Weak convergence of distribution functions

Central limit theorems

Let $\xi_1, \xi_2, \ldots, \xi_n$ be independent and identically distributed random variables with $\mathbb{E}[\xi_1] = 0$ and $\mathbb{E}[\xi_1^2] = 1$. Let

$$W_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i,$$

and let F_n be its distribution function. Then, by the central limit theorem,

$$F_n(x) \longrightarrow \Phi(x) \quad \text{for every } x \in \mathbb{R}.$$

We want to give a sufficient and necessary condition for weak convergence.

Definition of weak convergence

Definition 1 (Weak convergence)

Let $\{F_n\}$ be a sequence of distribution functions, and we say $\{F_n\}$ is weakly convergent if there exists a non-decreasing function F satisfying that

$$\lim_{n\to\infty} F_n(x) = F(x) \quad \text{for every continuity point } x \in \mathbb{R},$$

denoted by

$$F_n \xrightarrow{w} F.$$

Definition of weak convergence

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Remark

If $F_n \xrightarrow{w} F$, and let $X_n \sim F_n$ and $X \sim F$, then

$$X_n \xrightarrow{d} X.$$

- X_n and X may not be defined on the same probability space. [Example here.]
- However, we can construct X_n's and X on the same probability space, such that $X_n \sim F_n$, $X \sim F$, and

$$X_n \xrightarrow{a.s.} X.$$

Examples

Example 2

Let $X_1, X_2, \ldots$ be a sequence of random variables whose distribution functions are

$$F_n(x) = \begin{cases} 1 - (1 - \frac{1}{n})^{nx} & x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Show that $X_n \xrightarrow{d} \text{Exponential}(1)$.

Theorem 3

If $F_n \xrightarrow{w} F$, then there exist random variables $\{X_n\}$ and X defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that $X_n \sim F_n$ for every $n \geqslant 1$ and $X \sim F$ satisfying that

$$X_n \xrightarrow{a.s.} X.$$

Proof.

Recall the inverse functions of F_n and F are defined as

$$F_n^{-1}(t) = \inf\{x \in \mathbb{R} : F_n(x) \geqslant t\}, \quad F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geqslant t\}.$$

If $t \in (0,1)$ is a continuity point of F^{-1}, then

$$\lim_{n\to\infty} F_n^{-1}(t) = F^{-1}(t). \quad \text{(to be proved in the next page).}$$

Based on this result, consider the probability space $((0,1), \mathscr{B}(0,1), \mathbb{P})$, where $\mathbb{P}$ is the Lebesgue measure. For any $t \in (0,1)$, define

$$X_n(t) = F_n^{-1}(t), \quad X(t) = F^{-1}(t).$$

Then, (why?)

$$\mathbb{P}(\{t \in (0,1) : \lim_n X(t) \neq X(t)\}) = \mathbb{P}\{t \text{ is a dis-continuity point of } F^{-1}\} = 0.$$

Therefore, $X_n \xrightarrow{a.s.} X$. ∎

Proof.

Let $t \in (0,1)$ be a continuity point of F^{-1}. Since F has at most countably many discontinuities, we can find a continuity point x of F such that for any $\varepsilon > 0$,

$$F^{-1}(t) - \varepsilon < x < F^{-1}(t).$$

By the definition of F^{-1},

$$x < F^{-1}(t) \implies F(x) < t,$$

and by the weak convergence of F_n,

$$\lim_{n\to\infty} F_n(x) = F(x),$$

we have there exists $N \geqslant 1$ such that as long as $n \geqslant N$,

$$|F_n(x) - F(x)| < \frac{1}{2}(t - F(x)), \quad \implies \quad F_n(x) < t.$$

Therefore,

$$x \leqslant F_n^{-1}(t) \implies F^{-1}(t) - \varepsilon < x \leqslant F_n^{-1}(t).$$

By the definition of $\liminf$, we have shown that for any $\varepsilon > 0$, there exists $N \geqslant 1$ such that for any $n \geqslant N$,

$$F^{-1}(t) < F_n^{-1}(t) + \varepsilon,$$

which means that

$$F^{-1}(t) \leqslant \liminf_{n \to \infty} F_n^{-1}(t).$$

Similarly,

$$\limsup_{n \to \infty} F_n^{-1}(t) \leqslant F^{-1}(t).$$

■

Theorem 4

If $X_n \xrightarrow{a.s.} X$, and let g be a continuous function, then

$$g(X_n) \xrightarrow{a.s.} g(X).$$

Proof.

