Module 7

Limiting Distributions

Let $\underline{T} = (T_1, \dots, T_n)$ be a random vector having a probability density function/probability mass function (p.d.f./p.m.f.) $f_{\underline{T}}(\cdot)$ and let $h: \mathbb{R}^n \to \mathbb{R}$ be a Borel function. Suppose that the distribution of random variable $X_n = h(\underline{T})$ is desired. Very often it is not possible to derive the expression for distribution (i.e., p.d.f. or p.m.f.) of $X_n = h(\underline{T})$. To make this point clear let $T_1, \dots, T_n$ be a random sample from $\mathrm{Be}(a, b)$ distribution, where a and b are positive real constants, and suppose that the distribution (i.e., the distribution function or a p.d.f.) of the sample mean $\bar{T}_n = \frac{1}{n}\sum_{i=1}^{n} T_i$ is desired. The form of the p.d.f. (or distribution function) of $\bar{T}_n$ is so complicated (it involves multiple integrals which cannot be expressed in a closed form) that hardly anybody would be interested in using it. Therefore, it will be helpful if we can approximate the distribution of $\bar{T}_n$ by a distribution which is mathematically tractable. In this module we will develop a theory which will help us in approximating distributions of a sequence $\{X_n\}_{n\geq 1}$ of random variables for large values of n (say, as $n \to \infty$). Such approximations are quite useful in statistical inference problems.

1. Convergence in Distribution and Probability

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with corresponding sequence of distribution functions (d.f.s) as $\{F_n\}_{n\geq 1}$. Suppose that an approximation to the distribution of X_n (i.e., of F_n) is desired, for large values of n (say, as $n \to \infty$). It may be tempting to approximate $F_n(\cdot)$ by $F(x) = \lim_{n\to\infty} F_n(x), x \in \mathbb{R}$. However, as the following examples illustrate, $(x) = \lim_{n\to\infty} F_n(x), x \in \mathbb{R}$, may not be a d.f..

Examples 1.1

(i) Let $\{X_n\}_{n\geq 1}$ be sequence of random variables with $P(\{X_n = n\}) = 1,\ n = 1, 2, \dots$. Then the d.f. of X_n is given by

$$F_n(x) = \begin{cases} 0, & \text{if } x < n \\ 1, & \text{if } x \geq n \end{cases}, \qquad n = 1, 2, \dots.$$

We have $F(x) \stackrel{\text{def}}{=} \lim_{n\to\infty} F_n(x) = 0, \forall\ x \in \mathbb{R}$. Clearly F is not a d.f..

(ii) Let $X_n \sim U(-n, n), n = 1, 2, \dots$. Then the d.f. of X_n is

$$F_n(x) = \begin{cases} 0, & \text{if } x < -n \\ \dfrac{x+n}{2n}, & \text{if} -n \leq x < n, \\ 1, & \text{if } x \geq n \end{cases} \qquad n = 1, 2, \dots.$$

Clearly $F(x) \stackrel{\text{def}}{=} \lim_{n\to\infty} F_n(x) = \frac{1}{2}, \forall\, x \in \mathbb{R}$ and $F(\cdot)$ is not a d.f.. ■

The above examples illustrate that a sequence $\{F_n\}_{n\geq 1}$ of d.f.s on $\mathbb{R}$ may converge, at all points, but the limiting function $F(x) = \lim_{n\to\infty} F_n(x)\,, x \in \mathbb{R}$, may not be a d.f..

The following example illustrates that if a sequence $\{F_n\}_{n\geq 1}$ of d.f.s converges at every point then it may be too restrictive to require that $\{F_n\}_{n\geq 1}$ converges to a d.f. F at all points (i.e., to require that $\lim_{n\to\infty} F_n(x) = F(x), \forall x \in \mathbb{R}$, for some d.f. F).

Example 1.2

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with $P\left(\left\{X_n = \frac{1}{n}\right\}\right) = 1,\ n = 1, 2, \ldots$. Then the d.f. of X_n is

$$F_n(x) = \begin{cases} 0, & \text{if } x < \dfrac{1}{n} \\ 1, & \text{if } x \geq \dfrac{1}{n} \end{cases}, \qquad n = 1, 2, \ldots.$$

Clearly,

$$F(x) \stackrel{\text{def}}{=} \lim_{n\to\infty} F_n(x) = \begin{cases} 0, & \text{if } x \leq 0 \\ 1, & \text{if } x > 0 \end{cases}$$

is not a d.f. (it is not right continuous at $x = 0$). However, F can be converted into a distribution function

$$F^*(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases},$$

by changing its value at the point 0 (the point of discontinuity of F). Since $P\left(\left\{X_n = \frac{1}{n}\right\}\right) = 1, n = 1, 2, \ldots,$ and $\lim_{n\to\infty} \frac{1}{n} = 0$, a natural approximation of F_n seems to be the distribution function of a random variable X that is degenerate at 0 $(\text{i.e.}, P(\{X = 0\}) = 1)$. Note that F^* is the d.f. of random variables X that is degenerate at 0. The above discussion suggests that it is too restrictive to require

$$\lim_{n\to\infty} F_n\,(x) = F^*(x), \forall\, x \in \mathbb{R},$$

and that exceptions should be permitted at the points of discontinuities of F^*. ■

Definition 1.1

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables and let F_n be the d.f. of $X_n, n = 1, 2, \ldots$.

(i) Let X be a random variables with d.f. F. The sequence $\{X_n\}_{n\geq 1}$ is said to *converge in distribution* to X, as $n \to \infty$ (written as $X_n \xrightarrow{d} X, \text{as } n \to \infty$) if $\lim_{n\to\infty} F_n(x) = F(x), \forall\, x \in C_F$, where C_F is the set of continuity points of F. The d.f. F (or the corresponding p.d.f/p.m.f.) is called the *limiting distribution* of X_n , as $n \to \infty$.

(ii) Let $c \in \mathbb{R}$. The sequence $\{X_n\}_{n\geq 1}$ is said to converge in probability to c, as $n \to \infty$ (written as $X_n \xrightarrow{p} c$, as $n \to \infty$) if $X_n \xrightarrow{d} X$, as $n \to \infty$, where X is a random variable that is degenerate at c. ■

Remark 1.1

(i) Suppose that $X_n \xrightarrow{d} X, \text{as } n \to \infty$. Since the set $D_F = C_F^c = \mathbb{R} - C_F$ of discontinuity points of limiting d.f. F is at most countable we have $\lim_{n\to\infty} F_n(x) = F(x)$ everywhere except, possibly, at a countable number of points.

(ii) Note that the distribution function of a random variable degenerate at point $c \in \mathbb{R}$ is given by

$$F(x) = \begin{cases} 0, & \text{if } x < c \\ 1, & \text{if } x \geq c \end{cases}.$$

Thus we have

$$X_n \xrightarrow{p} c, \text{as } n \to \infty \Leftrightarrow \lim_{n\to\infty} F_n\,(x) = \begin{cases} 0, & \text{if } x < c \\ 1, & \text{if } x \geq c \end{cases}.$$

(iii) Suppose that $X_n \xrightarrow{d} X, \text{as } n \to \infty$. If the random variable X is of continuous type $(\text{i.e.}, \ C_F = \mathbb{R})$ then $\lim_{n\to\infty} F_n\,(x) = F(x), \forall\, x \in \mathbb{R}$.

(iv) Note that, for a real constant c, $X_n \xrightarrow{p} c$ if, and only if, $X_n - c \xrightarrow{p} 0$, as $n \to \infty$. ■

Example 1.3

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables such that $P(\{X_n = 0\}) = \frac{1}{n} = 1 - P\left(\left\{X_n = \frac{1}{n}\right\}\right)$, $n = 1, 2, \ldots$. Show that $X_n \xrightarrow{p} 0, \text{as } n \to \infty$.

Solution. Let F be the d.f. of a random variable degenerate at 0, i.e.,

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases}.$$

Since F is continuous everywhere except at point 0 (i.e., $C_F = \mathbb{R} - \{0\}$), we need to show that $\lim_{n\to\infty} F_n(x) = F(x), \forall\, x \in \mathbb{R} - \{0\}$, where $F_n(\cdot)$ is the d.f. of $X_n, n = 1, 2, \ldots$.

We have

$$F_n(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{n}, & \text{if } 0 \le x < \frac{1}{n}, \\ 1, & \text{if } x \ge \frac{1}{n} \end{cases} \qquad n = 1, 2, \ldots$$

$$\xrightarrow{n\to\infty} \begin{cases} 0, & \text{if } x \le 0 \\ 1, & \text{if } x > 0 \end{cases}.$$

Clearly $\lim_{n\to\infty} F_n(x) = F(x), \forall\, x \in \mathbb{R} - \{0\}$. ■

Example 1.4

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) $U(0, \theta)$ random variables, where $\theta > 0$. Let $X_{n:n} = \max\{X_1, \ldots, X_n\}$ and let $Y_n = n(\theta - X_{n:n}), n = 1, 2, \ldots$.

(i) Show that $X_{n:n} \xrightarrow{p} \theta$, as $n \to \infty$;

(ii) Find the limiting distribution of $\{Y_n\}_{n\ge1}$.

Solution.

(i) Let H_n be the d.f. of $X_{n:n}$, $n = 1, 2, \ldots$, and let

$$H(x) = \begin{cases} 0, & \text{if } x < \theta \\ 1, & \text{if } x \ge \theta \end{cases}$$

be the d.f. of random variable degenerate at θ. We need to show that $\lim_{n\to\infty} H_n(x) = H(x)$, $\forall\, x \in \mathbb{R} - \{\theta\}$.

