7 Fourier Transforms

7.1 Preliminaries

The exponential function $e^z = \exp(z)$ is *defined* by the power series[6] $e^z = \sum_{n=0}^{\infty} z^n / n!$. This series converges absolutely for every complex number z, and uniformly in every disk $|z| \leq R$ of finite radius $0 \leq R < \infty$. The following elementary properties of the exponential function will be taken as *known*.

Properties of the Exponential Function:

(1) $\exp(z) = \sum_{n=0}^{\infty} z^n / n!$
(2) $\exp(z + w) = \exp(z) + \exp(w)$ for all $z, w \not\mathbb{C}$.
(3) $\exp(i\theta) = \cos\theta + i \sin\theta$ for all $\theta \in \mathbb{R}$.
(4) $(d/dz)\exp(z) = \exp(z)$.
(5) $\exp(\log t) = \log(\exp(t)) = t$ for all $t > 0$.

The infinite series (1) converges absolutely and uniformly in the disk $|z| \leq R$, for any $R < \infty$. The identity (3) is *Euler's formula*; it implies that $|e^{i\theta}| = 1$ for any real θ, and also that the mapping $\theta \mapsto e^{i\theta}$ from $\mathbb{R}$ to the unit circle $\{|\xi| = 1\}$ is 2π−periodic. Euler's formula also implies that every complex number $z = x + iy$ has a *polar* representation $z = |z|e^{i\theta}$, where $|z| = \sqrt{x^2 + y^2}$ and $\theta = \arctan(y/x)$.

Exercise 7.1. Prove that for any sequence z_n of complex numbers,

$$\lim_{n\to\infty} z_n = z \quad \Longrightarrow \quad \lim_{n\to\infty} \left(1 + \frac{z_n}{n}\right)^n = e^z. \tag{7.1}$$

Exercise 7.2. The integral of a Borel-measurable complex-valued function $f : \Omega \to \mathbb{C}$ on a measure space $(\Omega, \mathscr{F}, \mu)$ is defined by

$$\int f\,d\mu = \int \Re f\,d\mu + i \int \Im f\,d\mu.$$

Show that the integral is well-defined for any function f such that the real-valued function $|f|$ is integrable, and that

$$\left|\int f\,d\mu\right| \leq \int |f|\,d\mu. \tag{7.2}$$

[6] The series should be easy for a probabilist or statistician to remember, because for *positive* values of z the equation $e^z = \sum_{n=0}^{\infty} z^n / n!$ is equivalent (after you divide both sides by e^z) to the fact that the Poisson distribution $n \mapsto z^n e^{-z} / n!$ with mean parameter z is actually a probability distribution on the nonnegative integers.

7.2 Characteristic Functions and Fourier Transforms

Definition 7.3. Let μ be a finite Borel measure on $\mathbb{R}$ and let $f \in L^1(\mathbb{R})$ be an integrable function (relative to Lebesgue measure). The *Fourier transforms* $\hat{\mu}$ and $\hat{f}$ of μ and f are the complex-valued functions

$$\hat{\mu}(\theta) := \int e^{i\theta x}\, d\mu(x) \quad \text{and} \tag{7.3}$$

$$\hat{f}(\theta) := \int_{\mathbb{R}} e^{i\theta x} f(x)\, dx. \tag{7.4}$$

The functions $\hat{\mu}$ and $\hat{f}$ are well-defined for all arguments $\theta \in \mathbb{R}$; in fact, the inequality (7.2) implies that they are bounded functions:

$$|\hat{\mu}(\theta)| \le \hat{\mu}(0) = \mu(\mathbb{R}) \quad \text{and} \quad |\hat{f}(\theta)| \le \|f\|_1.$$

Moreover, both $\hat{\mu}$ and $\hat{f}$ are *uniformly continuous* on $\mathbb{R}$, by a routine application of the dominated convergence theorem.

