

# STAT 7200

## Introduction to Advanced Probability

### Lecture 22

Taylor R. Brown

- 1 Characteristic Function
  - (Levy's) Continuity Theorem
  - Method of Moments

# Continuity Theorem

## Theorem 1 (The Continuity Theorem)

*Let  $\mu, \mu_1, \mu_2, \dots$  be probability measures with characteristic functions  $\phi, \phi_1, \phi_2, \dots$ . Then  $\mu_n$  converges weakly to  $\mu$  if and only if  $\phi_n(t) \rightarrow \phi(t)$  for all  $t \in \mathbb{R}$ . That is, the weak convergence is equivalent to the pointwise convergence of characteristic function.*

# Results So Far

## Theorem 2 (Fourier Uniqueness Theorem)

Let  $X, Y$  be random variables. Then  $\phi_X(t) = \phi_Y(t)$  if and only if  $\mathcal{L}(X) = \mathcal{L}(Y)$ .

- **Tightness of Measure** A collection of probability measure  $\{\mu_n\}$  on  $\mathbb{R}$  is **tight** if for all  $\varepsilon > 0$ , there are  $a < b$  with  $\mu_n([a, b]) \geq 1 - \varepsilon$  for all  $n$ . (Probability mass does not “escape of infinity”).

## Lemma 3

Let  $\{\mu_n\}$  be a sequence of probability measure on  $\mathbb{R}$ , and  $\{\phi_n(t)\}$  be the characteristic functions. If there is a function  $g$  that is continuous at 0, and  $\phi_n(t) \rightarrow g(t)$  for all  $|t| < t_0$  ( $t_0 > 0$ ), then  $\{\mu_n\}$  is tight.

## Tightness and Weak Convergence I

- Our discussion shows that,  $\{\mu_n\}$  in the Continuity Theorem is tight. Here we will introduce a key property of a tight sequence of measure. That is, you may always find a subsequence that converges weakly to a probability measure.

### Theorem 4 (11.1.10)

*Let  $\{\mu_n\}$  be a tight sequence of probability measure, then there must be a subsequence  $\{\mu_{n_k}\}$  and another probability measures  $\mu$ , such that  $\{\mu_{n_k}\}$  converges weakly to  $\mu$ .*

- **Proof:** Let  $F_n(x)$  represents the CDF of  $\mu_n$ . The Helly selection principle (Lemma 11.1.8 of the textbook) states that, for a sequence of CDFs  $\{F_n(x)\}$  (not necessarily tight), there is always a subsequence  $\{F_{n_k}(x)\}$  and a non-decreasing right-continuous function  $F$  ( $0 \leq F \leq 1$ ) so that  $F_n(x) \rightarrow F(x)$  at all continuity points of  $F$ .
- With the extra condition of tightness, we can show that, such  $F$  must be a proper CDF of a probability measure  $\mu$ .

## Tightness and Weak Convergence I: continued

- **Proof continued:** To show  $F$  is a proper CDF, we only need to show that  $\lim_{x \rightarrow \infty} F(x) = 1$  and  $\lim_{x \rightarrow -\infty} F(x) = 0$ .
- Since  $\{\mu_n\}$  is tight,  $\{\mu_{n_k}\}$  is also tight. For any  $\varepsilon > 0$ , we can find  $a < b$  so that  $a, b$  are the continuity points of  $F$  (the discontinuous points of  $F$  is at most countable) and  $\mu_{n_k}((a, b]) \geq 1 - \varepsilon$  for all  $n$ :

$$\begin{aligned}\lim_{x \rightarrow \infty} F(x) - \lim_{x \rightarrow -\infty} F(x) &\geq F(b) - F(a) \\ &= \lim_k [F_{n_k}(b) - F_{n_k}(a)] \\ &= \lim_k \mu_{n_k}((a, b]) \geq 1 - \varepsilon\end{aligned}$$

for all  $\varepsilon > 0$ , then  $\lim_{x \rightarrow \infty} F(x) - \lim_{x \rightarrow -\infty} F(x) \geq 1$  and we have  $\lim_{x \rightarrow \infty} F(x) = 1$  and  $\lim_{x \rightarrow -\infty} F(x) = 0$ .

- For a proper CDF  $F(x)$ , we can always defined measure  $\mu$  as  $\mu((a, b]) = F(b) - F(a)$ . Then by definition,  $\{\mu_{n_k}\}$  converges weakly to  $\mu$ .

