CHAPTER VIII

Laplace Transforms and Tauberian Theorem

Like the Fourier transform, the Laplace transform of a measure μ has a number of useful operational properties pertaining to moment generation, convolutions, and vague convergence. However, the main point of this chapter is to show that if μ is concentrated on a half-line, say $[0, \infty)$, then its Laplace transform can also be useful for obtaining the asymptotic behavior of $\mu[0, x]$ as $x \to \infty$.

Definition 8.1. Let μ be a measure on $[0, \infty)$. The **Laplace transform** $\hat{\mu}(\lambda)$ of μ is the real-valued function defined for $\lambda \geq c$ by

$$\hat{\mu}(\lambda) := \int_0^\infty e^{-\lambda x} \mu(dx), \quad \lambda > c, \tag{8.1}$$

where $c = \inf\{\lambda : \int_0^\infty e^{-\lambda x} \mu(dx) < \infty\}$.

Notice that by monotonicity of $e^{-\lambda x}$, $x \geq 0$, as a function of λ, the finiteness of the integral defining $\hat{\mu}(a)$ implies finiteness of $\int_0^\infty e^{-\lambda x} \mu(dx)$ for all $\lambda \geq a$. If μ is a finite measure, then $\hat{\mu}(\lambda)$ is defined at least for all $\lambda \geq 0$. On the other hand, one may also wish to view $\hat{\mu}(\lambda)$ as an extended real-valued, i.e., possibly infinite-valued, function defined for all $\lambda \in \mathbb{R}$, which is easy to do since the integrand is nonnegative. However, in general, the statement that the Laplace transform $\hat{\mu}(\lambda)$ exists is intended to mean that the defining integral is finite on some half-line.

Remark 8.1 *(Special Cases and Terminology)*. In the case that μ is absolutely continuous, say $\mu(dx) = g(x)dx$, then $\hat{g}(\lambda) := \hat{\mu}(\lambda)$ is also referred to as the Laplace transform of the (Radon–Nikodym derivative) function g. Also, if $\mu = P \circ X^{-1}$ is the distribution of a nonnegative random variable X defined on a probability space $(\Omega, \mathcal{F}, P)$, then $\hat{\mu}(\lambda)$ is also referred to as the Laplace transform of X,

$$\hat{\mu}(\lambda) = \mathbb{E}e^{-\lambda X}.$$

In the case that μ is a probability, the function $\hat{\mu}(-\lambda)$ is the **moment-generating function**.

Although the Laplace transform is an analytic tool, the theory to be developed is largely based on the probabilistic ideas already introduced in previous sections. This is made possible by the **exponential size-bias** transformation introduced in the treatment of large deviations, although in terms of the moment-generating function of a probability. Specifically, if μ is a measure on $[0,\infty)$ such that $\hat{\mu}(c) < \infty$ for some c, then one obtains a probability μ_c on $[0,\infty)$ by

$$\mu_c(dx) = \frac{1}{\hat{\mu}(c)} e^{-cx}\mu(dx). \tag{8.2}$$

Observe also that

$$\hat{\mu}_c(\lambda) = \frac{\hat{\mu}(c+\lambda)}{\hat{\mu}(c)}. \tag{8.3}$$

Just as with the Fourier transform one has the following basic operational calculus.

Proposition 8.1 *(Moment Generation).* If $\hat{\mu}$ exists on $(0,\infty)$, then $\hat{\mu}(\lambda)$ has derivatives of all orders $m = 1, 2, \ldots$ given by

$$\frac{d^m}{d\lambda^m}\hat{\mu}(\lambda) = (-1)^m \int_0^\infty x^m e^{-\lambda x}\mu(dx), \quad \lambda > 0.$$

In particular, μ has an mth order finite moment if and only if $\frac{d^m}{d\lambda^m}\hat{\mu}(0^+)$ exists and is finite.

Proof. For the first derivative one has for arbitrary $\lambda > 0$,

$$\lim_{h\to 0} \frac{\hat{\mu}(\lambda+h) - \hat{\mu}(\lambda)}{h} = \lim_{h\to 0} \int_0^\infty \left(\frac{e^{-hx} - 1}{h}\right) e^{-\lambda x}\mu(dx).$$

Since $|(e^{-hx}-1)/h| \le c(\delta)e^{\delta x}$ for some constant $c(\delta)$ if $|h| \le \delta/2$, where $\lambda - \delta > 0$, the limit may be passed under the integral sign by the dominated convergence theorem. The remainder of the proof of the first assertion follows by induction. For the final assertion, by the monotone convergence theorem,

$$\lim_{\lambda \downarrow 0} \int_0^\infty x^m e^{-\lambda x} \mu(dx) = \int_0^\infty x^m \mu(dx). \qquad \blacksquare$$

The proof of the following property is obvious, but its statement is important enough to record.

Proposition 8.2 *(Scale Change).* Let μ be a measure on $[0, \infty)$ with Laplace transform $\hat{\mu}(\lambda)$ for $\lambda > 0$. Define $\alpha : [0, \infty) \to [0, \infty)$ by $\alpha(x) = ax$, for an $a > 0$. Then one has $\widehat{\mu \circ \alpha^{-1}}(\lambda) = \hat{\mu}(a\lambda)$.

