CHARACTERISTIC FUNCTIONS

We work in 1-dimension (at first)

Defn: $\mu \in \mathcal{P}(\mathbb{R})$. Its characteristic function or Fourier transform, denoted φ_μ or $\hat{\mu}$ is a function from $\mathbb{R}$ to $\mathbb{C}$

$$\hat{\mu}(t) = \int_{\mathbb{R}} e^{itx}\, d\mu(x)$$

$$= \int \cos(tx)\, d\mu(x) + i \int \sin(tx)\, d\mu(x).$$

If $X \sim \mu$, we also write φ_X to mean φ_μ.

Remark: Earlier we defined the moment generating function of μ to be $\lambda \longrightarrow \mathbb{E}_\mu\left[e^{-\lambda X}\right]$ which may be represented as $\varphi_\mu(-i\lambda)$. For μ supported on $\mathbb{N} = \{0,1,2,\dots\}$, the probability generating function $F(t) = \sum_{n=0}^{\infty} \mu_n t^n$ can also be written as $\varphi_\mu(-i \log t)$ (at least for $t > 0$).

Thus c.f. is similar to these other 'integral transforms' of μ, but unlike the mgf or pgf, c.f exists $\forall\, \mu \in \mathcal{P}(\mathbb{R})$

Basic properties of c.f: Let $\mu \in \mathcal{P}(\mathbb{R})$. Then

$\hat{\mu}(0) = 1$; $|\hat{\mu}(t)| \leq 1 \;\forall t \in \mathbb{R}$; $t \to \hat{\mu}(t)$ is uniformly continuous.

Proof: The first two are obvious, since $|e^{itx}| = 1$ $(t \in \mathbb{R}, x \in \mathbb{R})$

For the last, note that

$$|\hat{\mu}(t+h) - \hat{\mu}(t)| \leq \int |e^{i(t+h)x} - e^{itx}|\, d\mu(x)$$

$$= \int |e^{ihx} - 1|\, d\mu(x)$$

As $h \to 0$, $|e^{ihx} - 1| \to 0 \;\forall x$. Further $|e^{ihx} - 1| \leq 2$ which is integrable w.r.t μ. By DCT, we get uniform continuity. ∎

The importance of characteristic functions stem from the following facts.

(A) Certain operations on measures such as shifting and scaling or convolutions, lead to simple operations in terms of their characteristic functions.

(B) The characteristic function determines the measure - i.e., $\hat{\mu} = \hat{\nu} \Rightarrow \mu = \nu$. (This fact alone is sufficient for our purposes, but there are 'explicit' formulas for recovering μ from $\hat{\mu}$, and more usefully, simple formulas to extract useful functionals such as mean and variance of μ from a knowledge of $\hat{\mu}$).

(C) $\hat{\mu}_n(t) \longrightarrow \hat{\mu}(t)$ pointwise $\forall t \in \mathbb{R}$ $\Leftrightarrow \mu_n \xrightarrow{d} \mu$.

This gives an alternate way to prove weak convergence. It is sometimes better than working with CDFs to prove $F_{\mu_n}(t) \to F_\mu(t)$ $\forall$ continuity points t (for example, if $\mu_n = \nu_1 * \nu_2 * \dots * \nu_n$, then F_{μ_n} is very complicated to write in terms of $F_{\nu_1}, \dots F_{\nu_n}$, but $\hat{\mu}_n(t) = \hat{\nu}_1(t)\hat{\nu}_2(t)\dots\hat{\nu}_n(t)$!!)

A) Proposition: Let X, Y be real-valued random variables on $(\Omega, \mathcal{F}, P)$

(a) $\forall \alpha, \beta \in \mathbb{R}$, $\psi_{\alpha X+\beta}(t) = e^{i\beta t}\psi_X(\alpha t)$

(b) If X, Y are independent, then

$\psi_{X+Y}(t) = \psi_X(t)\psi_Y(t)$.

