Chapter 6

Characteristic functions

We define the characteristic function of a random variable $X$ by $\varphi_X(t) = \mathbb{E} e^{itx}$ for $t \in \mathbb{R}$.

Note that $\varphi_X(t) = \int e^{itx} \mathbb{P}_X(dx)$. So if $X$ and $Y$ have the same law, they have the same characteristic function. Also, if the law of $X$ has a density, that is, $\mathbb{P}_X(dx) = f_X(x) \, dx$, then $\varphi_X(t) = \int e^{itx} f_X(x) \, dx$, so in this case the characteristic function is the same as (one definition of) the Fourier transform of $f_X$.

**Proposition 6.1** $\varphi(0) = 1$, $|\varphi(t)| \leq 1$, $\varphi(-t) = \overline{\varphi(t)}$, and $\varphi$ is uniformly continuous.

**Proof.** Since $|e^{itx}| \leq 1$, everything follows immediately from the definitions except the uniform continuity. For that we write

$$|\varphi(t + h) - \varphi(t)| = |\mathbb{E} e^{i(t+h)x} - \mathbb{E} e^{ix} | \leq \mathbb{E} |e^{itx} (e^{ihX} - 1)| = \mathbb{E} |e^{ihX} - 1|.$$ 

$|e^{ihX} - 1|$ tends to 0 almost surely as $h \to 0$, so the right hand side tends to 0 by dominated convergence. Note that the right hand side is independent of $t$. \hfill $\square$

**Proposition 6.2** $\varphi_{aX}(t) = \varphi_X(at)$ and $\varphi_{X+b}(t) = e^{itb} \varphi_X(t)$,
Proof. The first follows from $\mathbb{E} e^{it(aX)} = \mathbb{E} e^{i(at)X}$, and the second is similar.

Proposition 6.3 If $X$ and $Y$ are independent, then $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

Proof. From the multiplication theorem,

$$\mathbb{E} e^{it(X+Y)} = \mathbb{E} e^{itX}e^{itY} = \mathbb{E} e^{itX}\mathbb{E} e^{itY}.$$ 

Note that if $X_1$ and $X_2$ are independent and identically distributed, then

$$\varphi_{X_1-X_2}(t) = \varphi_{X_1}(t)\varphi_{-X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(-t) = \varphi_{X_1}(t)\overline{\varphi_{X_2}(t)} = |\varphi_{X_1}(t)|^2.$$

Let us look at some examples of characteristic functions.

(a) Bernoulli: By direct computation, this is $p e^{it} + (1-p) = 1 - p(1-e^{it})$.

(b) Coin flip: (i.e., $\mathbb{P}(X = +1) = \mathbb{P}(X = -1) = 1/2$) We have $\frac{1}{2}e^{it} + \frac{1}{2}e^{-it} = \cos t$.

(c) Poisson:

$$\mathbb{E} e^{itX} = \sum_{k=0}^{\infty} e^{itk}e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} = e^{-\lambda}e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}.$$ 

(d) Point mass at $a$: $\mathbb{E} e^{itX} = e^{ita}$. Note that when $a = 0$, then $\varphi \equiv 1$.

(e) Binomial: Write $X$ as the sum of $n$ independent Bernoulli r.v.s $B_i$. So

$$\varphi_X(t) = \prod_{i=1}^{n} \varphi_{B_i}(t) = [\varphi_{B_i}(t)]^n = [1 - p(1-e^{it})]^n.$$ 

(f) Geometric:

$$\varphi(t) = \sum_{k=0}^{\infty} p(1-p)^k e^{itk} = p \sum_{k=0}^{\infty} ((1-p)e^{it})^k = \frac{p}{1 - (1-p)e^{it}}.$$
6.1. INVERSION FORMULA

(g) Uniform on \([a, b]\):

\[ \varphi(t) = \frac{1}{b - a} \int_a^b e^{itx}dx = \frac{e^{itb} - e^{ita}}{(b - a)it}. \]

Note that when \(a = -b\) this reduces to \(\sin(bt)/bt\).

