

# Stat 709: Mathematical Statistics

## Lecture 13

Jun Shao

Department of Statistics  
University of Wisconsin  
Madison, WI 53706, USA

# Lecture 13: Weak convergence

## Tightness

A sequence  $\{P_n\}$  of probability measures on  $(\mathcal{R}^k, \mathcal{B}^k)$  is *tight* if for every  $\varepsilon > 0$ , there is a compact set  $C \subset \mathcal{R}^k$  such that  $\inf_n P_n(C) > 1 - \varepsilon$ .

## Proposition 1.17

Let  $\{P_n\}$  be a sequence of probability measures on  $(\mathcal{R}^k, \mathcal{B}^k)$ .

- (i) Tightness of  $\{P_n\}$  is a necessary and sufficient condition that for every subsequence  $\{P_{n_i}\}$  there exists a further subsequence  $\{P_{n_j}\} \subset \{P_{n_i}\}$  and a probability measure  $P$  on  $(\mathcal{R}^k, \mathcal{B}^k)$  such that  $P_{n_j} \rightarrow_w P$  as  $j \rightarrow \infty$ .
- (ii) If  $\{P_n\}$  is tight and if each subsequence that converges weakly at all converges to the same probability measure  $P$ , then  $P_n \rightarrow_w P$ .

## Proof

See Billingsley (1995, pp. 336-337).

# Lecture 13: Weak convergence

## Tightness

A sequence  $\{P_n\}$  of probability measures on  $(\mathcal{R}^k, \mathcal{B}^k)$  is *tight* if for every  $\varepsilon > 0$ , there is a compact set  $C \subset \mathcal{R}^k$  such that  $\inf_n P_n(C) > 1 - \varepsilon$ .

## Proposition 1.17

Let  $\{P_n\}$  be a sequence of probability measures on  $(\mathcal{R}^k, \mathcal{B}^k)$ .

- (i) Tightness of  $\{P_n\}$  is a necessary and sufficient condition that for every subsequence  $\{P_{n_i}\}$  there exists a further subsequence  $\{P_{n_j}\} \subset \{P_{n_i}\}$  and a probability measure  $P$  on  $(\mathcal{R}^k, \mathcal{B}^k)$  such that  $P_{n_j} \rightarrow_w P$  as  $j \rightarrow \infty$ .
- (ii) If  $\{P_n\}$  is tight and if each subsequence that converges weakly at all converges to the same probability measure  $P$ , then  $P_n \rightarrow_w P$ .

## Proof

See Billingsley (1995, pp. 336-337).

# Lecture 13: Weak convergence

## Tightness

A sequence  $\{P_n\}$  of probability measures on  $(\mathcal{R}^k, \mathcal{B}^k)$  is *tight* if for every  $\varepsilon > 0$ , there is a compact set  $C \subset \mathcal{R}^k$  such that  $\inf_n P_n(C) > 1 - \varepsilon$ .

## Proposition 1.17

Let  $\{P_n\}$  be a sequence of probability measures on  $(\mathcal{R}^k, \mathcal{B}^k)$ .

- (i) Tightness of  $\{P_n\}$  is a necessary and sufficient condition that for every subsequence  $\{P_{n_i}\}$  there exists a further subsequence  $\{P_{n_j}\} \subset \{P_{n_i}\}$  and a probability measure  $P$  on  $(\mathcal{R}^k, \mathcal{B}^k)$  such that  $P_{n_j} \rightarrow_w P$  as  $j \rightarrow \infty$ .
- (ii) If  $\{P_n\}$  is tight and if each subsequence that converges weakly at all converges to the same probability measure  $P$ , then  $P_n \rightarrow_w P$ .

## Proof

See Billingsley (1995, pp. 336-337).

## Remark

If  $\{X_n\}$  is a sequence of random  $k$ -vectors, then the tightness of  $\{P_{X_n}\}$  is the same as the boundedness of  $\{\|X_n\|\}$  in probability ( $\|X_n\| = O_p(1)$ ), i.e., for any  $\varepsilon > 0$ , there is a constant  $C_\varepsilon > 0$  such that  $\sup_n P(\|X_n\| \geq C_\varepsilon) < \varepsilon$ .

