

TAUBERIAN THEOREMS AND LARGE DEVIATIONS

N. H. BINGHAM

To Cindy Greenwood, on her seventieth birthday

Abstract. The link between Tauberian theorems and large deviations is surveyed, with particular reference to regular variation.

1. Introduction.

To this author, limit theorems in probability theory have always been irresistibly attractive, as well as centrally important. As Kolmogorov put it in his introduction to Bernoulli's book, "The cognitive value of probability theory lies in the establishment of strict regularities resulting from the combined effects of mass random phenomena. The very notion of mathematical probability would have been fruitless if it were not realized as the frequency of a certain result under repeated experimentation. That is why the works of Pascal and Fermat can be viewed as only the prehistory of probability, while its true history begins with J. Bernoulli's law of large numbers" ([Kol]; [Shir], 874-5). The weak law of large numbers gives us a limiting frequency in the presence of a mean; its refinement, the central limit theorem, gives us a limiting distribution in the further presence of a variance. Both phenomena can occur more generally. The necessary and sufficient condition is known in each case, and in each case involves regular variation (for details see e.g. [Bin2]). Thus probabilists are led inescapably towards regular variation (for which see e.g. [BGT], particularly Chapter 8). Furthermore, the role of the characteristic function, or Fourier transform, in translating averaging, or addition of independent random variables, into multiplication rather than convolution leads one to work with transforms rather than with distributions. But one needs eventually to translate the information thus obtained from transform to distribution language. This is the role of Tauberian theory. For a magisterial treatment of the whole subject, see [Kor]; for regular variation, see Chapter IV there, or [BGT] Chapter 4.

This long chapter (64 pages), entitled *Abelian and Tauberian theorems*, originally covered Mercerian theorems also. It is a pleasure to record a personal debt to Cindy Greenwood here. Only after repeated and urgent pleas from Cindy did I detach the Mercerian component, which became Chapter 5 (25 pages – with Chapter V in [Pitt], incidentally, the only textbook chapters I know, though see also [Kor] IV.13, 14, [PalWie]

IV.18). She pointed out with remorseless clarity that separated, the Mercerian material would be impossible to miss, but unseparated risked sinking without trace. I wish I had had the courage of her convictions and split the Abelian and Tauberian parts also.

Our starting point here is the section immediately before the Mercerian part, on Tauberian theorems of exponential type ([BGT], §4.12). Here our treatment closely followed that of Kasahara [Kas] (leading to three Tauberian theorems, named after de Bruijn, Kohlbecker and Kasahara). At that time, neither Yuji Kasahara nor I realized the link between this area and large deviations. This link has been subsequently elaborated, by Kasahara and Kosugi; see [Kos1], [Kos2], [KasKos1], [KasKos2]. It involves convex analysis and the Gibbs variational principle; for background, see [Wig], [Geor], [Ell], [DupEll]. The theme is exemplified by the relation displayed on the cover of [DupEll]:

$$-\frac{1}{n} \log \int_x e^{-nh} d\theta^n \rightarrow \inf_{x \in X} \{h(x) + I(x)\}$$

(I is the *rate function*).

For later use, we record here our notation for the *Fenchel dual* (or Legendre-Fenchel dual, or conjugate) of a convex function f :

$$f^*(x) := \sup_y \{xy - f(y)\}$$

([Rock] Sections 7, 12). Then f^* is convex, proper ($< +\infty$ somewhere, $> -\infty$ everywhere) iff f is proper; the closure (greatest lower semi-continuous minorant if f is nowhere $-\infty$, identically $-\infty$ otherwise) $\text{cl } f$ has the same dual as f , f^* is closed (equal to its closure), and $f^{**} = \text{cl } f$.

We shall only be concerned here with the case when all moments $\mu_n := E[X^n]$ of X are finite (in this case the moment-generating function (mgf) $M(t) := \sum_{n=0}^{\infty} \mu_n t^n / n!$ is defined, though its radius of convergence may be zero). For X non-negative (which we may suppose here without loss) the non-integer moments are then finite also, and one has the option of working with the *Mellin transform* (or Mellin-Stieltjes transform) $\mathcal{F}(s) := E[X^s]$ (for $s > 0$ in the first instance, but we may continue analytically to $\text{Res} > 0$). Just as the cf is the natural tool for handling *sums* of independent random variables, the Mellin transform is the natural tool for handling *products*. For a systematic treatment, see [Zol]; for some applications, see [BinTeu]. We return to Mellin transforms in Section 4.