Definition 7.4. The *characteristic function* of a real random variable X is the Fourier transform of its distribution, equivalently, it is the function $\varphi_X : \mathbb{R} \to \mathbb{C}$ defined by

$$\varphi_X(\theta) = Ee^{i\theta X}. \tag{7.5}$$

Proposition 7.5. *If $X_1, X_2, \dots$ are independent integer-valued random variables with partial sumes $S_n = \sum_{k=1}^n X_k$ then*

$$\varphi_{S_n}(\theta) = \prod_{k=1}^{n} \varphi_{X_k}(\theta). \tag{7.6}$$

Therefore, in particular, if the random variables $X_1, X_2, \dots$ are independent and identically distributed then

$$\varphi_{S_n}(\theta) = \varphi_{X_1}(\theta)^n. \tag{7.7}$$

This follows immediately from the multiplication rule for expectations of products of independent random variables. There are corresponding product rules for the convolutions of finite Borel measures and L^1 functions: in particular, for any two finite Borel measures μ, ν on $\mathbb{R}$ and any two L^1 functions f, g,

$$\widehat{\mu * \nu}(\theta) = \hat{\mu}(\theta)\hat{\nu}(\theta) \quad \text{and} \tag{7.8}$$

$$\widehat{f * g}(\theta) = \hat{f}(\theta)\hat{g}(\theta). \tag{7.9}$$

These both follow by routine application of Fubini's theorem.

Proposition 7.6. *Let μ be a finite Borel measure on $\mathbb{R}$ with finite kth moment for some integer $k \ge 1$. Then the Fourier transform $\hat{\mu}(\theta)$ has k bounded, continuous derivatives, and these are given by the formula*

$$\frac{d^k}{d\theta^k}\hat{\mu}(\theta) = \int (ix)^k e^{i\theta x}\, d\mu(x). \tag{7.10}$$

Proof. It suffices (by induction and the linearity of the integral) to prove this for the case $k = 1$. For this, take difference quotients and use the dominated convergence theorem. $\square$

In the special case where μ is a probability measure, the formula (7.10), specialized to $\theta = 0$, shows that the moments of a random variable X can be read off from the derivatives of the characteristic function $\phi_X(\theta)$ at 0:

$$i^k EX^k = \frac{d^k}{d\theta^k}\varphi_X(0). \tag{7.11}$$

Proposition 7.7. *Let $f : \mathbb{R} \to \mathbb{R}$ be a C^∞ function such that f and its first two derivative f' are integrable. Then*

$$\sup_{\theta\in\mathbb{R}} |\theta \hat{f}(\theta)| \le \|f'\|_1 \quad \textit{and} \quad \hat{f}(\theta) = \frac{-1}{i\theta}\widehat{f'}(\theta). \tag{7.12}$$

Proof. Exercise. (Hint: Integrate by parts. The smoothness hypothesis guarantees that the Fourier transforms of f and all its derivatives decay rapidly at ∞.) $\square$

Example 7.8. Let f be the standard normal probability density. The function f is infinitely differentiable, and all of its derivatives are integrable, so (7.12) applies. Now

$$\begin{aligned}
\hat{f}'(\theta) &= \int e^{i\theta x}(-xe^{-x^2/2})\,dx/\sqrt{2\pi} \\
&= \int xe^{i\theta x}(-e^{-x^2/2})\,dx/\sqrt{2\pi} \\
&= \int \left(-i\frac{d}{d\theta}e^{i\theta x}\right)(-e^{-x^2/2})\,dx/\sqrt{2\pi} \\
&= i\frac{d}{d\theta}\hat{f}(\theta).
\end{aligned}$$

Thus, (7.12) shows that the Fourier transform $\hat{f}$ satisfies the first-order differential equation

$$\hat{f}(\theta) = -\frac{1}{\theta}\frac{d}{d\theta}\hat{f}(\theta).$$

Together with the auxiliary condition $\hat{f}(0) = 1$, this differential equation implies that

$$\hat{f}(\theta) = e^{-\theta^2/2}. \tag{7.13}$$

Example 7.9. The Fourier transform of the double-exponential density $f(x) = \frac{1}{2}e^{-|x|}$ is gotten by a simple integration:

$$\hat{f}(\theta) = \frac{1}{1+\theta^2}.$$

This is proportional to the *Cauchy density*. Later we will see (by the Fourier inversion theorem) that the reverse is also true: the Fourier transform of the Cauchy density is the double exponential function.

7.3 $\mathbb{Z}-$ Valued Random Variables

If X is an integer-valued random variable then its characteristic function $\varphi_X(\theta)$ is given by the *Fourier series*

$$\varphi_X(\theta) = Ee^{i\theta X} = \sum_{-\infty}^{\infty} P(X = m)e^{i\theta m}. \tag{7.14}$$

This function is $2\pi-$periodic. In this case, the distribution of the random variable X can be recovered from the characteristic function by using the following *orthogonality relations* for the complex exponential functions $e^{im\theta}$.