By definition, let

$$D = \{\omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\},$$

then $\mathbb{P}(D) = 0$. Because g is continuous, then

$$\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\} \subset \{\omega : \lim_{n\to\infty} g(X_n(\omega)) = g(X(\omega))\},$$

and thus

$$\mathbb{P}\{\omega : g(X_n(\omega)) \neq g(X_n(\omega))\} \leqslant \mathbb{P}(D) = 0.$$

∎

Part 3: bounded and continuous function $g(X_n)$

Theorem 5

If $X_n \xrightarrow{a.s.} X$, then for any bounded and continuous function g, then

$$\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)].$$

Proof

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Proof.

We assume that $|g(x)| \leqslant M$ for some $M > 0$. As $g(X_n) \xrightarrow{a.s.} g(X)$, then for any $\varepsilon > 0$,

$$\mathbb{P}(|g(X_n) - g(X)| > \varepsilon) \to 0 \quad \text{as } n \to \infty.$$

Therefore,

$$\begin{aligned} |\mathbb{E}[g(X_n)] - \mathbb{E}[g(X)]| &\leqslant \mathbb{E}|g(X_n) - g(X)| \\ &= \mathbb{E}|g(X_n) - g(X)|\mathbf{1}_{|g(X_n)-g(X)|>\varepsilon} + \mathbb{E}|g(X_n) - g(X)|\mathbf{1}_{|g(X_n)-g(X)|\leqslant\varepsilon} \\ &\leqslant 2M\,\mathbb{P}(|g(X_n) - g(X)| > \varepsilon) + \varepsilon. \end{aligned}$$

Taking $n \to \infty$, and $\varepsilon \to 0$, we have the right hand side goes to 0. ∎

An important property of weak convergence

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Theorem 6

Let $\{F_n\}$ be a sequence of distribution functions and let F be a distribution function. Let $\{X_n\}$ be a sequence of random variables satisfying that $X_n \sim F_n$ and let X be such that $X \sim F$. The following conditions are equivalent:

(i) $F_n \xrightarrow{w} F$.

(ii) $X_n \xrightarrow{d} X$.

(iii) For any bounded and continuous function g, then

$$\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)].$$

Proof.

Conditions (i) and (ii) are equivalent by definition. We only prove that (i) and (iii) are equivalent.

(i) $\Longrightarrow$ (iii). As $F_n \xrightarrow{w} F$, by the Skorokhod's representation theorem (see Theorem 3), then there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ on which we can define $\{\widetilde{X}_n\}$ and $\widetilde{X}$ such that

$$\widetilde{X}_n \stackrel{d}{=} X_n, \quad \widetilde{X} \stackrel{d}{=} X$$

and

$$\widetilde{X}_n \xrightarrow{a.s.} \widetilde{X}.$$

Therefore,

$$\mathbb{E}[g(X_n)] = \mathbb{E}[g(\widetilde{X}_n)] \to \mathbb{E}[g(\widetilde{X})] = \mathbb{E}[g(X)], \quad \text{as } n \to \infty.$$

■

Proof.

Proof of $(iii) \implies (i)$. For any continuity point x of F, let $g_{x,\varepsilon}$ be defined as

$$g_{x,\varepsilon}(w) = \begin{cases} 1 & w \leqslant x \\ 0 & w > x + \varepsilon, \\ \text{linear} & x < w \leqslant x + \varepsilon. \end{cases}$$

We can see that $g_{x,\varepsilon}$ is continuous and bounded. If (iii) is true, then

$$\limsup_{n\to\infty} \mathbb{P}(X_n \leqslant x) \leqslant \limsup_{n\to\infty} \mathbb{E}[g_{x,\varepsilon}(X_n)] = \mathbb{E}[g_{x,\varepsilon}(X)] \leqslant \mathbb{P}(X \leqslant x + \varepsilon).$$

Letting $\varepsilon \to 0$ gives

$$\limsup_{n\to\infty} \mathbb{P}(X_n \leqslant x) \leqslant \mathbb{P}(X \leqslant x).$$

