PROBABILITY THEORY: LECTURE NOTES 3

CHANDRA NAIR

Disclaimer: These notes are created from a variety of great references in the field, most notably, Varadhan's lecture notes and Dembo's lecture notes.

3. Characteristic functions

Let $f : \Omega \to \mathbb{C}$ be a complex valued function. Let f_r, f_i denote its real and imaginary components. We say that f is measurable if f_r and f_i are measurable. Further if f_r and f_i are integrable, we define

$$\int f d\mathrm{P} = \int f_r d\mathrm{P} + i \int f_i d\mathrm{P}.$$

Exercise 3.1. Let $f : \Omega \to \mathbb{C}$ be a complex valued mesurable function. Then show that

$$\left| \int f d\mathrm{P} \right| \leq \int |f| d\mathrm{P}.$$

For any random variable X we define the characteristic function according to

$$\phi(t) = \int \exp[itx] d\alpha = E(\exp[itX]).$$

Using above exercise, note that $|\phi(t)| \leq 1$.

Theorem 3.1. *The characteristic function of any probability distribution is a uniformly continuous function of t that is positive definite, i.e. for any real numbers $t_1, .., t_n$ the matrix $M \equiv [\phi(t_i - t_j)]$ is non-negative semidefinite.*

Proof.

$$\begin{aligned}
\vec{\xi} M \vec{\xi}^* &= \sum_{i,j} \xi_i \phi(t_i - t_j) \xi_j^* \\
&= \sum_{i,j} \xi_i E(\exp[i(t_i - t_j)X]) \xi_j^* \\
&= E(\sum_{i,j} \xi_i \exp[i(t_i - t_j)X] \xi_j^*) \\
&= E(\sum_i \xi_i \exp[it_i X] \sum_j \exp[-it_j X] \xi_j^*) \\
&= E(|\sum_i \xi_i \exp[it_i X]|^2) \geq 0.
\end{aligned}$$

The equality holds if and only if $Y = \sum_i \xi_i \exp[it_i X] = 0$ almost surely (i.e. with probability 1).

Date: October 2, 2019.

To show uniform continuity observe that

$$\begin{aligned}|\phi(t)-\phi(s)| &= |E(e^{itX}-e^{isX})| \\ &\leq E(|e^{itX}-e^{isX}|) \\ &= E(|e^{i(t-s)X}-1|).\end{aligned}$$

Thus it suffices to show that for every $\epsilon > 0$ we can pick a $\delta > 0$ such that whenever $|t| < \delta$ we have $E(|e^{itX}-1|) < \epsilon$. Assume otherwise, i.e. for a sequence $\delta_n \downarrow 0$ there exists points $t_n, |t_n| \leq \delta_n$ such that $E(|e^{it_nX}-1|) \geq \epsilon$. Now $t_n \to 0$ and hence $Y_n = |e^{it_nX}-1| \to 0$ pointwise. Since Y_n is bounded we have from bounded convergence theorem that $E(Y_n) \to 0$, and this yields a contradiction. □

Lemma 3.2. *If $\int |X|dP < \infty$ then $\phi(t)$ is continuously differentiable and $\phi'(0) = i\int XdP$.*

Proof.

$$\frac{1}{\delta}E(e^{itX}-e^{i(t-\delta)X}) = E(e^{itX}\frac{1-e^{-i\delta X}}{i\delta X}iX)$$

Set $Y_\delta = e^{itX}\frac{1-e^{-i\delta X}}{i\delta X}iX$ and $Y = iXe^{itX}$. Clearly $Y_\delta \to Y$ pointwise.

Further $|Y_\delta| \leq c|X|$ when $c = \sup_x |\frac{1-e^{ix}}{x}| < \infty$ (see lemma below). Thus from dominated convergence theorem (since $E(|X|) < \infty$) we have that $E(Y_\delta) \to E(Y)$. Therefore $\phi'(t) = E(iXe^{itX})$ exists. The continuity of $\phi'(t)$ is left as an exercise. □

Lemma 3.3. $|\frac{1-e^{ix}}{x}| \leq 1, \ \forall x \in \mathbb{R}$.

Proof.

$$\left|\frac{1-e^{ix}}{x}\right| = \left|2\frac{\sin\left(\frac{x}{2}\right)}{x}\right| \leq 1.$$

□

Exercise 3.2. Show that if $E(|X|^r) < \infty$ then $\phi(t)$ is r times continuously differentiable.
Extra credit: If r is even, show that the converse holds.