2. Entire characteristic function

We know from the work of Paley and Wiener [PalWie] that the proper setting for the Fourier transform is the complex plane (or domain, in their terminology). Thus the

traditional ‘ i ’ in the definition $\phi(t) := \int e^{itx} dF(x)$ of the characteristic function (cf) ϕ of a probability distribution F is merely conventional, and may be omitted if we intend to let the argument t take complex values. One then needs to distinguish cases according to the nature of the cf, or mgf, as a function of a complex variable. It may be entire, analytic but not entire or non-analytic, according as the radius of convergence is infinite, positive and finite, or zero. We turn now to the entire case, for which see e.g. [Luk1], and [BGT] Section 7.6 (Levin-Pfluger theory: functions of completely regular growth). For the analytic case, see [Luk1] Ch. 7, [Luk2] Ch. 7. For the non-analytic case, see Section 3 below.

An entire cf (of a positive random variable X , with distribution function F) has order $\rho \in [1, \infty)$ ([Luk1], Th. 7.1.3). If for large t $\log E(e^{tX})/t^\rho$ has $\liminf a$ and $\limsup b$, with $0 < a \leq b < \infty$, we say that X has *very regular growth* ([Val] II.7). Davies [Dav] shows that in this case one can find upper and lower bounds on

$$-\log P(X > x)/x^{\rho/(\rho-1)},$$

in terms of a and b . If the limit exists, and $a = b = \alpha$, say, then we say that X has *completely regular growth*. The Levin-Pfluger theory of [Lev], [BGT] Section 7.6 then applies, and $-\log P(X > x)/x^{\rho/(\rho-1)}$ is regularly varying with index $p = \rho/(\rho - 1)$.

This last result is extended to the case of general regular variation by Geluk [Gel1]. Writing R_p for the class of functions regularly varying with index p , he links

$$\liminf_{x \rightarrow \infty} -\log P(X > x)/\phi(x) = 1$$

for $\phi \in R_p$ with

$$\limsup_{x \rightarrow \infty} [E(X^n)]^{1/n} \psi(n)/n = q^{1/q} e^{-1/p},$$

where $p > 1$, $q > 1$ are conjugate (in the usual sense of L_p spaces) and ψ is increasing and in $R_{1/q}$. Then

$$\psi^\rightarrow(x) \sim \phi^*(x),$$

where one has the inverse function of ψ on the left and the Fenchel dual of ϕ on the right, so both sides are in R_q (see e.g. [BGT] Section 1.8.4).

Example: The Mittag-Leffler distribution.

For $\alpha \in (0, 1)$, the *Mittag-Leffler law* with *index* (or parameter) α has mgf

$$ML_\alpha(t) := \sum_{n=0}^{\infty} t^n / \Gamma(1 + n\alpha);$$

for uses in probability theory, see e.g. [Bin1], [Fel] XIII.8, [BGT] Section 8.11. Thus

$$\mu_n = n!/\Gamma(1 + n\alpha) = \Gamma(n)/(\alpha\Gamma(n\alpha)),$$

and by Stirling's formula one finds, in the notation above,

$$\psi(x) = x^\alpha, \quad \psi^\rightarrow(x) \sim \phi^*(x) \sim x^{1/\alpha},$$

and so

$$-\log P(X > x) \sim \phi(x) \sim (1 - \alpha)\alpha^{\alpha/(1-\alpha)}x^{1/(1-\alpha)}.$$

This result is derived twice in [BinTeu], p. 350, 352, and again in [BGT], Th. 8.1.12, as an instance of Kasahara's Tauberian theorem. For the mgf itself, one has

$$ML_\alpha(t) \sim \exp\{t^{1/\alpha}\}/\alpha$$

(Cartwright [Car], p. 50). The Mittag-Leffler law of index α is closely connected to the stable subordinator (increasing stable process) of index α , and the result above translates into a tail estimate at zero for $X = X(1)$ in the stable subordinator:

$$-\log P(X < x) \sim (1 - \alpha)\alpha^{\alpha/(1-\alpha)}/x^{\alpha/(1-\alpha)} \quad (x \downarrow 0);$$

see [BinTeu] 349-350, [BGT] Section 8.3.4.