Proof: (a) $\psi_{\alpha X+\beta}(t) = E[e^{it(\alpha X+\beta)}]$

$= e^{it\beta} E[e^{i\alpha t X}]$

$= e^{it\beta}\psi_X(\alpha t)$.

(b) $\psi_{X+Y}(t) = E[e^{it(X+Y)}]$

$= E[e^{itX} \cdot e^{itY}]$

$= E[e^{itX}]\, E[e^{itY}]$ (independence)

$= \psi_X(t)\psi_Y(t)$. ■

Examples:

① $\mathrm{Ber}(p)$ has c.f. $pe^{it}+q$ (where $q = 1-p$).

$\mathrm{Bin}(p) \stackrel{d}{=} X_1 + \dots + X_n$ where $X_k \sim$ iid $\mathrm{Ber}(p)$,

hence by the proposition, $\mathrm{Bin}(p)$ has c.f. $(pe^{it}+q)^n$.

② $\mathrm{Pois}(\lambda)$ has c.f. $\sum_{n=0}^{\infty} e^{itk} e^{-\lambda}\frac{\lambda^k}{k!}$

$= e^{-\lambda} e^{\lambda e^{it}}$

$= e^{-\lambda+\lambda e^{it}}$

③ $\mathrm{Exp}(\lambda)$ has pdf $\lambda e^{-\lambda x}$ on $[0,\infty)$ and hence its c.f. is $\int_0^\infty \lambda e^{-\lambda x} e^{itx}\,dx$

$= \lambda \int_0^\infty e^{-(\lambda - it)x}\,dx$

$= \frac{\lambda}{\lambda - it}$. (why?)

$\Gamma(\nu,\lambda) \stackrel{d}{=} X_1 + \dots + X_\nu$ where $X_k \sim$ iid $\mathrm{Exp}(\lambda)$

when ν is an integer

and hence, in that case, $\Gamma(\nu,\lambda)$ has c.f. equal to $\left(\frac{\lambda}{\lambda - it}\right)^\nu$.

This is also valid for non-inger ν (exercise) ($\nu > 0$ is always assumed)

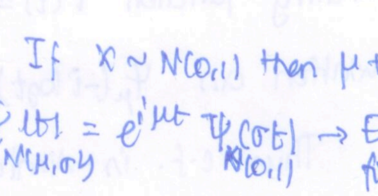
④ $N(\mu,\sigma^2)$. If $X \sim N(0,1)$ then $\mu + \sigma X \sim N(\mu,\sigma^2)$

Hence $\psi_{N(\mu,\sigma^2)}(t) = e^{i\mu t}\psi_{N(0,1)}(\sigma t)$ → Enough to find $\psi_{N(0,1)}$.

For $N(0,1)$,

$$\psi(t) = \int_{\mathbb{R}} e^{itx}\frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}\,dx = \int_{\mathbb{R}} e^{-\frac{1}{2}(x-it)^2}\frac{dx}{\sqrt{2\pi}}$$

$$= \lim_{A\to\infty} \underbrace{\int_{\gamma_A} e^{-\frac{1}{2}z^2}\frac{dz}{\sqrt{2\pi}}}_{\text{Contour integral along the curve } \gamma_A(u) = -A+it+u,\ u\in[-A,+A].}$$



By Cauchy's theorem

$$\int_{\gamma_A} e^{-z^2/2}\,dz = \int_{-A}^{A} e^{-x^2/2}\,dx + \int_{\eta_A} + \int_{\tilde{\eta}_A}$$

Exer: Show $\int_{\eta_A}, \int_{\tilde{\eta}_A} \to 0$ as $A \to \infty$

Thus conclude that

$$\psi(t) = e^{-t^2/2}$$

and also, $\psi_{N(\mu,\sigma^2)}(t) = e^{i\mu t - \frac{1}{2}\sigma^2 t^2}$.