Proposition 8.3 *(Convolution Products).* If μ and ν are measures on $[0, \infty)$ such that $\hat{\mu}(\lambda)$ and $\hat{\nu}(\lambda)$ both exist for $\lambda > 0$, then the convolution $\gamma = \mu * \nu$ has Laplace transform $\hat{\gamma}(\lambda) = \hat{\mu}(\lambda)\hat{\nu}(\lambda)$ for all $\lambda > 0$.

Proof. This is a consequence of the basic formula (Exercise 1)

$$\int_0^\infty g(x) \mu * \nu(dx) = \int_0^\infty \int_0^\infty g(x + y) \mu(dx) \nu(dy)$$

for bounded Borel-measurable functions g, using the nonnegativity and multiplicative property of the exponential function. $\blacksquare$

Theorem 8.4 *(Uniqueness & Inversion Formula).* Let μ, ν be two measures on $[0, \infty)$ such that $\int_0^\infty e^{-cx} \mu(dx) = \int_0^\infty e^{-cx} \nu(dx) < \infty$ for some c and

$$\hat{\mu}(\lambda) = \hat{\nu}(\lambda) < \infty, \qquad \forall \lambda \geq c.$$

Then one has $\mu = \nu$. Moreover if $\mu[0, \infty) < \infty$, then one also has the inversion formula

$$\mu[0, x] = \lim_{\lambda \to \infty} \sum_{j \leq \lambda x} \frac{(-\lambda)^j}{j!} \frac{d^j}{d\lambda^j} \hat{\mu}(\lambda)$$

at each continuity point x of the (distribution) function $x \to \mu([0, x])$.

Proof. Assume first that μ and ν are finite measures. In this case a probabilistic proof is made possible by the asserted inversion formula obtained as follows. Without loss of generality, assume that μ and ν are normalized to probabilities. For arbitrary fixed $x, z > 0$, consider the expression $\sum_{j \leq \lambda z} \frac{(-\lambda)^j}{j!} \frac{d^j}{d\lambda^j} \hat{\mu}(\lambda) = \sum_{j \leq \lambda z} \frac{(-1)^j \lambda^j}{j!} \frac{d^j}{d\lambda^j} \hat{\mu}(\lambda)$, along with the expected value

$$P(Y_{\lambda x} \le z) = \mathbb{E}h_z(Y_{\lambda x}) = \sum_{j=0}^{\infty} h_z(\frac{j}{\lambda})\frac{(\lambda x)^j}{j!}e^{-\lambda x},$$

where $Y_{\lambda x}$, $\lambda, x > 0$, is Poisson distributed on the lattice $\{0, 1/\lambda, 2/\lambda, \ldots\}$ with intensity λx, and $h_z(y) = \mathbf{1}_{[0,z]}(y), y \ge 0$. Note that $\mathbb{E}Y_{\lambda x} = x$, and $\mathrm{Var}(Y_{\lambda x}) = \frac{x}{\lambda} \to 0$ as $\lambda \to \infty$. Notice that in general, if $\{\mu_{t,a} : t \ge 0, a \in \mathbb{R}\}$ is a collection of probabilities on $\mathbb{R}$, such that $\mu_{t,a}$ has mean a and variance $\sigma^2(a) \to 0$ as $t \to \infty$, then $\mu_{t,a} \Rightarrow \delta_a$ as $t \to \infty$. In particular,

$$\lim_{\lambda\to\infty} P(Y_{\lambda x} \le z) = \begin{cases} 0, & \text{if } z < x, \\ 1 & \text{if } z > x, \end{cases} \tag{8.4}$$

Now, in view of the moment-generation formula $(-1)^j \frac{d^j}{d\lambda^j}\hat{\mu}(\lambda) = \int_0^\infty x^j e^{-\lambda x}\mu(dx)$, one has

$$\sum_{j\le\lambda z} \frac{(-\lambda)^j}{j!}\frac{d^j}{d\lambda^j}\hat{\mu}(\lambda) = \int_0^\infty P(Y_{\lambda x} \le z)\mu(dx).$$

The inversion formula and hence uniqueness follows in the limit $\lambda \to \infty$ by application of the dominated convergence theorem. The general uniqueness assertion follows by the exponential size-bias transformation. Specifically, since μ_c and ν_c are probabilities whose Laplace transforms agree, one has $\mu_c = \nu_c$. Since $\mu \ll \mu_c = \nu_c$ and $\nu_c \ll \nu$, it follows that $\mu \ll \nu$ and $\frac{d\mu}{d\nu} = \frac{d\mu}{d\mu_c}\frac{d\nu_c}{d\nu} = \frac{\hat{\mu}(c)}{e^{-cx}}\frac{e^{-cx}}{\hat{\nu}(c)} = 1$. ■

Recall from Chapter V that a sequence of measures $\mu_n (n \ge 1)$ on $[0, \infty)$ is said to **converge vaguely** to a measure μ if $\int_{[0,\infty)} g\,d\mu_n \to \int_{[0,\infty)} g\,d\mu$ for all continuous functions g vanishing at infinity, i.e., $g(x) \to 0$ as $x \to \infty$.

Theorem 8.5 ***(Continuity).*** Let $\mu_n, n \ge 1$, be a sequence of measures on $[0, \infty)$ with respective Laplace transforms $\hat{\mu}_n, n \ge 1$, defined on a common half-line $\lambda \ge c$.