(h) Exponential:

\[ \int_0^\infty \lambda e^{itx}e^{-\lambda x} dx = \lambda \int_0^\infty e^{(it-\lambda)x} dx = \frac{\lambda}{\lambda - it}. \]

(i) Standard normal:

\[ \varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{itx}e^{-x^2/2}dx. \]

This can be done by completing the square and then doing a contour integration. Alternately, \(\varphi'(t) = (1/\sqrt{2\pi}) \int_{-\infty}^\infty ixe^{itx}e^{-x^2/2}dx\). (do the real and imaginary parts separately, and use the dominated convergence theorem to justify taking the derivative inside.) Integrating by parts (do the real and imaginary parts separately), this is equal to \(-t\varphi(t)\). The only solution to \(\varphi'(t) = -t\varphi(t)\) with \(\varphi(0) = 1\) is \(\varphi(t) = e^{-t^2/2}\).

(j) Normal with mean \(\mu\) and variance \(\sigma^2\): Writing \(X = \sigma Z + \mu\), where \(Z\) is a standard normal, then \(\varphi_X(t) = e^{i\mu t}\varphi_Z(\sigma t) = e^{i\mu t - \sigma^2 t^2/2}\).

(k) Cauchy: We have

\[ \varphi(t) = \frac{1}{\pi} \int \frac{e^{itx}}{1 + x^2}dx. \]

This is a standard exercise in contour integration in complex analysis. The answer is \(e^{-|t|}\).

6.1 Inversion formula

We need a preliminary real variable lemma, and then we can proceed to the inversion formula, which gives a formula for the distribution function in terms of the characteristic function.
Lemma 6.4 (a) \( \int_0^N (\sin(Ax)/x) \, dx \to \text{sgn}(A) \pi/2 \) as \( N \to \infty \).

\[ \sup_a | \int_0^a (\sin(Ax)/x) \, dx | < \infty. \]

Proof. If \( A = 0 \), this is clear. The case \( A < 0 \) reduces to the case \( A > 0 \) by the fact that \( \sin \) is an odd function. By a change of variables \( y = Ax \), we reduce to the case \( A = 1 \). Part (a) is a standard result in contour integration, and part (b) comes from the fact that the integral can be written as an alternating series.

An alternate proof of (a) is the following. \( e^{-xy} \sin x \) is integrable on \( \{(x, y); 0 < x < a, 0 < y < \infty\} \). So

\[
\int_0^a \frac{\sin x}{x} \, dx = \int_0^a \int_0^\infty e^{-xy} \sin x \, dy \, dx \\
= \int_0^\infty \int_0^a e^{-xy} \sin x \, dx \, dy \\
= \int_0^\infty \left[ \frac{e^{-xy}}{y^2 + 1} (-y \sin x - \cos x) \right]_0^a \, dy \\
= \int_0^\infty \left[ \{ \frac{e^{-ay}}{y^2 + 1} (-y \sin a - \cos a) \} - \frac{-1}{y^2 + 1} \right] \, dy \\
= \frac{\pi}{2} - \sin a \int_0^\infty \frac{ye^{-ay}}{y^2 + 1} \, dy - \cos a \int_0^\infty \frac{e^{-ay}}{y^2 + 1} \, dy.
\]

The last two integrals tend to 0 as \( a \to \infty \) since the integrand is bounded by \((1 + y)e^{-y} \) if \( a \geq 1 \). \( \square \)

Theorem 6.5 (Inversion formula) Let \( \mu \) be a probability measure and let \( \varphi(t) = \int e^{itx} \mu(dx) \). If \( a < b \), then

\[
\lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) \, dt = \mu(a, b) + \frac{1}{2} \mu(\{a\}) + \frac{1}{2} \mu(\{b\}).
\]

The example where \( \mu \) is point mass at 0, so \( \varphi(t) = 1 \), shows that one needs to take a limit, since the integrand in this case is \( 2 \sin t/t \), which is not integrable.
6.1. INVERSION FORMULA

Proof. By Fubini,

\[ \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) \, dt = \int_{-T}^{T} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} e^{itx} \mu(dx) \, dt \]

\[ = \int \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} e^{itx} \mu(dx) \, dt. \]

To justify this, we bound the integrand by the mean value theorem.

