CHAPTER 3

Characteristic functions as tool for studying weak convergence

Defintions and basic properties

Definition 3.1. Let μ be a probability measure on $\mathbb{R}$. The function $\psi_\mu : \mathbb{R}^d \to \mathbb{R}$ define by $\psi_\mu(t) := \int_{\mathbb{R}} e^{itx} d\mu(x)$ is called the *characteristic function* or the *Fourier transform* of μ. If X is a random variable on a probability space, we sometimes say "characteristic function of X" to mean the c.f of its distribution. We also write $\hat{\mu}$ instead of ψ_μ.

There are various other "integral transforms" of a measure that are closely related to the c.f. For example, if we take $\psi_\mu(it)$ is the moment generating function of μ (if it exists). For μ supported on $\mathbb{N}$, its so called generating function $F_\mu(t) = \sum_{k\geq 0} \mu\{k\} t^k$ (which exists for $|t| < 1$ since μ is a probability measure) can be written as $\psi_\mu(-i \log t)$ (at least for $t > 0$!) etc. The characteristic function has the advantage that it exists for all $t \in \mathbb{R}$ and for all finite measures μ.

The following lemma gives some basic properties of a c.f.

Lemma 3.2. *Let $\mu \in \mathscr{P}(\mathbb{R})$. Then, $\hat{\mu}$ is a uniformly continuous function on $\mathbb{R}$ with $|\hat{\mu}(t)| \leq 1$ for all t with $\hat{\mu}(0) = 1$. (equality may be attained elsewhere too).*

PROOF. Clearly $\hat{\mu}(0) = 1$ and $|\hat{\mu}(t)| \leq 1$. U ■

The importance of c.f comes from the following facts.

(A) It transforms well under certain operations of measures, such as shifting a scaling and under convolutions.
(B) The c.f. determines the measure.
(C) $\hat{\mu}_n(t) \to \hat{\mu}(t)$ pointwise, if and only if $\mu_n \xrightarrow{d} \mu$.
(D) There exist necessary and sufficient conditions for a function $\psi : \mathbb{R} \to \mathbb{C}$ to be the c.f o f a measure. Because of this and part (B), sometimes one defines a measure by its characteristic function.

(A) Transformation rules

Theorem 3.3. *Let X, Y be random variables.*

(1) *For any $a, b \in \mathbb{R}$, we have $\psi_{aX+b}(t) = e^{ibt}\psi_X(at)$.*
(2) *If X, Y are independent, then $\psi_{X+Y}(t) = \psi_X(t)\psi_Y(t)$.*

PROOF. (1) $\psi_{aX+b}(t) = \mathbf{E}[e^{it(aX+b)}] = \mathbf{E}[e^{itaX}]e^{ibt} = e^{ibt}\psi_X(at)$.
(2) $\psi_{X+Y}(t) = \mathbf{E}[e^{it(X+Y)}] = \mathbf{E}[e^{itX}e^{itY}] = \mathbf{E}[e^{itX}]\mathbf{E}[e^{itY}] = \psi_X(t)\psi_Y(t)$.

■

Examples.

(1) If $X \sim \text{Ber}(p)$, then $\psi_X(t) = pe^{it} + q$ where $q = 1-p$. If $Y \sim \text{Binomial}(n,p)$, then, $Y \overset{d}{=} X_1 + \ldots + X_n$ where X_k are i.i.d Ber(p). Hence, $\psi_Y(t) = (pe^{it}+q)^n$.

(2) If $X \sim \text{Exp}(\lambda)$, then $\psi_X(t) = \int_0^\infty \lambda e^{-\lambda x} e^{itx} dx = \frac{1}{\lambda - it}$. If $Y \sim \text{Gamma}(\nu,\lambda)$, then if ν is an integer, then $Y \overset{d}{=} X_1 + \ldots + X_n$ where X_k are i.i.d Exp(λ). Therefore, $\psi_Y(t) = \frac{1}{(\lambda - it)^\nu}$.

(3) $Y \sim \text{Normal}(\mu,\sigma^2)$. Then, $Y = \mu + \sigma X$, where $X \sim N(0,1)$ and by the transofrmatin rules, $\psi_Y(t) = e^{i\mu t}\psi_X(\sigma t)$. Thus it suffices to find the c.f of $N(0,1)$.

$$\psi_X(t) = \frac{1}{\sigma\sqrt{2\pi}}\int_{\mathbb{R}} e^{itx} e^{-\frac{x^2}{2\sigma^2}} dx = e^{-\frac{\sigma^2 t^2}{2}}\left(\frac{1}{\sigma\sqrt{2\pi}}\int_{\mathbb{R}} e^{-\frac{(x-it)^2}{2\sigma^2}} dx\right).$$

It appears that the stuff inside the brackets is equal to 1, since it looks like the integral of a normal density with mean it and variance σ^2. But if the mean is complex, what does it mean?! I gave a rigorous proof that the stuff inside brackets is indeed equal to 1, in class using contour integration, which will not be repeated here. The final concusion is that $N(\mu,\sigma^2)$ has c.f $e^{it\mu - \frac{\sigma^2 t^2}{2}}$.