The limiting case $p = 1$ relates to entire cfs of infinite order. This is a boundary case of the above, and may be handled by using second-order theory of regular variation (de Haan theory: [BGT] Chapter 3). Let ϕ be a positive function with $\phi(t)/t \rightarrow \infty$ as $t \rightarrow \infty$, and similarly for a . Suppose the second-order condition

$$\frac{\phi(tx) - x\phi(t)}{ta(t)} \rightarrow x \log x \quad (t \rightarrow \infty) \quad \forall x > 0$$

holds. Then if the tail function is everywhere positive and the cf is entire, the following are equivalent:

$$\liminf_{x \rightarrow \infty} \frac{-\log P(|X| > x) - \phi(x)}{xa(x)} = 0$$

for some ϕ , a as above,

$$\liminf_{n \rightarrow \infty} \frac{[E(|X|^n)]^{-1/n} n!^{1/n} - \psi(n)}{\alpha(n)} = 0$$

for some ψ and some $\alpha \in \Pi$ (the de Haan class Π ; see e.g. [BGT] Ch. 3). Then

$$\phi^*(x) \sim (ex)^{-1} \int_0^x \psi^\rightarrow(u) du$$

and $\alpha \sim a$. We refer to Geluk [Gel2], Geluk, de Haan and Stadtmüller [GdHS] for details.

3. Non-analytic characteristic function

A cf is defined in the first instance on a line in the complex plane (the real axis if one includes the ‘ i ’ in the definition, the imaginary axis if, as here, one does not), but can be continued analytically to a strip in the complex plane (which may be the whole plane, a half-plane, a strip of positive finite width or the imaginary axis). A cf is analytic iff all moments μ_n are finite, and $\limsup_{n \rightarrow \infty} [|\mu_n|/n!]^{1/n} = 1/R$ is finite (R is then the radius of convergence of the Taylor series of the cf, or mgf). Equivalently, it is analytic iff for some $R > 0$ the tail function $T(x) := P(|X| > x)$ is $O(e^{-rx})$ for all $r < R$, and then the strip of regularity contains the strip $|\operatorname{Im} z| < R$ ([Luk1], Th. 7.2.1). We proceed to consider results analogous to those above, but with the tails heavy enough to cause non-analyticity – that is, with the log-tail decaying more slowly than a multiple of x .

A recent result of this kind, which may serve as a prototype, is due to König and Mörters ([KoMo], Lemma 2.3): for X non-negative, $p > 1$ and c real,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log E[X^n / (n!)^p] = -c$$

implies

$$\lim_{x \rightarrow \infty} x^{-1/p} \log P(X > x) = -pe^{c/p},$$

with a similar implication with limsup in place of limit. (In their setting, $p = 1, 2, \dots$, and they consider the intersection of p independent Brownian motions in d dimensions. We exclude here the case $p = 1$, which falls under analyticity above, and is not relevant to their intersection problem, but there is no need here to restrict to p integer.) Their proof is simple, using only Markov’s inequality, Stirling’s formula and reduction of the limit to a suitable sequence. On the other hand, the setting of their main results makes heavy use of large-deviation theory.

4. Complements

1. *Laplace’s method.* The essence of Tauberian theorems of exponential type, and of large deviations, is that growth is dominated by behaviour in the neighbourhood of the maximum. This is simply a reflection of how overwhelmingly rapid exponential growth (or decay) is. The prototype in analysis is Laplace’s method (for which see e.g. de Bruijn [dBr]), which goes back to Laplace’s *Théorie analytique des probabilités* of 1820 and of which the most telling instance is Stirling’s formula of 1730. Its applicability in statistical mechanics rests ultimately on the enormous magnitude of Avogadro’s constant, $c. 6 \times 10^{23}$.

