# **REGULAR VARIATION AND PROBABILITY: THE EARLY YEARS**

### N. H. BINGHAM

#### §1. Introduction.

It is a pleasure for 'B of BGT' to write in appreciation of 'T of BGT', on the occasion of Jef Teugels' retirement, and also to remind myself of the promise we made each other – all those years ago, in the early seventies – to write the book that regular variation so obviously required. The theme has continued to attract my interest, Jef's and that of his pupils since. As for the book ([BGT] below), it continues to be my most cited work, and to find its place in the working library of probabilists.

It is a pleasure also to return to the theme of Bingham (1990a), with the benefit of another fifteen years' worth of hindsight.

### $\S$ **2.** Early History: Analysis.

Anyone with an interest in the history of mathematics and in regular variation would be well advised to read Hardy's Cambridge Tract Orders of infinity (Hardy (1910)). This venerable work is fascinating for a variety of reasons – not least its extensive list of references, some going back to the 1700s. We find (§4.41), for example, where Karamata found the name 'regular growth' (the original term for 'regular variation'): it was used by Borel – not of course in Karamata's sense - in a way suggested by the theory of integral functions. We see here the Hardy of Pure Mathematics, interested in convergence tests and the like, and of the early Hardy-Littlewood papers on Tauberian theorems. Much emphasis is given to functions of 'logarithmico-exponential' type - functions that can be built up from products of powers of logarithms and their iterates, and exponentials and their iterates.

One of the earliest results in what came to be the theory of regular variation goes back to Landau (1911). Motivated by analytic number theory, Landau worked with monotone functions, and made the observation that if for a positive monotone function  $\ell$ on  $\mathbb{R}_+$  one has

$$\ell(\lambda x)/\ell(x) \to 1 \qquad (x \to \infty)$$

for one  $\lambda \neq 1$ , one has it for all  $\lambda > 0$ , and so  $\ell$  is slowly varying (in modern terminology).

Pólya (1917), also motivated by analytic number theory, assumed that  $\ell$  is positive

and

$$\ell(2x)/\ell(x) \to 1 \qquad (x \to \infty).$$

He shows that for f Riemann integrable on [0, 1] and p the primes,

$$\frac{\log x}{x} \sum_{p \le x} f(p/x) \to \int_0^1 f(t) dt \qquad (x \to \infty),$$

and similarly with the primes p replaced by other sequences q whose counting functions satisfy

$$\sum_{q \le x} 1 \sim x/\ell(x).$$

Pólya's proof is noteworthy, both for anticipating the use of approximation above and below by step functions exploited so well later by Karamata, Wielandt and others, and for providing (again, in modern terminology) an Abelian theorem for a Mellin convolution of Stieltjes form. Pólya's work was continued in the books Pólya & Szegö (1925), which influenced Karamata, the key figure in the field, to whom we now turn.

The modern period of regular variation in analysis begins with Karamata (1930). This famous classic led to the Hardy-Littlewood-Karamata theorem next year, Karamata (1931), to a succession of other contributions by Karamata himself, and to the work of the 'Jugoslav school' of Karamata's pupils, notably Aljančić, Bojanic and Tomić. Some of the best and most important work here was published in preprint form by the Mathematical Research Center, then supported at Madison, Wisconsin by the US Army, rather than in regular journals, where it would have been more easily accessible and might have made an earlier impact.

One thing that puzzled me when I began work in this area was why Karamata and his co-workers had turned aside from this promising line of work in 1963. I asked Ranko Bojanic this when I met him at Ohio State University in 1988. He replied that 'they hadn't known what it was good for'! This illustrates beautifully, both the supreme importance of applications to theory and the crucial role played in this field by probability theory, to which we turn next.

#### $\S$ **3.** Extreme-Value Theory

The formal beginning of the field of Extreme-Value Theory (EVT) may be taken to be the period 1927-28. In Fréchet (1927), two of the three kinds of extreme-value distribution (the Fréchet and the Weibull, in modern terminology) are obtained. All three are obtained in the classic paper by Fisher & Tippett (1928). This presents the extremevalue distributions as – to within type, or to within location and scale – a one-parameter family, split into three by the value zero of the parameter. These have become known since as the Fréchet (heavy-tailed), Gumbel (light-tailed) and Weibull (bounded tail), after Maurice Fréchet (1878-1973), French mathematician, Emil Julius Gumbel (1891-1966), German statistician and Waloddi Weibull (1887-1979), Swedish engineer.