- **a.** If $\mu_n, n \ge 1$, converges vaguely to μ, and if $\{\hat{\mu}_n(c) : n \ge 1\}$ is a bounded sequence of real numbers, then $\lim_n \hat{\mu}_n(\lambda) = \hat{\mu}(\lambda)$ for all $\lambda > c$. Conversely, if for a sequence of measures $\mu_n(n \ge 1)$, $\hat{\mu}_n(\lambda) \to \varphi(\lambda) > 0\ \forall\ \lambda > c$ as $n \to \infty$, then φ is the Laplace transform of a measure, μ and μ_n converges vaguely to μ.
- **b.** Suppose $c = 0$ in (a), $\varphi(0^+) = 1$, and $\mu_n, n \ge 1$, is a sequence of probabilities. Then μ is a probability and $\mu_n \Rightarrow \mu$ as $n \to \infty$.

Proof. We will prove part (b) first. For this we use the Helly selection principle (Corollary 5.6) to select a weakly convergent subsequence $\{\mu_{n_m} : m \ge 1\}$ to a measure μ with $\mu(\mathbb{R}) \le 1$ on $[0, \infty)$. Since $x \mapsto e^{-\lambda x}$ is continuous and vanishes at infinity on

$[0,\infty)$, $\hat{\mu}_{n_m}(\lambda) \to \hat{\mu}(\lambda)$ as $m \to \infty$ for each $\lambda > 0$. Thus μ is the unique measure on $[0,\infty)$ with Laplace transform φ. In particular, there can be only one (vague) limit point. Since $\varphi(0^+) = 1$ it follows that μ is a probability.

We now turn to part (a). Assume that μ_n, $n \geq 1$, converges vaguely to μ, and first suppose that $\lim_n \hat{\mu}_n(c) = m$ exists. Apply exponential size-biasing to obtain for bounded continuous functions f vanishing at infinity that

$$\lim_{n\to\infty} \int_0^\infty f(x)\frac{e^{-cx}}{\hat{\mu}_n(c)}\mu_n(dx) = \int_0^\infty f(x)\frac{e^{-cx}}{m}\mu(dx) = \int_0^\infty f(x)\mu_c(dx),$$

for some measure μ_c. For $\lambda > c$, take $f(x) = e^{-(\lambda-c)x}, x \geq 0$, to see that $\lim_n \hat{\mu}_n(\lambda) = \hat{\mu}(\lambda), \lambda > c$. Assuming only that $\{\hat{\mu}_n(c) : n \geq 1\}$ is bounded, consider any convergent subsequence $\lim_{n'} \hat{\mu}_{n'}(c) = m'$. Since the limit $\lim_n \hat{\mu}_{n'}(\lambda) = \hat{\mu}(\lambda)$ does not depend on the subsequence, $\hat{\mu}_n(\lambda) \to \hat{\mu}(\lambda)$.

For the converse part of (a) suppose that $\hat{\mu}_n(\lambda) \to \varphi(\lambda)$ for all $\lambda > c$. For any fixed $\lambda' > c$, note that $\frac{\hat{\mu}_n(\lambda+\lambda')}{\hat{\mu}_n(\lambda')}$ is the Laplace transform of the exponentially size-biased probability $\mu'_n(dx) = \frac{1}{\hat{\mu}_n(\lambda')}e^{-\lambda' x}\mu_n(dx)$. By part (b), $\mu'_n, n \geq 1$, converges vaguely to a finite measure μ', and therefore μ_n converges vaguely to $\mu(dx) = \varphi(c)e^{cx}\mu'(dx)$. ■

Definition 8.2. A function φ on $(0,\infty)$ is said to be **completely monotone** if it possesses derivatives of all orders $m = 1, 2, \ldots$ on $(0,\infty)$ and $(-1)^m \frac{d^m}{d\lambda^m}\hat{\mu}(\lambda) \geq 0$ for each $\lambda > 0$.

It follows from the moment generation theorem that $\hat{\mu}(\lambda)$ is completely monotone. In fact, we will now see that this property characterizes Laplace transforms of measures on $[0,\infty)$. We preface this with two lemmas characterizing the range of generating functions (combinatorial) originally due to S. Bernstein, while the proofs here are along the lines of those given in Feller.[1]

For a given continuous function g on $[0,1]$, the **Bernstein polynomials** arise naturally in the Weierstrass approximation theorem (see Appendix B) and are defined by

$$B_n(t) = \sum_{k=0}^{n} g(\frac{k}{n})\binom{n}{k}t^k(1-t)^{n-k}, \quad 0 \leq t \leq 1.$$

Lemma 1 ***(Finite Differences and Bernstein Polynomials).*** The following is an equivalent representation of the Bernstein polynomials for a given continuous function

[1]See Feller, W. (1970).

g on $[0,1]$ in terms of the difference operator $\Delta_h g(t) = \frac{g(t+h)-g(t)}{h}$:

$$B_n(t) = \sum_{k=0}^{n} \binom{n}{k} (\frac{t}{n})^k \Delta_{\frac{1}{n}}^k g(0),$$

where $\Delta_h^1 = \Delta_h, \Delta_h^k = \Delta_h(\Delta_h^{k-1})$, $k \geq 1$, and Δ_h^0 is the identity operator.

