3 Characteristic Functions

Characteristic functions (a straightforward generalization of Fourier transforms) yield a very convenient way to deal with many properties of distributions; in particular, they allow to describe weak convergence.

We use the notation $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ for the euclidean inner product and norm. Recall that $\mathfrak{M}(\mathbb{R}^k)$ denotes the class of all probability measures on $(\mathbb{R}^k, \mathfrak{B}_k)$. Further, $\imath$ denotes the imaginary unit. Let $\mu \in \mathfrak{M}(\mathbb{R}^k)$.

Definition 1. $f : \mathbb{R}^k \to \mathbb{C}$ is *μ-integrable* if $\Re f$ and $\Im f$ are μ-integrable, in which case

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

A bounded continuous function $f : \mathbb{R}^k \to \mathbb{C}$ is obviously μ–integrable. A special class of such functions is given by $f_y(x) := \exp(\imath\langle x, y\rangle)$ with $y \in \mathbb{R}^k$.

Definition 2. The mapping $\widehat{\mu} : \mathbb{R}^k \to \mathbb{C}$ with

$$\widehat{\mu}(y) = \int \exp(\imath\langle x, y\rangle) \, \mu(dx), \qquad y \in \mathbb{R}^k,$$

is called the *Fourier transform* of μ.

Example 1.

(i) For μ discrete,

$$\mu = \sum_{j=1}^{\infty} \alpha_j \cdot \delta_{x_j}$$

we have

$$\widehat{\mu}(y) = \sum_{j=1}^{\infty} \alpha_j \cdot \exp(\imath\langle x_j, y\rangle).$$

For instance, if $\mu = Pois(\lambda)$ is the Poisson distribution with parameter $\lambda > 0$, then

$$\begin{aligned}\widehat{\mu}(y) &= \exp(-\lambda) \cdot \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \cdot \exp(\imath j y) = \exp(-\lambda) \cdot \exp(\lambda \cdot \exp(\imath y)) \\ &= \exp(\lambda \cdot (\exp(\imath y) - 1)).\end{aligned}$$

(ii) If $\mu = f \cdot \lambda_k$ then

$$\widehat{\mu}(y) = \int \exp(\imath\langle x, y\rangle) \cdot f(x) \, \lambda_k(dx).$$

For any λ_k-integrable function f, the right-hand side defines its *Fourier transform*, see Analysis III. For instance, if μ is the k-dimensional standard normal distribution, i.e.,

$$f(x) = (2\pi)^{-k/2} \cdot \exp(-\|x\|^2/2),$$

then

$$\widehat{\mu}(y) = \exp(-\|y\|^2/2).$$

See Bauer (1996, p. 187) for the case $k = 1$. Use Fubini's Theorem for $k > 1$.

Theorem 1.

(i) $\widehat{\mu}$ is uniformly continuous on $\mathbb{R}^k$,

(ii) $|\widehat{\mu}(y)| \leq 1 = \widehat{\mu}(0)$ for $y \in \mathbb{R}^k$,

(iii) for $n \in \mathbb{N}$, $a_1, \ldots, a_n \in \mathbb{C}$, and $y_1, \ldots, y_n \in \mathbb{R}^k$,

$$\sum_{j,\ell=1}^{n} a_j \cdot \overline{a_\ell} \cdot \widehat{\mu}(y_j - y_\ell) \geq 0$$

(*positive semi-definite*).