2. *The Wiman-Valiron method.* In the discrete analogue of Laplace's method (as in [BinTeu], 346), one examines the growth of an entire function (an entire cf in our case) by localizing at the term of maximum modulus in its Taylor series. For surveys of the Wiman-Valiron method, see Hayman [Hay], Fenton [Fen].
3. *Esscher transforms.* For F a law with mgf $M(t)$, the law $F_t(dx) := e^{tx}dF(x)/M(t)$ is called the *associated* law by Feller [Fel], XVI.7, or the *exponentially tilted* law. Its use goes back to F. Esscher in 1932, and later to Cramér in 1938; it is accordingly called the *Esscher transform*. It is widely used in ruin theory and insurance and actuarial mathematics; see e.g. Mikosch [Mik], Section 4.2. In brief, one may use it to change at will the point in whose neighbourhood one wishes to probe the asymptotic behaviour.
4. *Mellin transforms.* As mentioned earlier, with all moments finite as here, one may consider the Mellin transform $\mathcal{F}(s) := E[X^s]$ of X , and seek to link its behaviour for large s with the tail behaviour of X . One early example is the work of Wagner [Wag], also considered in [BinTeu] (the beginning of my collaboration with Jef Teugels, which led to [BGT], the origin of Fenchel duality in regular variation, and one source, together with [BGdH], of smooth variation – see [BGT] Section 1.8).
5. *Branching processes.* Tauberian theorems of exponential type are also useful in branching processes; see e.g. [BigBin] (near-constancy phenomena). Again, the theme is that when populations grow, they grow so rapidly that the contribution near the maximum is all that counts.
6. *Random potentials.* We close by mentioning the wealth of related literature on random Schrödinger operators in models of disordered media, etc. For a recent survey, with many references, see [LMW]. As Cindy says, all good mathematicians are also interested in physics.

Acknowledgement. I thank Peter Mörters for rekindling my interest in this subject.

References

- [BGdH] A. A. BALKEMA, J. L. GELUK and L. de HAAN, An extension of Karamata's Tauberian theorem and its connection with complementary convex functions, *Quart. J. Math.* **30** (1979), 385-416.
- [BigBin] J. D. BIGGINS and N. H. BINGHAM, Large deviations in the supercritical branching process, *Adv. Appl. Probab.* **25** (1993), 759-772.
- [Bin1] N. H. BINGHAM, Limit theorems for occupation times for Markov processes, *Z. Wahrschein.* **17** (1971), 1-22.

- [Bin2] N. H. BINGHAM, Regular variation and probability: The early years, *J. Computational and Applied Mathematics* **200** (2007), 357-363.
- [BGT] N. H. BINGHAM, C. M. GOLDIE and J. L. TEUGELS, *Regular variation*, 2nd ed., Cambridge University Press, 1989 (1st ed. 1987).
- [BinTeu] N. H. BINGHAM and J. L. TEUGELS, Duality for regularly varying functions, *Quart. J. Math.* (3) **26** (1975), 333-353.
- [dBr] N. G. de BRUIJN, *Asymptotic methods in analysis*, North-Holland, 1958.
- [Car] M. L. CARTWRIGHT, *Integral functions*, Cambridge Tracts Math. **44**, Cambridge University Press, 1956.
- [Dav] P. L. DAVIES, Tail behaviour for positive random variables with entire characteristic functions and completely regular growth, *Z. ang. Math. Mech.* **56** (1976), T334-336.
- [DupEll] P. DUPUIS and R. S. ELLIS, *A weak convergence approach to the theory of large deviations*, Wiley, New York, 1997.
- [Ell] R. S. ELLIS, *Entropy, large deviations and statistical mechanics*, Grundle. math. Wiss. **271**, Springer, 1985.
- [Fel] W. FELLER, *An introduction to probability theory and its applications*, Volume 2, 2nd ed., Wiley, 1971.
- [Fen] P. C. FENTON, A glance at Wiman-Valiron theory. *Contemporary Math.* **382** (2005), 131-140.
- [Gel1] J. L. GELUK, On the relation between the tail probability and the moments of a random variable, *Indagationes Math. A* **87** (1984), 4091-405.
- [Gel2] J. L. GELUK, On a theorem in entire function theory and its applications in probability theory, *J. Math. Anal. Appl.* **118** (1986), 165-172.
- [GdHS] J. L. GELUK, L. de HAAN and U. STADTMÜLLER, A Tauberian theorem of exponential type, *Canad. J. Math.* **38** (1986), 697-718.
- [Geor] H.-O. GEORGII, *Gibbs measures and phase transitions*, Walter de Gruyter, 1988.
- [Hay] W. K. HAYMAN, The local growth of power series: a survey of the Wiman-Valiron method, *Canad. Math. Bull.* **17** (1974), 317-358.
- [Kas] Y. KASAHARA, Tauberian theorems of exponential type, *J. Math. Kyoto Univ.* **18** (1978), 209-219.
- [KasKos1] Y. KASAHARA and N. KOSUGI, Large deviations around the origin for sums of non-negative i.i.d. random variables, *Natur. Sci. Rep. Ochanomizu Univ.* **51** (2002), 27-31.
- [KasKos2] Y. KASAHARA and N. KOSUGI, Remarks on Tauberian theorem of exponential type and Fenchel-Legendre transform, *Osaka J. Math.* **39** (2002), 613-620.
- [Kol] A. N. KOLMOGOROV, Preface to *The Law of Large Numbers*, Nauka, Moscow,