Expanding \( e^{-itb} \) and \( e^{-ita} \) using Euler’s formula, and using the fact that \( \cos \) is an even function and \( \sin \) is odd, we are left with

\[ \int 2 \left[ \int_{0}^{T} \frac{\sin(t(x-a))}{t} \, dt - \int_{0}^{T} \frac{\sin(t(x-b))}{t} \, dt \right] \mu(dx). \]

Using Lemma 6.4 and dominated convergence, this tends to

\[ \int [\pi \text{sgn}(x-a) - \pi \text{sgn}(x-b)] \mu(dx). \]

\[ \square \]

Theorem 6.6 If \( \int |\varphi(t)| \, dt < \infty \), then \( \mu \) has a bounded density \( f \) and

\[ f(y) = \frac{1}{2\pi} \int e^{-ity} \varphi(t) \, dt. \]

Proof.

\[ \mu(a, b) + \frac{1}{2} \mu(\{a\}) + \frac{1}{2} \mu(\{b\}) \]

\[ = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) \, dt \]

\[ = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) \, dt \]

\[ \leq \frac{b-a}{2\pi} \int |\varphi(t)| \, dt. \]

Letting \( b \to a \) shows that \( \mu \) has no point masses.
We now write
\[ \mu(x, x + h) = \frac{1}{2\pi} \int e^{-itx} \frac{e^{-it(x+h)}}{it} \varphi(t) dt \]
\[ = \frac{1}{2\pi} \int \left( \int_x^{x+h} e^{-ity} dy \right) \varphi(t) dt \]
\[ = \int_x^{x+h} \left( \frac{1}{2\pi} \int e^{-ity} \varphi(y) dy \right) dt. \]

So \( \mu \) has density \( (1/2\pi) \int e^{-ity} \varphi(t) dt \). As in the proof of Proposition 6.1, we see \( f \) is continuous.

A corollary to the inversion formula is the uniqueness theorem.

**Theorem 6.7** If \( \varphi_X = \varphi_Y \), then \( \mathbb{P}_X = \mathbb{P}_Y \).

**Proof.** Let \( a \) and \( b \) be such that \( \mathbb{P}(X = a) = \mathbb{P}(X = b) = 0 \), and the same when \( X \) is replaced by \( Y \). We have

\[ F_X(b) - F_X(a) = \mathbb{P}(a < X \leq b) = \mathbb{P}_X((a, b]) = \mathbb{P}_X((a, b)) \]

and the same when \( X \) is replaced by \( Y \). By the inversion theorem, \( \mathbb{P}_X((a, b]) = \mathbb{P}_Y((a, b]) \). Now let \( a \to -\infty \) along a sequence of points that are continuity points for \( F_X \) and \( F_Y \). We obtain \( F_X(b) = F_Y(b) \) if \( b \) is a continuity point for \( X \) and \( Y \). Since \( F_X \) and \( F_Y \) are right continuous, \( F_X(x) = F_Y(x) \) for all \( x \).

The following proposition can be proved directly, but the proof using characteristic functions is much easier.

**Proposition 6.8** (a) If \( X \) and \( Y \) are independent, \( X \) is a normal with mean \( a \) and variance \( b^2 \), and \( Y \) is a normal with mean \( c \) and variance \( d^2 \), then \( X + Y \) is normal with mean \( a + c \) and variance \( b^2 + d^2 \).
(b) If \( X \) and \( Y \) are independent, \( X \) is Poisson with parameter \( \lambda_1 \), and \( Y \) is Poisson with parameter \( \lambda_2 \), then \( X + Y \) is Poisson with parameter \( \lambda_1 + \lambda_2 \).
(c) If \( X_i \) are i.i.d. Cauchy, then \( S_n/n \) is Cauchy.
6.2. CONTINUITY THEOREM

Proof. For (a),
\[
\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t) = e^{iat-b^2t^2/2}e^{ict-c^2t^2/2} = e^{i(a+c)t-(b^2+d^2)t^2/2}.
\]
Now use the uniqueness theorem.