(B) Inversion formulas

Theorem 3.4. *If $\hat{\mu} = \hat{\nu}$, then $\mu = \nu$.*

PROOF. Let θ_σ denote the N$(0,\sigma^2)$ distribution and let $\phi_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$ and $\Phi_\sigma(x) = \int_{-\infty}^x \phi_\sigma(u)du$ and $\hat{\theta}_\sigma(t) = e^{-\sigma^2 t^2/2}$ denote the density and cdf and characteristic functions, respectively. Then, by Parseval's identity, we have for any α,

$$\begin{aligned}\int e^{-i\alpha t}\hat{\mu}(t)d\theta_\sigma(t) &= \int \hat{\theta}_\sigma(x-\alpha)d\mu(x) \\ &= \frac{\sqrt{2\pi}}{\sigma}\int \phi_{\frac{1}{\sigma}}(\alpha - x)d\mu(x)\end{aligned}$$

where the last line comes by the explicit Gaussian form of $\hat{\theta}_\sigma$. Let $f_\sigma(\alpha) := \frac{\sigma}{\sqrt{2\pi}}\int e^{-i\alpha t}\hat{\mu}(t)d\theta_\sigma(t)$ and integrate the above equation to get that for any finite $a < b$,

$$\begin{aligned}\int_a^b f_\sigma(\alpha)d\alpha &= \int_a^b \int_{\mathbb{R}} \phi_{\frac{1}{\sigma}}(\alpha - x)d\mu(x)d\mu(x) \\ &= \int_{\mathbb{R}}\int_a^b \phi_{\frac{1}{\sigma}}(\alpha - x)d\alpha d\mu(x) \qquad \text{(by Fubini)} \\ &= \int_{\mathbb{R}} \left(\Phi_{\frac{1}{\sigma}}(\alpha - a) - \Phi_{\frac{1}{\sigma}}(\alpha - b)\right)d\mu(x).\end{aligned}$$

Now, we let $\sigma \to \infty$, and note that

$$\Phi_{\frac{1}{\sigma}}(u) \to \begin{cases} 0 & \text{if } u < 0. \\ 1 & \text{if } u > 0. \\ \frac{1}{2} & \text{if } u = 0.\end{cases}$$

Further, $\Phi_{\sigma^{-1}}$ is bounded by 1. Hence, by DCT, we get

$$\lim_{\sigma\to\infty}\int_a^b f_\sigma(\alpha)d\alpha = \int\left[\mathbf{1}_{(a,b)}(x) + \frac{1}{2}\mathbf{1}_{\{a,b\}}(x)\right]d\mu(x) = \mu(a,b) + \frac{1}{2}\mu\{a,b\}.$$

Now we make two observations: (a) that f_σ is determined by $\hat{\mu}$, and (b) that the measure μ is determined by the values of $\mu(a,b)+\frac{1}{2}\mu\{a,b\}$ for all finite $a<b$. Thus, $\hat{\mu}$ determines the measure μ. ■

Corollary 3.5 (Fourier inversion formula). *Let $\mu \in \mathscr{P}(\mathbb{R})$.*

(1) *For all finite $a<b$, we have*

$$\mu(a,b)+\frac{1}{2}\mu\{a\}+\frac{1}{2}\mu\{b\} = \lim_{\sigma\to\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\frac{e^{-iat}-e^{-ibt}}{it}\hat{\mu}(t)e^{-\frac{t^2}{2\sigma^2}}dt \tag{3.1}$$