Proposition 7.10. *For $m \in \mathbb{Z}$,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{im\theta}\,d\theta = \delta_{m,0}$$

where $\delta_{m,0}$ is the Kronecker delta function (0 if $m \neq 0$ and 1 is $m = 0$).

This is proved by a trivial integration. Despite the easy proof, the result is of fundamental importance, as it implies that the complex exponential functions $e^{im\theta}$ (also known as the *group characters* for the additive group $\mathbb{Z}$) are *orthonormal* in $L^2(d\theta/2\pi)$.

Corollary 7.11. *For an integer-valued random variable X with characteristic function $\varphi_X(\theta)$ the probability distribution is given by*

$$P(X = m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-im\theta}\varphi_X(\theta)\,d\theta.$$

7.4 Fourier Inversion

We have seen that certain of the important features of a measure are captured by its Fourier transform: for instance, if μ is a probability measure then its moments (when they are finite) can be determined from the derivatives of $\hat{\mu}$ at $\theta = 0$. In this section we will address more systematically the general problem of recovering information about a measure μ from its Fourier transform. Following is a list of the main theorems to be proved.

Theorem 7.12. *(Uniqueness) Finite Borel measures are uniquely determined by their Fourier transforms, that is, if two Borel measures have the same Fourier transform then they are the same measure.*

Theorem 7.12 will follow immediately from the following theorem, which gives an explicit formula showing how the the integral of a bounded, continuous function f against a finite measure μ can be recovered from the Fourier transforms of f and μ.

Theorem 7.13. *(Plancherel-Parseval Formula) For any finite Borel measure μ and any bounded, continuous function $f : \mathbb{R} \to \mathbb{R}$ with compact support,*

$$\int f(x)\, d\mu(x) = \lim_{\varepsilon \to 0} \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\theta)\hat{\mu}(-\theta) e^{-\varepsilon^2\theta^2/2}\, d\theta. \tag{7.15}$$

The hypothesis that f have compact support is needed to guarantee that the Fourier transform $\hat{f}$ is well-defined. The factor $e^{-\varepsilon^2\theta^2/2}$ in the integral is needed because in general the Fourier transforms $\hat{f}$ and $\hat{\mu}$, although bounded, are not integrable. As $\varepsilon \to 0$ these factors increase to 1, so if the product $\hat{f}(\theta)\hat{\mu}(-\theta)$ happens to be integrable then it will follow, by the dominated convergence theorem, that

$$\int f(x)\, d\mu(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\theta)\hat{\mu}(-\theta)\, d\theta. \tag{7.16}$$

Corollary 7.14. *If $f : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable and has compact support then for every finite Borel measure μ the identity* (7.16) *holds.*

Proof. If f is C^2 and has compact support then its Fourier transform decays at least as fast as C/θ^2 as $|\theta| \to \infty$. (This can be seen by doing two successive integrations by parts in the integral defining $\hat{f}$.) Consequently, $\hat{f}$ is integrable, and since the Fourier transform $\hat{\mu}$ of a finite Borel measure is always bounded, it then follows that the product $\hat{f}(\theta)\hat{\mu}(-\theta)$ is integrable. □

If the measure μ has an integrable Fourier transform $\hat{\mu}$ then the identity (7.16) will hold for every continuous function f with compact support, regardless of smoothness. In this case, even more can be said:

Theorem 7.15. *(Existence of Densities) If the Fourier transform $\hat{\mu}(\theta)$ of a finite Borel measure is an integrable function (with respect to Lebesgue measure on $\mathbb{R}$) then the measure μ has a bounded, uniformly continuous density g given by*

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\mu}(\theta) e^{-i\theta x}\, d\theta. \tag{7.17}$$