We have, for $x \in \mathbb{R}$,

$$\begin{aligned} H_n(x) &= P(\{X_{n:n} \le x\}) \\ &= P(\{\max\{X_1, \ldots, X_n\} \le x\}) \\ &= P(\{X_i \le x,\ i = 1, \ldots, n\}) \\ &= \prod_{i=1}^{n} P(\{X_i \le x\}) \qquad \text{(since } X_i\text{s are independent)} \\ &= [F(x)]^n, \qquad n = 1,2,\ldots, \quad \text{(since } X_i\text{s are identically distributed),} \end{aligned}$$

where

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{x}{\theta}, & \text{if } 0 \le x < \theta \\ 1, & \text{if } x \ge \theta \end{cases}$$

is the common distribution function of $X_1, X_2, \ldots$.

Thus

$$H_n(x) = \begin{cases} 0, & \text{if } x < 0 \\ \left(\frac{x}{\theta}\right)^n, & \text{if } 0 \le x < \theta \\ 1, & \text{if } x \ge \theta \end{cases}$$

$$\xrightarrow{n\to\infty} \begin{cases} 0, & \text{if } x < \theta \\ 1, & \text{if } x \ge \theta \end{cases}$$

$$= H(x), \forall\, x \in \mathbb{R}.$$

(ii) For $y \in \mathbb{R}$, we have

$$\begin{aligned} F_{Y_n}(y) &= P(\{Y_n \le y\}) \\ &= P\left(\left\{X_{n:n} \ge \theta - \frac{y}{n}\right\}\right) \\ &= 1 - H_n\left(\left(\theta - \frac{y}{n}\right) -\right) \\ &= 1 - H_n\left(\theta - \frac{y}{n}\right) \qquad \text{(since } H_n \text{ is continuous)} \\ &= \begin{cases} 0, & \text{if } y \le 0 \\ 1 - \left(1 - \frac{y}{n\theta}\right)^n, & \text{if } 0 < y \le n\theta, \quad n = 1, 2, \ldots. \\ 1, & \text{if } y > n\theta \end{cases} \end{aligned}$$

$$\xrightarrow{n\to\infty} \begin{cases} 0, & \text{if } y \le 0 \\ 1 - e^{-\frac{y}{\theta}}, & \text{if } y > 0 \end{cases}$$

$$= G(y), \text{ say.}$$

Note that $G(\cdot)$ is the d.f. of $\text{Exp}(\theta)$ random variable. Thus $Y_n \xrightarrow{d} Y \sim \text{Exp}(\theta)$, as $n \to \infty$. ■

In the above example we saw that $X_{n:n} \xrightarrow{p} \theta$, as $n \to \infty$, and $n(\theta - X_{n:n}) \xrightarrow{d} Y \sim \text{Exp}(\theta)$, as $n \to \infty$, i.e., the limiting distribution of $X_{n:n}$ is degenerate (at θ) and, to get a non-degenerate limiting distribution, we needed normalized version $Y_n = n(\theta - X_{n:n})$ of $X_{n:n}, n = 1, 2, \ldots$. This phenomenon is observed quite commonly. Generally, we will have a sequence $\{X_n\}_{n\geq 1}$ of random variables, such that $X_n \overset{p}{\to} c, \text{as } n \to \infty$ for some real constant c (i.e., the limiting distribution of X_n is degenerate at c). In order to get a non-degenerate limiting distribution a normalized version $Z_n = n^r(X_n - c)\ \left(\text{or } Z_n = n^r(c - X_n)\right),\ r > 0, \text{of } X_n,\ n = 1, 2, \ldots$ is considered. Typically there is a choice of $r > 0$ such that the limiting distribution of Z_n is non-degenerate.

Theorem 1.1

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables such that $X_n \overset{d}{\to} X, \text{as } n \to \infty$, for some random variable X. Let F_n and F denote the d.f.s of X_n $(n = 1, 2, \ldots)$ and X, respectively. Then

$$\lim_{n\to\infty} F_n\,(x\,-) = F(x\,-) = F(x) = \lim_{n\to\infty} F_n\,(x), \forall\, x \in C_F,$$

where C_F is the set of continuity points of F.

Proof. We are given that

$$\lim_{n\to\infty} F_n\,(x) = F(x), \forall\, x \in C_F \qquad \left(\text{since } X_n \overset{d}{\to} X, \text{as } n \to \infty\right).$$

Moreover $F(x\,-) = F(x), \forall\, x \in C_F$. Thus it suffices to show that $\lim_{n\to\infty} F_n\,(x\,-) = F(x\,-), \forall\, x \in C_F$. Let $d \in C_F$ so that $F(d\,-) = F(d)$. Fix $m \in \mathbb{N} = \{1, 2, \ldots\}$. Since the set $C_F^c = \mathbb{R} - C_F$ of discontinuity points of F is countable and the interval $\left(d - \frac{1}{m}, d\right)$ is uncountable there exists a $d_m \in \left(d - \frac{1}{m}, d\right) \cap C_F$. Then we have $\lim_{n\to\infty} F_n\,(d_m) = F(d_m)$ and $\lim_{n\to\infty} F_n\,(d) = F(d)$. Moreover

$$F_n(d_m) \leq F_n(d\,-) \leq F_n(d), n = 1, 2, \ldots$$

$$\Rightarrow \lim_{n\to\infty} F_n\,(d_m) \leq \lim_{n\to\infty} F_n\,(d\,-) \leq \lim_{n\to\infty} F_n\,(d)$$

$$\Rightarrow F(d_m) \leq \lim_{n\to\infty} F_n\,(d\,-) \leq F(d) = F(d\,-). \tag{1.1}$$

Since $d_m \in \left(d - \frac{1}{m}, d\right)$, we have

$$\lim_{m\to\infty} F(d_m) = F(d\,-) = F(d). \tag{1.2}$$

On taking $m \to \infty$ in (1.1) we get

$$\lim_{m\to\infty} F(d_m) \leq \lim_{m\to\infty} F_n(d\,-) \leq F(d\,-)$$

$$\Rightarrow F(d-) \leq \lim_{n\to\infty} F_n(d-) \leq F(d-) \qquad \text{(using (1.2))}$$

$$\Rightarrow \lim_{n\to\infty} F_n(d-) = F(d-) \cdot \blacksquare$$

Corollary 1.1

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with corresponding sequence of d.f.s as $\{F_n\}_{n\geq 1}$. Further let X be another random variable having the d.f. F.

(i) If $X_n \xrightarrow{d} X$, as $n \to \infty$, and X is of continuous type then $\lim_{n\to\infty} F_n(x) = F(x), \forall\, x \in \mathbb{R}$ and $\lim_{n\to\infty} F_n(x-) = F(x-), \forall\, x \in \mathbb{R}$.

(ii) Suppose that $P(\{X_n \in \{0,1,2,\ldots\}\}) = P(\{X \in \{0,1,2,\ldots\}\}) = 1$ and $X_n \xrightarrow{d} X$, as $n \to \infty$. Then $\lim_{n\to\infty} F_n(x) = F(x),\ \forall\, x \in \mathbb{R}$ and $\lim_{n\to\infty} F_n(x-) = F(x-), \forall\, x \in \mathbb{R}$.

(iii) Under the assumptions of (ii), let of f and f_n be the p.m.f.s of X and X_n, respectively, $n = 1,2,\ldots$. Then

$$X_n \xrightarrow{d} X, \text{as } n \to \infty \Leftrightarrow \lim_{n\to\infty} f_n(x) = f(x), \forall\, x \in \{0,1,2,\ldots\}.$$

Proof.

(i) Since X is of continuous type we have $C_F = \mathbb{R}$, where C_F is the set of continuity points of F. The assertion now follows from Theorem 1.1

(ii) Fix $x \in \mathbb{R}$. If $P(\{X = x\}) = 0$ then $x \in C_F$ and, therefore, by Theorem 1.1

$$\lim_{n\to\infty} F_n(x) = F(x), \qquad \text{and} \qquad \lim_{n\to\infty} F_n(x-) = F(x-).$$

Now suppose that $P(\{X = x\}) > 0$. Then $x \in \{0,1,2,\ldots\}$ and $P(\{X = x + 0.5\}) = P(\{X = x - 0.5\}) = 0$. Consequently $x \pm 0.5 \in C_F$,

$$F_n(x) = F_n(x + 0.5) \quad \text{and} \quad F_n(x-) = F_n(x - 0.5),\ n = 1,2,\ldots$$

$$\Rightarrow \lim_{n\to\infty} F_n(x) = F(x+0.5) = F(x) \quad \text{and} \quad \lim_{n\to\infty} F_n(x-) = F(x-0.5) = F(x-).$$

It follows that

$$\lim_{n\to\infty} F_n(x) = F(x) \text{ and } \lim_{n\to\infty} F_n(x-) = F(x-), \forall x \in \mathbb{R}.$$

(iii) First suppose that $X_n \xrightarrow{d} X$, as $n \to \infty$. Then, for $x \in \{0,1,2,\ldots\}$,

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} P(\{X_n = x\})$$

$$= \lim_{n\to\infty}[F_n(x) - F_n(x-)]$$

$$= F(x) - F(x-) \qquad \text{(using (ii))}$$

$$= P(\{X = x\})$$

$$= f(x).$$

Conversely suppose that $\lim_{n\to\infty} f_n(x) = f(x), \forall\, x \in \{0, 1, 2, \dots\}$. Then, for $x \in \mathbb{R}$,

$$F_n(x) = P(\{X_n \le x\})$$

$$= \sum_{k=0}^{[x]} P(\{X_n = k\})$$

$$= \sum_{k=0}^{[x]} f_n(k)$$

$$\xrightarrow{n\to\infty} \sum_{k=0}^{[x]} f(k)$$

$$= F(x),$$

where $[x]$ denotes the largest integer not exceeding x. It follows that $X_n \xrightarrow{d} X$, as $n \to \infty$. ■

For the random variables of absolutely continuous type we state the following theorem without providing its proof.