Theorem 1.9 (useful sufficient and necessary conditions for convergence in distribution)

Let  $X, X_1, X_2, \dots$  be random  $k$ -vectors.

- (i)  $X_n \rightarrow_d X$  is equivalent to any one of the following conditions:
  - (a)  $E[h(X_n)] \rightarrow E[h(X)]$  for every bounded continuous function  $h$ ;
  - (b)  $\limsup_n P_{X_n}(C) \leq P_X(C)$  for any closed set  $C \subset \mathcal{R}^k$ ;
  - (c)  $\liminf_n P_{X_n}(O) \geq P_X(O)$  for any open set  $O \subset \mathcal{R}^k$ .
- (ii) (Lévy-Cramér continuity theorem). Let  $\phi_X, \phi_{X_1}, \phi_{X_2}, \dots$  be the ch.f.'s of  $X, X_1, X_2, \dots$ , respectively.  
 $X_n \rightarrow_d X$  iff  $\lim_{n \rightarrow \infty} \phi_{X_n}(t) = \phi_X(t)$  for all  $t \in \mathcal{R}^k$ .
- (iii) (Cramér-Wold device).  $X_n \rightarrow_d X$  iff  $c^\tau X_n \rightarrow_d c^\tau X$  for every  $c \in \mathcal{R}^k$ .

## Remark

If  $\{X_n\}$  is a sequence of random  $k$ -vectors, then the tightness of  $\{P_{X_n}\}$  is the same as the boundedness of  $\{\|X_n\|\}$  in probability ( $\|X_n\| = O_p(1)$ ), i.e., for any  $\varepsilon > 0$ , there is a constant  $C_\varepsilon > 0$  such that  $\sup_n P(\|X_n\| \geq C_\varepsilon) < \varepsilon$ .

## Theorem 1.9 (useful sufficient and necessary conditions for convergence in distribution)

Let  $X, X_1, X_2, \dots$  be random  $k$ -vectors.

- (i)  $X_n \rightarrow_d X$  is equivalent to any one of the following conditions:
  - (a)  $E[h(X_n)] \rightarrow E[h(X)]$  for every bounded continuous function  $h$ ;
  - (b)  $\limsup_n P_{X_n}(C) \leq P_X(C)$  for any closed set  $C \subset \mathcal{R}^k$ ;
  - (c)  $\liminf_n P_{X_n}(O) \geq P_X(O)$  for any open set  $O \subset \mathcal{R}^k$ .
- (ii) (Lévy-Cramér continuity theorem). Let  $\phi_X, \phi_{X_1}, \phi_{X_2}, \dots$  be the ch.f.'s of  $X, X_1, X_2, \dots$ , respectively.  
 $X_n \rightarrow_d X$  iff  $\lim_{n \rightarrow \infty} \phi_{X_n}(t) = \phi_X(t)$  for all  $t \in \mathcal{R}^k$ .
- (iii) (Cramér-Wold device).  $X_n \rightarrow_d X$  iff  $c^\tau X_n \rightarrow_d c^\tau X$  for every  $c \in \mathcal{R}^k$ .

## Proof of Theorem 1.9(i)

First, we show  $X_n \rightarrow_d X$  implies (a).

By Theorem 1.8(iv) (Skorohod's theorem), there exists a sequence of random vectors  $\{Y_n\}$  and a random vector  $Y$  such that  $P_{Y_n} = P_{X_n}$  for all  $n$ ,  $P_Y = P_X$  and  $Y_n \rightarrow_{a.s.} Y$ .

For bounded continuous  $h$ ,  $h(Y_n) \rightarrow_{a.s.} h(Y)$  and, by the dominated convergence theorem,  $E[h(Y_n)] \rightarrow E[h(Y)]$ .

(a) follows from  $E[h(X_n)] = E[h(Y_n)]$  for all  $n$  and  $E[h(X)] = E[h(Y)]$ .

Next, we show (a) implies (b).