Example 7.16. Let μ be the probability measure on $\mathbb{R}$ with the *double-exponential* density $g(x) = e^{-|x|}/2$. A routine integration shows that

$$\hat{\mu}(\theta)(= \hat{g}(\theta)) = \frac{1}{1+\theta^2}.$$

Consequently, the inversion formula (7.17) implies that if ν is the probability measure with the *Cauchy density* $\pi^{-1}(1+x^2)^{-1}$ then

$$\hat{\nu}(\theta) = e^{-|\theta|}.$$

This example is interesting (and important) for a number of reasons. First, it shows that the Fourier transform of a probability distribution with infinite first moment need no be differentiable at $\theta = 0$. Second, it shows that if $X_1, X_2, \ldots, X_n$ are independent, identically distributed, all with the Cauchy distribution ν, then the sample average S_n/n also has the Cauchy distribution. To see this, observe that by the multiplication rule for characteristic functions of independent random variables, the characteristic function of S_n/n is $e^{-|\theta|}$; hence, by the uniqueness theorem (Theorem 7.12), the distribution of S_n/n must be the same as that of X_1.

Proofs of the Main Theorems

Fourier analysis on the real line $\mathbb{R}$ is complicated by the fact that Fourier transforms need not be integrable, and so it is often necessary to multiply by "convergence factors", such as $e^{-\theta^2\varepsilon^2/2}$, to obtain convergent Fourier integrals. The choice of the normal density is useful not only because it has such rapid decrease at ∞ but because *the normal density is its own Fourier transform* (apart from the always annoying factors of 2π): if Z is standard normal, then

$$Ee^{i\theta Z} = e^{-\theta^2/2}.$$

Thus, if

$$\varphi_\varepsilon(x) := \frac{1}{\sqrt{2\pi}\varepsilon} e^{-x^2/2\varepsilon^2} \quad \text{then} \quad \hat{\varphi}_\varepsilon(\theta) = e^{-\varepsilon^2\theta^2/2}.$$

Lemma 7.17. *If $f : \mathbb{R} \to \mathbb{R}$ is a continuous function with compact support, then for every $\varepsilon > 0$*

$$f * \varphi_\varepsilon(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f * \varphi_\varepsilon}(\theta) e^{-i\theta x}\, d\theta. \tag{7.18}$$

Remark 7.18. This equality asserts that every function of the form $f * \varphi_\varepsilon$ can be recovered from its Fourier transform by *Fourier inversion*. It can be shown that the same is true for any C^2 function with compact support. Since we will not need this fact, we won't prove it.

Proof of Lemma 7.17. Fubini. □

Lemma 7.19. *For any bounded, continuous function $f : \mathbb{R} \to \mathbb{R}$,*

$$\lim_{\varepsilon \to 0} f * \varphi_\varepsilon(x) = f(x). \tag{7.19}$$

Proof. Let Z be a random variable defined on some probability space $(\Omega, \mathscr{F}), P$ with standard Gaussian distribution $N(0,1)$; then

$$f * \varphi_\varepsilon(x) = Ef(x + \varepsilon Z).$$

Hence, the convergence (7.19) follows by the dominated convergence theorem. □

Proof of the Plancherel-Parseval Formula (Theorem 7.13). By Lemma 7.19 and the dominated convergence theorem,

$$\int f(x)\,d\mu(x) = \lim_{\varepsilon\to 0}\int (f*\varphi_\varepsilon)(x)\,d\mu(x).$$

But by Lemma 7.17 and Fubini's theorem,

$$\begin{aligned}\int (f*\varphi_\varepsilon)(x)\,d\mu(x) &= \frac{1}{2\pi}\iint \widehat{f*\varphi_\varepsilon}(\theta)e^{-i\theta x}\,d\theta\,d\mu(x)\\ &= \frac{1}{2\pi}\iint \widehat{f*\varphi_\varepsilon}(\theta)e^{-i\theta x}\,d\mu(x)\,d\theta\\ &= \frac{1}{2\pi}\int \widehat{f*\varphi_\varepsilon}(\theta)\hat{\mu}(-\theta)\,d\theta\\ &= \frac{1}{2\pi}\int \hat{f}(\theta)e^{-\varepsilon^2\theta^2/2}\hat{\mu}(-\theta)\,d\theta.\end{aligned}$$