Theorem 1.2

Let $\{X_n\}_{n\ge1}$ be a sequence of random variables of absolutely continuous type with corresponding sequence of p.d.f.s as $\{f_n\}_{n\ge1}$. Further let X be another random variable of absolutely continuous type with p.d.f. f. Suppose that $\lim_{n\to\infty} f_n(x) = f(x), \forall\, x \in \mathbb{R}$. Then $X_n \xrightarrow{d} X$, as $n \to \infty$. ■

The following example demonstrates that if $X_n \xrightarrow{d} X$, as $n \to \infty$, then $\lim_{n\to\infty} F_n(x-) = F(x-)$ may not hold; here F_n and F are d.f.s of X_n $(n = 1, 2, \dots)$ and X, respectively.

Example 1.5

Let $X_n \sim N\left(0, \frac{1}{n}\right), n = 1,2,\ldots,$ and let X be a random variable degenerate at 0 (i.e., $P(\{X = 0\}) = 1$). Then, for $x \in \mathbb{R}$,

$$F(x) = P(\{X \le x\}) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \ge 0 \end{cases}$$

$$\begin{aligned} F_n(x) &= P(\{X_n \le x\}) \\ &= \Phi(\sqrt{n}x) \\ &\xrightarrow{n\to\infty} \begin{cases} 0, \text{if } x < 0 \\ \frac{1}{2}, \text{if } x = 0 \\ 1, \text{if } x > 0 \end{cases} \end{aligned}$$

Clearly $\lim_{n\to\infty} F_n(x) = F(x), \forall\, x \in C_F = \mathbb{R} - \{0\}$ and, therefore, $X_n \overset{d}{\to} X$ $\left(\text{equivalenty } X_n \overset{p}{\to} 0\right)$, as $n \to \infty$. However

$$\lim_{n\to\infty} F_n(0-) = \lim_{n\to\infty} F_n(0) = \frac{1}{2} \neq F(0-) = 0. \blacksquare$$

The following example illustrates that, in general, the limiting distribution cannot be obtained by taking the limit of p.m.f.s/p.d.f.s.

Example 1.6

Let $\{X_n\}_{n\ge1}$ be a sequence of random variables such that

$$P\left(\left\{X_n = \frac{1}{2n}\right\}\right) = P\left(\left\{X_n = \frac{1}{n}\right\}\right) = \frac{1}{2}, n = 1,2,\ldots,$$

and let X be another random variable with $P(\{X = 0\}) = 1$. Then it is easy to verify that $X_n \overset{d}{\to} X$, as $n \to \infty$. The p.m.f. of X_n is

$$f_n(x) = \begin{cases} \frac{1}{2}, & \text{if } x \in \left\{\frac{1}{2n}, \frac{1}{n}\right\}, \\ 0, & \text{otherwise} \end{cases}$$

and the p.m.f. of X is

$$f(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise} \end{cases}.$$

We have

$$\lim_{n\to\infty} f_n(x) = 0 \neq f(x), \forall\, x \in \mathbb{R}. \blacksquare$$

The following theorem provides a characterization of converge in probability.

Theorem 1.3

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables and let c be a real constant. Then

$$X_n \xrightarrow{p} c, \text{as } n \to \infty \iff \forall\, \varepsilon > 0, \lim_{n\to\infty} P(\{|X_n - c| \geq \varepsilon\}) = 0.$$

Proof. Let F_n denote the d.f. of X_n ($n = 1, 2, \ldots$) and let F denote the d.f. of random variable degenerate at c. First suppose that $X_n \xrightarrow{p} c, \text{as } n \to \infty$. Then, for $x \in \mathbb{R} - \{c\}$,

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} P(\{X_n \leq x\})$$

$$= \begin{cases} 0, & \text{if } x < c \\ 1, & \text{if } x > c \end{cases} = F(x).$$

Fix $\varepsilon > 0$. Then $c \pm \varepsilon \in C_F$ and therefore, using Theorem 1.1,

$$\begin{aligned}
\lim_{n\to\infty} P(\{|X_n - c| \geq \varepsilon\}) &= \lim_{n\to\infty} [P(\{X_n \leq c - \varepsilon\}) + P(\{X_n \geq c + \varepsilon\})] \\
&= \lim_{n\to\infty} \big[F_n(c-\varepsilon) + 1 - F_n\big((c+\varepsilon)-\big)\big] \qquad (1.3) \\
&= [F(c-\varepsilon) + 1 - F(c+\varepsilon)] \\
&= 0.
\end{aligned}$$

Conversely, suppose that

$$\lim_{n\to\infty} P(\{|X_n - c| \geq \varepsilon\}) = 0, \forall\, \varepsilon > 0.$$

Then, using (1.3),

$$\lim_{n\to\infty} \big[F_n(c-\varepsilon) + 1 - F_n\big((c+\varepsilon)-\big)\big] = 0, \forall\, \varepsilon > 0,$$

$$\Rightarrow \lim_{n\to\infty} F_n(c-\varepsilon) = \lim_{n\to\infty} \big[1 - F_n\big((c+\varepsilon)-\big)\big] = 0, \forall\, \varepsilon > 0$$

$$\big(\text{since } F_n(c-\varepsilon) \geq 0 \text{ and } 1 - F_n\big((c+\varepsilon)-\big) \geq 0, \forall\, n \geq 1\big)$$

$$\Rightarrow \lim_{n\to\infty} F_n(x) = 0, \forall\, x < c \text{ and } \lim_{n\to\infty} F_n(y-) = 1, \forall\, y > c$$

$$\Rightarrow \lim_{n\to\infty} F_n(x) = 0, \forall\, x < c \text{ and } \lim_{n\to\infty} F_n(y) = 1, \forall\, y > c$$

$$(\text{since } 1 \geq F_n(y) \geq F_n(y-), n = 1,2, \ldots).$$

Thus, for all $x \in \mathbb{R} - \{c\}$,

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0, & \text{if } x < c \\ 1, & \text{if } x > c \end{cases} = F(x)$$

$$\Rightarrow \quad X_n \xrightarrow{p} c, \text{as } n \to \infty. \blacksquare$$

In many situations the above theorem in conjunction with Markov's inequality (see Corollary 5.1, Module 3) turns out to be quite useful in proving convergence in probability.

Theorem 1.4

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with $E(X_n) = \mu_n \in (-\infty, \infty)$, and $\text{Var}(X_n) = \sigma_n^2 \in (0, \infty), n = 1, 2, \ldots$. Suppose that $\lim_{n\to\infty} \mu_n = \mu \in \mathbb{R}$ and $\lim_{n\to\infty} \sigma_n^2 = 0$. Then $X_n \xrightarrow{p} \mu$, as $n \to \infty$.

Proof. Fix $\varepsilon > 0$. Using the Markov inequality we have

$$0 \leq P(\{|X_n - \mu| \geq \varepsilon\}) \leq \frac{E(|X_n - \mu|^2)}{\varepsilon^2} = \frac{E((X_n - \mu)^2)}{\varepsilon^2}.$$

Also,

$$\begin{aligned} E((X_n - \mu)^2) &= E((X_n - \mu_n + \mu_n - \mu)^2) \\ &= E((X_n - \mu_n)^2) + (\mu_n - \mu)^2 \\ &= \sigma_n^2 + (\mu_n - \mu)^2 \cdot \end{aligned}$$

Therefore,

$$0 \leq P(\{|X_n - \mu| \geq \varepsilon\}) \leq \frac{\sigma_n^2 + (\mu_n - \mu)^2}{\varepsilon^2}$$

$$\xrightarrow{n\to\infty} 0.$$

$$\Rightarrow \lim_{n\to\infty} P(\{|X_n - \mu| \geq \varepsilon\}) = 0, \qquad \forall\, \varepsilon > 0.$$

$$\Rightarrow X_n \xrightarrow{p} \mu, \text{as } n \to \infty \qquad \text{(using Theorem 1.3).} \blacksquare$$

Example 1.7

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. $U(0, \theta)$ random variables, where $\theta > 0$. Let $X_{n:n} = \max\{X_1, X_2, \ldots, X_n\}, n = 1, 2, \ldots$. For any real constant s, show that $X_{n:n}^s \xrightarrow{p} \theta^s$, as $n \to \infty$.