Let  $C$  be a closed set and  $f_C(x) = \inf\{\|x - y\| : y \in C\}$ .

Then  $f_C$  is continuous.

For  $j = 1, 2, \dots$ , define  $\varphi_j(t) = I_{(-\infty, 0]} + (1 - jt)I_{(0, j^{-1}]}$ .

Then  $h_j(x) = \varphi_j(f_C(x))$  is continuous and bounded,  $h_j \geq h_{j+1}$ ,

$j = 1, 2, \dots$ , and  $h_j(x) \rightarrow I_C(x)$  as  $j \rightarrow \infty$ .

Hence  $\limsup_n P_{X_n}(C) \leq \lim_{n \rightarrow \infty} E[h_j(X_n)] = E[h_j(X)]$  for each  $j$  (by (a)).

By the dominated convergence theorem,

$E[h_j(X)] \rightarrow E[I_C(X)] = P_X(C)$ .

This proves (b).

## Proof of Theorem 1.9(i)

First, we show  $X_n \rightarrow_d X$  implies (a).

By Theorem 1.8(iv) (Skorohod's theorem), there exists a sequence of random vectors  $\{Y_n\}$  and a random vector  $Y$  such that  $P_{Y_n} = P_{X_n}$  for all  $n$ ,  $P_Y = P_X$  and  $Y_n \rightarrow_{a.s.} Y$ .

For bounded continuous  $h$ ,  $h(Y_n) \rightarrow_{a.s.} h(Y)$  and, by the dominated convergence theorem,  $E[h(Y_n)] \rightarrow E[h(Y)]$ .

(a) follows from  $E[h(X_n)] = E[h(Y_n)]$  for all  $n$  and  $E[h(X)] = E[h(Y)]$ .

Next, we show (a) implies (b).

Let  $C$  be a closed set and  $f_C(x) = \inf\{\|x - y\| : y \in C\}$ .

Then  $f_C$  is continuous.

For  $j = 1, 2, \dots$ , define  $\varphi_j(t) = I_{(-\infty, 0]} + (1 - jt)I_{(0, j^{-1}]}$ .

Then  $h_j(x) = \varphi_j(f_C(x))$  is continuous and bounded,  $h_j \geq h_{j+1}$ ,  $j = 1, 2, \dots$ , and  $h_j(x) \rightarrow I_C(x)$  as  $j \rightarrow \infty$ .

Hence  $\limsup_n P_{X_n}(C) \leq \lim_{n \rightarrow \infty} E[h_j(X_n)] = E[h_j(X)]$  for each  $j$  (by (a)).

By the dominated convergence theorem,

$$E[h_j(X)] \rightarrow E[I_C(X)] = P_X(C).$$

This proves (b).

## Proof of Theorem 1.9(i) (continued)

For any open set  $O$ ,  $O^c$  is closed.

Hence, (b) is equivalent to (c).

To complete the proof we now show that (b) and (c) imply  $X_n \rightarrow_d X$ .

For  $x = (x_1, \dots, x_k) \in \mathcal{R}^k$ , let  $(-\infty, x] = (-\infty, x_1] \times \dots \times (-\infty, x_k]$  and  $(-\infty, x) = (-\infty, x_1) \times \dots \times (-\infty, x_k)$ .

From (b) and (c),

$$\begin{aligned} P_X((-\infty, x)) &\leq \liminf_n P_{X_n}((-\infty, x)) \leq \liminf_n F_{X_n}(x) \\ &\leq \limsup_n F_{X_n}(x) = \limsup_n P_{X_n}((-\infty, x]) \leq P_X((-\infty, x]) = F_X(x). \end{aligned}$$

If  $x$  is a continuity point of  $F_X$ , then  $P_X((-\infty, x)) = F_X(x)$ .

This proves  $X_n \rightarrow_d X$ .

## Proof of Theorem 1.9(ii)

From (a) of part (i),  $X_n \rightarrow_d X$  implies  $\phi_{X_n}(t) \rightarrow \phi_X(t)$ , since

$e^{\sqrt{-1}t^T x} = \cos(t^T x) + \sqrt{-1} \sin(t^T x)$  and  $\cos(t^T x)$  and  $\sin(t^T x)$  are bounded continuous functions for any fixed  $t$ .