Proof

Observe that

$$\liminf_{n\to\infty} \mathbb{P}(X_n \leqslant x) \geqslant \liminf_{n\to\infty} \mathbb{E}[g_{x-\varepsilon,\varepsilon}(X_n)] = \mathbb{E}[g_{x-\varepsilon,\varepsilon}(X)] \geqslant \mathbb{P}(X \leqslant x-\varepsilon).$$

Letting $\varepsilon \longrightarrow 0$ gives

$$\liminf_{n\to\infty} \mathbb{P}(X_n \leqslant x) \geqslant \mathbb{P}(X < x).$$

If x is a continuity point of F, then $\mathbb{P}(X < x) = \mathbb{P}(X \leqslant x)$, which implies that

$$\mathbb{P}(X \leqslant x) \leqslant \liminf_{n\to\infty} \mathbb{P}(X_n \leqslant x) \leqslant \limsup_{n\to\infty} \mathbb{P}(X_n \leqslant x) \leqslant \mathbb{P}(X \leqslant x),$$

and this proves the result. ■

An equivalent definition of weak convergence

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 7

We say F_n convergence weakly to F if

$$\int_{-\infty}^{\infty} g(x)dF_n(x) \to \int_{-\infty}^{\infty} g(x)dF(x) \quad \text{as } n \to \infty$$

for all bounded and continuous function g. Or, let $X_n \sim F_n$ and $X \sim F$, F_n converges weakly to F if

$$\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)] \quad \text{as } n \to \infty.$$

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Theorem 8

Consider the sequence $X_1, X_2, \ldots$ and the random variable X. Assume that X_n's and X are supported on $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, and their pmf are p_n's and p, respectively. Then, $X_n \xrightarrow{d} X$ if and only if

$$\lim_{n\to\infty} p_n(k) = p(k) \quad \text{for all } k \in \mathbb{N}_0.$$

Proof.

Since X is integer-valued, its CDF, $F_X(x)$, is continuous at all $x \in \mathbb{R} - \{0, 1, 2, \ldots\}$. If $X_n \xrightarrow{d} X$, then

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x), \qquad \text{for all } x \in \mathbb{R} - \{0, 1, 2, \ldots\}.$$

Thus, for $k = 0, 1, 2, \cdots$, we have

$$\begin{aligned}
\lim_{n\to\infty} p_n(k) &= \lim_{n\to\infty} \left[F_{X_n}\left(k + \frac{1}{2}\right) - F_{X_n}\left(k - \frac{1}{2}\right) \right] \qquad (X_n \text{ 's are integer-valued}) \\
&= \lim_{n\to\infty} F_{X_n}\left(k + \frac{1}{2}\right) - \lim_{n\to\infty} F_{X_n}\left(k - \frac{1}{2}\right) \\
&= F_X\left(k + \frac{1}{2}\right) - F_X\left(k - \frac{1}{2}\right) \qquad \left(\text{since } X_n \xrightarrow{d} X\right) \\
&= p(k) \qquad (\text{since } X \text{ is integer-valued}).
\end{aligned}$$

To prove the converse, assume that we know

$$\lim_{n\to\infty} p_n(k) = p(k), \qquad \text{for } k = 0, 1, 2, \cdots .$$

Then, for all $x \in \mathbb{R}$, we have

$$\begin{aligned}\lim_{n\to\infty} F_{X_n}(x) &= \lim_{n\to\infty} P(X_n \leqslant x) \\ &= \lim_{n\to\infty} \sum_{k=0}^{\lfloor x \rfloor} p_n(k),\end{aligned}$$

where $\lfloor x \rfloor$ shows the largest integer less than or equal to x.

Since for any fixed x, the set $\{0, 1, \cdots, \lfloor x \rfloor\}$ is a finite set, we can change the order of the limit and the sum, so we obtain

$$\begin{aligned}
\lim_{n\to\infty} F_{X_n}(x) &= \sum_{k=0}^{\lfloor x \rfloor} \lim_{n\to\infty} p_n(k) \\
&= \sum_{k=0}^{\lfloor x \rfloor} p(k) \quad \text{(by assumption)} \\
&= P(X \leqslant x) = F_X(x). \qquad \blacksquare
\end{aligned}$$

Example 9

Let $X_1, X_2, X_3, \cdots$ be a sequence of random variable such that

$$X_n \sim \text{Binomial}\left(n, \frac{\lambda}{n}\right), \quad \text{for } n \in \mathbb{N}, n > \lambda,$$

where $\lambda > 0$ is a constant. Show that X_n converges in distribution to $\mathrm{Poisson}(\lambda)$.