Richard von Mises (1883-1953) studied EVT in 1936, giving in particular the von Mises conditions – sufficient conditions on the hazard rate (assuming the density exists) in order to give a situation in which EVT behaviour occurs, leading to one of the above three types of limit law – that is, giving an extremal domain of attraction D(G) for the extreme-value distribution G. The domains of attraction of the Fréchet laws  $\Phi_{\alpha}$  and the Weibull laws  $\Psi_{\alpha}$  (to use one of the common notations) were given by Gnedenko (1943), and progress towards the domain of attraction  $D(\Lambda)$  of the Gumbel law  $\Lambda$  was made by Mejzler (1949). This was later completed by Marcus & Pinsky (1969).

Meanwhile, mathematics was overtaken by reality. On the night of 31 January to 1 February 1953, a storm surge in the North Sea caused extensive flooding and many deaths. In the UK, 307 were killed; in the low-lying Netherlands, 1,783 people were killed (over 1,800 on some counts). The author, then a schoolboy of seven, remembers the public shock at the time very well. The Netherlands Government immediately gave top priority to understanding the causes of such tragedies with a view to preventing them if possible. Since it is the maximum sea level which is the danger, EVT is immediately relevant, and thus EVT became a Netherlands scientific priority. One outcome of this was the doctoral thesis of Laurens de Haan, de Haan (1970), written under the supervision of Professor J. Th. Runnenberg. EVT has continued to be the leading theme of de Haan's scientific work – witness for instance the title of de Haan (1990), *Fighting the arch-enemy with mathematics*.

De Haan's definitive work gave rise to a variant, or refinement, of Karamata's regular variation. One may regard *Karamata theory* as the study of asymptotic relations of the form

$$f(\lambda x)/f(x) \to g(\lambda) \qquad (x \to \infty) \qquad \forall \lambda > 0$$

(leading to  $g(\lambda) = \lambda^{\rho}$  for some  $\rho$ : g is regularly varying with index  $\rho$ ,  $g \in R_{\rho}$ ), and de Haan theory as the study of more complicated relations of the form

$${f(\lambda x) - f(x)}/{g(x)} \to h(\lambda) \qquad (x \to \infty) \qquad \forall \lambda > 0$$

(leading to  $g \in R_0$  slowly varying,  $h(\lambda) = c \log \lambda$  for some c, the *g*-index of f). We then say that f belongs to  $\Pi$ , the *de Haan class*. See e.g. BGT Ch. 1,2 for Karamata theory, BGT Ch. 3, de Haan (1970) (in the context of EVT), Geluk & de Haan (1987) for de Haan theory. To summarise:

1.  $F \in D(\Phi_{\alpha})$  iff  $1 - F \in R_{-\alpha}$ ;

2.  $F \in D(\Psi_{\alpha})$  iff the upper end-point  $x_{+}$  of F is finite, and  $1 - F(x_{-} - 1/x) \in R_{-\alpha}$ ;

3.  $F \in D(\Lambda)$  iff  $U := (1 - F)^{-1} \in \Pi_+$ , the class of f as above with  $g = \ell$  slowly varying and with positive  $\ell$ -index.

Extreme-Value Theory has flourished ever since, both as a branch of probability, pure and applied, and as an area of statistics. The textbook literature includes Leadbetter, Lindgren & Rootzén (1983), Resnick (1987) and Embrechts, Klüppelberg & Mikosch (1997). The specialist journal *Extremes* began publication in 1999.

### $\S4$ . Sums of random variables.

Even more central to probability than the *maximum* of random variables as in EVT is *sums* of random variables, in random walks - and in forming sample means etc. in statistics.