## Proof of Theorem 1.9(i) (continued)

For any open set  $O$ ,  $O^c$  is closed.

Hence, (b) is equivalent to (c).

To complete the proof we now show that (b) and (c) imply  $X_n \rightarrow_d X$ .

For  $x = (x_1, \dots, x_k) \in \mathcal{R}^k$ , let  $(-\infty, x] = (-\infty, x_1] \times \dots \times (-\infty, x_k]$  and  $(-\infty, x) = (-\infty, x_1) \times \dots \times (-\infty, x_k)$ .

From (b) and (c),

$$\begin{aligned} P_X((-\infty, x)) &\leq \liminf_n P_{X_n}((-\infty, x)) \leq \liminf_n F_{X_n}(x) \\ &\leq \limsup_n F_{X_n}(x) = \limsup_n P_{X_n}((-\infty, x]) \leq P_X((-\infty, x]) = F_X(x). \end{aligned}$$

If  $x$  is a continuity point of  $F_X$ , then  $P_X((-\infty, x)) = F_X(x)$ .

This proves  $X_n \rightarrow_d X$ .

## Proof of Theorem 1.9(ii)

From (a) of part (i),  $X_n \rightarrow_d X$  implies  $\phi_{X_n}(t) \rightarrow \phi_X(t)$ , since

$e^{\sqrt{-1}t^\tau x} = \cos(t^\tau x) + \sqrt{-1} \sin(t^\tau x)$  and  $\cos(t^\tau x)$  and  $\sin(t^\tau x)$  are bounded continuous functions for any fixed  $t$ .

## Proof of Theorem 1.9(ii) (continued)

Suppose that  $k = 1$  and that  $\phi_{X_n}(t) \rightarrow \phi_X(t)$  for every  $t \in \mathcal{R}$ .

We want to show that  $X_n \rightarrow_d X$ .

We first show that  $\{P_{X_n}\}$  is tight.

By Fubini's theorem,

$$\begin{aligned} \frac{1}{u} \int_{-u}^u [1 - \phi_{X_n}(t)] dt &= \int_{-\infty}^{\infty} \left[ \frac{1}{u} \int_{-u}^u (1 - e^{\sqrt{-1}tx}) dt \right] dP_{X_n}(x) \\ &= 2 \int_{-\infty}^{\infty} \left( 1 - \frac{\sin ux}{ux} \right) dP_{X_n}(x) \\ &\geq 2 \int_{\{|x| > 2u^{-1}\}} \left( 1 - \frac{1}{|ux|} \right) dP_{X_n}(x) \\ &\geq P_{X_n} \left( (-\infty, -2u^{-1}) \cup (2u^{-1}, \infty) \right) \end{aligned}$$

for any  $u > 0$ .

Since  $\phi_X$  is continuous at 0 and  $\phi_X(0) = 1$ , for any  $\varepsilon > 0$  there is a  $u > 0$  such that  $u^{-1} \int_{-u}^u [1 - \phi_X(t)] dt < \varepsilon/2$ .

## Proof of Theorem 1.9(ii) (continued)

Since  $\phi_{X_n} \rightarrow \phi_X$ , by the dominated convergence theorem,

$$\sup_n \left\{ u^{-1} \int_{-u}^u [1 - \phi_{X_n}(t)] dt \right\} < \varepsilon.$$

Hence,

$$\inf_n P_{X_n} \left( [-2u^{-1}, 2u^{-1}] \right) \geq 1 - \sup_n \left\{ \frac{1}{u} \int_{-u}^u [1 - \phi_{X_n}(t)] dt \right\} \geq 1 - \varepsilon,$$

i.e.,  $\{P_{X_n}\}$  is tight.

Let  $\{P_{X_{n_j}}\}$  be any subsequence that converges to a probability measure  $P$ .