Parts (b) and (c) are proved similarly. 

\[\square\]

6.2 Continuity theorem

Lemma 6.9 Suppose \( \varphi \) is the characteristic function of a probability \( \mu \). Then
\[
\mu([-2A, 2A]) \geq A \left| \int_{-1/A}^{1/A} \varphi(t) \, dt \right| - 1.
\]

Proof. Note
\[
\frac{1}{2T} \int_{-T}^{T} \varphi(t) \, dt = \frac{1}{2T} \int_{-T}^{T} e^{itx} \mu(dx) \, dt
\]
\[
= \int \int \frac{1}{2T} 1_{[-T,T]}(t)e^{itx} \, dt \mu(dx)
\]
\[
= \int \frac{\sin Tx}{Tx} \mu(dx).
\]
Since \(|\sin(Tx)| \leq 1\), then \(|(\sin(Tx))/Tx| \leq 1/2TA\) if \(|x| \geq 2A\). Since \(|(\sin(Tx))/Tx| \leq 1\), we then have
\[
\left| \int \frac{\sin Tx}{Tx} \mu(dx) \right| \leq \mu([-2A,2A]) + \int_{[-2A,2A]} \frac{1}{2TA} \mu(dx)
\]
\[
= \mu([-2A,2A]) + \frac{1}{2TA} (1 - \mu([-2A,2A]))
\]
\[
= \frac{1}{2TA} + \left(1 - \frac{1}{2TA}\right) \mu([-2A,2A]).
\]
Setting \( T = 1/A \),
\[
\left| \frac{A}{2} \int_{-1/A}^{1/A} \varphi(t) \, dt \right| \leq \frac{1}{2} + \frac{1}{2} \mu([-2A,2A]).
\]
Now multiply both sides by 2.

\[\square\]
Proposition 6.10 If \( \mu_n \) converges weakly to \( \mu \), then \( \varphi_n \) converges to \( \varphi \) uniformly on every finite interval.

**Proof.** Let \( \varepsilon > 0 \) and choose \( M \) large so that \( \mu([-M, M]^c) < \varepsilon \). Define \( f \) to be 1 on \([−M, M] \), 0 on \([-M − 1, M + 1]^c \), and linear in between. Since \( \int f \, d\mu_n \to \int f \, d\mu \), then if \( n \) is large enough,

\[
\int (1 - f) \, d\mu_n \leq 2\varepsilon.
\]

We have

\[
|\varphi_n(t + h) - \varphi_n(t)| \leq \int |e^{ihx} - 1| \mu_n(dx)
\]
\[
\leq 2 \int (1 - f) \, d\mu_n + h \int |x| f(x) \mu_n(dx)
\]
\[
\leq 2\varepsilon + h(M + 1).
\]

So for \( n \) large enough and \( |h| \leq \varepsilon/(M + 1) \), we have

\[
|\varphi_n(t + h) - \varphi_n(t)| \leq 3\varepsilon,
\]

which implies that the \( \varphi_n \) are equicontinuous. Therefore the convergence is uniform on finite intervals. \( \square \)

The interesting result of this section is the converse, Lévy’s continuity theorem.

**Theorem 6.11** Suppose \( \mu_n \) are probabilities, \( \varphi_n(t) \) converges to a function \( \varphi(t) \) for each \( t \), and \( \varphi \) is continuous at 0. Then \( \varphi \) is the characteristic function of a probability \( \mu \) and \( \mu_n \) converges weakly to \( \mu \).