Solution. It is easy to verify that a p.d.f. of $X_{n:n}$ is

$$f_n(x) = \begin{cases} \dfrac{nx^{n-1}}{\theta^n}, & \text{if } 0 < x < \theta \\ 0, & \text{otherwise} \end{cases}.$$

Then

$$E(X_{n:n}^s) = \frac{n}{n+s}\theta^s, \qquad n > -s$$

$$\to \theta^s, \text{as } n \to \infty \cdot$$

Also,

$$\begin{aligned} \mathrm{Var}(X_{n:n}^s) &= E(X_{n:n}^{2s}) - \left(E(X_{n:n}^s)\right)^2 \\ &= \frac{n}{n+2s}\,\theta^{2s} - \left(\frac{n}{n+s}\,\theta^s\right)^2, \qquad n > \max(-s, -2s) \\ &\to 0, \qquad \text{as } n \to \infty. \end{aligned}$$

Now, using Theorem 1.4, it follows that $X_{n:n}^s \xrightarrow{p} \theta^s$, as $n \to \infty$. ■

Example 1.8

Let $X_n \sim \mathrm{Bin}\,(n, \theta), n = 1, 2, \dots, \theta \in (0,1)$. If $Y_n = \frac{X_n}{n}, n = 1, 2, \dots$, show that $Y_n \xrightarrow{p} \theta$, as $n \to \infty$.

Solution. We have

$$E(Y_n) = E\left(\frac{X_n}{n}\right) = \theta, n = 1, 2, \dots,$$

and

$$\mathrm{Var}(Y_n) = \mathrm{Var}\left(\frac{X_n}{n}\right) = \frac{\mathrm{Var}(X_n)}{n^2} = \frac{\theta(1-\theta)}{n} \to 0, \text{as } n \to \infty \cdot$$

Using Theorem 1.4 it follows that $Y_n \xrightarrow{p} \theta$, as $n \to \infty$. ■

Remark 1.2

Theorem 1.3 provides an interpretation of the concept of convergence in probability. Theorem 1.3 suggests that if $X_n \xrightarrow{p} c$, as $n \to \infty$, then X_n is stochastically (in probability) very close to c for large values of n. Such an interpretation does not hold for the concept of convergence in distribution. Specifically, if $X_n \xrightarrow{d} X$, as $n \to \infty$, (where X is some non-degenerate random variable) then it cannot be inferred that X_n is getting close to X, for large values of n, in any sense. All we know in that case is that, for large values of n, the distribution of X_n is getting close to that of X. ■

The following example demonstrates that convergence in probability may not imply convergence of moments.

Example 1.9

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with

$$1 - P(\{X_n = 0\}) = P(\{X_n = n\}) = \frac{1}{n}, n = 1,2, \dots.$$

Then the d.f. of X_n is

$$F_n(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - \frac{1}{n}, & \text{if } 0 \leq x < n, n = 1, 2, \dots. \\ 1, & \text{if } x \geq n \end{cases}$$

$$\xrightarrow{n\to\infty} \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases}.$$

Thus $X_n \xrightarrow{p} 0$, as $n \to \infty$. However, for $r \in \{1, 2, \dots\}$

$$E(X_n^r) = E(|X_n|^r) = n^{r-1} \nrightarrow 0, \text{as } n \to \infty. ■$$

The following example illustrates that convergence in distribution to a non-degenerate random variable also does not imply convergence of moments.

Example 1.10

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables with p.m.f.s

$$f_n(x) = \begin{cases} \frac{1}{2} - \frac{1}{2n}, & \text{if } x \in \left\{0, \frac{1}{2}\right\} \\ \frac{1}{n}, & \text{if } x = n, \quad n = 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

and let X be a random variable with p.m.f.

$$f(x) = \begin{cases} \frac{1}{2}, & \text{if } x \in \left\{0, \frac{1}{2}\right\}. \\ 0, & \text{otherwise} \end{cases}$$

Then the distribution function of X is

$$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{2}, & \text{if } 0 \ \leq x < \frac{1}{2}, \\ 1, & \text{if } x \geq \frac{1}{2} \end{cases}$$

and the distribution function of X_n is

$$F_n(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{2} - \frac{1}{2n}, & \text{if } 0 \leq x < \frac{1}{2} \\ 1 - \frac{1}{n}, & \text{if } \frac{1}{2} \leq x < n \\ 1, & \text{if } x \geq n \end{cases}, n = 1, 2, \ldots$$

$$\xrightarrow{n\to\infty} \begin{cases} 0, & \text{if } x < 0 \\ \frac{1}{2}, & \text{if } 0 \leq x < \frac{1}{2}. \\ 1, & \text{if } x \geq \frac{1}{2} \end{cases}$$

It follows that $X_n \overset{d}{\to} X$, as $n \to \infty$. Moreover $E(X) = \frac{1}{4}$ and

$$E(X_n) = \frac{1}{2}\left[\frac{1}{2} - \frac{1}{2n}\right] + 1 \xrightarrow{n\to\infty} \frac{5}{4} \neq E(X). \blacksquare$$

We know that, for a real constant $c, X_n \overset{p}{\to} c$, as $n \to \infty \Leftrightarrow X_n - c \overset{p}{\to} 0$, as $n \to \infty$. The following example illustrates that $X_n \overset{d}{\to} X$, as $n \to \infty$ may not imply that $X_n - X \overset{p}{\to} 0$, as $n \to \infty$ or, equivalently, $X_n \overset{d}{\to} X$, as $n \to \infty$, does not imply that $X_n - X$ will converge in distribution to a random variable degenerate at 0 (also see Remark 1.2).

Example 1.11

Let $\{X_n\}_{n\geq 1}$ and X be as defined in Example 1.10. Further suppose that, for each $n \in \{1, 2, \ldots\}$, X_n and X are independent. Then $X_n \overset{d}{\to} X$, as $n \to \infty$. However, for $0 < \varepsilon < \frac{1}{2}$

$$P(\{|X_n - X| \geq \varepsilon\}) = \frac{1}{2}\left[P(\{|X_n| \geq \varepsilon\}) + P\left(\left\{\left|X_n - \frac{1}{2}\right| \geq \varepsilon\right\}\right)\right]$$

$$= \frac{1}{2}\left[\left(\frac{1}{2} - \frac{1}{2n}\right) + \frac{1}{n} + \left(\frac{1}{2} - \frac{1}{2n}\right) + \frac{1}{n}\right]$$

$$\xrightarrow{n\to\infty} \frac{1}{2},$$

implying that $X_n - X$ does not converge in distribution to a random variable degenerate at 0. ■

Definition 1.2

A sequence $\{X_n\}_{n\geq 1}$ of random variables is said to be *bounded in probability* if there exists a positive real constant M (not depending on n) such that

$$P\left(\bigcap_{n=1}^{\infty}\{|X_n| \leq M\}\right) = 1. ■$$

The following theorem relates convergence in distribution of a sequence $\{X_n\}_{n\geq 1}$ of random variables to the convergence of corresponding sequence of moment generating functions (m.g.f.s). We shall not provide the proof of the theorem as it is slightly involved.

Theorem 1.5

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables and let X be another random variable. Suppose that there exists an $h > 0$ such that the m.g.f.s $M(\cdot)$, $M_1(\cdot)$, $M_2(\cdot), \ldots$ of X, X_1, $X_2, \ldots$, respectively, are finite on $(-h, h)$.

(i) If $\lim_{n\to\infty} M_n(t) = M(t), \ \forall\, t \in (-h, h)$, then $X_n \xrightarrow{d} X$, as $n \to \infty$;

(ii) If $X_1, X_2, \ldots$ are bounded in probability and $X_n \xrightarrow{d} X$, as $n \to \infty$, then $\lim_{n\to\infty} M_n(t) = M(t), \ \forall\, t \in (-h, h)$. ■

The following example demonstrates that the conclusion of Theorem 1.5 (ii) may not hold if $X_1, X_2, \ldots$ are not bounded in probability.

Example 1.12

Let $\{X_n\}_{n\geq 1}$ and X be as defined in Example 1.10. Then the m.g.f. of X is

$$M(t) = \frac{1 + e^{\frac{t}{2}}}{2}, t \in \mathbb{R},$$

and the m.g.f. of X_n is

$$M_n(t) = \left(\frac{1}{2} - \frac{1}{2n}\right)\left(1 + e^{\frac{t}{2}}\right) + \frac{e^{nt}}{n}$$

$$\xrightarrow{n\to\infty} \begin{cases} \dfrac{1 + e^{\frac{t}{2}}}{2}, & \text{if } t \le 0 \\ \infty, & \text{if t} > 0 \end{cases}$$

$$\ne M(t), \qquad \forall\, t \in \mathbb{R}.$$

However, $X_n \xrightarrow{d} X$, as $n \to \infty$. ■

Proposition 1.1

Let $\{c_n\}_{n\ge1}$ be a sequence of positive real numbers such that $\lim_{n\to\infty} c_n = c \in \mathbb{R}$. Then

$$\lim_{n\to\infty}\left(1 + \frac{c_n}{n}\right)^n = e^c.$$

Proof. We know that

$$x - \frac{x^2}{2} \le \ln(1 + x) \le x, \forall\, x > 0$$

$$\Rightarrow c_n - \frac{{c_n}^2}{2n} \le n \ln\left(1 + \frac{c_n}{n}\right) \le c_n,\ \ n = 1, 2, \dots$$

$$\Rightarrow \lim_{n\to\infty}\left[n \ln\left(1 + \frac{c_n}{n}\right)\right] = c \text{ (on taking limits on both sides)}$$

$$\Rightarrow \lim_{n\to\infty}\left[\ln\left(1 + \frac{c_n}{n}\right)^n\right] = c$$

$$\Rightarrow \lim_{n\to\infty}\left(1 + \frac{c_n}{n}\right)^n = e^c. ■$$

Example 1.13 (Poisson Approximation to Binomial distribution)

Let $X_n \sim \text{Bin}\,(n, \theta_n)$, $n = 1, 2, \dots$, where $\theta_n \in (0, 1), n = 1, 2, \dots$, and $\lim_{n\to\infty}(n\theta_n) = \theta > 0$. Show that $X_n \xrightarrow{d} X$, as $n \to \infty$, where $X \sim \text{P}(\theta)$, the Poisson distribution with mean θ.