By the first part of the proof,  $\phi_{X_{n_j}} \rightarrow \phi$ , which is the ch.f. of  $P$ .

By the convergence of  $\phi_{X_n}$ ,  $\phi = \phi_X$ .

By the uniqueness theorem,  $P = P_X$ .

By Proposition 1.17(ii),  $X_n \rightarrow_d X$ .

## Proof of Theorem 1.9(ii) (continued)

Consider now the case where  $k \geq 2$  and  $\phi_{X_n} \rightarrow \phi_X$ .

Let  $Y_{nj}$  be the  $j$ th component of  $X_n$  and  $Y_j$  be the  $j$ th component of  $X$ .

Then  $\phi_{Y_{nj}} \rightarrow \phi_{Y_j}$  for each  $j$ .

By the proof for the case of  $k = 1$ ,  $Y_{nj} \rightarrow_d Y_j$ .

By Proposition 1.17(i),  $\{P_{Y_{nj}}\}$  is tight,  $j = 1, \dots, k$ .

This implies that  $\{P_{X_n}\}$  is tight (why?).

Then the proof for  $X_n \rightarrow_d X$  is the same as that for the case of  $k = 1$ .

## Proof of Theorem 1.9(iii)

Note that  $\phi_{c^\tau X_n}(u) = \phi_{X_n}(uc)$  and  $\phi_{c^\tau X}(u) = \phi_X(uc)$  for any  $u \in \mathcal{R}$  and any  $c \in \mathcal{R}^k$ .

Hence, convergence of  $\phi_{X_n}$  to  $\phi_X$  is equivalent to convergence of  $\phi_{c^\tau X_n}$  to  $\phi_{c^\tau X}$  for every  $c \in \mathcal{R}^k$ .

Then the result follows from part (ii).

## Proof of Theorem 1.9(ii) (continued)

Consider now the case where  $k \geq 2$  and  $\phi_{X_n} \rightarrow \phi_X$ .

Let  $Y_{nj}$  be the  $j$ th component of  $X_n$  and  $Y_j$  be the  $j$ th component of  $X$ .

Then  $\phi_{Y_{nj}} \rightarrow \phi_{Y_j}$  for each  $j$ .

By the proof for the case of  $k = 1$ ,  $Y_{nj} \rightarrow_d Y_j$ .

By Proposition 1.17(i),  $\{P_{Y_{nj}}\}$  is tight,  $j = 1, \dots, k$ .

This implies that  $\{P_{X_n}\}$  is tight (why?).

Then the proof for  $X_n \rightarrow_d X$  is the same as that for the case of  $k = 1$ .

## Proof of Theorem 1.9(iii)

Note that  $\phi_{c^\tau X_n}(u) = \phi_{X_n}(uc)$  and  $\phi_{c^\tau X}(u) = \phi_X(uc)$  for any  $u \in \mathcal{R}$  and any  $c \in \mathcal{R}^k$ .

Hence, convergence of  $\phi_{X_n}$  to  $\phi_X$  is equivalent to convergence of  $\phi_{c^\tau X_n}$  to  $\phi_{c^\tau X}$  for every  $c \in \mathcal{R}^k$ .

Then the result follows from part (ii).

## Example 1.28

Let  $X_1, \dots, X_n$  be independent random variables having a common c.d.f. and  $T_n = X_1 + \dots + X_n$ ,  $n = 1, 2, \dots$

Suppose that  $E|X_1| < \infty$ .

It follows from a result in calculus that the ch.f. of  $X_1$  satisfies

$$\phi_{X_1}(t) = \phi_{X_1}(0) + \sqrt{-1}\mu t + o(|t|)$$

as  $|t| \rightarrow 0$ , where  $\mu = EX_1$ .

Then, the ch.f. of  $T_n/n$  is

$$\phi_{T_n/n}(t) = \left[ \phi_{X_1}\left(\frac{t}{n}\right) \right]^n = \left[ 1 + \frac{\sqrt{-1}\mu t}{n} + o\left(\frac{t}{n}\right) \right]^n \rightarrow e^{\sqrt{-1}\mu t}$$

for any  $t \in \mathcal{R}$  as  $n \rightarrow \infty$ , because  $(1 + c_n/n)^n \rightarrow e^c$  for any complex sequence  $\{c_n\}$  satisfying  $c_n \rightarrow c$ .