Proof. Ad (i): Observe that

$$\left|\exp(\imath\langle x, y_1\rangle) - \exp(\imath\langle x, y_2\rangle)\right| \leq \|x\| \cdot \|y_1 - y_2\|.$$

For $\varepsilon > 0$ take $r > 0$ such that $\mu(B) \geq 1 - \varepsilon$, where $B = \{x \in \mathbb{R}^k : \|x\| \leq r\}$. Then

$$\begin{aligned}\left|\widehat{\mu}(y_1) - \widehat{\mu}(y_2)\right| &\leq \int_B \left|\exp(\imath\langle x, y_1\rangle) - \exp(\imath\langle x, y_2\rangle)\right| \mu(dx) + 2 \cdot \varepsilon \\ &\leq r \cdot \|y_1 - y_2\| + 2 \cdot \varepsilon.\end{aligned}$$

Properties (ii) and (iii) are easily verified. □

Remark 1. *Bochner's Theorem* states that every continuous, positive semi-definite function $\varphi : \mathbb{R}^k \to \mathbb{C}$ with $\varphi(0) = 1$ is the Fourier transform of a probability measure on $(\mathbb{R}^k, \mathfrak{B}_k)$. See Bauer (1996, p. 184) for references.

In the sequel: $X, Y, \ldots$ are k-dimensional random vectors on a probability space $(\Omega, \mathfrak{A}, P)$.

Definition 3. The *characteristic function* of X is given by

$$\varphi_X = \widehat{P_X}.$$

Remark 2. Due to Theorem II.9.1

$$\varphi_X(y) = \int_{\mathbb{R}^k} \exp(\imath\langle x, y\rangle)\, P_X(dx) = \mathrm{E}\, e^{\imath\langle X, y\rangle}\ .$$

Theorem 2.

(i) For every linear mapping $T : \mathbb{R}^k \to \mathbb{R}^\ell$

$$\varphi_{T \circ X} = \varphi_X \circ T^t.$$

(ii) For independent random vectors X and Y

$$\varphi_{X+Y} = \varphi_X \cdot \varphi_Y.$$

In particular, for $a \in \mathbb{R}^k$,

$$\varphi_{X+a} = \exp(\imath\langle a, \cdot\rangle) \cdot \varphi_X.$$

Proof. Ad (i): Let $z \in \mathbb{R}^\ell$. Use $P_{T\circ X} = T(P_X)$ to obtain

$$\varphi_{T\circ X}(z) = \int_{\mathbb{R}^k} \exp(\imath\langle T(x), z\rangle)\, P_X(dx) = \varphi_X(T^t(z)).$$

Ad (ii): Let $z \in \mathbb{R}^k$. Fubini's Theorem and Theorem III.5.5 imply

$$\varphi_{X+Y}(z) = \int_{\mathbb{R}^{2k}} \exp(\imath\langle x+y, z\rangle)\, P_{(X,Y)}(d(x,y)) = \varphi_X(z) \cdot \varphi_Y(z).$$

□

Corollary 1 (Convolution Theorem). For probability measures $\mu_j \in \mathfrak{M}(\mathbb{R})$,

$$\widehat{\mu_1 * \mu_2} = \widehat{\mu}_1 \cdot \widehat{\mu}_2.$$

Proof. Use Theorem 2.(ii) and Theorem III.5.8. □

Example 2. For $\mu = N(m, \sigma^2)$ with $\sigma \geq 0$ and $m \in \mathbb{R}$

$$\widehat{\mu}(y) = \exp(\imath m y) \cdot \exp(-\sigma^2 y^2/2).$$

See Example 1.(ii) and Theorem 2.

Lemma 1. For $z \in \mathbb{R}$ and $\sigma > 0$

$$\int \exp(-\imath yz/\sigma^2) \cdot \widehat{\mu}(y/\sigma^2)\, N(0,\sigma^2)(dy) = \int \exp\bigl(-(z-x)^2/(2\sigma^2)\bigr)\, \mu(dx).$$

Proof. We note first that the function $\varphi : \mathbb{R} \to \mathbb{C}$,

$$\varphi(r) := \int_{\mathbb{R}} e^{-(y-\imath r)^2/2\sigma^2}\, dy = \int_{\mathfrak{R}} e^{-(y^2-r^2)/2\sigma^2 + \imath yr/\sigma^2}\, dy$$

is constant. (This can be seen by computing φ' by switching integration and differentiation, or by Cauchy's integral theorem.) The left hand side in the lemma equals

$$\int\int e^{-iyz/\sigma^2} e^{iyx/\sigma^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-y^2/(2\sigma^2)}\, dy\, d\mu(x) = \int e^{-(z-x)^2/(2\sigma^2)} \cdot \Bigl(\frac{1}{\sqrt{2\pi\sigma^2}} \int e^{-(y^2-(x-z)^2)/(2\sigma^2)+iy(x-z)}\, dy$$