Proof. Insert the binomial expansion of $(1-t)^{n-k} = \sum_{j=0}^{n-k} \binom{n-k}{j} (-1)^{n-k-j} t^{n-k-j}$ in the definition of $B_n(t)$, to obtain

$$B_n(t) = \sum_{j=0}^{n} \sum_{k=0}^{j} g\left(\frac{k}{n}\right) \binom{n}{k} \binom{n-k}{n-j} (-1)^{j-k} t^j.$$

For any finite or infinite sequence $a_0, a_1, \ldots$ of real numbers, the difference notation $\Delta_1 a_m := a_{m+1} - a_m$ is also used. For notational convenience we simply write $\Delta := \Delta_1$, i.e., $h = 1$. Upon iteration of $\Delta a_m = a_{m+1} - a_m$, one inductively arrives at

$$\Delta^k a_m = \sum_{j=0}^{k} \binom{k}{j} (-1)^{k-j} a_{m+j}. \tag{8.5}$$

For another sequence $b_0, b_1, \ldots$, multiply this by $\binom{n}{k} b_k$ and sum over $k = 0, \ldots, n$. Then making a change in the order of summation, the coefficient of a_{m+j} may be read off as

$$\begin{aligned} \sum_{k=j}^{n} \binom{n}{k} \binom{k}{j} (-1)^{k-j} b_k &= (-1)^{n-j} \binom{n}{j} \sum_{l=0}^{n-j} \binom{n-j}{l} (-1)^{n-j-l} b_{l+j} \\ &= \binom{n}{j} (-1)^{n-j} \Delta^{n-j} b_j. \end{aligned}$$

The first equality is by a change of order of summation and writing $(-1)^l = (-1)^{n-j}(-1)^{n-j-l}$, and the last equality is by (8.5) applied to the sequence $b_0, b_1, \ldots$. Thus one has the so-called *general reciprocity formula* relating differences $\Delta^k a_m$ and $\Delta^k b_m$ for two arbitrary sequences $\{a_m : m = 0, 1, \ldots\}$ and $\{b_m : m = 0, 1, \ldots\}$ in a "summation by parts" form

$$\sum_{k=0}^{n} b_k \binom{n}{k} \Delta^k a_m = \sum_{j=0}^{n} a_{m+j} \binom{n}{j} (-1)^{n-j} \Delta^{n-j} b_j. \tag{8.6}$$

For $0 < t < 1$, applying (8.6) to $b_m = t^m$ using (8.5), one has $\Delta^k b_m = t^m(1-t)^k(-1)^k$. Thus, applying (8.6) yields the identity

$$\sum_{m=0}^{n} t^m \binom{n}{m} \Delta^m a_k = \sum_{j=0}^{n} a_{k+j} \binom{n}{j} t^j (1-t)^{n-j}. \tag{8.7}$$

Now fix $h = 1/n$ and consider the difference ratios $\Delta_h^m a_k \equiv h^{-m} \Delta^m a_k$ of the sequence $a_k = g(\frac{k}{n})$, $k = 0, 1, \ldots, n$. The asserted difference representation of the Bernstein polynomials for g now follows directly from (8.7). ■

The dual representation of Bernstein polynomials can be used to characterize power series with positive coefficients as follows.

Lemma 2. Let g be a function on $[0, 1)$. Then the following are equivalent: (a) $g(t) = \sum_{n=0}^{\infty} c_n t^n$, $0 \le t < 1$ with $c_n \ge 0$, $\forall n$; (b) $g^{(n)}(t) \equiv \frac{d^n}{dt^n} g(t)$ exists at each $t \in (0, 1)$ and is nonnegative for every $n = 0, 1, 2, \ldots$; (c) $\Delta_{\frac{1}{n}}^{k} g(0) \ge 0$, for $k = 0, 1, \ldots, n-1$, $n \ge 1$. Such functions g are said to be **absolutely monotone**.

Proof. That $(a) \Rightarrow (b)$ follows from the analyticity of Taylor series expansion and term by term differentiation (see Exercise 7, Chapter IV). Also $(b) \Rightarrow (c)$ since monotonicity of g implies $\Delta_{\frac{1}{n}} g(t) \ge 0$, and monotonicity of g' then implies monotonicity of $\Delta_{\frac{1}{n}} g(t)$, so that $\Delta_h^2 g(t) \ge 0$. Iterating this argument, one arrives at (c) from (b). In fact, $\Delta_{\frac{1}{n}}^{n} g(0) \ge 0$ as well. For $(c) \Rightarrow (a)$, first consider the case that g satisfies (c) for $k = 0, 1, \ldots, n$ and is continuously defined on the closed interval $[0, 1]$ with $g(1) = 1$. In view of the Weierstrass approximation theorem, the Bernstein polynomials

$$B_n(t) = \sum_{k=0}^{n} g\left(\frac{k}{n}\right) \binom{n}{k} t^k (1-t)^{n-k}, \quad 0 \le t \le 1,$$

converge uniformly to g on $[0, 1]$ as $n \to \infty$ (see Appendix B). From (c) one sees using Lemma 1 that the coefficients $p_{j,n} = \sum_{k=0}^{j} g(\frac{k}{n}) \binom{n}{k} \binom{n-k}{n-j} (-1)^{j-k}$, $j = 0, 1 \ldots, n$, are nonnegative and $\sum_{j=0}^{n} p_{j,n} = B_n(1) = 1$. Thus $B_n(e^{-\lambda})$ is the Laplace transform of the probability μ_n defined by $\{p_{j,n} : j = 0, 1, \ldots, n\}$, i.e., $\mu_n = \sum_{j=0}^{n} p_{j,n} \delta_{\{j\}}$. It follows from the Weierstrass approximation and the continuity theorem for Laplace transforms that there is a probability μ such that $\mu_n \Rightarrow \mu$, and μ has the desired Laplace transform $g(e^{-\lambda}) = \lim_{n \to \infty} B_n(e^{-\lambda})$. Take $\lambda = \log t$ to complete the proof of (a) for the case in which g continuously extends to $[0, 1]$. If $g(1^-) = \infty$, fix an arbitrary $0 < \delta < 1$ and define $g_\delta(t) = \frac{g(\delta t)}{g(\delta)}$, for $0 \le t \le 1$. Then g_δ satisfies (c) and the above proof applied to g_δ yields an expansion (in $s = \delta t$)

$$g(s) = g(\delta) \sum_{n=0}^{\infty} d_n(\delta) s^n, \quad 0 \le s < \delta.$$