One should begin at the beginning, with the weak law of large numbers. For  $X_1, X_2, \ldots$  independent and identically distributed random variables,  $S_n := \sum_{i=1}^{n} X_k$  the partial sums, one can ask for conditions for the convergence

$$S_n/n \to c \qquad (n \to \infty),$$

in probability (weak law of large numbers), or almost surely (strong law of large numbers). Necessary and sufficient conditions for the above are, with  $\phi(.)$  the characteristic function of the distribution F of the  $X_n$ ,

(a)  $\phi(.)$  is differentiable at the origin and  $\phi'(0) = ic$ ,

(b)  $xP(|X| > x) \to 0$  and  $\int_{-x}^{x} y dF(y) \to c \ (x \to \infty)$ 

(of course, if the mean  $\mu$  exists, then  $c = \mu$ , but (a), (b) can hold even if the mean does not exist). See Ehrenfeucht & Fisz (1960), Feller (1971), XVII.2a, VII.7.

One can then generalize, and ask for

$$S_n/a_n \to 1 \qquad (n \to \infty)$$

in probability, for some sequence of constants  $a_n$ . This is called *relative stability*; the necessary and sufficient condition for it is

$$xP(|X| > x) / \int_{-x}^{x} y dF(y) \to 0 \qquad (x \to \infty)$$
(1)

(Gnedenko & Kolmogorov (1954), §28; Rogozin (1976), Maller (1978), (1979)).

In the corresponding question regarding the central limit theorem, one asks for conditions for which

$$(S_n - a_n)/b_n \to G$$
 in distribution  $(n \to \infty)$ , (\*)

for suitable centring sequences  $a_n$ , norming sequences  $b_n$  and limit distributions G. Then the possible limit laws G are the *stable* laws, whose most important parameter is the *index*  $\alpha \in (0, 2]$ . For  $\alpha = 2$ , the limit law G is the normal or Gaussian (which one may take without loss of generality to be the standard normal,  $\Phi$ , by suitably adjusting  $a_n, b_n$ ).

Let us take the normal case  $\alpha = 2$  first. The necessary and sufficient condition for a normal limit above – that is, for F to belong to the domain of attraction  $D(\Phi)$  of the standard normal – is

$$x^{2}P(|X| > x) / \int_{-x}^{x} y^{2} dF(y) \to 0 \qquad (x \to \infty)$$
 (2)

(Lévy (1937), §36). The result was discovered independently by Lévy (1935), Feller (1935) and Khintchine (1935).

#### §5. Gnedenko; Gnedenko & Kolmogorov.

One is first struck by the similarity between (1) and (2) above. Thus one might suspect some link between relative stability and convergence of sums (suitably centred and scaled) to a Gaussian limit. This link in indeed there. It was shown by Gnedenko (1939) that one has relative stability of the partial sums  $S_n$  of the  $X_n$  iff the  $X_n$ , when centred at means and *squared*, are in the domain of attraction of the Gaussian,  $D(\Phi)$ . This remarkable result entered the textbook literature in Gnedenko & Kolmogorov (1949/54), §28. Of course, it immediately gives the link between (1) and (2) above.

For (\*) to hold for a non-Gaussian stable limit law G - that is, for  $0 < \alpha < 2$  - the necessary and sufficient condition is

(i) tail-balance:

$$P(X < -x)/P(|X| > x) \to q,$$
  $P(X > x)/P(|X| >) \to p$   $(p+q=1)$   $(x \to \infty),$  (3a)

(ii)

 $P(|X| > \lambda x) / P(|X| > x) \to \lambda^{\alpha} \qquad (x \to \infty), \qquad \forall \lambda > 0.$ (3b)

This result, which is due to Gnedenko (1939) and Doeblin (1940), appears in the textbook literature in Gnedenko & Kolmogorov (1949/1954), §35.

This year marks the half-centenary of the English version of Gnedenko & Kolmogorov (1949/54), translated by K.-L. Chung. The book is an enduring classic, and has been very influential. It has long since been a bibliographical rarity, and copies of it are hard to come by and treasured (my record here is 'given one, inherited one, given one away'). To quote Bingham (1990b), §2:

'Something of the power and scope of [the book], as well as its style, is apply summarized

by its translator, Chung in his preface: '... a certain amount of mathematical maturity, perhaps a touch of single-minded perfectionism, is needed to penetrate and appreciate the classic beauty of this definitive work' '. Of its central theme, Chung remarks again, in the preface to his own book Chung (1974), that it 'has been called the "central problem" of classical probability theory. Time has marched on and the centre of the stage has shifted, but this topic remains without doubt a crowning achievement' '.

