins his preoccupation with the Lévy Canonical Representation of infinitely divisible laws, of which we have spoken earlier. He states [our translation into English]:

> In the following, without claims to novelty, we will give a direct *analytical* proof, in essence, related to the proof of Kolmogorov's special case. Such a generalization is not without interest...

Feller's proof is not, however, itself yet free of the stochastic process (Markovian) context of Lévy. In the later (Serbo-Croatian) version of the paper there is an additional footnote indicating that Feller had just seen another new [1938] derivation by Khinchin of Lévy's formula.

Feller came to the U.S. in 1939. Since in the North American context the name "Willy" had a funny connotation, he decided to use "William". The surname was always "Feller", contrary to myth (Rota, 1989). His father was Eugen Viktor Feller, and his mother's maiden name was Ida Oehmichen. In 1990 Feller's sister (a retired medical doctor) Dr. Zora Feller, was still living in Germany.

Acknowledgements. I am grateful to Professors Herbert Busemann, Kai Lai Chung, Joanne Elliott, Loren Pitt and Simon Tavaré for details on W. Feller's family and early biography, in correspondence of 1990.

References

[1] — (1970), *William Feller, 1906-1970*, Ann. Math. Statist. **41**, iv–xiii.

[2] R. Bojanić and E. Seneta (1971), *Slowly varying functions and asymptotic relations*, J. Math. Anal. Appl. **34**, 302–315.

[3] W. Doeblin (1940), *Sur l' ensemble de puissances d'une loi de probabilité*, Studia Math. **9**, 71–96.

[4] W. Feller (1936), *Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung*, Math. Zeitschrift **40**, 521–559.

[5] W. Feller (1937), *Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung II*, Math. Zeitschrift **42**, 301–312.

[6] Vilim (W.) Feller (1938/1939), *O Kolmogoroff–P. Lévyjevu predočivanju beskonačno djeljivih funkcija reparticije*, Rad. Acad. Yougoslave Sci. Beaux-Arts Zagreb **263**, 95–112 (In Serbo-Croatian).

[7] W. Feller (1939), *Neuer Beweis für die Kolmogoroff-P. Lévysche Charakterisierung der unbeschränkt teilbaren Verteilungsfunktionen, (Auszug aus "Rad" Bd. 263, S. 95–112.)*, Bull. Internat. Acad. Yougoslave Sci. Beaux-Arts. Cl. Sci. Math. Natur. **32**, 1–8.

[8] W. Feller (1966), *An Introduction to Probability Theory and its Applications, Vol. 2*, Wiley, New York.

[9] W. Feller (1969), *One-sided analogues of Karamata's regular variation*, L'Enseignement Math. **15**, 107–121.

[10] B. V. Gnedenko (1939), *On the theory of domains of attraction of stable laws*, (In Russian) Uchenie Zapiski Moskov. Gos. Univ., Matematika **30**, 61–81.

[11] B. V. Gnedenko (1943), *Sur la distribution limite du terme maximum d'une série aléatoire*, Annals of Math. **44**, 423–453.

[12] B. V. Gnedenko and A. N. Kolmogorov (1968), *Limit Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, Mass. (revised; 1st English edition 1954; Russian edition 1949).

[13] J. Karamata (1930), *Sur un mode de croissance régulière des fonctions*, Mathematica(Cluj) **4**, 38–53.

[14] J. Karamata (1933), *Sur un mode de croissance régulière. Théorèmes fondamentaux*, Bull. Soc. Math. France **61**, 55–62.

[15] A. Khintchine [Khinchin] (1935), *Sul dominio di attrazione della legge di Gauss*, Giorn. Ist. Ital. Attuari **6**, 371–393.

[16] Ya. A. Khinchin (1938), *Predel'nie Zakoni dlya Summ Nezavisimikh Sluchaĭnikh Velichin, [Limit Theorems for Sums of Independent Random Variables.]* Red. Tekhn.-Teor. Lit., Moscow–Leningrad.

[17] P. Lévy (1935), *Propriétés asymptotiques des sommes des variables aléatoires indépendantes ou enchainées*, J. Math. Pures Appl. (Ser. 9) **14**, 347–402.

[18] E. Lukacs (1960), *Characteristic Functions*, Griffin, London.

[19] G.-C. Rota (1989), *Fine Hall in its golden age: Remembrances of Princeton in the early fifties*, In: *A Century of Mathematics in America. Part II*, American Math. Soc., Providence, R.I., pp. 223–236.

[20] E. Seneta (1973), *An interpretation of some aspects of Karamata's theory of regular variation*, Publ. Inst. Math. Acad. Serbe Sci. **15**, 111–119.

[21] E. Seneta (1976), *Regularly Varying Functions*, Lecture Notes in Mathematics **508**, Springer-Verlag, Berlin.

School of Mathematics and Statistics FO7
University of Sydney, NSW 2006 Australia