By uniqueness of coefficients in a series expansion of $g(s)$ on an interval $[0,\delta)$, the coefficients $c_n = g(\delta)d_n(\delta)$ do not depend on δ, and the expansion (a) is therefore valid on $[0,\delta)$ for δ arbitrarily close to 1, i.e., valid on $[0,1)$. ■

Theorem 8.6 *(Range of Laplace Transforms).* A function φ on $(0,\infty)$ is completely monotone if and only if there is a measure μ on $[0,\infty)$ such that

$$\varphi(\lambda) = \int_0^\infty e^{-\lambda x}\mu(dx), \quad \lambda > 0.$$

In particular, μ is a probability if and only if $\varphi(0^+) = 1$.

Proof. In the case that μ is a finite measure, the necessity of complete monotonicity follows directly from the previous moment-generation formula. For general measures μ on $[0,\infty)$ for which $\hat{\mu}(\lambda)$ exists for $\lambda > 0$, it then follows from exponential size-biasing that $\frac{\hat{\mu}(\lambda+c)}{\hat{\mu}(c)}$ is completely monotone as a function of $\lambda > 0$ for any fixed $c > 0$. Thus, the necessity is proven.

Suppose that φ is a completely monotone function on $(0,\infty)$. For arbitrary fixed $h > 0$, define a measure μ_h by

$$\mu_h = \sum_{n=0}^\infty \frac{(-h)^n}{n!}\frac{d^n}{d\lambda^n}\varphi(h)\delta_{\{\frac{n}{h}\}}.$$

Then by linearity of the Laplace transform and the continuity theorem applied to the limit of the partial sums,

$$\hat{\mu}_h(\lambda) = \sum_{n=0}^\infty \frac{(-h)^n}{n!}\frac{d^n}{d\lambda^n}\varphi(h)e^{-\lambda\frac{n}{h}}.$$

Since $c_n := \frac{1}{n!}\frac{d^n}{d\lambda^n}\varphi(h(1-t))|_{t=0} = \frac{(-h)^n}{n!}\frac{d^n}{d\lambda^n}\varphi(h) \geq 0$ for each n, it follows from the preceding lemma that $\varphi(h(1-t))$ has the power series expansion

$$\varphi(h(1-t)) := \sum_{n=0}^\infty \frac{(-h)^n}{n!}\frac{d^n}{d\lambda^n}\varphi(h)t^n, \quad 0 \leq t < 1 \tag{8.8}$$

(also see Exercise 10). Thus $g_h(\lambda) := \varphi(h(1-e^{-\frac{\lambda}{h}}))$, $\lambda > 0$, is the Laplace transform of μ_h. Since $g_h(\lambda)$ converges to $\varphi(\lambda)$ on $(0,\infty)$ as $h \to \infty$, it follows from the continuity theorem that there exists a measure μ on $[0,\infty)$ having Laplace transform φ. ■

Already the condition that the Laplace transform $\hat{\mu}(\lambda)$ exists at some $\lambda \geq 0$, readily implies that for any bounded interval $J = (a,b)$, $\mu(J) \leq e^{\lambda b}\hat{\mu}(\lambda) < \infty$; finiteness of $\mu(J)$ for all bounded intervals J is referred to as the **Radon property**. As the following theorem illustrates, much more on the asymptotic behavior of μ may be obtained from that of its Laplace transform near zero, and vice versa. For the proofs of results it will be convenient to use the distribution function G_μ of a (Radon) measure μ on $[0,\infty)$, defined by

$$G_\mu(x) = \mu[0,x], \qquad x \geq 0.$$

Theorem 8.7 ***(****Karamata Tauberian Theorem****).*** Let μ be a measure on $[0,\infty)$ whose Laplace transform exists for $\lambda > 0$. Then for $\theta \geq 0$,

$$\lim_{\alpha \downarrow 0} \frac{\hat{\mu}(\alpha\lambda)}{\hat{\mu}(\alpha)} = \lambda^{-\theta} \quad \text{if and only if} \quad \lim_{a\to\infty} \frac{\mu[0,ax]}{\mu[0,a]} = x^\theta.$$

In particular, either of these implies for $\alpha \downarrow 0$, $a = \alpha^{-1} \to \infty$, that

$$\hat{\mu}(\alpha) \sim \mu[0,a]\Gamma(\theta+1),$$

where $\Gamma(r) = \int_0^\infty x^{r-1}e^{-x}dx$, $r > 0$, is the gamma function.