How do we get back the distribution function from the characteristic function?

Theorem 3.4. *When a, b are continuity points of $F(x) := P(X \leq x)$, then*

$$F(b)-F(a) = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\phi(t)\frac{\exp[-itb]-\exp[-ita]}{-it}dt.$$

Proof.

$$
\begin{aligned}
&\lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T}\phi(t)\frac{\exp[-itb]-\exp[-ita]}{-it}dt \\
&= \lim_{T\to\infty} \frac{1}{2\pi}\int_{-T}^{T}\frac{\exp[-itb]-\exp[-ita]}{-it}\int e^{itX}dPdt \\
&\overset{Fub}{=} \lim_{T\to\infty} \frac{1}{2\pi}\int\int_{-T}^{T}\frac{\exp[it(X-b)]-\exp[it(X-a)]}{-it}dtdP \\
&= \lim_{T\to\infty} \frac{1}{2\pi}\int\int_{-T}^{T}\frac{\sin t(X-a)-\sin t(X-b)}{t}dtdP \\
&= \lim_{T\to\infty} \frac{1}{\pi}\int\int_{0}^{T}\frac{\sin t(X-a)-\sin t(X-b)}{t}dtdP \\
&= \lim_{T\to\infty}\int u(T,X-a)-u(T,X-b)dP \\
&\overset{d.c}{=} \int\lim_{T\to\infty}\left(u(T,X-a)-u(T,X-b)\right)dP \\
&= \int \frac{1}{2}1_{X>a}-\frac{1}{2}1_{X<a}-\frac{1}{2}1_{X>b}+\frac{1}{2}1_{X<b}dP \\
&= \frac{1}{2}\left(P(X<b)-P(X>b)-P(X<a)+P(X>a)\right) \\
&= F(b)-F(a)-\frac{1}{2}(P(X=b)-P(X=a))
\end{aligned}
$$

Note that from (below) the definition of
Here

$$u(T,x)=\int_0^T\frac{\sin tx}{\pi t}dt=\int_0^{\frac{T}{\pi}}\frac{\sin \pi sx}{\pi sx}ds.$$

We know (see below) that $\sup_{T,x}|u(T,x)|\leq C$ and

$$\lim_{T\to\infty}u(T,x)=\begin{cases}\frac{1}{2} & \text{if } x>0\\ -\frac{1}{2} & \text{if } x<0\\ 0 & \text{if } x=0\end{cases}.$$

Hence one can find the distribution function from the characteristic function. $\square$

Exercise 3.3. Prove that: If two distribution functions agree on their points of continuity then they agree everywhere.
Hint: Show that the points of discontinuity are countable. Then use right continuity of the distribution functions.

Lemma 3.5. *Consider the Dirichlet integral defined according to*

$$u(T,x)=\int_0^T\frac{\sin tx}{\pi t}dt.$$

Then the following holds:

(*i*) $\sup_{T,x}|u(T,x)|\leq C$, *($C=2$ works).*

(ii)

$$\lim_{T\to\infty} u(T,x) = \begin{cases} \frac{1}{2} & \text{if } x>0 \\ -\frac{1}{2} & \text{if } x<0 \\ 0 & \text{if } x=0 \end{cases} .$$

Proof. Proof of (i): Without loss of generality, let us assume $x>0$. Further let k be such that $T \in (2k\frac{\pi}{x}, 2(k+1)\frac{\pi}{x}]$. If $k=0$ then

$$\left|\int_0^T \frac{\sin tx}{\pi t}dt\right| \leq \int_0^T \frac{x}{\pi}dt = \frac{Tx}{\pi} \leq 2.$$

For $k \geq 1$, we express

$$\begin{aligned} u(T,x) &= \sum_{j=0}^{k-1} \int_{2j\frac{\pi}{x}}^{(2j+1)\frac{\pi}{x}} \left(\frac{\sin tx}{\pi t} + \frac{\sin x(t+\frac{\pi}{x})}{\pi(t+\frac{\pi}{x})}\right)dt + \int_{2k\frac{\pi}{x}}^{T} \frac{\sin tx}{\pi t}dt \\ &= \sum_{j=0}^{k-1} \int_{2j\frac{\pi}{x}}^{(2j+1)\frac{\pi}{x}} \frac{\sin tx}{t(tx+\pi)}dt + \int_{2k\frac{\pi}{x}}^{T} \frac{\sin tx}{\pi t}dt. \end{aligned}$$