$e^{\sqrt{-1}\mu t}$  is the ch.f. of the point mass probability measure at  $\mu$ .

By Theorem 1.9(ii),  $T_n/n \rightarrow_d \mu$ .

From Theorem 1.8(vii), this also shows that  $T_n/n \rightarrow_p \mu$ .

## Example 1.28

Let  $X_1, \dots, X_n$  be independent random variables having a common c.d.f. and  $T_n = X_1 + \dots + X_n$ ,  $n = 1, 2, \dots$

Suppose that  $E|X_1| < \infty$ .

It follows from a result in calculus that the ch.f. of  $X_1$  satisfies

$$\phi_{X_1}(t) = \phi_{X_1}(0) + \sqrt{-1}\mu t + o(|t|)$$

as  $|t| \rightarrow 0$ , where  $\mu = EX_1$ .

Then, the ch.f. of  $T_n/n$  is

$$\phi_{T_n/n}(t) = \left[ \phi_{X_1}\left(\frac{t}{n}\right) \right]^n = \left[ 1 + \frac{\sqrt{-1}\mu t}{n} + o\left(\frac{t}{n}\right) \right]^n \rightarrow e^{\sqrt{-1}\mu t}$$

for any  $t \in \mathcal{R}$  as  $n \rightarrow \infty$ , because  $(1 + c_n/n)^n \rightarrow e^c$  for any complex sequence  $\{c_n\}$  satisfying  $c_n \rightarrow c$ .

$e^{\sqrt{-1}\mu t}$  is the ch.f. of the point mass probability measure at  $\mu$ .

By Theorem 1.9(ii),  $T_n/n \rightarrow_d \mu$ .

From Theorem 1.8(vii), this also shows that  $T_n/n \rightarrow_p \mu$ .

## Example 1.28 (continued)

Similarly,  $\mu = 0$  and  $\sigma^2 = \text{var}(X_1) < \infty$  imply

$$\phi_{T_n/\sqrt{n}}(t) = \left[ 1 - \frac{\sigma^2 t^2}{2n} + o\left(\frac{t^2}{n}\right) \right]^n \rightarrow e^{-\sigma^2 t^2/2}$$

for any  $t \in \mathcal{R}$  as  $n \rightarrow \infty$ .

$e^{-\sigma^2 t^2/2}$  is the ch.f. of  $N(0, \sigma^2)$ .

Hence  $T_n/\sqrt{n} \rightarrow_d N(0, \sigma^2)$ .

If  $\mu \neq 0$ , a transformation of  $Y_i = X_i - \mu$  leads to

$$(T_n - n\mu)/\sqrt{n} \rightarrow_d N(0, \sigma^2).$$

Suppose now that  $X_1, \dots, X_n$  are random  $k$ -vectors and  $\mu = EX_1$  and  $\Sigma = \text{var}(X_1)$  are finite.

For any fixed  $c \in \mathcal{R}^k$ , it follows from the previous discussion that  $(c^\tau T_n - nc^\tau \mu)/\sqrt{n} \rightarrow_d N(0, c^\tau \Sigma c)$ .

From Theorem 1.9(iii) and a property of the normal distribution (Exercise 81), we conclude that

$$(T_n - n\mu)/\sqrt{n} \rightarrow_d N_k(0, \Sigma).$$

## Example 1.28 (continued)

Similarly,  $\mu = 0$  and  $\sigma^2 = \text{var}(X_1) < \infty$  imply

$$\phi_{T_n/\sqrt{n}}(t) = \left[ 1 - \frac{\sigma^2 t^2}{2n} + o\left(\frac{t^2}{n}\right) \right]^n \rightarrow e^{-\sigma^2 t^2/2}$$

for any  $t \in \mathcal{R}$  as  $n \rightarrow \infty$ .