Proof. Suppose $\lim_{\alpha\downarrow 0} \frac{\hat{\mu}(\alpha\lambda)}{\hat{\mu}(\alpha)} = \lambda^{-\theta}$. Observe that the limit on the left side applies to Laplace transforms of measures μ_a obtained from μ by scale changes of the form $G_{\mu_a}(x) = \frac{G_\mu(ax)}{\hat{\mu}(\alpha)}$, where $a = \alpha^{-1}$. On the other hand, the right side is the Laplace transform $\hat{\gamma}(\lambda) = \lambda^{-\theta}$ of the measure $\gamma(dx) = \frac{1}{\Gamma(\theta)}x^{\theta-1}dx$ on $[0,\infty)$. Thus, by the continuity theorem for Laplace transforms, μ_a converges vaguely to γ as $a \to \infty$. Since γ is absolutely continuous with respect to Lebesgue measure, it follows that the (improper) distribution function converges at all points $x \geq 0$. That is,

$$G_{\mu_a}(x) \to G_\gamma(x) = \frac{1}{\Gamma(\theta+1)}x^\theta$$

as $a \to \infty$. Take $x = 1$ to get

$$\hat{\mu}(\alpha) \sim G_\mu(a)\Gamma(\theta+1) = \mu[0,a]\Gamma(\theta+1).$$

With this it also follows that

$$\lim_{a\to\infty} \frac{\mu[0,ax]}{\mu[0,a]} = x^\theta.$$

For the converse, assume that $\lim_{a\to\infty} \frac{\mu[0,ax]}{\mu[0,a]} = x^\theta$. The Laplace transform of the measure μ_a with distribution function $G_{\mu_a}(x) = \frac{\mu[0,ax]}{\mu[0,a]}$ is $\frac{\hat{\mu}(\alpha\lambda)}{G_\mu(a)}$, and that of $G_\gamma(x) =$

x^θ is $\Gamma(\theta+1)\lambda^{-\theta}$. Thus, in view of the continuity theorem, if one can show that $\frac{\hat{\mu}(\alpha c)}{G_\mu(a)}$ is bounded for some $c > 0$, then it will follow that

$$\frac{\hat{\mu}(\alpha\lambda)}{G_\mu(a)} \to \Gamma(\theta+1)\lambda^{-\theta}$$

as $\alpha \to 0$, $a = \alpha^{-1}$. From here the converse assertions follow as above. So it suffices to prove the boundedness of $\frac{\hat{\mu}(\alpha c)}{G_\mu(a)}$ for some $c > 0$. For this, first observe that the assumption $\lim_{a\to\infty} \frac{\mu[0,ax]}{\mu[0,a]} = x^\theta$ implies that there is a $c > 1$ such that $G_\mu(2x) \leq 2^{\theta+1}G_\mu(x)$ for $x > c$. Thus, with $a = \alpha^{-1}$,

$$\begin{aligned}
\hat{\mu}(\alpha c) \leq \hat{\mu}(\alpha) &= \int_0^a e^{-\alpha x}\mu(dx) + \sum_{n=0}^{\infty} \int_{2^n a}^{2^{n+1}a} e^{-\alpha x}\mu(dx) \\
&\leq G_\mu(a) + \sum_{n=0}^{\infty} e^{-2^n} G_\mu(2^{n+1}a) \\
&\leq G_\mu(a)\left\{1 + \sum_{n=0}^{\infty} 2^{(n+1)(\theta+1)}e^{-2^n}\right\},
\end{aligned}$$

for all $a > c > 1$. In particular, this establishes a desired bound to complete the proof. ∎

Definition 8.3. A function L on $[0,\infty)$ is said to be **slowly varying at infinity** if for each fixed $x > 0$, one has $\lim_{a\to\infty} \frac{L(ax)}{L(a)} = 1$.

The following corollary is essentially just a reformulation of the statement of the Tauberian theorem. The proof is left as Exercise 6.

Corollary 8.8. For L slowly varying at infinity and $0 \leq \theta < \infty$ one has

$$\hat{\mu}(\lambda) \sim \lambda^{-\theta} L\left(\frac{1}{\lambda}\right) \quad \text{as} \quad \lambda \downarrow 0$$

if and only if

$$\mu[0,x] \sim \frac{1}{\Gamma(\theta+1)} x^\theta L(x) \quad \text{as} \quad x \to \infty.$$

Remark 8.2. It is to be noted that asymptotic relations in the Tauberian theorem are also valid with the roles of α and a reversed, i.e., for $\alpha \to \infty$ and $a \to 0$.

In the case that $\mu(dx) = g(x)dx$ has a density f one may obtain a "differentiated form" of the asymptotic relation under sufficient regularity in g. One such condition[2] is the following:

Definition 8.4. A function g on $[0, \infty)$ is said to be **ultimately monotone** if it is monotone on some $[x_0, \infty)$ for some $x_0 \geq 0$.

Lemma 3 *(Monotone Density Lemma)*. Suppose that $\mu(dx) = g(x)dx$ has an ultimately monotone density g. If $G_\mu(x) \sim \frac{1}{\Gamma(\theta+1)} x^\theta L(x)$ as $x \to \infty$, then $g(x) \sim \frac{x^{\theta-1}}{\Gamma(\theta)} L(x) \sim \theta G_\mu(x)/x$ as $x \to \infty$.