Thus we have (the second integral below only appears when $k \geq 2$)

$$\begin{aligned} |u(T,x)| &\leq \int_0^{\frac{\pi}{x}} \frac{\sin tx}{t(tx+\pi)}dt + \int_{2\frac{\pi}{x}}^{(2k-1)\frac{\pi}{x}} \frac{1}{t^2 x}dt + \int_{2k\frac{\pi}{x}}^{T} \frac{1}{\pi t}dt \\ &\leq \int_0^{\frac{\pi}{x}} \frac{x}{(tx+\pi)}dt + \frac{1}{2\pi} - \frac{1}{(2k-1)\pi} + \frac{1}{\pi}\ln\left(\frac{Tx}{2k\pi}\right) \\ &\leq \ln 2 + \frac{1}{2\pi} + \ln\left(\frac{k+1}{k}\right) \leq \ln 4 + \frac{1}{2\pi}. \end{aligned}$$

This establishes part (i). We used $\sin(tx) \leq |tx|$ in the first inequality.

Proof of (ii): As before, w.l.o.g., let us assume $x>0$. We first write the integral of interest as a complex line integral

$$u(T,x) = \frac{1}{2}\int_{L:(-T,0)\to(T,0)} \frac{e^{izx}}{i\pi z}dz.$$

For every $T > \epsilon > 0$, we consider almost semi-circular closed contour consisting of the following parts: a line from $(-T,0) \to (-\epsilon,0)$, a clockwise semicircle (center at origin and above the real axis) from $(-\epsilon,0) \to (\epsilon,0)$, a line from $(\epsilon,0) \to (T,0)$ and finally a counter-clockwise semi-circle from $(T,0)$ to $(-T,0)$. Since the closed contour does not have any poles in its interior, and the function $\frac{e^{izx}}{i\pi z}$ is analytic in the interior of the contour, we have

$$u(T,x) - \frac{1}{2}\int_{L:(-\epsilon,0)\to(\epsilon,0)} \frac{e^{izx}}{i\pi z}dz + \int_\pi^0 \frac{1}{2\pi}e^{i\epsilon e^{i\theta}x}d\theta + \int_0^\pi \frac{1}{2\pi}e^{iTe^{i\theta}x}d\theta = 0.$$

We consider the three integrals separately. Note that

$$\begin{aligned} &\left|\frac{1}{2}\int_{L:(-\epsilon,0)\to(\epsilon,0)} \frac{e^{izx}}{i\pi z}dz\right| = \left|\frac{1}{2}\int_{L:(-\epsilon,0)\to(\epsilon,0)} \frac{\sin(zx)}{i\pi z}dz\right| \\ &\quad \leq \frac{1}{2}\int_{L:(-\epsilon,0)\to(\epsilon,0)} \frac{|x|}{\pi}dz = 2\epsilon\frac{|x|}{\pi}. \end{aligned}$$

Observe that

$$\int_{\pi}^{0} \frac{1}{2\pi} e^{i\epsilon e^{i\theta} x} d\theta = -\frac{1}{2} + \int_{\pi}^{0} \frac{1}{2\pi} (e^{i\epsilon e^{i\theta} x} - 1) d\theta.$$

Now

$$\begin{aligned}
\left| \int_{\pi}^{0} \frac{1}{2\pi} (e^{i\epsilon e^{i\theta} x} - 1) d\theta \right| &= \left| \int_{\pi}^{0} \frac{1}{2\pi} (e^{-\epsilon \sin(\theta) x} e^{i\epsilon \cos(\theta) x} - 1) d\theta \right| \\
&= \left| \int_{\pi}^{0} \frac{1}{2\pi} (e^{-\epsilon \sin(\theta) x} (e^{i\epsilon \cos(\theta) x} - 1) + e^{-\epsilon \sin(\theta) x} - 1) d\theta \right| \\
&\leq \int_{0}^{\pi} \frac{1}{2\pi} e^{-\epsilon \sin(\theta) x} \left| (2 \sin(\epsilon \cos(\theta) x/2) \right| d\theta \\
&\quad + \int_{0}^{\pi} \frac{1}{2\pi} \left| 1 - e^{-\epsilon \sin(\theta) x} \right| d\theta \\
&\leq \frac{\epsilon x}{2\pi} \int_{0}^{\pi} |\cos\theta| d\theta + \frac{\epsilon x}{2\pi} \int_{0}^{\pi} \sin\theta d\theta = \frac{2\epsilon x}{\pi}
\end{aligned}$$

The last inequality uses $|\sin(a)| \leq |a|$ and $|1 - e^{-a}| \leq a, a > 0$.

