Lecture 19: December 5

Lecturer: Persi Diaconis *Scribes: Kenneth Tay*

19.1 Characteristic Functions

Definition 19.1 *The **characteristic function** or **Fourier transform** of μ at $t \in \mathbb{R}^n$ is*

$$\phi(t) := \mathbb{E}(e^{it \cdot X}) = \int_{\mathbb{R}^n} \cos(t \cdot x) + i \sin(t \cdot x) \mu(dx).$$

The goal of this lecture is to prove the Continuity Theorem for characteristic functions. We will need to develop some machinery to do that.

Lemma 19.2 (Cantor's Diagonal Argument) *Let*

$$\begin{matrix} x_{11} & x_{12} & \dots \\ x_{21} & x_{22} & \dots \\ \vdots & \vdots & \ddots \end{matrix}$$

be an array of real numbers such that each row is bounded. Then there exist a subsequence $\{n_k\}_{k=1}^\infty$ of $\mathbb{N}$ and $\{l_r\}_{r=1}^\infty$ such that $x_{rn_k} \to l_r$ as $k \to \infty$ for all $r \in \mathbb{N}$.

Proof: Since $\{x_{1n}\}$ is bounded, by the Heine-Borel Theorem, there is a convergent subsequence n_{1i} and real number l_1 so that $X_{1n_{1i}} \to l_1$ as $i \to \infty$.

Next, look at $\{x_{2n_{1i}}\}$. This is bounded, and so there is a convergent subsequence n_{2i} of $\{n_{1i}\}$, and a real number l_2 such that $x_{2n_{2i}} \to l_2$.

By continuing this procedure, there exists subsequence $\{n_{ri}\}$ of $\{n_{(n-1)i}\}$ such that $x_{rn_{ri}} \to l_r$.

Finally, consider the "diagonal sequence" $\{n_{kk}\}$. $x_{rn_{kk}} \to l_r$ for all r as $k \to \infty$, as required. ∎

Proposition 19.3 (Helly Selection Theorem) *If $\{F_n\}_{n=1}^\infty$ are any distribution functions on $\mathbb{R}$, then there exist monotone, right-continuous F and a subsequence $n_k \nearrow \infty$ such that $F_{n_k}(x) \to F(x)$ for all points of continuity x of F.*

Proof: Let $\{r_i\}_{i=1}^\infty$ be an enumeration of the rationals $\mathbb{Q}$. Make an array

$$\begin{matrix} F_1(r_1) & F_2(r_1) & \dots \\ F_2(r_1) & F_2(r_2) & \dots \\ \vdots & \vdots & \ddots \end{matrix}$$

Since each row is bounded, by Cantor's diagonal argument, there exists a subsequence $n_k \nearrow \infty$ and $\{G(r) : r \in \mathbb{Q}\}$ such that $F_{n_k}(r) \to G(r)$ for all $r \in \mathbb{Q}$. (Note that for $r < s$, $F_{n_k}(r) \leq F_{n_k}(s)$ for all n_k, which implies that $G(r) \leq G(s)$.)

Define $F(x) =: \inf_{r>x} G(r)$. It is clear from the definition that F is increasing. Given x and $\varepsilon > 0$, choose $r > x$ such that $G(r) < F(x) + \varepsilon$. If $x < y < r$, then $F(x) \leq F(y) \leq G(r) < F(x) + \varepsilon$. Thus, F is also right-continuous.

It remains to show that if x is a continuity point of F, then $F_{n_k}(x) \to F(x)$. Given $\varepsilon > 0$, choose $y < x$ such that $F(x) - \varepsilon < F(y)$.

Next, choose rationals r, s so that $y < r < x < s$ and $G(s) < F(x) + \varepsilon$. We have

$$F(x) - \varepsilon < F(y) \leq G(r) \leq G(s) < F(x) + \varepsilon.$$

Since $F_{n_k}(r) \leq F_{n_k}(x) \leq F_{n_k}(s)$ for all n,

$$G(r) = \lim F_{n_k}(r) \leq \liminf F_{n_k}(x) \leq \limsup F_{n_k}(x) \leq \lim F_{n_k}(s) = G(s),$$

so $\liminf F_{n_k}(x)$ and $\limsup F_{n_k}(x)$ are within ε of $F(x)$ Since ε was arbitary, we must have $\lim F_{n_k}(x) = F(x)$.

■

Note: The limit F need not be a distribution function (does not need to have $F(-\infty) = 0$ and $F(\infty) = 1$).

- e.g. $F_n = \delta_n$ (point mass at n). $F_n(x) \to 0$ for all x.
- e.g. $F_n = \delta_{-n}$ (point mass at $-n$). $F_n(x) \to 1$ for all x.

When is the limit F a distribution function? It is when the sequence $\{F_n\}$ is "tight" (which we define next).

19.1.1 Tightness

Definition 19.4 *A family of probabilities $\{\mu_n\}$ on $\mathbb{R}$ is* ***tight*** *if for every $\varepsilon > 0$, there exist $a < b$ so that $\mu_n(a, b] > 1 - \varepsilon$ for all n. We say that μ is* ***almost compactly supported****.*

Theorem 19.5 *A necessary and sufficient condition for tightness is: for every subsequence $\{n_k\}_{k=1}^{\infty}$, there exists a further subsequence $\{n_{k_i}\}_{i=1}^{\infty}$ and a probability μ so that $\mu_{n_{k_i}} \Rightarrow \mu$.*

Proof: We only prove necessity (which is the direction we need later). Suppose $\{\mu_n\}$ is tight (on $\mathbb{R}$). Given a subsequence $\{n_k\}$, consider $\{F_{n_k}\}$. By Helly's Selection Theorem, there exists a further subsequence $\{n_{k_i}\}$ and F increasing, right-continuous such that $F_{n_{k_i}} \Rightarrow F$.