□

Proof of Theorem 7.15. Assume that the Fourier transform $\hat{\mu}(\theta)$ is an integrable function, and define $g(x)$ by (7.17). To prove that μ has density g it suffices to prove that for any continuous function $f:\mathbb{R}\to\mathbb{R}$ with compact support,

$$\int f(x)\,d\mu(x) = \int f(x)g(x)\,dx.$$

The function g is bounded and continuous, by the usual arguments, but *a priori* we do not know that it is integrable. Nevertheless, if f is any bounded continuous function with compact support then fg is integrable, since g is bounded, and furthermore, for any $\varepsilon>0$ the function $f*\varphi_\varepsilon$ is integrable (Fubini) and so $g(x)(f*\varphi_\varepsilon)(x)$ is also integrable. By definition of g and Fubini, for any $\varepsilon>0$,

$$\begin{aligned}\int f*\varphi_\varepsilon(x)g(x)\,dx &= \frac{1}{2\pi}\iint f*\varphi_\varepsilon(x)e^{-i\theta x}\hat{\mu}(\theta)\,d\theta\,dx\\ &= \frac{1}{2\pi}\iint f*\varphi_\varepsilon(x)e^{-i\theta x}\hat{\mu}(\theta)\,dx\,d\theta\\ &= \frac{1}{2\pi}\int \hat{f}(-\theta)\hat{\varphi}_\varepsilon(-\theta)\hat{\mu}(\theta)\,d\theta.\end{aligned}$$

As $\varepsilon\to 0$ the last integral converges to $\int f\,d\mu$. Therefore, to complete the proof it suffices to show that

$$\lim_{\varepsilon\to 0}\int f*\varphi_\varepsilon(x)g(x)\,dx = \int f(x)g(x)\,dx.$$

But by Fubini, for any $\varepsilon > 0$

$$\begin{aligned}
\int f * \varphi_\varepsilon(x) g(x)\, dx &= \iint f(y)\varphi_\varepsilon(x-y) g(x)\, dy dx \\
&= \iint f(y)\varphi_\varepsilon(x-y) g(x)\, dx dy \\
&= \int f(y)(g * \varphi_\varepsilon)(y)\, dy.
\end{aligned}$$

Lemma 7.19 implies that as $\varepsilon \to 0$ the integrand converges to $f(y)g(y)$, and because f is integrable and $\|g * \varphi_\varepsilon\|_\infty \le \|g\|_\infty < \infty$ the dominated convergence theorem implies that

$$\lim_{\varepsilon \to 0} \int f(y)(g * \varphi_\varepsilon)(y)\, dy = \int f(y) g(y)\, dy.$$

□

7.5 Fourier Transforms and Weak Convergence

Recall that a sequence $\{\mu_n\}_{n\ge 1}$ of Borel probability measures on $\mathbb{R}$ converges weakly to a Borel probability measure μ (written $\mu_n \Longrightarrow \mu$) if for every continuous function $u : \mathbb{R} \to \mathbb{R}$ with compact support,

$$\lim_{n\to\infty} \int u\, d\mu_n = \int u\, d\mu. \tag{7.20}$$

Recall also that a *sufficient* condition for this is that relation (7.20) holds for all C^∞ functions u with compact support. By the Parseval relation (7.16) (cf. Corollary 7.14),

$$\begin{aligned}
\int u(x)\, d\mu(x) &= \frac{1}{2\pi} \int \hat{u}(\theta)\hat{\mu}(\theta)\, d\theta \quad \text{and} \\
\int u(x)\, d\mu_n(x) &= \frac{1}{2\pi} \int \hat{u}(\theta)\hat{\mu_n}(\theta)\, d\theta;
\end{aligned}$$

consequently, since the Fourier transform $\hat{u}$ decays rapidly at ∞ (cf. Homework), the convergence (7.20) will hold for any compactly supported C^∞ function u if

$$\lim_{n\to\infty} \hat{\mu_n}(\theta) = \hat{\mu}(\theta) \quad \text{for every } \theta \in \mathbb{R}. \tag{7.21}$$

This proves

Theorem 7.20. *(Lévy's Continuity Theorem) A sequence μ_n of Borel probability measures on $\mathbb{R}$ converge weakly to a Borel probability μ if and only if for every $\theta \in \mathbb{R}$,*

$$\lim_{n\to\infty} \hat{\mu}_n(\theta) = \hat{\mu}(\theta). \tag{7.22}$$

One can use Lévy's theorem as the basis for another proof of the central limit theorem.