$e^{-\sigma^2 t^2/2}$  is the ch.f. of  $N(0, \sigma^2)$ .

Hence  $T_n/\sqrt{n} \rightarrow_d N(0, \sigma^2)$ .

If  $\mu \neq 0$ , a transformation of  $Y_i = X_i - \mu$  leads to

$$(T_n - n\mu)/\sqrt{n} \rightarrow_d N(0, \sigma^2).$$

Suppose now that  $X_1, \dots, X_n$  are random  $k$ -vectors and  $\mu = EX_1$  and  $\Sigma = \text{var}(X_1)$  are finite.

For any fixed  $c \in \mathcal{R}^k$ , it follows from the previous discussion that  $(c^\tau T_n - nc^\tau \mu)/\sqrt{n} \rightarrow_d N(0, c^\tau \Sigma c)$ .

From Theorem 1.9(iii) and a property of the normal distribution (Exercise 81), we conclude that

$$(T_n - n\mu)/\sqrt{n} \rightarrow_d N_k(0, \Sigma).$$

## Example 1.29

Let  $X_1, \dots, X_n$  be independent random variables having a common Lebesgue p.d.f.  $f(x) = (1 - \cos x)/(\pi x^2)$ .

Then the ch.f. of  $X_1$  is  $\max\{1 - |t|, 0\}$  (Exercise 73) and the ch.f. of  $T_n/n = (X_1 + \dots + X_n)/n$  is

$$\left( \max \left\{ 1 - \frac{|t|}{n}, 0 \right\} \right)^n \rightarrow e^{-|t|}, \quad t \in \mathcal{R}.$$

Since  $e^{-|t|}$  is the ch.f. of the Cauchy distribution  $C(0, 1)$  (Table 1.2), we conclude that  $T_n/n \rightarrow_d X$ , where  $X$  has the Cauchy distribution  $C(0, 1)$ .

- Does this result contradict the first result in Example 1.28?
- Other examples are given in Exercises 135-140.

The next result can be used to check whether  $X_n \rightarrow_d X$  when  $X$  has a p.d.f.  $f$  and  $X_n$  has a p.d.f.  $f_n$ .

## Example 1.29

Let  $X_1, \dots, X_n$  be independent random variables having a common Lebesgue p.d.f.  $f(x) = (1 - \cos x)/(\pi x^2)$ .

Then the ch.f. of  $X_1$  is  $\max\{1 - |t|, 0\}$  (Exercise 73) and the ch.f. of  $T_n/n = (X_1 + \dots + X_n)/n$  is

$$\left( \max \left\{ 1 - \frac{|t|}{n}, 0 \right\} \right)^n \rightarrow e^{-|t|}, \quad t \in \mathcal{R}.$$

Since  $e^{-|t|}$  is the ch.f. of the Cauchy distribution  $C(0, 1)$  (Table 1.2), we conclude that  $T_n/n \rightarrow_d X$ , where  $X$  has the Cauchy distribution  $C(0, 1)$ .

- Does this result contradict the first result in Example 1.28?
- Other examples are given in Exercises 135-140.

The next result can be used to check whether  $X_n \rightarrow_d X$  when  $X$  has a p.d.f.  $f$  and  $X_n$  has a p.d.f.  $f_n$ .

## Example 1.29

Let  $X_1, \dots, X_n$  be independent random variables having a common Lebesgue p.d.f.  $f(x) = (1 - \cos x)/(\pi x^2)$ .

Then the ch.f. of  $X_1$  is  $\max\{1 - |t|, 0\}$  (Exercise 73) and the ch.f. of  $T_n/n = (X_1 + \dots + X_n)/n$  is

$$\left( \max \left\{ 1 - \frac{|t|}{n}, 0 \right\} \right)^n \rightarrow e^{-|t|}, \quad t \in \mathcal{R}.$$

Since  $e^{-|t|}$  is the ch.f. of the Cauchy distribution  $C(0, 1)$  (Table 1.2), we conclude that  $T_n/n \rightarrow_d X$ , where  $X$  has the Cauchy distribution  $C(0, 1)$ .