The power and probabilistic importance of the results quoted above is clear. What is strikingly lacking in them is any explicit use of the language, viewpoint and results of regular variation, although this had been available in the works of Karamata and his school since 1930.

#### $\S 6.$ Regular variation and probability theory: Sakovich and Feller.

Credit for making the link between Karamata's regular variation and the probability limit theorems above explicit belongs to Sakovich (1956), writing – appropriately enough – in the first volume of the then new Soviet journal *Theory of Probability and its Applications*.

For some reason, Sakovich's work was overlooked, and had to be rediscovered later. In 1966, Feller published his sequel An introduction to probability theory and its applications, Volume II, Feller (1966/71), to his earlier book Feller (1950/57/68). In Feller (1966/71), VIII.8 one finds a definition of regular variation, in VIII.9 one finds Karamata's theorem, and in IX.8 domains of attraction. Here one finds the equivalence of condition (2) for  $D(\Phi)$  with

the truncated variance 
$$V(x) := \int_{-x}^{x} y^2 dF(y)$$
 is slowly varying, (4)

that (2) and (4) are equivalent being an instance of Karamata's theorem ([BGT],  $\S1.6$ ,  $\S8.3$ ). Here one can also find (3b) in its natural form:

the tail-sum 
$$T(x) := P(|X| > x)$$
 is regularly varying with index  $-\alpha, T \in R_{-\alpha}$ . (5)

These passages are perhaps the most used and most quoted of Feller's book. The book has many general virtues – a generation of probabilists, including myself, were brought up on it. But one would hesitate to use it for instructional purposes today. Its great virtue – lots of beautiful examples – carries the drawback of length, and the danger of not seeing the wood for the trees, especially for the young or inexperienced. The structural weaknesses of avoiding both measure theory and continuous time – a third volume on stochastic processes was planned, but Feller died before he could write it – are plain to see. The next generation of books – Breiman (1968), Chung (1968/74), Billingsley (1979/86) and their successors

- were so excellent that 'Feller vol. II' became a work of reference rather than a text to learn from - *except* for the precious passages on regular variation referred to above.

# $\S7: BGT and after.$

Such was the situation when BGT was planned and written. In addition to Feller, there was de Haan (1970), written from the point of view of EVT, followed by Geluk and de Haan (1987), a brief (132-page) treatment from the point of view of analysis and Tauberian theory. Following the penetrating studies Seneta (1969), (1974) on regular variation and branching processes came Seneta (1976), a brief (112-page) analytic treatment of the basic theory. The Bibliographic Notes and Discussion sections in Seneta (1976) are of historic interests, as too are Seneta (1990), (2002).

So far as the later literature on regular variation is concerned, one may perhaps subdivide things by theme.

Analysis and Tauberian theory.

The book by Korevaar (2004), Tauberian theory – A century of developments, contains a wealth of results, including – Chapter 4, Part 2 – a thorough treatment of the role of regular variation within Tauberian theory.

Extreme-Value theory.

The book by Resnick (1987) – published in the same year as BGT – provides a monograph account of the role of EVT and regular variation in the context of point processes. Regular variation is used extensively in the book by Embrechts, Klüppelberg & Mikosch (1997), where the motivation is its use in EVT and applications in the mathematics of finance, insurance, actuarial science and the like.

Higher dimensions.

The most obvious limitation of BGT is its restriction to one dimension. The multivariate theory is interesting, and necessary for many applications. For a monograph treatment, see Meerschaert & Scheffler (2001), especially Part II (Chapters 4-6), Multivariate regular variation.

## Heavy tails.

Heavy tails in the broad sense underlie most of the work above: non-Gaussian domains of attraction show regularly varying tail-decay – 'Pareto tails' – which is extremely slow compared to the ultra-fast – log-quadratic – tail decay in the Gaussian case. A great impetus to the study of heavy tails was provided by the study of *long-range dependence* (Beran (1994)), motivated by such things as the Hurst effect in hydrology and the study of statistics of internet traffic. A broad account of heavy tails in theory and practice is given by Adler, Feldman and Taqqu (1998).