Solution. Note that the m.g.f. of X is

$$M(t) = e^{\theta(e^t - 1)}, t \ \in \mathbb{R},$$

and the m.g.f. of X_n is

$$M_n(t) = (1 - \theta_n + \theta_n e^t)^n$$
$$= \left(1 + \frac{c_n(t)}{n}\right)^n, \qquad t \in \mathbb{R},$$

where $c_n(t) = n\theta_n(e^t - 1), t \in \mathbb{R}, n = 1, 2, \ldots$. Clearly $\lim_{n\to\infty} c_n(t) = \theta(e^t - 1), \forall\, t \in \mathbb{R}$. Now using Proposition 1.1 we get

$$\lim_{n\to\infty} M_n(t) = e^{\theta(e^t - 1)} = M(t), \qquad \forall\, t \in \mathbb{R}.$$

Using Theorem 1.5 (i) we conclude that $X_n \xrightarrow{d} X \sim P(\theta)$, as $n \to \infty$. ■

2. The Weak Law of Large Numbers (WLLN) and the Central Limit Theorem (CLT)

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables and let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, n = 1,2, \ldots$, be the sequence of sample means. In this section we will study the convergence behavior of the sequence $\{\bar{X}_n\}_{n\geq 1}$ of sample means.

Theorem 2.1

Let $\{X_n\}_{n\geq 1}$ be a sequence of i.i.d. random variables and let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i, n = 1,2, \ldots$.

(i) (**WLLN**) Suppose that $E(X_1) = \mu$ is finite. Then $\bar{X}_n \xrightarrow{p} \mu$, as $n \to \infty$.

(ii) (**CLT**) suppose that $0 < \mathrm{Var}(X_1) = \sigma^2 < \infty$. Then

$$Z_n \stackrel{\text{def}}{=} \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1), \text{ as } n \to \infty.$$

Proof.

(i) As the proof for the case $\mathrm{Var}(X_1) = \infty$ is quite involved, we assume that $\mathrm{Var}(X_1) = \sigma^2 < \infty$. Then

$$E(\bar{X}_n) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = E(X_1) = \mu$$

and

$$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}\,(X_i) = \frac{\sigma^2}{n} \xrightarrow{n\to\infty} 0.$$

Using Theorem 1.4 it follows that $\bar{X}_n \xrightarrow{p} \mu$, as $n \to \infty$.

(ii) For simplicity we will assume that the common m.g.f. $M(\cdot)$ of $X_1, X_2, \ldots$ is finite in an interval $(-a, a)$ for some $a > 0$. Then, by Theorem 3.4, Module 3, $\mu_r' = E(X_1^r)$ is finite for each $r \in \{1, 2, \cdots\}$ and $\mu_r' = E(X_1^r) = M^{(r)}(0) = \left[\frac{d^r}{dt^r} M(t)\right]_{t=0}, r = 1, 2, \ldots$. Let $Y_i = \frac{X_i - \mu}{\sigma}, i = 1, \ldots, n$. Then $Y_1, Y_2, \ldots$ are i.i.d. random variables with mean 0 and variance 1. Let $M_Y(\cdot)$ denote the common m.g.f. of $Y_1, Y_2, \ldots$, so that

$$M_Y(t) = e^{-\frac{\mu t}{\sigma}} M\left(\frac{t}{\sigma}\right), \qquad -a\sigma < t < a\sigma,$$

$$M_Y^{(1)}(0) = -\frac{\mu}{\sigma} + \frac{M^{(1)}(0)}{\sigma} = 0 = E(Y_1)$$

$$\text{and} \quad M_Y^{(2)}(0) = \left(\frac{\mu}{\sigma}\right)^2 M(0) - \frac{2\mu}{\sigma^2} M^{(1)}(0) + \frac{1}{\sigma^2} M^{(2)}(0) = 1 = E(Y_1^2).$$

Let $\psi_2: (-a\sigma, a\sigma) \to \mathbb{R}$ be such that

$$M_Y(t) = M_Y(0) + t\, M_Y^{(1)}(0) + \frac{t^2}{2}\left(M_Y^{(2)}(0) + \psi_2(t)\right), \qquad t \in (-a\sigma, a\sigma) \qquad (2.1)$$

i.e., $$\psi_2(t) = \frac{M_Y(t) - M_Y(0) - t\, M_Y^{(1)}(0)}{t^2/2} - M_Y^{(2)}(0), \qquad t \in (-a\sigma, a\sigma), \;\; t \neq 0.$$

Using L' Hospital rule $(0/0$ form$)$ we get

$$\begin{aligned} \lim_{t \to 0} \psi_2(t) &= \lim_{t \to 0} \frac{M_Y^{(1)}(t) - M_Y^{(1)}(0)}{t} - M_Y^{(2)}(0) \\ &= M_Y^{(2)}(0) - M_Y^{(2)}(0) \\ &= 0. \end{aligned} \qquad (2.2)$$

The m.g.f. of $Z_n = \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} = \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$ is

$$\begin{aligned} M_n(t) &= E\left(e^{\frac{t}{\sqrt{n}} \sum_{i=1}^n Y_i}\right) \\ &= E\left(\prod_{i=1}^n e^{\frac{tY_i}{\sqrt{n}}}\right) \end{aligned}$$

$$= \prod_{i=1}^{n} E\left(e^{\frac{tY_i}{\sqrt{n}}}\right) \qquad (Y_i\text{s are independent})$$

$$= \left[M_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n \qquad (Y_i\text{s are i.i.d.})$$

$$= \left[M_Y(0) + \frac{t}{\sqrt{n}} M_Y^{(1)}(0) + \frac{t^2}{2n}\left(M_Y^{(2)}(0) + \psi_2\left(\frac{t}{\sqrt{n}}\right)\right)\right]^n \qquad (\text{using } (2.1))$$

$$= \left[1 + \frac{t^2}{2n}\left(1 + \psi_2\left(\frac{t}{\sqrt{n}}\right)\right)\right]^n, t \in (-\sqrt{n}a\sigma, \sqrt{n}a\sigma), n = 1, 2, \ldots.$$

Now using (2.2) and Proposition 1.1 we get

$$\lim_{n\to\infty} M_n(t) = e^{\frac{t^2}{2}} = K(t), \text{ say}, \qquad t \in \mathbb{R}.$$

Note that $K(t), t \in \mathbb{R}$, is the m.g.f. of $Z \sim N(0,1)$. Using Theorem 1.5 (i) we conclude that $Z_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$. ■

Remark 2.1

(i) The WLLN implies that the sample mean, based on a random sample from any parent distribution, can be made arbitrary close to population mean in probability by choosing sufficiently large sample size.

(ii) The CLT states that, irrespective of the nature of the parent distribution, the probability distribution of a normalized version of the sample mean, based on a random sample of large size, is approximately normal. For this reason the normal distribution is quite important in the field of Statistics. ■

Example 2.1 (Random Walk)

Consider a drunkard, who having missed his bus from the bus stand, starts walking towards his residence. Every second he either moves half a meter forward or half a meter backward from his current position, each with probability $1/2$. Assuming that steps are taken independently, find the (approximate) probability that after fifteen minutes the drunkard will be within 30 meters form the bus stand.

Solution. Note that in 15 minutes ($=$ 900 seconds) the drunkard will take 900 steps. Let Y_i be the size (in meters) of the i-th step, $i = 1, 2, \ldots, 900$. Then $Y_1, Y_2, \ldots$ are i.i.d. random variables with

$$P\left(\left\{Y_1 = -\frac{1}{2}\right\}\right) = P\left(\left\{Y_1 = \frac{1}{2}\right\}\right) = \frac{1}{2},$$

and $Y = \sum_{i=1}^{900} Y_i$ is the position of the drunkard after 15 minutes. The desired probability is

$$P(\{|Y| \leq 30\}) = P\left(\left\{-\frac{1}{30} \leq \bar{Y}_{900} \leq \frac{1}{30}\right\}\right),$$

where $\bar{Y}_{900} = \frac{1}{900}\sum_{i=1}^{900} Y_i = \frac{Y}{900}$. Note that $E(Y_1) = 0$ and $\mathrm{Var}(Y_1) = E(Y_1^2) = \frac{1}{4} = \sigma^2$, say. By the CLT

$$Z_{900} = \frac{\sqrt{900}\,(\bar{Y}_{900} - 0)}{1/2} \overset{\text{approx.}}{\sim} N(0,1),$$

$$\text{i.e.}, \; Z_{900} = 60\,\bar{Y}_{900} \overset{\text{approx.}}{\sim} N(0,1).$$

The desired probability is

$$\begin{aligned} P(\{|Y| \leq 30\}) &= P(\{-2 \leq Z_{900} \leq 2\}) \\ &\overset{\text{approx.}}{=} \Phi(2) - \Phi(-2) \\ &= 2\Phi(2) - 1 \\ &= 2 \times .9772 - 1 \\ &= .9544. \blacksquare \end{aligned}$$

Example 2.2 (Justification of Relative Frequency Method of Assigning Probabilities)

Suppose that we have independent repetitions of a random experiment under identical conditions. Further suppose that we are interested in assigning probability, say $P(E)$, to an event E. To do this we repeat the random experiment a large (say N) number of times. Define

$$Y_i = \begin{cases} 1, & \text{if i − th trial reuslts in occurrence of } E \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \dots, N.$$

Then $Y_1, Y_2, \dots$ are i.i.d. random variables with common mean $\mu = E(Y_1) = P(E)$. Also

$$f_N(E) = \text{number of times event } E \text{ occurs in first } N \text{ trials}$$

$$= \sum_{i=1}^{N} Y_i$$

and the relative frequency of event E in first N trials is

$$r_N(E) = \frac{f_N(E)}{N} = \frac{1}{N}\sum_{i=1}^{N} Y_i = \bar{Y}_N \text{, say} \cdot$$

The WLLN implies that

$$r_N(E) = \bar{Y}_N \xrightarrow{p} \mu = P(E), \text{ as } N \to \infty.$$

Thus the WLLN justifies the relative frequency approach to assign probabilities. ■

3. Some Preservation Results

In this section we will investigate that under what algebraic operations convergence in probability and/or convergence in distribution is preserved.