(2) *If $\int_{\mathbb{R}}|\hat{\mu}(t)|dt<\infty$, then μ has a continuous density given by*

$$f(x) := \frac{1}{2\pi}\int_{\mathbb{R}}\hat{\mu}(t)e^{-ixt}dt.$$

PROOF. (1) Recall that the left hand side of (3.1) is equal to $\lim_{\sigma\to\infty}\int_a^b f_\sigma$ where $f_\sigma(\alpha) := \frac{\sigma}{\sqrt{2\pi}}\int e^{-i\alpha t}\hat{\mu}(t)d\theta_\sigma(t)$. Writing out the density of θ_σ we see that

$$\begin{aligned}\int_a^b f_\sigma(\alpha)d\alpha &= \frac{1}{2\pi}\int_a^b\int_{\mathbb{R}}e^{-i\alpha t}\hat{\mu}(t)e^{-\frac{t^2}{2\sigma^2}}dt\,d\alpha\\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\int_a^b e^{-i\alpha t}\hat{\mu}(t)e^{-\frac{t^2}{2\sigma^2}}d\alpha\, dt \qquad \text{(by Fubini)}\\ &= \frac{1}{2\pi}\int_{\mathbb{R}}\frac{e^{-iat}-e^{-ibt}}{it}\hat{\mu}(t)e^{-\frac{t^2}{2\sigma^2}}dt.\end{aligned}$$

Thus, we get the first statement of the corollary.

(2) With f_σ as before, we have $f_\sigma(\alpha) := \frac{1}{2\pi}\int e^{-i\alpha t}\hat{\mu}(t)e^{-\frac{t^2}{2\sigma^2}}dt$. Note that the integrand converges to $e^{-i\alpha t}\hat{\mu}(t)$ as $\sigma\to\infty$. Further, this integrand is bounded by $|\hat{\mu}(t)|$ which is assumed to be integrable. Therefore, by DCT, for any $\alpha\in\mathbb{R}$, we conclude that $f_\sigma(\alpha)\to f(\alpha)$ where $f(\alpha) := \frac{1}{2\pi}\int e^{-i\alpha t}\hat{\mu}(t)dt$.

Next, note that for any $\sigma>0$, we have $|f_\sigma(\alpha)|\le C$ for all α where $C = \int|\hat{\mu}$. Thus, for finite $a<b$, using DCT again, we get $\int_a^b f_\sigma\to\int_a^b f$ as $\sigma\to\infty$. But the proof of Theorem 3.4 tells us that

$$\lim_{\sigma\to\infty}\int_a^b f_\sigma(\alpha)d\alpha = \mu(a,b)+\frac{1}{2}\mu\{a\}+\frac{1}{2}\mu\{b\}.$$

Therefore, $\mu(a,b)+\frac{1}{2}\mu\{a\}+\frac{1}{2}\mu\{b\} = \int_a^b f(\alpha)d\alpha$. Fixing a and letting $b\downarrow a$, this shows that $\mu\{a\}=0$ and hence $\mu(a,b)=\int_a^b f(\alpha)d\alpha$. Thus f is the density of μ.

The proof that a c.f. is continuous carries over verbatim to show that f is continuous (since f is the Furier trnasform of $\hat{\mu}$, except for a change of sign in the exponent). ■

An application of Fourier inversion formula Recall the Cauchy distribution μ with with density $\frac{1}{\pi(1+x^2)}$ whose c.f is not easy to find by direct integration (Residue theorem in complex analysis is a way to compute this integral).

Consider the seemingly unrelated p.m ν with density $\frac{1}{2}e^{-|x|}$ (a symmetrized exponential, this is also known as Laplace's distribution). Its c.f is easy to compute and we get

$$\hat{nu}(t) = \frac{1}{2}\int_0^\infty e^{itx-x}dx+\frac{1}{2}\int_{-\infty}^0 e^{itx+x}dx = \frac{1}{2}\left(\frac{1}{1-it}+\frac{1}{1+it}\right) = \frac{1}{1+t^2}.$$

By the Fourier inversion formula (part (b) of the corollary), we therefore get

$$\frac{1}{2}e^{-|x|} = \frac{1}{2\pi}\int \hat{\nu}(t)e^{itx}dt = \frac{1}{2\pi}\int \frac{1}{1+t^2}e^{itx}dt.$$

This immediately shows that the Cauchy distribution has c.f. $e^{-|t|}$ without having to compute the integral!!

(C) Continuity theorem

Theorem 3.6. *Let $\mu_n, \mu \in \mathscr{P}(\mathbb{R})$.*

(1) *If $\mu_n \xrightarrow{d} \mu$ then $\hat{\mu}_n(t) \to \hat{\mu}(t)$ pointwise for all t.*
(2) *If $\hat{\mu}_n(t) \to \psi(t)$ pointwise for all t, then $\psi = \hat{\mu}$ for some $\mu \in \mathscr{P}(\mathbb{R})$ and $\mu_n \xrightarrow{d} \mu$.*

PROOF. (1) If $\mu_n \xrightarrow{d} mu$, then $\int f d\mu_n \to \int f d\mu$ for any $f \in C_b(\mathbb{R})$ (bounded continuous function). Since $x \to e^{itx}$ is a bounded continuous function for any $t \in \mathbb{R}$, it follows that $\hat{\mu}_n(t) \to \hat{\mu}(t)$ pointwise for all t.