The inner integral equals $\varphi(x-z) = \varphi(0) = 1$. □

Lemma 2. For $\sigma_n > 0$ with $\lim_{n\to\infty} \sigma_n = 0$,

$$N(0,\sigma_n^2) * \mu \xrightarrow{\text{w}} \mu.$$

Proof. Consider independent random variables X_n and Y such that $X_n \sim N(0, \sigma_n^2)$ and $Y \sim \mu$. Then $X_n \xrightarrow{\mathfrak{L}^2} 0$, and therefore $X_n + Y \xrightarrow{\mathfrak{L}^2} Y$, which implies

$$X_n + Y \xrightarrow{\mathrm{d}} Y.$$

□

Theorem 3 (Uniqueness Theorem). For probability measures $\mu_j \in \mathfrak{M}(\mathbb{R}^k)$,

$$\mu_1 = \mu_2 \qquad \Leftrightarrow \qquad \widehat{\mu_1} = \widehat{\mu_2}.$$

Proof. '$\Rightarrow$' holds by definition. '$\Leftarrow$': See Bauer (1996, Thm. 23.4) or Billingsley (1979, Sec. 29) for the case $k > 1$. Here: the case $k = 1$. For $\sigma > 0$ and $A \in \mathfrak{B}$, Theorem III.5.9 yields

$$N(0, \sigma^2) * \mu_j(A) = (2\pi\sigma^2)^{-1/2} \cdot \int_A \int \exp(-(z-x)^2/(2\sigma^2))\, \mu_j(dx)\, \lambda_1(dz).$$

From Lemma 1 we conclude that

$$\forall\, \sigma > 0: \ N(0, \sigma^2) * \mu_1 = N(0, \sigma^2) * \mu_2.$$

Then, by Lemma 2 and Corollary III.3.1, $\mu_1 = \mu_2$. □

Example 3. For independent random variables X_1 and X_2 with $X_j \sim \pi(\lambda_j)$ we have $X_1 + X_2 \sim \pi(\lambda_1 + \lambda_2)$.

Proof: Theorem 2 and Example 1.(i) yield

$$\begin{aligned} \varphi_{X_1+X_2}(y) &= \exp(\lambda_1 \cdot (\exp(\imath y) - 1)) \cdot \exp(\lambda_2 \cdot (\exp(\imath y) - 1)) \\ &= \exp((\lambda_1 + \lambda_2) \cdot (\exp(\imath y) - 1)). \end{aligned}$$

Use Theorem 3.

Lemma 3. For every $\varepsilon > 0$ and every probability measure $\mu \in \mathfrak{M}(\mathbb{R})$,

$$\mu(\{x \in \mathbb{R} : |x| \geq 1/\varepsilon\}) \leq 7/\varepsilon \cdot \int_0^\varepsilon (1 - \Re\widehat{\mu}(y))\, dy.$$

Proof. Clearly

$$\Re\widehat{\mu}(y) = \int_{\mathbb{R}} \cos(xy)\, \mu(dx).$$

Hence, with the convention $\sin(0)/0 = 1$,

$$\begin{aligned} 1/\varepsilon \cdot \int_0^\varepsilon (1 - \Re\widehat{\mu}(y))\, dy &= 1/\varepsilon \cdot \int_{[0,\varepsilon]} \int_{\mathbb{R}} (1 - \cos(xy))\, \mu(dx)\, \lambda_1(dy) \\ &= \int_{\mathbb{R}} \left(1/\varepsilon \cdot \int_0^\varepsilon (1 - \cos(xy))\, dy \right) \mu(dx) \\ &= \int_{\mathbb{R}} (1 - \sin(\varepsilon x)/(\varepsilon x))\, \mu(dx) \\ &\geq \inf_{|z| \geq 1} (1 - \sin(z)/z) \cdot \mu(\{x \in \mathbb{R} : |\varepsilon x| \geq 1\}). \end{aligned}$$

Finally,

$$\inf_{|z| \geq 1} (1 - \sin(z)/z) \geq 1/7.$$

□

Theorem 4 (Continuity Theorem, Lévy).