**Proof.** Let \( \varepsilon > 0 \). Since \( \varphi \) is continuous at 0, choose \( \delta \) small so that

\[
\left| \frac{1}{2\delta} \int_{-\delta}^{\delta} \varphi(t) \, dt - 1 \right| < \varepsilon.
\]

Using the dominated convergence theorem, choose \( N \) such that

\[
\frac{1}{2\delta} \int_{-\delta}^{\delta} |\varphi_n(t) - \varphi(t)| \, dt < \varepsilon
\]
6.2. CONTINUITY THEOREM

if \( n \geq N \). So if \( n \geq N \),

\[
\left| \frac{1}{2\delta} \int_{-\delta}^{\delta} \varphi_n(t) \, dt \right| \geq \frac{1}{2\delta} \int_{-\delta}^{\delta} \varphi(t) \, dt \geq \frac{1}{2\delta} \int_{-\delta}^{\delta} |\varphi_n(t) - \varphi(t)| \, dt \\
\geq 1 - 2\varepsilon.
\]

By Lemma 6.9 with \( A = 1/\delta \), for such \( n \),

\[
\mu_n[-2/\delta, 2/\delta] \geq 2(1 - 2\varepsilon) - 1 = 1 - 4\varepsilon.
\]

This shows the \( \mu_n \) are tight.

Let \( n_j \) be a subsequence such that \( \mu_{n_j} \) converges weakly, say to \( \mu \). Then \( \varphi_{n_j}(t) \to \varphi_\mu(t) \), hence \( \varphi(t) = \varphi_\mu(t) \), or \( \varphi \) is the characteristic function of a probability \( \mu \). If \( \mu' \) is any subsequential weak limit point of \( \mu_n \), then \( \varphi_{\mu'}(t) = \varphi(t) = \varphi_\mu(t) \); so \( \mu' \) must equal \( \mu \). Hence \( \mu_n \) converges weakly to \( \mu \). \( \Box \)

We need the following estimate on moments.

**Proposition 6.12** If \( \mathbb{E} |X|^k < \infty \) for an integer \( k \), then \( \varphi_X \) has a continuous derivative of order \( k \) and

\[
\varphi^{(k)}(t) = \int (ix)^k e^{itx} \mathbb{P}_X(dx).
\]

In particular, \( \varphi^{(k)}(0) = i^k \mathbb{E} X^k \).

**Proof.** Write

\[
\frac{\varphi(t + h) - \varphi(t)}{h} = \int \frac{e^{i(t+h)x} - e^{itx}}{h} \mathbb{P}(dx).
\]

The integrand is bounded by \( |x| \). So if \( \int |x| \mathbb{P}_X(dx) < \infty \), we can use dominated convergence to obtain the desired formula for \( \varphi'(t) \). As in the proof of Proposition 6.1, we see \( \varphi'(t) \) is continuous. We do the case of general \( k \) by induction. Evaluating \( \varphi^{(k)} \) at 0 gives the particular case. \( \Box \)

Here is a converse.
Proposition 6.13 If \( \phi \) is the characteristic function of a random variable \( X \) and \( \phi''(0) \) exists, then \( \mathbb{E}|X|^2 < \infty \).

Proof. Note
\[
\frac{e^{ihx} - 2 + e^{-ihx}}{h^2} = -2\frac{1 - \cos hx}{h^2} \leq 0
\]
and \( 2(1 - \cos hx)/h^2 \) converges to \( x^2 \) as \( h \to 0 \). So by Fatou’s lemma,
\[
\int x^2 \mathbb{P}_X(dx) \leq 2 \liminf_{h \to 0} \int \frac{1 - \cos hx}{h^2} \mathbb{P}_X(dx)
= -\limsup_{h \to 0} \frac{\phi(h) - 2\phi(0) + \phi(-h)}{h^2} = \phi''(0) < \infty.
\]

One nice application of the continuity theorem is a proof of the weak law of large numbers. Its proof is very similar to the proof of the central limit theorem, which we give in the next section. Another nice use of characteristic functions and martingales is the following.