By tightness of μ_n, for every $\varepsilon > 0$, there exist continuity points a, b of F with $a < b$ so that $F_{n_k}(b) - F_{n_k}(a) > 1 - \varepsilon$ for all k. Taking limits, we have $F(b) - F(a) \leq 1 - \varepsilon$. This is enough to show that F is a distribution function.

■

Corollary 19.6 *If $\{F_n\}$ is tight, then there exists a distribution function F and a subsequence $\{n_k\}_{k=1}^{\infty}$ such that $F_{n_k} \Rightarrow F$.*

19.1.2 Aside

The tools that we have developed all work for general spaces (e.g. $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ complete separable metric spaces). We say that probabilities $\{\mu_n\}$ on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ converge to μ if $\int f d\mu_n \to \int f d\mu$ for all bounded continuous $f : \mathcal{X} \to \mathbb{R}$.

In this setting,

1. $\mu_n \Rightarrow \mu$ iff for all A with $\mu(\partial A) = 0$, $\mu_n(A) \to \mu(A)$.
2. (Skrohod's Theorem is true) If $\mu_n \Rightarrow \mu$, then there exist $(\Omega, \mathcal{F}, P)$ and random variables $X, X_n : \Omega \to \mathcal{X}$ such that $X_n(\omega) \to X(\omega)$ for all ω.
3. We say that μ is **tight** if for all $\varepsilon > 0$, there exist a compact $K \subseteq \mathcal{X}$ such that $\mu_n(K) > 1 - \varepsilon$ for all n.
4. If $\{\mu_n\}$ is tight, then there exist a probability μ and a subsequence $\{n_k\}$ such that $\mu_{n_k} \Rightarrow \mu$.

References for this material:

- Billingsley, "Convergence of Probability Measures."
- Kallenberg, "Probability Theory."
- Dudley, "Real Analysis and Probability."

19.1.3 Continuity Theorem

We are now ready to prove the continuity theorem:

Theorem 19.7 *Let $\{F_n\}$, F be distribution functions on $\mathbb{R}$ with characteristic functions $\{\phi_n\}$, ϕ. Then*

$$F_n \Rightarrow F \quad \Leftrightarrow \quad \phi_n(t) \to \phi(t) \text{ for all } t.$$

Proof: If $F_n \Rightarrow F$, then $\phi_n(t) \to \phi(t)$ for all t using Skorohod's Theorem.

Now, assume $\phi_n(t) \to \phi(t)$ for all t. We claim that $\{\mu_n\}$ is tight.

Consider $\frac{1}{u}\int_{-u}^{u} (1 - \phi(t))dt$. (This integral is real-valued for all t.)

$$\begin{aligned}
\frac{1}{u}\int_{-u}^{u} (1 - \phi(t))dt &= \int_{-\infty}^{\infty} \frac{1}{u}\left[\int_{-u}^{u} (1 - e^{itx})dt\right]\mu(dx) \\
&= 2\int_{-\infty}^{\infty} \frac{1}{ux}[1 - \sin(ux)]\mu(dx) \\
&\geq 2\int_{\{|x|>2/u\}} \left(1 - \frac{1}{ux}\right) dx \\
&\geq \mu\left\{x : |x| > \frac{2}{u}\right\}.
\end{aligned}$$

Note that ϕ is continuous and $\phi(0) = 1$. Thus, for every $\varepsilon > 0$, there exists a sufficiently small u such that

$$\frac{1}{u}\int_{-u}^{u}(1-\phi(t))dt < \varepsilon.$$

Since $\phi_n(t) \to \phi(t)$, by the Bounded Convergence Theorem, there is an n_0 such that

$$\mu_n\left\{x : |x| > \frac{2}{u}\right\} \leq \frac{1}{u}\int_{-u}^{u}(1-\phi_n(t))dt < 2\varepsilon$$

for $n > n_0$.

Increasing $\dfrac{2}{u}$ sufficiently so that $\mu_i\left\{x : |x| > \dfrac{2}{u}\right\} < 2\varepsilon$ for $i = 1, \ldots, n_0 - 1$, we conclude that $\{\mu_n\}$ is tight.

If $\mu_n \Rightarrow \mu$ is false, then there is a point of continuity x of F such that $\mu_n(-\infty, x]$ does not converge to $\mu(-\infty, x]$. Pick a subsequence $\{n_k\}$ such that $|\mu_{n_k}(-\infty, x] - \mu(-\infty, x]| > \varepsilon$. By Corollary 19.6, there exists a ν and further subsequence $\{n_{k_i}\}$ such that $\mu_{n_{k_i}} \Rightarrow \nu$. This means that $\phi_{n_k}(t) \to \phi_\nu(t)$ for all t, i.e. $\mu = \nu$ by uniqueness of characteristic functions. Contradiction!

Thus, we must have $\mu_n \Rightarrow \mu$.

$\blacksquare$