Risk management.

The benchmark model of mathematical finance is the Black-Scholes-Merton model, where the underlying noise driving price processes is Gaussian. Experience has shown that the tails of financial data sets are typically much heavier than in the Gaussian case. Since financial crises are triggered by exceptional large losses, much emphasis is placed nowadays on quantifying the probability of such exceptionally large losses. The usual measure is value at risk (VaR), but the problem area of *risk management* is much broader. The area is vast; for background and references, see e.g. Bassi, Embrechts & Kafetzaki (1998), Bingham & Kiesel (2002), Bingham, Kiesel & Schmidt (2003).

#### References.

Adler, R. J., Feldman, R. E. & Taqqu, M. S. (1998): A practical guide to heavy tails, Birkhäuser, Boston, Basel, Berlin.

Bassi, F., Embrechts, P. & Kafetzaki, M. (1998): Risk management and quantile estimation. In Adler et al. (1998), 111-130.

Beran, J. 1994): Statistics of long-memory processes, Chapman & Hall, New York.

Billingsley, P. (1979): Probability and measure, Wiley, New York (2nd ed. 1986).

Bingham, N. H. (1990a): Regular variation in probability theory, *Publ. Inst. Math.* Beograd **48** (62), 169-180.

Bingham, N. H. (1990b): Kolmogorov's work on probability, particularly limit theorems. Bull. London Math. Soc. 22, 51-58 [in D. G. Kendall et al.: Obituary, Andrei Nikolaevich Kolmogorov, ibid., 31-100].

Bingham, N. H. Goldie, C. M & Teugels, J. L. (1987): *Regular Variation*, Cambridge University Press, Cambridge.

Bingham, N. H. & Kiesel, R. (2002): Semi-parametric modelling in finance: Theoretical foundations, *Quantitative Finance* **2**, 241-250.

Bingham, N. H., Kiesel, R. & Schmidt, R. (2003): A semi-parametric approach to risk management. *Quantitative Finance* **3**, 426-441.

Breiman, L. (1968): Probability, Addison-Wesley, Reading, MA.

Chung, K.-L. (1968): A course in probability theory, Harcourt, Brace, Jovanovich, New York (2nd ed., Academic Press, New York, 1974, 3rd ed. 2001).

Doeblin, W. (1940): Sur l'ensemble des puissances d'une loi de probabilité, *Studia Math.* 9, 71-96.

Ehrenfeucht, A. & Fisz, M. (1960): A necessary and sufficient condition for the validity of the weak law of large numbers, *Bull. Acad, Polon. Sci.* **9**, 17-19.

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997): *Modelling extremal events*, Springer-Verlag, Berlin-Heidelberg, New York. Feller, W. (1935): Uber den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung, Math. Z. 40, 521-559.

Feller, W. (1968): An introduction to probability theory and its applications, Volume I, 3rd ed., Wiley, (1st ed. 1950, 2nd ed. 1957).

Feller, W. (1971): An introduction to probability theory and its applications, Volume II, second ed., Wiley, (1st ed. 1966).

Fisher, R. A. & Tippett, L. H. C. (1928): Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Phil. Soc.* **24**, 180-190 (reprinted in R. A. Fisher, *Contributions to Mathematical Statistics*, Wiley, 1950, and *Collected Papers of R. A. Fisher, Volume II, 1925-31* (ed. J. H. Bennett), The University of Adelaide, 1972, 208-219).

Fréchet, M. (1937): Sur la loi de probabilité de l'écart maximum, Ann. Soc. Math. Polon. 6, 93-116.

Geluk, J. L. & de Haan, L. (1987): Regular variation, extensions and Tauberian theorems, CWI Tracts **40**, Stichting Mathematisch Centrum, Amsterdam.

Gnedenko, B. V. (1939): On the theory of domains of attraction for stable laws (Russian), Gos. Univ. Uch. Zap. Mat. **30**, no. 3, 61-81.

Gnedenko, B. V. (1943): Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math. 44, 423-453.