From these two estimates, setting $\epsilon \to 0$ we see that we have

$$u(T, x) - \frac{1}{2} + \int_{0}^{\pi} \frac{1}{2\pi} e^{iTe^{i\theta} x} d\theta = 0.$$

(Note: one can use this relation to get a better upper bound on part (i), if desired).

The remaining integral is dealt with as follows:

$$\begin{aligned}
\int_{0}^{\pi} \frac{1}{2\pi} e^{iTe^{i\theta} x} d\theta &= \frac{1}{2\pi} \int_{0}^{\pi} e^{-T \sin\theta} e^{iT \cos\theta} d\theta \\
&= \frac{1}{2\pi} \int_{0}^{\sin^{-1} \frac{1}{\sqrt{T}}} e^{-T \sin\theta} e^{iT \cos\theta} d\theta + \frac{1}{2\pi} \int_{\sin^{-1} \frac{1}{\sqrt{T}}}^{\pi} e^{-T \sin\theta} e^{iT \cos\theta} d\theta.
\end{aligned}$$

Bounding each integral separately (the first integrand by 1 and the second integrand by $e^{-\sqrt{T}}$) we obtain that

$$\left| \int_{0}^{\pi} \frac{1}{2\pi} e^{iTe^{i\theta} x} d\theta \right| \leq \frac{1}{2\pi} \sin^{-1} \left(\frac{1}{\sqrt{T}} \right) + \frac{1}{2} e^{-\sqrt{T}}.$$

Thus when $x > 0$ we have

$$\lim_{T \to \infty} u(T, x) = \frac{1}{2}.$$

□

3.1. **Weak Convergence.**

Definition 3.1. A sequence P_n of probability distributions on $(R, \mathcal{F}_R)$ is said to converge *weakly* to a probability distribution P if

$$\lim_{n} P_n(I) = P(I),$$

where $I = [a, b]$ is any interval such that $P(\{a\}) = P(\{b\}) = 0$.

Exercise 3.4. Show that the following is an alternate definition of weak convergence: Let $F_n(x)$ be the distribution functions associated with P_n and $F(x)$ be the distribution function associated with P. Then $P_n \Rightarrow P$ if $\lim_n F_n(x) = F(x)$ at every continuity point of F.

Theorem 3.6. *(Levy-Cramer Continuity Theorem) The following are equivalent.*

(i) $P_n \Rightarrow P$ *or* $F_n \Rightarrow F$.

(ii) *For every bounded continuous function* $f(x)$ *on* R

$$\lim_n \int f(x)dP_n = \int f(x)dP.$$

(iii) *Let* $\phi_n(t)$ *be the characteristic function of* P_n *and* $\phi(t)$ *the characteristic function of* P. $\phi_n(t) \to \phi(t)$ *pointwise.*

Proof. We shall show the equivalence by showing that $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$.

- $(i) \Rightarrow (ii)$: Let $a < b$ be continuity points of F and $F(a) \leq \epsilon, F(b) \geq 1-\epsilon$. For large enough n, $F_n(a) \leq 2\epsilon$ and $F_n(b) \geq 1-2\epsilon$.

  Pick a $\delta > 0$. Divide the interval $(a,b]$ to finite number N_δ of subintervals $\mathcal{X}_j := (a_j, a_{j+1}]$ $a = a_1 < a_2 < ... < a_{N_\delta+1} = b$ such that all end points are continuity points of F and the fluctuation of f in each $\mathcal{X}_j$ is less than δ (We can do this since any continuous function f is uniformly continuous in a compact interval).

  Define $\hat{f} = \sum_{j=1}^{N_\delta} f(a_j)\mathbf{1}_{\mathcal{X}_j}$. Since $\lim_n F_n(a_j) = F(a_j)$ for all $1 \leq j \leq N$

$$\int \hat{f} dP_n = \sum_{j=1}^{N_\delta} f(a_i)(F_n(a_{j+1}) - F_n(a_j))$$

  and taking $n \to \infty$ we obtain that

$$\begin{aligned}
\lim_{n\to\infty} \int \hat{f} dP_n &= \lim_{n\to\infty} \sum_{j=1}^{N_\delta} f(a_i)(F_n(a_{j+1}) - F_n(a_j)) \\
&= \sum_{j=1}^{N_\delta} f(a_i)(F(a_{j+1}) - F(a_j)) = \int \hat{f} dP.
\end{aligned}$$