## Tightness and Weak Convergence II

- A tight sequence always has a subsequence that converges weakly. The following corollary shows that, if there is only one possible weak limit, the full sequence would converge weakly as well.

### Corollary 5 (11.1.11)

*Let  $\{\mu_n\}$  be a tight sequence of probability measures. Suppose that for any subsequence  $\{\mu_{n_k}\}$ , if it converges weakly, it must converge to probability measure  $\mu$ . Then  $\{\mu_n\}$  must converge to  $\mu$ .*

- **Proof:** Suppose that  $\{\mu_n\}$  does not converge to  $\mu$ , then  $F_n(x)$  would not converge to  $F(x)$  at some  $x$  with  $\mu(\{x\}) = 0$ . Hence, we can find  $x \in \mathbb{R}$ ,  $\varepsilon > 0$  and a subsequence  $\{n_k\}$  so that for all  $k$ ,  
 $|F_{n_k}(x) - F(x)| \geq \varepsilon$ .
- As  $\{\mu_{n_k}\}$  is tight, too, there must be a further subsequence of  $\{\mu_{n_{k_j}}\}$  that converges weakly. However, according to the condition, this subsequence of  $\{\mu_{n_{k_j}}\}$  must converge weakly to  $\mu$ . This is a contradiction.

## Back to Continuity Theorem

- Now we have enough results to prove the “if” part of continuity theorem, that is, if  $\phi_n(t) \rightarrow \phi(t)$  for all  $t \in \mathbb{R}$  then  $\{\mu_n\}$  converges weakly to  $\mu$ . The argument runs as following:
  - (a) As  $\phi_n(t) \rightarrow \phi(t)$  for all  $t \in \mathbb{R}$ , and  $\phi(t)$  is continuous at 0, then  $\{\mu_n\}$  is tight (Theorem 11.1.13).
  - (b) As  $\{\mu_n\}$  is tight, there must be a subsequence  $\{\mu_{n_k}\}$  that converges weakly to a probability measure  $\nu$  (Theorem 11.1.10).
  - (c) As  $\{\mu_{n_k}\}$  converges weakly to a probability measure  $\nu$ , the corresponding characteristic functions  $\phi_{n_k}(t) \rightarrow \phi_\nu(t)$  for all  $t \in \mathbb{R}$  (easy half of this proof we did last class).
  - (d) But  $\phi_{n_k}(t) \rightarrow \phi(t)$  as well, then  $\phi(t) = \phi_\nu(t)$ . By the Fourier uniqueness theorem (Corollary 11.1.7),  $\nu = \mu$ .
  - (e) Then  $\mu$  is the only possible weak limit of  $\{\mu_n\}$ . Thus,  $\{\mu_n\}$  converges weakly to  $\mu$  (Corollary 11.1.11).

QED.

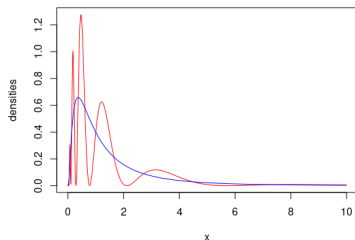


# Moment and Weak Convergences

- The continuity theorem demonstrates that the weak convergence is equivalent to the convergence of characteristic functions. On the other hand, the convergence of moments (even all the moments) does not necessarily indicate the weak convergence. Still, under the following condition, we can still determine the weak convergence based on the convergence of moments.
- **Determined by the Moments:** We say a distribution  $\mu$  is determined by its moments if: 1) For all  $k \in \mathbb{N}$ ,  $\int |x^k| \mu(dx) < \infty$ . 2) If there is another distribution  $\nu$  so that  $\int x^k \mu(dx) = \int x^k \nu(dx)$  for all  $k \in \mathbb{N}$ , then  $\mu = \nu$ .

# Moment and Weak Convergences

**Remark** Not all distributions are determined by their moments. Example: a log-normal distribution. Let  $g(x)$  be the density function of a log-normal distribution. Then define  $h(x) = g(x)[1 + \sin(2\pi \log x)]$ , then we can show that  $g(x)$  is a proper density function, and the distribution determined by  $g(x)$  has the same moments as the log-normal distribution.