Theorem 3.1

Let $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ be sequences of random variables and let X be another random variable.

(i) Let $g: \mathbb{R} \to \mathbb{R}$ be continuous at $c \in \mathbb{R}$ and let $X_n \xrightarrow{p} c$, as $n \to \infty$. Then $g(X_n) \xrightarrow{p} g(c)$, as $n \to \infty$.

(ii) Let $h: \mathbb{R}^2 \to \mathbb{R}$ be continuous at $(c_1, c_2), \in \mathbb{R}^2$ and let $X_n \xrightarrow{p} c_1$, $Y_n \xrightarrow{p} c_2$, as $n \to \infty$. Then $h(X_n, Y_n) \xrightarrow{p} h(c_1, c_2)$, as $n \to \infty$.

(iii) Let $g: \mathbb{R} \to \mathbb{R}$ be continuous on a support S_X of X and let $X_n \xrightarrow{d} X$, as $n \to \infty$. Then $g(X_n) \xrightarrow{d} g(X)$, as $n \to \infty$.

(iv) Let $h: \mathbb{R}^2 \to \mathbb{R}$ be continuous at all points in $D = \{(x, b): x \in S_X\}$, where b is a fixed real constant and S_X is a support of X. if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{p} b$, as $n \to \infty$, then $h(X_n, Y_n) \xrightarrow{d} h(X, b)$, as $n \to \infty$.

Proof. We shall not attempt to prove assertions (iii) and (iv) here as their proofs are slightly involved.

(i) Fix $\varepsilon > 0$. Since $g: \mathbb{R} \to \mathbb{R}$ is continuous at $c \in \mathbb{R}$, there exists a $\delta \equiv \delta(\varepsilon, c)$ such that

$$|x-c|<\delta \Rightarrow |g(x)-g(c)|<\varepsilon \cdot$$

or equivalently $$|g(x)-g(c)| \geq \varepsilon \Rightarrow |x-c| \geq \delta.$$

Therefore,

$$0 \leq P(\{|g(X_n)-g(c)| \geq \varepsilon\}) \leq P(\{|X_n-c| \geq \delta\}) \xrightarrow{n\to\infty} 0 \qquad (\text{since } X_n \xrightarrow{p} c)$$

$$\Rightarrow \quad \lim_{n\to\infty} P(\{|g(X_n)-g(c)| \geq \varepsilon\}) = 0$$

$$\Rightarrow \quad g(X_n) \xrightarrow{p} g(c), \text{ as } n \to \infty.$$

(ii) Fix $\varepsilon > 0$. Since $h: \mathbb{R}^2 \to \mathbb{R}$ is continuous at $(c_1, c_2) \in \mathbb{R}^2$, there exists a $\delta = \delta(\varepsilon, c_1, c_2)$ such that

$$|x-c_1|<\delta \text{ and } |y-c_2|<\delta \;\Rightarrow\; |h(x,y)|-h(c_1,c_2)|<\varepsilon,$$

or equivalently

$$|h(x,y)-h(c_1,c_2)| \geq \varepsilon \Rightarrow |x-c_1| \geq \delta \text{ or } |y-c_2| \geq \delta.$$

Therefore,

$$P(\{|h(X_n,Y_n)-h(c_1,c_2)| \geq \varepsilon\}) \leq P(\{|X_n-c_1| \geq \delta\} \cup \{|Y_n-c_2| \geq \delta\})$$

$$\leq P(\{|X_n-c_1| \geq \delta\} + P\{|Y_n-c_2| \geq \delta\}) \text{ (using Boole's inequality)}$$

$$\xrightarrow{n\to\infty} 0+0=0 \left(\text{since } X_n \xrightarrow{p} c_1 \text{ and } Y_n \xrightarrow{p} c_2\right)$$

$$\Rightarrow \quad \lim_{n\to\infty} P\,(\{|h(X_n,Y_n)-h(c_1,c_2)| \geq \varepsilon\}) = 0$$

$$\Rightarrow \quad h(X_n,Y_n) \xrightarrow{p} h(c_1,c_2), \text{as } n \to \infty. \blacksquare$$

Throughout, we shall use the following convention. If, for a real constant c, we write $X_n \xrightarrow{d} c, \text{as } n \to \infty$, then it would mean that X_n converges in distribution, as $n \to \infty$, to a random variable degenerate at c (i.e., $X_n \xrightarrow{p} c, \text{as } n \to \infty$). Similarly, for a random variable $X, 0 \times X$ will be treated as a random variable degenerate at 0.

Now we provide the following useful lemma whose proof, being straight forward, is left as an exercise.

Lemma 3.1

(i) Let X and Y be random variables and let c be a real constant. If $P(\{Y = c\}) = 1$ then $X + Y \overset{d}{=} X + c$ and $XY \overset{d}{=} cX$, where $0 \times X$ is treated as a random variable degenerate at 0.

(ii) Let $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ be sequences of real numbers such that $X_n \overset{d}{=} Y_n, n = 1, 2, \ldots$. If, for some real constant c, $X_n \overset{p}{\to} c$, as $n \to \infty$, then $Y_n \overset{p}{\to} c$, as $n \to \infty$.

(iii) Let $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ be sequences of real numbers such that $X_n \overset{d}{=} Y_n, n = 1, 2, \ldots$. If, for some random variable , $X_n \overset{d}{\to} X$, as $n \to \infty$, then $Y_n \overset{d}{\to} X$, as $n \to \infty$.

(iv) Let $\{a_n\}_{n\geq 1}$ be sequence of real numbers such that $\lim_{n\to\infty} a_n = a \in \mathbb{R}$ and let $\{X_n\}_{n\geq 1}$ be a sequence of random variables such that X_n is degenerate at $a_n, n = 1, 2, \ldots$. Then $X_n \overset{p}{\to} a$, as $n \to \infty$. ■

Theorem 3.2

Let $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ be sequences of random variables and let $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ be sequences of real numbers such that $\lim_{n\to\infty} a_n = a$ and $\lim_{n\to\infty} b_n = b$.

(i) Suppose that, for some real constants c_1 and c_2, $X_n \overset{p}{\to} c_1$ and $Y_n \overset{p}{\to} c_2$, as $n \to \infty$. Then, as $n \to \infty$, $X_n + Y_n \overset{p}{\to} c_1 + c_2$, $X_n - Y_n \overset{p}{\to} c_1 - c_2$ and $X_n Y_n \overset{p}{\to} c_1\, c_2$. Moreover, if $c_2 \neq 0$, then $\frac{X_n}{Y_n} \overset{p}{\to} \frac{c_1}{c_2}$, as $n \to \infty$.

(ii) Suppose that, for a real constant c and a random variable X, $X_n \overset{d}{\to} X$ and $Y_n \overset{p}{\to} c$, as $n \to \infty$. Then, as $n \to \infty$, $X_n + Y_n \overset{d}{\to} X + c$, $X_n - Y_n \overset{d}{\to} X - c$ and $X_n Y_n \overset{d}{\to} cX$. Moreover, if $c \neq 0$, then $\frac{X_n}{Y_n} \overset{d}{\to} \frac{X}{c}$, as $n \to \infty$.

(iii) Suppose that, for a real constant c, $X_n \overset{p}{\to} c$, as $n \to \infty$. Then $a_n X_n + b_n \overset{p}{\to} ac + b$, as $n \to \infty$.

(iv) Suppose that, for a random variable X, $X_n \overset{d}{\to} X$, as $n \to \infty$. Then $a_n X_n + b_n \overset{d}{\to} aX + b$, as $n \to \infty$.

Proof. (i) and (ii). Follow from Theorem 3.1 (ii) and (iv) as $h_1(x, y) = x + y, h_2(x, y) = x - y$ and $h_3(x, y) = xy$ are continuous functions on $\mathbb{R}^2$, and $h_4(x, y) = \frac{x}{y}$ is continuous on $D = \{(s, t) \in \mathbb{R}^2 : t \neq 0\}$.