  Since f is bounded by M and $\hat{f} = 0$ on $(-\infty, a] \cup (b, \infty)$

$$\begin{aligned}
\left| \int f dP_n - \int \hat{f} dP_n \right| &\leq \left| \int_{[a,b]} f dP_n - \int_{[a,b]} \hat{f} dP_n \right| + 4M\epsilon \\
&\leq \int_{[a,b]} \left| f - \hat{f} \right| dP_n + 4M\epsilon \leq \delta + 4M\epsilon.
\end{aligned}$$

  Similarly

$$\left| \int f dP - \int \hat{f} dP \right| \leq \delta + 2M\epsilon$$

  and by triangle inequality we conclude

$$\limsup | \int f dP_n - \int f dP | \leq 2\delta + 6M\epsilon.$$

  Since $\epsilon, \delta > 0$ are arbitrary, we are done.

- $(ii) \Rightarrow (iii)$: Consider the bounded continuous function, $f = e^{itx}$.
- $(iii) \Rightarrow (i)$: This is the most interesting part of the Levy-Cramer Theorem.

  First, we prove a stronger version with a lesser assumption on $\phi(t)$.

  Let $\phi_n(t)$ *be the characteristic function of* P_n, *for all* $n \geq 1$. *Assume* $\phi_n(t) \to \phi(t)$ *for all real* t *and* $\phi(t)$ *is continuous at* $t = 0$. *Then* $\phi(t)$ *is the* *characteristic function of some probability distribution P and $P_n \stackrel{W}{\Longrightarrow} P$.*

Step 1: Let $r_1, r_2, ...$ be some enumeration of rationals and F_n is the distribution function corresponding to ϕ_n. Since $F_n(r_1)$ is a bounded sequence hence there exists a convergent subsequence $F_{n_k^{(1)}}(r_1)$. Again since $F_{n_k^{(1)}}(r_2)$ is a bounded sequence, there is a convergent subsequence (a subsubsequence of F_n), $F_{n_k^{(2)}}(r)$ that converges at both r_1 and r_2. By induction proceed to create subsequences of previously defined subsequences that also converge at the next rational point. Hence $F_{n_k^{(j)}}(r)$ will converge pointwise at all points $r_1, .., r_j$. Now define $G_k(x) = F_{n_k^{(k)}}(x)$. Observe that this sequence converges at all rational points r_i. (This is called the diagonalization argument.) Call this limit function on rationals to be $G_\infty(r)$.

Step 2: From $G_\infty(r)$, which is defined on rationals, define the function $G(x)$ on the real line as $G(x) = \inf_{\substack{r>x \\ r\in\mathbb{Q}}} G_\infty(r)$. From definition, $G(x)$ is clearly non-decreasing.
If $x_n \downarrow x$ for any rational $r > x$, for large enough n, $r > x_n$ which allows to conclude $G_\infty(r) \geq \inf_n G(x_n) \geq G(x)$. Taking infimum over $r > x$ we get $G(x) \geq \inf_n G(x_n) \geq G(x)$, establishing right continuity.

Step 3: Here we show that at every continuity point of $G(x)$, $\lim_k G_k(x) = G(x)$. For any rational $r > x$ note that $G_k(r) \geq G_k(x)$; hence

$$G_\infty(r) = \lim_k G_k(r) \leq \limsup_k G_k(x).$$

Taking infimum over $r > x$, we obtain

$$G(x) \geq \limsup_k G_k(x).$$

On the other hand, for any $y < x$, take a rational r such that $y < r < x$. Then

$$\liminf_k G_k(x) \geq \lim_k G_k(r) = G_\infty(r) \geq G(y).$$

Since x is a point of continuity of $G()$, letting $y \uparrow x$ yields that

$$G(x) \geq \limsup_k G_k(x) \geq \liminf_k G_k(x) \geq \lim_{y\uparrow x} G(y) = G(x).$$

Thus $F_{n_k^{(k)}}(x)$ converges pointwise to a right continuous, non-decreasing function, $G(x)$ at all continuity points of $G(x)$. Note that $0 \leq G(-\infty) \leq G(\infty) \leq 1$.