# Determined by the Moments and Weak Convergence

## Theorem 6 (11.4.1)

*If  $\mu$  is determined by its moments. Let  $\{\mu_n\}$  be a sequence of distributions whose moments are all finite, and  $\lim_n \int x^k \mu_n(dx) = \int x^k \mu(dx)$  for each  $k \in \mathbb{N}$ . Then  $\{\mu_n\}$  converges weakly to  $\mu$ .*

- **Proof:** First we will show that  $\{\mu_n\}$  is tight
  - Note that as  $\lim_n \int x^k \mu_n(dx) = \int x^k \mu(dx) < \infty$ , for each  $k$ , we can find finite constants  $M_k$  such that  $|\int x^k \mu_n(dx)| \leq M_k$ .
  - Let  $Y_n \sim \mu_n$ , then

$$\mu_n([-R, R]) = 1 - P(|Y_n|^2 > R^2) \geq 1 - \frac{E(|Y_n|^2)}{R^2} \geq 1 - \frac{M_2}{R^2},$$

- for all  $n$ . Then we can always choose  $R$  big enough, so that  $\mu_n([-R, R]) \geq 1 - \frac{M_2}{R^2} \geq 1 - \varepsilon$ . Thus  $\{\mu_n\}$  is tight.

# Determined by the Moments and Weak Convergence

- **continued:** By Theorem 11.1.10, every tight sequence of measures has a weakly convergent subsequence.
- **continued:** Recall

## Corollary 7 (11.1.11)

*Let  $\{\mu_n\}$  be a tight sequence of probability measures. Suppose that for any subsequence  $\{\mu_{n_k}\}$ , if it converges weakly, it must converge to probability measure  $\mu$ . Then  $\{\mu_n\}$  must converge to  $\mu$ .*

- If we can verify the antecedent, then we can prove our final result.

## Determined by the Moments and Weak Convergence

- **continued:** Suppose we can find a subsequence  $\{\mu_{n_r}\}$  that converges weakly to some measure  $\nu$ . By Skorohod's representation theorem, we can find random variables  $Y$  and  $\{Y_r\}$  such that  $Y \sim \nu$ , and  $Y_r \sim \mu_{n_r}$  and  $Y_r \xrightarrow{\text{a.s.}} Y$  almost surely. Furthermore,  $Y_r^k \xrightarrow{\text{a.s.}} Y^k$ .
- Next

$$E(|Y_r|^k 1_{|Y_r|^k \geq \alpha}) \leq E\left(\frac{|Y_r|^{2k}}{\alpha} 1_{|Y_r|^k \geq \alpha}\right) \leq \frac{1}{\alpha} E(|Y_r|^{2k}) \leq \frac{M_{2k}}{\alpha},$$

for  $\alpha > 0$ .

- Thus, as  $\alpha \rightarrow \infty$ ,

$$\sup_r E(|Y_r|^k 1_{|Y_r|^k \geq \alpha}) \leq \frac{M_{2k}}{\alpha} \rightarrow 0.$$

Hence  $\{Y_r^k\}$  are uniformly integrable.

- By the uniform integrability convergence theorem  $E[Y_r^k] \rightarrow E[Y^k]$ , so  $\lim_n \int x^k \mu_{n_r}(dx) = \int x^k \nu(dx)$ .

# Determined by the Moments and Weak Convergence

- Now we want to show that  $\nu$  must equal  $\mu$ . By hypothesis, all the moments converge to moments of  $\mu$ . And by the previous slide  $\lim_n \int x^k \mu_{n_r}(dx) = \int x^k \nu(dx)$ .
- We have  $\lim_n \int x^k \mu_{n_r}(dx) = \int x^k \mu(dx)$ . Then  $\int x^k \mu(dx) = \int x^k \nu(dx)$  for all  $k$ . As  $\mu$  is determined by its moments, we have  $\mu = \nu$ . That is,  $\mu$  is the only weak limit of  $\{\mu_n\}$ .
- So we can invoke Corollary 11.1.11, and thus we have  $\mu_n \Rightarrow \mu$ .

# Determined by the Moments

We will skip the proof of the following result:

## Theorem 8 (11.4.3)

*Let  $X$  be a random variable. If for some  $s_0 > 0$ ,  $M_X(s) < \infty$  when  $|s| < s_0$ , then  $\mathcal{L}(X)$  is determined by its moments.*