Solution.

By Theorem 8, it suffices to show that

$$\lim_{n\to\infty} p_n(k) = P_X(k), \qquad \text{for all } k = 0, 1, 2, \cdots .$$

We have

$$\begin{aligned}
\lim_{n\to\infty} p_n(k) &= \lim_{n\to\infty} \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \lambda^k \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{1}{n^k}\right)\left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \left(\left[\frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}\right] \left[\left(1 - \frac{\lambda}{n}\right)^n\right] \left[\left(1 - \frac{\lambda}{n}\right)^{-k}\right] \right).
\end{aligned}$$

Solution

Note that for a fixed k, we have

$$\lim_{n\to\infty} \frac{n(n-1)(n-2)\dots(n-k+1)}{n^k} = 1,$$

$$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{-k} = 1,$$

$$\lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{n} = e^{-\lambda}.$$

Thus, we conclude

$$\lim_{n\to\infty} p_n(k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

Continuous mapping theorem

Theorem 10

Let h be a continuous function. If $X_n \xrightarrow{d} X$, then $h(X_n) \xrightarrow{d} h(X)$.

Proof.

To show that $h(X_n) \xrightarrow{d} h(X)$, it suffices to show that for any bounded and continuous function g,

$$\mathbb{E}[g(h(X_n))] \to \mathbb{E}[g(h(X))] \quad \text{as } n \to \infty,$$

but this is true because $g \circ h$ is also a bounded and continuous function. ∎

A further condition to show weak convergence

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

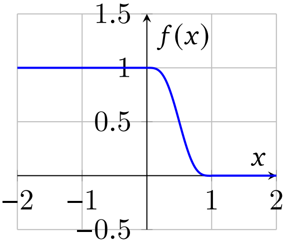
Theorem 11

Let g be a bounded function having three bounded and continuous derivatives, that is,

$$g, g', g'', g''' \quad \text{are bounded and continuous.}$$

If $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$, then $X_n \xrightarrow{d} X$.

A smooth function with bounded h'''



$$h(x) = \begin{cases} 1 & \text{if } x \leqslant 0, \\ 0 & \text{if } x > 1, \\ 1 - 140\left(\frac{1}{4}x^4 - \frac{3}{5}x^5 + \frac{1}{2}x^6 - \frac{1}{7}x^7\right) & \text{if } 0 < x \leqslant 1. \end{cases}$$

Theorem 12 (Weak convergence implies convergence of ch.f.)

If $\{F_n\}$ weakly converges to F, then

$$\varphi_n(t) \to \varphi(t) \quad \text{pointwise.}$$

Here,

$$\varphi_n(t) = \int_{-\infty}^{\infty} e^{itx} dF_n(x), \quad \varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x).$$

In other words, if $X_n \sim F_n$ and $X \sim F$, then

$$\varphi_n(t) = \mathbb{E}[e^{itX_n}], \quad \varphi(t) = \mathbb{E}[e^{itX}].$$

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY 南方科技大学

Proof.

Note that the function $g(x) = e^{itx} = \cos(tx) + i\sin(tx)$, and both $\cos(tx)$ and $\sin(tx)$ are bounded and continuous for all $x \in \mathbb{R}$ and for all $t \in \mathbb{R}$. Therefore, for any $t \in \mathbb{R}$, by Theorem 6, we have

$$\varphi_n(t) \longrightarrow \varphi(t).$$

∎

Convergence of ch.f. also implies weak convergence

Theorem 13

If a sequence of characteristic functions $\{\varphi_n(t)\}$ converges to $\varphi(t)$, and if $\varphi(t)$ is continuous at $t = 0$. Then, the corresponding distribution functions $\{F_n\}$ converges weakly to F, where $\varphi(t)$ is the characteristic function F.

Proof

Proof.

The proof is omitted. You can find a proof from Theorem 5.2.5 in Xianping Li's book.
At this level, you can just believe that this theorem is true. ∎

Theorem 14

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} c$, then

$$X_n + Y_n \xrightarrow{d} X + c, \quad X_n Y_n \xrightarrow{d} cX.$$

Remark

This theorem is not true for general $Y_n \xrightarrow{d} Y$.