Remark: $\psi_{N(0,\sigma^2)}(t) = e^{-\frac{1}{2}\sigma^2 t^2}$

$N(0,\sigma^2)$ density $= \dfrac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi}}$

If $\sigma \downarrow 0$, the density localizes, but ψ becomes more spread out. This is an instance of the 'uncertainty principle'.

B) Proposition: If $\hat{\mu} = \hat{\nu}$ then $\mu = \nu$.

We ~~need to~~ shall ~~prove~~ use

Parseval's identity: If $\mu, \nu \in \mathcal{P}(\mathbb{R})$, then

1. $E_\mu \psi_\nu = E_\nu \psi_\mu$

i.e., $\int \psi_\nu(t)\, d\mu(t) = \int \psi_\mu(t)\, d\nu(t)$

2. $\int e^{-i\alpha t} \psi_\nu(t)\, d\mu(t) = \int \psi_\mu(t-\alpha)\, d\nu(t)$.

Proof: Let $f(s,t) = e^{ist} : \mathbb{R}^2 \to \mathbb{C}$

Clearly f is integrable w.r.t the probability measure $\mu \otimes \nu$.

Hence, by Fubini's

$$\int_{\mathbb{R}^2} f(s,t)\, d(\mu\otimes\nu)(s,t)$$

$$\int_{\mathbb{R}}\int_{\mathbb{R}} e^{ist}\, d\mu(s)\, d\nu(t) \qquad \int_{\mathbb{R}}\int_{\mathbb{R}} e^{ist}\, d\nu(t)\, d\mu(s)$$

$$= \int_{\mathbb{R}} \psi_\mu(t)\, d\nu(t). \qquad = \int_{\mathbb{R}} \psi_\nu(s)\, d\mu(s). \quad \blacksquare$$

(b) - exercise.

Proof of the proposition: Let $\theta_\sigma = N(0,\sigma^2)$.

By the fact that

$$\psi_{\theta_\sigma}(t) = e^{-\sigma^2 t^2/2} = \frac{\sqrt{2\pi}}{\sigma} \frac{1}{\sqrt{2\pi}\frac{1}{\sigma}} e^{-t^2/2(1/\sigma^2)}$$

$$= \frac{\sqrt{2\pi}}{\sigma} \cdot \theta_{1/\sigma}(t)$$

where we use $\theta_{1/\sigma}$ to denote the density of ~~N(0,~~ $\theta_{1/\sigma} = N(0, \frac{1}{\sigma^2})$.

Apply Parseval's to μ and θ_σ to get

$$\int \hat{\mu}(t)\, e^{-i\alpha t}\, d\theta_\sigma(t) = \int \hat{\theta}_\sigma(t-\alpha)\, d\mu(t)$$

$$= \frac{\sqrt{2\pi}}{\sigma} \int \theta_{1/\sigma}(t-\alpha)\, d\mu(t)$$

$$= \frac{\sqrt{2\pi}}{\sigma} \underbrace{(\theta_{1/\sigma} * \mu)(\alpha)}_{\text{density of the convolution } \theta_{1/\sigma} * \mu}.$$

Now, as $\sigma \to 0$, $\frac{1}{\sigma} \to \infty$ and hence $\theta_{1/\sigma} \xrightarrow{d} \delta_0$

Conclude that $\theta_{1/\sigma} * \mu \xrightarrow{d} \delta_0 * \mu = \mu$
(Why?)

Hence for any x = cty. point of F_μ,

$$F_\mu(x) = \lim_{\sigma \to 0} \int_{-\infty}^{x} (\theta_{1/\sigma} * \mu)(\alpha)\, d\alpha$$

$$= \lim_{\sigma \downarrow 0} \int_{-\infty}^{x} \frac{\sigma}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{\mu}(t)\, e^{-i\alpha t}\, d\theta_\sigma(t)$$

= a function of $\hat{\mu}$ alone!

Thus we can recover $F_\mu(x)$ from $\hat{\mu}$ (for a.e. x and hence for all x, by right continuity of F_μ).

Thus $\hat{\mu}$ determines μ. $\blacksquare$