(iii) Let Y_n be a random variable that is degenerate at a_n and let Z_n be a random variable that is degenerate at $b_n, n = 1, 2, \ldots$. Then $Y_n \overset{p}{\to} a$ and $Z_n \overset{p}{\to} b$, as $n \to \infty$ (Lemma 3.1

(iv)). Now using (i) we get $X_nY_n + Z_n \xrightarrow{p} ac + b$, as $n \to \infty$. Since $a_n X_n + b_n \stackrel{d}{=} X_nY_n + Z_n,\ n = 1,2,\cdots$, (Lemma 3.1 (i)), the assertion follows on using Lemma 3.1 (ii).

(iv) Let Y_n and Z_n be as defined in (iii). Then $Y_n \xrightarrow{p} a$ and $Z_n \xrightarrow{p} b$, as $n \to \infty$. Using (ii) we get $X_nY_n + Z_n \xrightarrow{d} aX + b$, as $n \to \infty$. Since $a_n X_n + b_n \stackrel{d}{=} X_nY_n + Z_n,\ n = 1, 2, \ldots,$ the assertion follows on using Lemma 3.1 (iii). ■

Remark 3.1

The CLT asserts that if $X_1, X_2, \ldots$ are i.i.d. random variables with mean μ and finite variance $\sigma^2 > 0$, then

$$Z_n \stackrel{\text{def}}{=} \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty,$$

where $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Since $\frac{\sigma}{\sqrt{n}} \to 0$, as $n \to \infty$, using Theorem 3.2 (iv) we get

$$\bar{X}_n - \mu = \frac{\sigma}{\sqrt{n}}\, Z_n \xrightarrow{d} 0 \times Z, \text{as } n \to \infty.$$

Note that $0 \times Z$ is a random variable degenerate at 0. Thus it follows that

$$\begin{aligned} & \bar{X}_n - \mu \xrightarrow{d} 0, && \text{as } n \to \infty \\ \Leftrightarrow\ & \bar{X}_n - \mu \xrightarrow{p} 0, && \text{as } n \to \infty \\ \Leftrightarrow\ & \bar{X}_n \xrightarrow{p} \mu, && \text{as } n \to \infty. \end{aligned}$$

The above discussion suggests that, under the finiteness of second moment (or variance), the CLT is a stronger result than the WLLN. ■

Example 3.1

Let $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ be sequences of random variables.

(i) If $X_n \xrightarrow{p} \ln 4$ and $Y_n \xrightarrow{p} 2$, as $n \to \infty$, show that $X_n + \ln Y_n \xrightarrow{p} \ln 8$ and $e^{X_n} \ln Y_n \to \ln 16$, as $n \to \infty$;

(ii) If $X_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$, show that and $X_n^2 \xrightarrow{d} Q_1 \sim \chi_1^2$ (the chi-square distribution with one degree of freedom), as $n \to \infty$.

(iii) If $X_n \xrightarrow{d} Z \sim N(0,1)$, and $Y_n \xrightarrow{p} 3$, as $n \to \infty$, show that $X_nY_n \xrightarrow{d} V \sim N(0, 9)$ and $2X_n + 3Y_n \xrightarrow{d} Q_2 \sim N(9, 4)$, as $n \to \infty$.

(iv) For a given $\theta > 0$, if $X_1, X_2, \ldots$ are i.i.d. $U(0,\theta)$ random variables and $X_{n:n} = \max\{X_1, \ldots, X_n\}$, $n = 1, 2, \ldots$, show that $e^{X_{n:n}} \xrightarrow{p} e^{\theta}, X_{n:n}^2 + X_{n:n} + 1 \xrightarrow{p} \theta^2 + \theta + 1$ and $e^{-\frac{n(\theta - X_{n:n})}{\theta}} \xrightarrow{d} U \sim U(0,1)$, as $n \to \infty$.

Solution.

(i) Since $h_1(x) = \ln x\,, x \in (0,\infty)$ is a continuous function, using Theorem 3.1 (i) it follows that $\ln Y_n \xrightarrow{p} \ln 2$, as $n \to \infty$. Now on using Theorem 3.2 (i) we get $X_n + \ln Y_n \xrightarrow{p} \ln 4 + \ln 2 = \ln 8$, as $n \to \infty$. Also, since $h_2(x) = e^x, x \in \mathbb{R}$, is a continuous function on $\mathbb{R}$, on using Theorem 3.1 (i), we get $e^{X_n} \xrightarrow{p} e^{\ln 4} = 4$, as $n \to \infty$. Now on using Theorem 3.2 (i) it follows that $e^{X_n} \ln Y_n \xrightarrow{p} 4 \ln 2 = \ln 16$, as $n \to \infty$.

(ii) Since $h_3(x) = x^2, x \in \mathbb{R}$, is a continuous function on $\mathbb{R}$, using Theorem 3.1 (iii) we get $X_n^2 \xrightarrow{d} Z^2$, as $n \to \infty$. Let $Q_1 = Z^2$. Since $Z \sim N(0,1)$, we have $Q_1 \sim \chi_1^2$ (Theorem 4.1 (ii), Module 5). Consequently $X_n^2 \xrightarrow{d} Q_1 \sim \chi_1^2$, as $n \to \infty$.

(iii) Using Theorem 3.2 (ii) we get $X_n Y_n \xrightarrow{d} 3Z$, as $n \to \infty$. Let $V = 3Z$. Since $Z \sim N(0,1)$ we have $V = 3Z \sim N(0,9)$ (Theorem 4.2 (ii) Module 5) and, therefore, $X_n Y_n \xrightarrow{d} V \sim N(0,9)$, as $n \to \infty$. Using theorem 3.2 (iii) and (iv) we get $2X_n \xrightarrow{d} 2Z$ and $3Y_n \xrightarrow{p} 9$, as $n \to \infty$. Now using Theorem 3.2 (ii) we also conclude that $2X_n + 3Y_n \xrightarrow{d} 2Z + 9$, as $n \to \infty$. Let $Q_2 = 2Z + 9$. Since $Z \sim N(0,1)$, we have $Q_2 \sim N(9,4)$ (Theorem 4.2 (ii), Module 5).

(iv) From Example 1.4 we have $X_{n:n} \xrightarrow{p} \theta$, as $n \to \infty$, and $Y_n = n\,(\theta - X_{n:n}) \xrightarrow{d} Y \sim \mathrm{Exp}(\theta)$, as $n \to \infty$. Since $h_4(x) = e^x, x \in \mathbb{R}$, $h_5(x) = x^2 + x + 1$, $x \in \mathbb{R}$, and $h_6(x) = e^{-\frac{x}{\theta}}$, $x \in \mathbb{R}$, are continuous functions on $\mathbb{R}$, using Theorem 3.1 (i) and (ii), we get $e^{X_{n:n}} \xrightarrow{p} e^{\theta}, X_{n:n}^2 + X_{n:n} + 1 \xrightarrow{p} \theta^2 + \theta + 1$ and $e^{-\frac{Y_n}{\theta}} \xrightarrow{d} e^{-\frac{Y}{\theta}}$, as $n \to \infty$. Let $U = e^{-\frac{Y}{\theta}}$. Since $Y \sim \mathrm{Exp}\,(\theta)$, it is easy to verify that $U \sim U(0,1)$. Consequently, $e^{-\frac{n(\theta - X_{n:n})}{\theta}} = e^{-\frac{Y_n}{\theta}} \xrightarrow{d} U \sim U(0,1)$, as $n \to \infty$. ■

Theorem 3.3

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with finite mean μ. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$, $n = 2, 3, \ldots$, be sequences of sample means and sample variances, respectively. Define $T_n = \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{S_n}$, $n = 2, 3, \ldots$.

(i) If $\sigma^2 = \mathrm{Var}(X_1) \in (0,\infty)$, then $S_n^2 \xrightarrow{p} \sigma^2$, $S_n \xrightarrow{p} \sigma$ and $T_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$;

(ii) Suppose that the kurtosis $\gamma_1 = \frac{E((X_1-\mu)^4)}{\sigma^4} < \infty$. Then $\sqrt{n}\,(S_n^2 - \sigma^2) \xrightarrow{d} W \sim N(0, (\gamma_1 - 1)\sigma^4)$, as $n \to \infty$.

Proof.