Proof. Assume that g is ultimately nondecreasing. Then, for arbitrary $0 < c < d < \infty$, for all x sufficiently large one may bound $\frac{G_\mu(dx)-G_\mu(cx)}{x^\theta L(x)} = \frac{\int_{cx}^{dx} g(y)dy}{x^\theta L(x)}$ above and below by

$$\frac{(d-c)xg(cx)}{x^\theta L(x)} \leq \frac{G_\mu(dx) - G_\mu(cx)}{x^\theta L(x)} \leq \frac{(d-c)xg(dx)}{x^\theta L(x)}.$$

Thus,

$$\begin{aligned}
\limsup_{x\to\infty} \frac{g(cx)}{x^{\theta-1}L(x)} &\leq \limsup_{x\to\infty} \frac{G_\mu(dx) - G_\mu(cx)}{(d-c)x^\theta L(x)} \\
&= \limsup_{x\to\infty} \left\{ \frac{G_\mu(dx)}{(dx)^\theta (d-c) L(dx)} d^\theta \frac{L(dx)}{L(x)} - \frac{G_\mu(cx)}{(cx)^\theta (d-c) L(cx)} c^\theta \frac{L(cx)}{L(x)} \right\} \\
&\to \frac{(d^\theta - c^\theta)}{d-c}.
\end{aligned}$$

Take $c = 1$ and let $d \downarrow 1$ to get the desired upper bound on the limsup. The same lower bound on the liminf is obtained by the same considerations applied to the other inequality. Finally, the case in which g is nonincreasing follows by the same argument but with reversed estimates for the upper and lower bounds. ■

The Tauberian theorem together with the monotone density lemma immediately yields the following consequence.

[2]A treatment of the problem with less-stringent conditions can be found in the more comprehensive monograph Bingham, N.H., C.M. Goldie, J.L. Teugels (1987).

Corollary 8.9. Suppose that $\mu(dx) = g(x)dx$ has an ultimately monotone density g. For L slowly varying at infinity and $0 \leq \theta < \infty$ one has

$$\hat{\mu}(\lambda) \sim \lambda^{-\theta} L\left(\frac{1}{\lambda}\right) \quad \text{as} \quad \lambda \downarrow 0 \quad \text{if and only if} \quad g(x) \sim \frac{1}{\Gamma(\theta)} x^{\theta-1} L(x) \quad \text{as} \quad x \to \infty.$$

Finally, for discrete measures one has the following asymptotic behavior conveniently expressed in terms of *(combinatorial) generating functions*, i.e., with $t = e^{-\lambda}$.

Corollary 8.10. Let $\tilde{\mu}(t) = \sum_{n=0}^{\infty} \mu_n t^n$, $0 \leq t < 1$, where $\{\mu_n\}_{n=0}^{\infty}$ is a sequence of nonnegative numbers. For L slowly varying at infinity and $0 \leq \theta < \infty$ one has

$$\hat{\mu}(t) \sim (1-t)^{-\theta} L\left(\frac{1}{1-t}\right) \quad \text{as} \quad t \uparrow 1$$

if and only if

$$\sum_{j=0}^{n} \mu_j \sim \frac{1}{\Gamma(\theta)} n^{\theta} L(n) \quad \text{as} \quad n \to \infty.$$

Moreover, if the sequence $\{\mu_n\}_{n=0}^{\infty}$ is ultimately monotone and $0 < \theta < \infty$, then equivalently,

$$\mu_n \sim \frac{1}{\Gamma(\theta)} n^{\theta-1} L(n) \quad \text{as} \quad n \to \infty.$$

Proof. Let $\mu(dx) = \sum_{n=0}^{\infty} \mu_n \mathbf{1}_{[n,n+1)}(x)dx$, with (improper) distribution function G_μ. Then $G_\mu(n) = \sum_{j=0}^{n} \mu_j$. Also

$$\hat{\mu}(\lambda) = \frac{1-e^{-\lambda}}{\lambda} \sum_{n=0}^{\infty} \mu_n e^{-n\lambda} = \frac{1-e^{-\lambda}}{\lambda} \tilde{\mu}(e^{-\lambda}).$$

The assertions now follow immediately from the Tauberian theorem and previous corollary. ■

EXERCISES

Exercise Set VIII

1. Establish the formula $\int_0^\infty g(x)\mu * \nu(dx) = \int_0^\infty \int_0^\infty g(x+y)\mu(dx)\nu(dy)$ for bounded Borel-measurable functions g used in the proof of the convolution property of Laplace transforms.
2. Show that size-biasing a Gaussian distribution corresponds to a shift in the mean.

3. Show that $g(t) = \frac{1}{1-t}$, $0 \le t < 1$, is absolutely monotone and $\varphi(\lambda) = \frac{e^\lambda}{e^\lambda - 1}$, $\lambda > 0$, is completely monotone. Calculate the measure μ with Laplace transform $\varphi(\lambda)$.

4. Show that if g is absolutely monotone on $[0,1]$ with $g(1) = 1$, then $p_{j,n} = \sum_{k=0}^{j} g\left(\frac{k}{n}\right)\binom{n}{k}\binom{n-k}{n-j}(-1)^{j-k}$ is a probability.

5. Show that (i) $|\log x|^r$, $x > 0$, is slowly varying at infinity and at 0 for any exponent r; (ii) $\log\log x, x > 1$, is slowly varying at ∞; (iii) $(1 + x^{-s})^r$, $x > 0$, is slowly varying at ∞ for any exponents r and $s > 0$.