(i) Let $\mu, \mu_n \in \mathfrak{M}(\mathbb{R}^k)$ for $n \in \mathbb{N}$. Then

$$\mu_n \xrightarrow{\text{w}} \mu \qquad \Rightarrow \qquad \forall\, y \in \mathbb{R}^k : \lim_{n\to\infty} \widehat{\mu_n}(y) = \widehat{\mu}(y).$$

(ii) Let $\mu_n \in \mathfrak{M}(\mathbb{R}^k)$ for $n \in \mathbb{N}$, and let $\varphi : \mathbb{R}^k \to \mathbb{C}$ be continuous at 0 with $\varphi(0) = 1$. Then

$$\forall\, y \in \mathbb{R}^k : \lim_{n\to\infty} \widehat{\mu_n}(y) = \varphi(y) \qquad \Rightarrow \qquad \exists\, \mu \in \mathfrak{M}(\mathbb{R}^k) : \widehat{\mu} = \varphi \wedge \mu_n \xrightarrow{\text{w}} \mu.$$

Proof. Ad (i): Note that $x \mapsto \exp(\imath\langle x, y\rangle)$ is bounded and continuous on $\mathbb{R}^k$.

Ad (ii): See Bauer (1996, Thm. 23.8) or Billingsley (1979, Sec. 29) for the case $k > 1$. Here: the case $k = 1$.

We first show that

$$\{\mu_n : n \in \mathbb{N}\} \text{ is tight.} \tag{1}$$

By Lemma 3

$$\mu_n(\{x \in \mathbb{R} : |x| \geq 1/\varepsilon\}) \leq 7/\varepsilon \cdot c_n(\varepsilon)$$

with

$$c_n(\varepsilon) = \int_0^\varepsilon (1 - \Re\widehat{\mu_n}(y))\, dy.$$

Since $|\Re\widehat{\mu_n}| \leq |\widehat{\mu_n}| \leq 1$, we have by dominated convergence

$$\lim_{n\to\infty} c_n(\varepsilon) = c(\varepsilon)$$

with

$$c(\varepsilon) = \int_0^\varepsilon (1 - \Re\varphi(y))\, dy\ .$$

Now we exploit the assumptions on φ; given $\delta > 0$ take $\varepsilon > 0$ such that

$$7/\varepsilon \cdot c(\varepsilon) \leq \delta/2.$$

Furthermore, take $n_0 \in \mathbb{N}$ such that, for every $n \geq n_0$,

$$|c_n(\varepsilon) - c(\varepsilon)| \leq \varepsilon/7 \cdot \delta/2.$$

Hence, for $n \geq n_0$,

$$\mu_n(\{x \in \mathbb{R} : |x| \geq 1/\varepsilon\}) \leq \delta,$$

and hereby we get (1). Thus, by Prohorov's Theorem, $\{\mu_n : n \in \mathbb{N}\}$ is relatively compact; on the other hand, for any point of accumulation μ of the sequence μ_n, Part (i) implies that $\varphi = \widehat{\mu}$. Hence, there is exactly one point μ of accumulation, and for this μ we have $\widehat{\mu} = \varphi$. Finally, Remark 4.4 reveals that $\mu_n \xrightarrow{\text{w}} \mu$. □

Corollary 2. Weak convergence in $\mathfrak{M}(\mathbb{R}^k)$ is equivalent to pointwise convergence of Fourier transforms.