Gnedenko, B. V. & Kolmogorov, A. N. (1949): Limit distributions for sums of independent random variables (Russian), Moscow-Leningrad; English translation Addison-Wesley, Cambridge, MA, 1954.

de Haan, L. (1970): On regular variation and its application to the weak convergence of sample extremes, Math. Centre Tracts **32**, Mathematisch Centrum Amsterdam.

de Haan (1990): Fighting the arch-enemy with mathematics, *Statistica Neerlandica* **44**, 45-68.

Hardy, G. H. (1910): Orders of infinity: The 'infinitärcalcül' of Paul du Bois-Reymond, Cambridge Tracts in Math. **12**, Cambridge Univ. Press, Cambridge (2nd ed. 1924).

Karamata, J. (1930): Sur un mode de croissance régulière des fonctions, Mathematica (Cluj) 4, 38-53.

Karamata, J. (1931): Neuer Beweis und Verallgemeinerung der Tauberschen Sätze, welche die Laplacessche und Stieltjessche Transformationen betreffen, J. reine ang. Math. 164, 27-39.

Khintchine, A. Ya. (1935): Sul dominio di attrazione della legge di Gauss, Giorn. Ist. Ital. Attuari 6, 371-393.

Korevaar, J. (2004): Tauberian theory. A century of developments, Springer-Verlag.

Landau, E. (1911): Sur les valeurs moyennes de certaines fonctions arithmétiques, *Bull.* Acad. Roy. Belgique, 443-72.

Leadbetter, M. R., Lindgren, G. & Rootzén, H. (1983): Extremes and related properties of random sequences and processes, Springer-Verlag, New York, Heidelberg, Berlin.

Lévy, P. (1935): Propriétés asymptotiques des sommes des variables aléatoires indépendantes ou enchainées, J. Math. Pures Appl. (8) **14**, 347-402.

Lévy, P. (1937): Théorie de l'addition des variables aléatoires, Gauthier-Villars, Paris (2nd ed. 1954).

Maller, R. A. (1978): Relative stability and the strong law of large numbers, Z. Wahrschein. **43**, 141-148.

Maller, R. A. (1979): Relative stability, characteristic functions and stochastic compactness, J. Australian Math. Soc. A 28, 499-509.

Marcus, M. B. & Pinsky, M. A. (1969): On the domain of attraction of  $\exp(-e^{-x})$ . J. Math. Anal. Appl. 28, 440-9.

Meerschaert, M. M. & Scheffler, H.-P. (2001): Limit distributions for sums of independent random vectors. Heavy tails in theory and practice. Wiley, New York.

Mejzler, D. G. (1949): On a theorem of B. V. Gnedenko (Russian), Sborn. Trud. Inst. Mat. Akad. Nauk Ukrain. RSR 12, 31-5.

Pólya, G. (1917): Über eine neue Weise, bestimmte Integrale in der analytischen Zahlentheorie zu gebrauchen, *Göttinger Nachrichten*, 149-59.

Pólya, G. & Szegö, G. (1925): Aufgaben und Lehrsätze aus der Analysis, I, II, Springer-Verlag, Berlin.

Resnick, S. I. (1987): Extreme values, regular variation and point processes, Springer-Verlag, New York, Berlin, Heidelberg.

Rogozin, B. A. (1976): Relatively stable random walks, Th. Probab. Appl. 16, 575-595.

Sakovich, G. N. (1956): A single form for the conditions for attraction to stable laws. Theory of Probability and Applications 1, 322-325.

Seneta, E. (1969): Functional equations and the Galton-Watson process, Advances Appl. Probab. 1, 1-42.

Seneta, E. (1974): Regularly varying functions in the theory of simple branching processes, Advances Appl. Probab. 6, 408-20.

Seneta, E. (1976): Regularly varying functions, Lecture Notes in Math. 508, Springer-Verlag.

Seneta, E. (1990): Karamata's iteration theory and normed regularly varying sequences in a historical light. *Publ. Inst. Math. Beograd* **48** (62), 45-51.

Seneta, E. (2002): Karamata's characterization theorem, Feller, and regular variation in

probability theory. Publ. Inst. Math. Beograd 71 (85), 79-89.

Department of Probability and Statistics, University of Sheffield, Sheffield S3 7RH, UK nick.bingham@sheffield.ac.uk