Step 4

$$\begin{aligned}
\frac{1}{2T}\int_{-T}^{T} \phi_n(t)dt &= \frac{1}{2T}\int_{-T}^{T}\left(\int e^{itx} dP_n(x)\right) dt \\
&= \int \left(\frac{1}{2T}\int_{-T}^{T} e^{itx} dt\right) dP_n(x) \qquad \text{(Fubini)} \\
&= \int \frac{\sin(Tx)}{Tx} dP_n(x).
\end{aligned}$$

Observe that

$$\begin{aligned}\int \left|\frac{\sin(Tx)}{Tx} dP_n(x)\right| &\leq \int_{x\in(-l,l]} \left|\frac{\sin(Tx)}{Tx} dP_n(x)\right| + \int_{x\notin(-l,l]} \left|\frac{\sin(Tx)}{Tx} dP_n(x)\right| \\ &\leq \int_{x\in(-l,l]} dP_n(x) + \frac{1}{Tl}\int_{x\notin(-l,l]} dP_n(x) \\ &\leq F_n(l) - F_n(-l) + \frac{1}{Tl}(1 - F_n(l) + F_n(-l)) \\ &= (F_n(l) - F_n(-l))\left(1 - \frac{1}{Tl}\right) + \frac{1}{Tl}.\end{aligned}$$

Thus

$$\left|\frac{1}{2T}\int_{-T}^{T} \phi_n(t)dt\right| \leq (F_n(l) - F_n(-l))\left(1 - \frac{1}{Tl}\right) + \frac{1}{Tl}.$$

In particular

$$\left|\frac{1}{2T}\int_{-T}^{T} \phi_{n_k^{(k)}}(t)dt\right| \leq \left(F_{n_k^{(k)}}(l) - F_{n_k^{(k)}}(-l)\right)\left(1 - \frac{1}{Tl}\right) + \frac{1}{Tl}.$$

Again, applying Bounded convergence theorem (to interchange limit and integration) and taking $k \to \infty$ and observing that $l \in \mathbb{N} \subseteq Q$, we get

$$\left|\frac{1}{2T}\int_{-T}^{T} \phi(t)dt\right| = (G_\infty(l) - G_\infty(-l))\left(1 - \frac{1}{Tl}\right) + \frac{1}{Tl}.$$

Let $T = \frac{1}{\sqrt{l}}$ and letting $l \to \infty$ we obtain (from the continuity of $\phi(t)$ at $t = 0$) and definition, non-decreasingness of $G_\infty(r), G(x)$

$$1 = G(\infty) - G(-\infty),$$

implying that $G(x)$ is a distribution function.

Thus $F_{n_k^{(k)}} \Rightarrow G$; however $\phi_{n_k^{(k)}}(t) \to \phi(t)$. Thus $\phi(t)$ is the characteristic function of $G(x)$; and further G is uniquely determined by $\phi(t)$.

Step 5 To complete the argument, we need to show that $F_k \Rightarrow G$, i.e. the entire sequence converges pointwise at all continuity points of G. Assume note, then one can find a subsequence F_{k_n} and a continuity point x_0, of $G(x)$, such that $\lim_n |F_{k_n}(x_0) - G(x_0)| > \epsilon$, for some $\epsilon > 0$. Starting with this subsequence $F_{k_n}(x_0)$ we further find a sub-subsequence that converges to a distribution function; however since $\phi_{k_n}(t) \to \phi(t)$, that distribution function must be $G(x)$, yielding a contradiction.

□

Theorem 3.7 (Portmanteau). *If $P_n \Rightarrow P$, then*

(1) for all closed C, $\limsup_n P_n(C) \to P(C)$.
(2) for all open sets C, $\liminf_n P_n(C) \geq P(C)$.
(3) for all continuity sets of P, i.e. open sets C such that $P(closure(C) \setminus C) = 0$, $\lim_n P_n(C) = P(C)$.

Proof. (1) Let C be closed. Define

$$\hat{d}(x, C) = \inf_{y\in C} |x - y|.$$

Let $f_k(x) = \left(\frac{1}{1+\hat{d}(x,C)}\right)^k$. For every $k \geq$

$$1_C \leq f_k(x), \quad \implies P_n(C) \leq \int f_k(x) dP_n.$$

Since $f_k(x)$ is bounded continuous function,

$$\limsup_n P_n(C) \leq \lim_n \int f_k(x) dP_n = \int f_k dP.$$

Since f_k is bounded and decreases point wise to 1_C which is bounded, monotone convergence theorem (as k goes to infinity) yields

$$\limsup_n P_n(C) \leq \lim_k \int f_k dP \overset{(MCT)}{=} \int f dP = P(C).$$

(2) Taking complements yields this part.
(3) Combining the two yields the third part.

$\square$