(2) Now suppose $\hat{\mu}_n(t) \to \hat{\mu}(t)$ pointwise for all t. We first claim that the sequence $\{\mu_n\}$ is tight. Assuming this, the proof can be completed as follows.

Let μ_{n_k} be any subsequence that converges in distribution, say to ν. By tightness, $nu \in \mathscr{P}(\mathbb{R})$. Therefore, by part (a), $\hat{\mu}_{n_k} \to \hat{\nu}$ pointwise. But obviously, $\hat{\mu}_{n_k} \to \hat{\mu}$ since $\hat{\mu}_n \to \hat{\mu}$. Thus, $\hat{\nu} = \hat{\mu}$ which implies that $\nu = \mu$. That is, any convergent subsequence of $\{\mu_n\}$ converges in distribution to μ. This shows that $\mu_n \xrightarrow{d} \mu$ (because, if not, then there is some subsequence $\{n_k\}$ and some $\epsilon > 0$ such that the Lévy distance between μ_{n_k} and μ is at least ϵ. By tightness, $\mu_{\mathbf{n}_k}$ must have a subsequence that converges to some p.m ν which cannot be equal to μ contradicting what we have shown!).

It remains to show tightness. From Lemma 3.7 below, as $n \to \infty$,

$$\mu_n\left([-2/\delta, 2/\delta]^c\right) \quad \le \quad \frac{1}{\delta}\int_{-\delta}^{\delta}(1-\hat{\mu}_n(t))dt \longrightarrow \frac{1}{\delta}\int_{-\delta}^{\delta}(1-\hat{\mu}(t))dt$$

where the last implication follows by DCT (since $1-\hat{\mu}_n(t) \to 1-\hat{\mu}(t)$ for each t and also $|1-\hat{\mu}_n(t)| \le 2$ for all t. Further, as $\delta \downarrow 0$, we get $\frac{1}{\delta}\int_{-\delta}^{\delta}(1-\hat{\mu}(t))dt \to 0$ (because, $1-\hat{\mu}(0) = 0$ and $\hat{\mu}$ is continuous at 0).

Thus, given $\epsilon > 0$, we can find $\delta > 0$ such that $\limsup_{n\to\infty}\mu_n([-2/\delta, 2/\delta]^c) < \epsilon$. This means that for some finite N, we have $\mu_n([-2/\delta, 2/\delta]^c) < \epsilon$ for all $n \ge N$. Now, find $A > 2/\delta$ such that for any $n \le N$, we get $\mu_n([-2/\delta, 2/\delta]^c) < \epsilon$. Thus, for any $\epsilon > 0$, we have produced an $A > 0$ so that $\mu_n([-A, A]^c) < \epsilon$ for all n. This is the definition of tightness. ∎

Lemma 3.7. *Let $\mu \in \mathscr{P}(\mathbb{R})$. Then, for any $\delta > 0$, we have*

$$\mu\left([-2/\delta, 2/\delta]^c\right) \le \frac{1}{\delta}\int_{-\delta}^{\delta}(1-\hat{\mu}(t))dt.$$

PROOF. We write

$$\begin{aligned}\int_{-\delta}^{\delta}(1-\hat{\mu}(t))dt &= \int_{-\delta}^{\delta}\int_{\mathbb{R}}(1-e^{itx})d\mu(x)dt \\ &= \int_{\mathbb{R}}\int_{-\delta}^{\delta}(1-e^{itx})dtd\mu(x) \\ &= \int_{\mathbb{R}}\left(2\delta-\frac{\sin(x\delta)}{x}\right)d\mu(x) \\ &= 2\delta\int_{\mathbb{R}}\left(1-\frac{\sin(x\delta)}{2x\delta}\right)d\mu(x).\end{aligned}$$

When $|x|\delta > 2$, we have $\frac{\sin(x\delta)}{2x\delta} \le \frac{1}{2}$ (since $\sin(x\delta) \le 1$). Therefore, the integrand is at least $\frac{1}{2}$ when $|x| > \frac{2}{\delta}$ and the integrand is always non-negative since $|\sin(x)| \le |x|$. Therefore we get

$$\int_{-\delta}^{\delta}(1-\hat{\mu}(t))dt \ge \frac{1}{2}\mu\left([-2/\delta, 2/\delta]^c\right).$$

■