6. Complete the proofs of the corollaries to the Tauberian theorem. [*Hint*: Note that $\frac{G_\mu(ax)}{G_\mu(a)} \sim x^\theta$ as $a \to \infty$ if and only if $L(x) = \frac{G_\mu(x)}{x^\theta}$ is slowly varying at infinity, and $\frac{\hat{\mu}(\alpha\lambda)}{\hat{\mu}(\alpha)} \sim \lambda^{-\theta}$ as $\alpha \to 0$ if and only if $\lambda^\theta \hat{\mu}(\lambda)$ varies slowly at 0.]

7. (*Renewal Equation Asymptotics*) Let μ be a probability on $[0,\infty)$ not concentrated at $\{0\}$, and suppose g is a nonnegative measurable function on $[0,\infty)$. Show that $u(t) = g * \mu(t) := \int_0^t g(t-s)\mu(ds)$, $t \ge 0$, satisfies the **renewal equation** $u(t) = g(t) + \int_0^t u(t-s)\mu(ds)$, $t \ge 0$. Show that if g is integrable on $[0,\infty)$ and μ has finite first moment m, then $u(t) \sim \{\frac{1}{m}\int_0^\infty g(s)ds\}t$ as $t \to \infty$. [*Hint*: Use the Tauberian theorem.]

8. (*Branching with Geometric Offspring*) Let $Y_{n,1}, Y_{n,2}, \ldots, Y_{n,n}$ be a triangular array of i.i.d. random variables having geometric (offspring) distribution $P(Y_{n,j} = k) = qp^k, k = 0, 1, 2 \ldots$. Recursively define $X_{n+1} = \sum_{j=1}^{X_n} Y_{n,j}\mathbf{1}_{[X_n \ge 1]}$, for $n \ge 0$, with $X_0 = 1$. Then X_{n+1} may be viewed as the number of offspring in the $(n+1)$st generation produced by ancestors in the nth generation. The geometric offspring assumption makes various explicit calculations possible that are otherwise impossible. Let $g_n(t) = \mathbb{E}t^{X_n}$, and $g_1(t) = g(t) = \mathbb{E}t^{Y_{n,j}}$ the generating function of the offspring distribution.
   (i) Show that $g_{n+1}(t) = g(g_n(t))$.
   (ii) For $p \ne q$ show that $g_n(t) = q\frac{p^n - q^n - (p^{n-1} - q^{n-1})pt}{p^{n+1} - q^{n+1} - (p^n - q^n)pt}$.
   (iii) For $p < \frac{1}{2}$, consider the *total progeny* defined by $N = \sum_{n=0}^\infty X_n$. Show that $P(N < \infty) = 1$. [*Hint*: Consider $P(X_n = 0) = g_n(0)$ and $[X_n = 0] \subseteq [X_{n+1} = 0]$.]
   (iv) For $p < \frac{1}{2}$, let $h(t) = \mathbb{E}t^N$ be the generating function for the total progeny. Show that $h(t) = tg(h(t)) = \frac{qh(t)}{1 - ph(t)}$, $0 < t < 1$. [*Hint*: Consider the generating functions $h_n(t) = \mathbb{E}t^{\sum_{j=0}^n X_j}$, $n \ge 0$, in the limit as $n \to \infty$.]
   (v) For $p < \frac{1}{2}$, show that $h(t) = \frac{1 - \sqrt{1 - 4pqt}}{2p}$, $0 < t < 1$. [*Hint*: Solve the quadratic equation implied by the preceding calculation.]
   (vi) Show that $\sum_{k=1}^n \frac{k}{(4pq)^k} P(N = k) \sim \frac{1}{p\sqrt{\pi}} n^{\frac{1}{2}}$ as $n \to \infty$. [*Hint*: Apply the Tauberian theorem to $h'(\frac{t}{4pq})$ and use properties of the gamma function: $\Gamma(x+1) = x\Gamma(x), \Gamma(\frac{1}{2}) = \sqrt{\pi}$.]

9. Show that under the hypothesis of Theorem 8.5, the sequence of probabilities $\{\mu_n : n \ge 1\}$ is tight. [*Hint*: Given $\varepsilon > 0$ there exists $\lambda_\varepsilon > 0$ such that $\hat{\mu}_n(\lambda_\varepsilon) \ge 1 - \frac{\varepsilon}{2}$ for all n. Now find $M = M_\varepsilon$ such that $e^{-\lambda_\varepsilon M} < \frac{\varepsilon}{2}$. Then $\mu_n[0, M] \ge 1 - \varepsilon$ for all n.]

10. Show that the series (8.8) converges uniformly on $[0, a]$ for all $a < 1$. [*Hint*: Check that the series increases monotonically to $\varphi(h(1-t))$ and apply Dini's theorem from advanced calculus.]

11. (i) Show that under the hypothesis of part (b) of Theorem 8.5 the sequence of probabilities $\{\mu_n : n \ge 1\}$ is tight. [*Hint*: Given $\varepsilon > 0$ there exists $\lambda_\varepsilon > 0$ such that $\hat{\mu}_n(\lambda_\varepsilon) \ge 1 - \frac{\varepsilon}{2}$

for all n. Now find $M = M_\varepsilon$ such that $e^{-\lambda_\varepsilon M} < \varepsilon/2$. Then $\mu_n[0, M] \geq 1 - \varepsilon$ for all n.]
(ii) Give an example to show that the boundedness of $\hat{\mu}_n(c)$ is necessary in part (a). [*Hint*: Consider point-mass measures μ_n at positive integers n.]