Theorem 7.21. *Let $X_1, X_2, \dots$ be independent, identically distributed random variables with $EX_i = 0$ and $EX_i^2 = 1$, and write $S_n = \sum_{i=1}^n X_i$. Then the random variables $S_n/\sqrt{n}$ converge in distribution to the standard normal distribution, that is, if μ_n is the distribution of $S_n/\sqrt{n}$ then*

$$\mu_n \Longrightarrow \mu \quad \textit{where} \quad \mu(B) = \frac{1}{\sqrt{2\pi}} \int_B e^{-x^2/2}\, dx. \tag{7.23}$$

Proof. It suffices to show that $Ee^{i\theta S_n/\sqrt{n}} \to e^{-\theta^2/2}$ for every θ. The random variables X_i have mean 0 and variance 1, so by Taylor's theorem their characteristic function $\varphi(\theta) := Ee^{i\theta X_n}$ satisfies

$$\varphi(\theta) = 1 - \frac{\theta^2}{2} \varphi''(\beta(\theta))$$

where $\beta(\theta)$ is a point of the real line intermediate between 0 and θ. Consequently, for any $\theta \in \mathbb{R}$,

$$\varphi(\theta/\sqrt{n}) = 1 - \frac{\theta^2}{2n}(1 + o(1)).$$

Therefore,

$$Ee^{i\theta S_n/\sqrt{n}} = (1 - \frac{\theta^2}{2n}(1 + o(1)))^n \longrightarrow e^{-\theta^2/2}.$$

□

7.6 Lévy's Continuity Theorem

Lévy actually proved a much more useful extension of the Continuity Theorem 7.20 that can be used, in certain situations, to show that a given function is the Fourier transform of a Borel probability measure on $\mathbb{R}$.

Theorem 7.22. *The pointwise limit of a sequence of characteristic functions is a characteristic function if it is continuous at $\theta = 0$.*

Proof. This will be based on the Helly selection principle (Theorem 6.16). The strategy is as follows. Let $\varphi_n(\theta) = \hat{\mu}_n(\theta)$, where each μ_n is a Borel probability measure, and suppose that $\varphi_n(\theta) \to \varphi(\theta)$ for each $\theta \in \mathbb{R}$. If we knew that the sequence μ_n were tight, then Helly's theorem would imply that for any subsequence there is a weakly convergent subsequence – and in particular, the weak limit would be a Borel probability measure. But any such weak limit must have Fourier transform $\varphi(\theta)$; thus, by the Uniqueness Theorem, *there is only one possible weak limit* μ, and its Fourier transform must be $\varphi(\theta)$.

The gap in this argument is the tightness of the sequence μ_n. It is here that we will need the hypothesis that the limit function $\varphi(\theta)$ is continuous at $\theta = 0$. To make use of this hypothesis, we will find it convenient to utilize another characterization of tightness.

Lemma 7.23. *A sequence of probability measures μ_n on $\mathbb{R}$ is tight if and only if for every $\varepsilon > 0$ there exists $\sigma^2 < \infty$ such that*

$$\int_{-\infty}^{\infty} e^{-x^2/2\sigma^2}\, d\mu_n(x) > 1 - \varepsilon. \tag{7.24}$$

Proof. Exercise. □

Corollary 7.24. *Let μ_n be a sequence of Borel probability measures on $\mathbb{R}$ whose Fourier transforms $\hat{\mu}_n(\theta)$ converge pointwise to a function $\varphi(\theta)$ that is* continuous *at $\theta = 0$. Then the sequence μ_n is tight.*

Proof. Recall that the Plancherel-Parseval identity (7.16) holds for any bounded, continuous function f whose Fourier transform $\hat{f}(\theta)$ is integrable. This is the case for the function $f_\sigma(x) = \exp\{-x^2/2\sigma^2\}$: its Fourier transform is

$$\hat{f}_\sigma(\theta) = \sqrt{2\pi}\sigma \exp\{-\theta^2\sigma^2/2\},$$

which is certainly integrable. Thus, for any $\sigma > 0$ and each $n = 1, 2, \ldots,$

$$\int_{-\infty}^{\infty} e^{-x^2/2\sigma^2}\, d\mu_n(x) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\sigma^2\theta^2/2} \hat{\mu}_n(-\theta)\, d\theta = E\hat{\mu}_n(Z/\sigma),$$

where Z is a standard normal random variable.