Proposition 6.14 Suppose \( X_i \) is a sequence of independent r.v.s and \( S_n \) converges weakly. Then \( S_n \) converges almost surely.

Proof. We first rule out the possibility that \( |S_n| \to \infty \) with positive probability. If this happens with positive probability \( \varepsilon \), given \( M \) there exists \( N \) depending on \( M \) such that if \( n \geq N \), then \( \mathbb{P}(|S_n| > M) > \varepsilon \). Then the limit law of \( \mathbb{P}_{S_n} \), say, \( \mathbb{P}_{\infty} \), will have \( \mathbb{P}_{\infty}([-M, M]^c) \geq \varepsilon \) for all \( M \), a contradiction. Let \( N_1 \) be the set of \( \omega \) for which \( |S_n(\omega)| \to \infty \).

Suppose \( S_n \) converges weakly to \( W \). Then \( \varphi_{S_n}(t) \to \varphi_W(t) \) uniformly on compact sets by Proposition 6.10. Since \( \varphi_W(0) = 1 \) and \( \varphi_W \) is continuous, there exists \( \delta \) such that \( |\varphi_W(t) - 1| < 1/2 \) if \( |t| < \delta \). So for \( n \) large, \( |\varphi_{S_n}(t)| \geq 1/4 \) if \( |t| < \delta \).

Note
\[
\mathbb{E}[e^{itS_n} | X_1, \ldots, X_{n-1}] = e^{itS_{n-1}}\mathbb{E}[e^{itX_n} | X_1, \ldots, X_{n-1}] = e^{itS_{n-1}}\varphi_{X_n}(t).
\]
Since $\varphi_{S_n}(t) = \prod \varphi_{X_i}(t)$, it follows that $e^{itS_n}/\varphi_{S_n}(t)$ is a martingale.

Therefore for $|t| < \delta$ and $n$ large, $e^{itS_n}/\varphi_{S_n}(t)$ is a bounded martingale, and hence converges almost surely. Since $\varphi_{S_n}(t) \to \varphi_W(t) \neq 0$, then $e^{itS_n}$ converges almost surely if $|t| < \delta$.

Let $A = \{ (\omega, t) \in \Omega \times (-\delta, \delta) : e^{itS_n(\omega)} \text{ does not converge} \}$. For each $t$, we have almost sure convergence, so $\int 1_A(\omega, t) \mathbb{P}(d\omega) = 0$. Therefore $\int_{-\delta}^{\delta} \int 1_A d\mathbb{P} dt = 0$, and by Fubini, $\int \int_{-\delta}^{\delta} 1_A dt d\mathbb{P} = 0$. Hence almost surely, $\int 1_A(\omega, t) dt = 0$. This means, there exists a set $N_2$ with $\mathbb{P}(N_2) = 0$, and if $\omega \notin N_2$, then $e^{itS_n(\omega)}$ converges for almost every $t \in (-\delta, \delta)$. Call the limit, when it exists, $L(t)$.

Fix $\omega \notin N_1 \cup N_2$.

Suppose one subsequence of $S_n(\omega)$ converges to $Q$ and another subsequence to $R$. Then $L(a) = e^{iaR} = e^{iaQ}$ for a.e. $a$ small, which implies that $R = Q$.

Suppose one subsequence converges to $R \neq 0$ and another to $\pm \infty$. By integrating $e^{itS_n}$ and using dominated convergence, we have

$$\int_0^a e^{itS_n} dt = \frac{e^{iaS_n} - 1}{iS_n}$$

converges. We then obtain

$$\frac{e^{iaR} - 1}{iR} = 0$$

for a.e. $a$ small, which implies $R = 0$.

Finally suppose one subsequence converges to 0 and another to $\pm \infty$. Since $(e^{iaS_n} - 1)/iS_n \to a$, we obtain $a = 0$ for a.e. $a$ small, which is impossible.

We conclude $S_n$ must converge.