- Does this result contradict the first result in Example 1.28?
- Other examples are given in Exercises 135-140.

The next result can be used to check whether  $X_n \rightarrow_d X$  when  $X$  has a p.d.f.  $f$  and  $X_n$  has a p.d.f.  $f_n$ .

## Proposition 1.18 (Scheffé's theorem)

Let  $\{f_n\}$  be a sequence of p.d.f.'s on  $\mathcal{R}^k$  w.r.t. a measure  $\nu$ .  
Suppose that  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  a.e.  $\nu$  and  $f(x)$  is a p.d.f. w.r.t.  $\nu$ .  
Then  $\lim_{n \rightarrow \infty} \int |f_n(x) - f(x)| d\nu = 0$ .

### Proof

Let  $g_n(x) = [f(x) - f_n(x)] I_{\{f \geq f_n\}}(x)$ ,  $n = 1, 2, \dots$

Then

$$\int |f_n(x) - f(x)| d\nu = 2 \int g_n(x) d\nu.$$

Since  $0 \leq g_n(x) \leq f(x)$  for all  $x$  and  $g_n \rightarrow 0$  a.e.  $\nu$ , the result follows from the dominated convergence theorem.

As an example, consider the Lebesgue p.d.f.  $f_n$  of the t-distribution  $t_n$  (Table 1.2),  $n = 1, 2, \dots$

One can show (exercise) that  $f_n \rightarrow f$ , where  $f$  is the p.d.f. of  $N(0, 1)$ .  
This is an important result in statistics.

## Proposition 1.18 (Scheffé's theorem)

Let  $\{f_n\}$  be a sequence of p.d.f.'s on  $\mathcal{R}^k$  w.r.t. a measure  $\nu$ .  
Suppose that  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  a.e.  $\nu$  and  $f(x)$  is a p.d.f. w.r.t.  $\nu$ .  
Then  $\lim_{n \rightarrow \infty} \int |f_n(x) - f(x)| d\nu = 0$ .

## Proof

Let  $g_n(x) = [f(x) - f_n(x)]I_{\{f \geq f_n\}}(x)$ ,  $n = 1, 2, \dots$

Then

$$\int |f_n(x) - f(x)| d\nu = 2 \int g_n(x) d\nu.$$

Since  $0 \leq g_n(x) \leq f(x)$  for all  $x$  and  $g_n \rightarrow 0$  a.e.  $\nu$ , the result follows from the dominated convergence theorem.

As an example, consider the Lebesgue p.d.f.  $f_n$  of the t-distribution  $t_n$  (Table 1.2),  $n = 1, 2, \dots$

One can show (exercise) that  $f_n \rightarrow f$ , where  $f$  is the p.d.f. of  $N(0, 1)$ .  
This is an important result in statistics.

## Proposition 1.18 (Scheffé's theorem)

Let  $\{f_n\}$  be a sequence of p.d.f.'s on  $\mathcal{R}^k$  w.r.t. a measure  $\nu$ .  
Suppose that  $\lim_{n \rightarrow \infty} f_n(x) = f(x)$  a.e.  $\nu$  and  $f(x)$  is a p.d.f. w.r.t.  $\nu$ .  
Then  $\lim_{n \rightarrow \infty} \int |f_n(x) - f(x)| d\nu = 0$ .

## Proof

Let  $g_n(x) = [f(x) - f_n(x)]I_{\{f \geq f_n\}}(x)$ ,  $n = 1, 2, \dots$

Then

$$\int |f_n(x) - f(x)| d\nu = 2 \int g_n(x) d\nu.$$

Since  $0 \leq g_n(x) \leq f(x)$  for all  $x$  and  $g_n \rightarrow 0$  a.e.  $\nu$ , the result follows from the dominated convergence theorem.

As an example, consider the Lebesgue p.d.f.  $f_n$  of the t-distribution  $t_n$  (Table 1.2),  $n = 1, 2, \dots$

One can show (exercise) that  $f_n \rightarrow f$ , where  $f$  is the p.d.f. of  $N(0, 1)$ .  
This is an important result in statistics.