By hypothesis, the functions $\hat{\mu}_n(\theta)$ converge pointwise to $\varphi(\theta)$, and the function $\varphi(\theta)$ is continuous at $\theta = 0$. Now since $\hat{\mu}_n(0) = 1$ for every n, it must also be the case that $\varphi(0) = 1$; and since $|\hat{\mu}_n(\theta)| \le 1$ for every $\theta \in \mathbb{R}$, it must also be the case that $|\varphi(\theta)| \le 1$. Hence, by the bounded convergence theorem,

$$\lim_{\sigma\to\infty} E\varphi(Z/\sigma) = \varphi(0) = 1,$$

and so for any $\varepsilon > 0$ there exists $\sigma < \infty$ such that

$$E\varphi(Z/\sigma) \ge 1 - \varepsilon.$$

But the pointwise convergence $\hat{\mu}_n(\theta) \to \varphi(\theta)$, together with another application of the bounded convergence theorem, implies that

$$\lim_{n\to\infty} E\hat{\mu}_n(Z/\sigma) = E\varphi(Z/\sigma) \ge 1 - \varepsilon.$$

Therefore, for all sufficiently large n,

$$E\hat{\mu}_n(Z/\sigma) > 1 - 2\varepsilon,$$

and so by Lemma 7.23, the sequence μ_n is tight. □

□

7.7 The Symmetric Stable Laws

Theorem 7.25. *For every real number $\alpha \in (0,2)$ the function $\varphi_\alpha(\theta) = \exp\{-|\theta|^\alpha\}$ is a characteristic function.*

Proof. Fix α, and let$X_1, X_2, \dots$ be independent, identically distributed random variables with probability density $f_\alpha(x) = \frac{1}{2}\alpha|x|^{-\alpha-1}\mathbf{1}_{[1,\infty)}(|X|)$, and set $S_n = \sum_{i=1}^n X_i$. We will show that the characteristic functions of the random variables $S_n/n^{1/\alpha}$ converge pointwise to $\varphi_\alpha(\theta)$; Theorem 7.22 will then imply that φ_α is a characteristic function.

The characteristic function of X_i is the Fourier transform of the density f_α:

$$\varphi_{X_i}(\theta) = \hat{f}_\alpha(\theta) = \alpha \int_1^\infty \frac{\cos(|\theta|x|)}{x^{\alpha+1}}\, dx.$$

Consequently, as $|\theta| \to 0$,

$$1 - \varphi_{X_i}(\theta) = \alpha \int_1^\infty \frac{1-\cos(|\theta|x|)}{x^{\alpha+1}}\, dx = \alpha|\theta|^\alpha \int_{|\theta|}^\infty \frac{1-\cos y}{y^{1+\alpha}}\, dy \sim C|\theta|^\alpha,$$

where

$$C = \int_0^\infty \frac{1-\cos y}{y^{1+\alpha}}\, dy;$$

the relation $\sim$ follows from the dominated convergence theorem, using the fact that the (dominating) function $(1-\cos y)y^{1+\alpha}$ is nonnegative and integrable on $(0,\infty)$. (This uses the hypothesis that $0 < \alpha < 2$.) Therefore, for any $\theta \in \mathbb{R}$ and any $n \geq 1$,

$$\varphi_{S_n/n^{1/\alpha}}(\theta) = \varphi_{X_1}(\theta/n^{1/\alpha})^n = \left(1 - \frac{C|\theta|^\alpha}{n}\right)^n,$$

and as $n \to \infty$ this converges to $\exp\{-|\theta|^{1\alpha}\}$. □

Exercise 7.26. Let $g(x)$ be any even probability density on $\mathbb{R}$ such that $g(x) = f_\alpha(x)$ for all x outside some compact interval $[-A, A]$. Show that if $Y_1, Y_2, \dots$ are independent, identically distributed with density g then

$$n^{-1/\alpha} \sum_{i=1}^n Y_i \Longrightarrow \mu_\alpha$$

where μ_α is the symmetric stable law of exponent $\alpha \in (0,2)$, that is, the unique probability measure with Fourier transform $\varphi_\alpha(\theta) = \exp\{|\theta|^\alpha\}$. HINT: Show that the random variables Y_i can be constructed on a probability space that supports independent, identically distributed random variables X_i with density f_α in such a way that $X_i = Y_i$ on the event that $|X_i| > A$ or $|Y_i| > A$.