1986 (in Russian).

- [KoMo] W. KÖNIG and P. MÖRTERS, Brownian intersection local times: upper tail asymptotics and thick points, *Ann. Probab.* **30** (2002), 1605-1656.
- [Kor] J. KOREVAAR, *Tauberian theory: a century of developments*, Grundle. math. Wiss. **329**, Springer, 2004.
- [Kos1] N. KOSUGI, Tauberian theorem of exponential type and its application to multiple convolution, *J. Math. Kyoto Univ.* **39** (1999), 331-346.
- [Kos2] N. KOSUGI, *Tauberian theorem of exponential type on limits of oscillation*, *J. Math. Kyoto Univ.* **39** (1999), 783-792, Correction, *ibid.* **40** (2000), 303.
- [LMW] H. LESCHKE, P. MÜLLER and S. WARZEL, A survey of rigorous results on random Schrödinger operators for amorphous solids, *arXiv:cond-mat/0210708 v3 11 Dec 2003* (to appear, *Markov Processes and Related Fields*).
- [Luk1] E. LUKACS, *Characteristic functions*, 2nd ed., Griffin, 1970 (1st ed. 1960).
- [Luk2] E. LUKACS, *Developments in characteristic function theory*, Griffin, 1983.
- [Lev] B. Ja. LEVIN, *Distribution of zeros of entire functions*, Translations of Math. Monographs **5**, American Math. Soc., Providence RI, 1964.
- [Mik] T. MIKOSCH, *Non-life insurance mathematics: an introduction with stochastic processes*, Universitext, Springer, 2004.
- [PalWie] R. E. A. C. PALEY and N. WIENER, *Fourier transforms in the complex domain*, AMS Colloquium Publications **XIX**, Amer. Math. Soc., 1934.
- [Pitt] H. R. PITT, *Tauberian theorems*, Oxford University Press, 1958.
- [Rock] R. T. ROCKAFELLER, *Convex analysis*, Princeton University Press, 1970.
- [Shir] A. N. SHIRYAEV, Kolmogorov: Life and creative activities, *Ann. Probab.* **17** (1989), 866-944.
- [Val] G. VALIRON, *Lectures on the general theory of integral functions*, Dover, 1949.
- [Wag] E. WAGNER, Taubersche Satz für die Mellin-Transformation, *Wiss. Z. Martin-Luther Univ. Halle-Wittenberg Math.-Natur. Reihe* **15** (1966), 617-624.
- [Wig] A. S. WIGHTMAN, Convexity and the notion of equilibrium state in thermodynamics and statistical mechanics. Introduction (ix-lxxxv) in R. B. Israel, *Convexity in the theory of lattice gases*, Princeton University Press, 1979.
- [Zol] V. M. ZOLOTAREV, Mellin-Stieltjes transform in probability theory, *Th. Probab. Appl.* **2** (1957), 433-460.

Mathematics Department, Imperial College, London SW7 2AZ
n.bingham@ic.ac.uk nick.bingham@btinternet.com