(i) We have

$$\begin{aligned} S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 \\ &= \frac{1}{n-1}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\,\bar{X}_n^2 \\ &= \frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \frac{n}{n-1}\,\bar{X}_n^2, \qquad n = 2, 3, \ldots. \end{aligned}$$

Let $Y_i = X_i^2, i = 1, 2, \ldots$ and let $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i\,, n = 2, 3, \ldots$. Then

$$S_n^2 = \frac{n}{n-1}(\bar{Y}_n - \bar{X}_n^2),$$

and $Y_1, Y_2, \ldots$ is a sequence of i.i.d. random variables with mean $E(Y_1) = E(X_1^2) = \sigma^2 + \mu^2$. By WLLN

$$\bar{Y}_n \xrightarrow{p} \sigma^2 + \mu^2, \qquad \text{as } n \to \infty$$

$$\text{and} \qquad \bar{X}_n \xrightarrow{p} \mu, \qquad \text{as } n \to \infty.$$

Using the continuity of function $h(x) = x^2,\ x \in \mathbb{R}$, and Theorem 3.1 (i) we have $\overline{X}_n^2 \xrightarrow{p} \mu^2$, as $n \to \infty$. Since $\frac{n}{n-1} \to 1$, on using Theorem 3.2 (i) and (iii) we get

$$S_n^2 = \frac{n}{n-1}\left(\overline{Y}_n - \overline{X}_n^2\right) \xrightarrow{p} \sigma^2, \qquad \text{as } n \to \infty.$$

Since $f(x) = \sqrt{x},\ x \in (0, \infty)$, is a continuous function, it follows that $S_n \xrightarrow{p} \sigma,$ as $n \to \infty$, and therefore $\frac{\sigma}{S_n} \xrightarrow{p} 1$, as $n \to \infty$. Using the CLT we have

$$Z_n \stackrel{\text{def}}{=} \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1), \qquad \text{as } n \to \infty$$

$$\Rightarrow \quad T_n = \frac{\sigma}{S_n}\,Z_n \xrightarrow{d} Z \sim N(0,1), \ \text{ as } n \to \infty, \ \text{(using Theorem 3.2 (iii))}.$$

(ii) Let $T_i = \frac{X_i - \mu}{\sigma}, i = 1, \ldots, n$, so that $T_1, T_2, \ldots$ are i.i.d. random variables with mean 0 and variance 1. Moreover $X_i = \mu + \sigma T_i,\ i = 1, 2, \ldots, \bar{X}_n = \mu + \sigma \bar{T}_n, \bar{T}_n = \frac{1}{n}\sum_{i=1}^{n} T_i$ and

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

$$= \frac{\sigma^2}{n-1}\sum_{i=1}^{n}(T_i - \bar{T}_n)^2$$

$$= \frac{n}{n-1}\,\sigma^2\left[\frac{1}{n}\sum_{i=1}^{n} T_i^2 - \bar{T}_n^2\right]$$

$$= \frac{n}{n-1}\,\sigma^2\left[\frac{1}{n}\sum_{i=1}^{n} Y_i - \bar{T}_n^2\right]$$

$$= \frac{n}{n-1}\,\sigma^2[\bar{Y}_n - \bar{T}_n^2],$$

where $Y_i = T_i^2, i = 1, 2, \ldots$ and $\bar{Y}_n = \frac{1}{n}\sum_{i=1}^{n} Y_i\,, n = 2, 3, \ldots.$ Then $Y_1, Y_2, \ldots$ are i.i.d. random variables with mean $E(Y_1) = E(T_1^2) = 1$ and $\mathrm{Var}(Y_1) = E(T_1^4) - \left(E(T_1^2)\right)^2 = \gamma_1 - 1$. By the CLT

$$U_n \stackrel{\text{def}}{=} \frac{\sqrt{n}\,(\bar{Y}_n - 1)}{\sqrt{\gamma_1 - 1}} \stackrel{d}{\to} U \sim N(0,1), \qquad \text{as } n \to \infty$$

$$\text{and} \qquad V_n = \sqrt{n}\,\bar{T}_n \stackrel{d}{\to} V \sim N(0,1), \qquad \text{as } n \to \infty.$$

Also,

$$\sqrt{n}\,(S_n^2 - \sigma^2) = \frac{n}{n-1}\sigma^2\sqrt{\gamma_1 - 1}\,U_n + \frac{\sqrt{n}}{n-1}\sigma^2 - \frac{\sqrt{n}}{n-1}\sigma^2 V_n^2,\ \ n = 2, 3, \ldots.$$

Using continuity of function $h(x) = x^2,\ x \in (0, \infty)$, and Theorem 3.1 (iii) we have $V_n^2 \stackrel{d}{\to} V^2$, as $n \to \infty$. Since, as $n \to \infty, \frac{n}{n-1}\,\sigma^2\sqrt{\gamma_1 - 1} \to \sigma^2\sqrt{\gamma_1 - 1}$ and $\frac{\sqrt{n}}{n-1}\,\sigma^2 \to 0$, using Theorem 3.2 we conclude that

$$\sqrt{n}(S_n^2 - \sigma^2) \stackrel{d}{\to} W \sim N(0, (\gamma_1 - 1)\sigma^4), \qquad \text{as } n \to \infty,$$

where $W = \sigma^2\sqrt{\gamma_1 - 1}\,U \sim N(0, (\gamma_1 - 1)\sigma^4)$. ■

Corollary 3.1. (Nomal Approximation to the Student-t Distribution)

Let $\{T_n\}_{n\geq 1}$ be a sequence of random variables such that $T_n \sim t_n$, the Student-t distribution with n degrees of freedom. Then $T_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$.

Proof. Let $Z_1, Z_2, \dots$ be a sequence of i.i.d. $N(0,1)$ random variables. Let $\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$ and $S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Z_i - \bar{Z}_n)^2$, $n = 2, 3, \dots$. Define

$$V_n = \frac{\sqrt{n}\,\bar{Z}_n}{S_n}, \qquad n = 2, 3, \dots.$$

By Corollary 11.1, Module 6, $V_n \stackrel{d}{=} T_{n-1}, n = 2, 3, \dots$. By Theorem 3.3 (i) we have

$$V_n \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty$$

$$\Rightarrow \ T_{n-1} \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty$$

$$\Rightarrow \ T_n \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty. \blacksquare$$

4. The Delta-Method

Generally we have a sequence $\{X_n\}_{n\geq 1}$ of random variables such that, for real constants c and $b > 0$, $X_n \xrightarrow{p} c$, and $n^b(X_n - c) \xrightarrow{d} X$, as $n \to \infty$, where X is some random variable. Then, for any continuous function $g(\cdot)$, we know that $g(X_n) \xrightarrow{p} g(c)$, as $n \to \infty$. The Delta-method is a tool for providing a non-degenerate limiting distribution to a normalized version of $g(X_n), n = 1, 2, \dots$.

Theorem 4.1 (The Delta-Method)

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables such that, for some real constants $b > 0$ and c and some random variable $X, n^b(X_n - c) \xrightarrow{d} X$, as $n \to \infty$. Let $g: \mathbb{R} \to \mathbb{R}$ be a function that is differentiable at c. Then

$$n^b\big(g(X_n) - g(c)\big) \xrightarrow{d} g^{(1)}(c)X, \text{as } n \to \infty,$$

where $g^{(1)}(c)$ is the derivative of $g(\cdot)$ at the point c.

Proof. Let $\Psi_1: \mathbb{R} \to \mathbb{R}$ be such that $\Psi_1(c) = 0$ and

$$g(x) = g(c) + (x - c)\left(g^{(1)}(c) + \Psi_1(x)\right), x \in \mathbb{R},$$

i.e.,

$$\Psi_1(x) = \begin{cases} \dfrac{g(x) - g(c)}{x - c} - g^{(1)}(c), & \text{if } x \in \mathbb{R} - \{c\} \\ 0, & \text{if } x = c \end{cases}.$$

Then $\lim_{x \to c} \Psi_1(x) = g^{(1)}(c) - g^{(1)}(c) = 0 = \Psi_1(c)$ (i.e., $\Psi_1(\cdot)$ is continuous at c) and

$$n^b\big(g(X_n) - g(c)\big) = g^{(1)}(c) n^b (X_n - c) + \Psi_1(X_n) n^b (X_n - c), n = 1, 2, \ldots.$$

Also, by Theorem 3.2 (iv),

$$X_n = n^{-b}\Big(n^b(X_n - c)\Big) + c \overset{d}{\to} 0 \times X + c, \text{as } n \to \infty$$

$$\Rightarrow \quad X_n \overset{p}{\to} c, \text{as } n \to \infty$$

$$\Rightarrow \quad \Psi_1(X_n) \overset{p}{\to} \Psi_1(c) = 0, \text{as } n \to \infty \ \ (\text{since } \Psi_1 \text{ is continuous at } c)$$

$$\Rightarrow \quad \Psi_1(X_n) n^b (X_n - c) \overset{p}{\to} 0, \ \text{as } n \to \infty \quad \text{(Theorem 3.2 (ii))}$$

$$\Rightarrow \quad n^b\big(g(X_n) - g(c)\big) = g^{(1)}(c) n^b (X_n - c) + \Psi_1(X_n) n^b (X_n - c)$$

$$\overset{d}{\to} g^{(1)}(c)\, X, \text{as } n \to \infty \qquad \text{(Theorem 3.2).} \ \blacksquare$$

Remark 4.1

Note that, in the above theorem, if we have $g^{(1)}(c) = 0$ then we conclude that

$$n^b\big(g(X_n) - g(c)\big) \overset{d}{\to} 0, \text{as } n \to \infty$$

$$\text{i.e.,} \qquad n^b\big(g(X_n) - g(c)\big) \overset{p}{\to} 0, \text{as } n \to \infty,$$

and we get a degenerate limiting distribution. Now suppose that $g^{(1)}(c) = 0$ and $g(\cdot)$ is twice differentiable at c with first and second derivatives at the point c given by $g^{(1)}(c)$ and $g^{(2)}(c)$, repectively. Define $\Psi_2 \colon \mathbb{R} \to \mathbb{R}$ by

$$\Psi_2(x) = \begin{cases} \dfrac{g(x) - g(c)}{(x - c)^2 / 2} - g^{(2)}(c), & \text{if } x \neq c \\ 0, & \text{if } x = c \end{cases}.$$

The, using L' Hospital rule (0/0 form), we have

$$\lim_{x \to c} \Psi_2(x) = \lim_{x \to c} \frac{g^{(1)}(x)}{x - c} - g^{(2)}(c)$$

$$= \lim_{x \to c} \frac{g^{(1)}(x) - g^{(1)}(c)}{x - c} - g^{(2)}(c) \ \ (\text{since } g^{(1)}(c) = 0)$$

$$= g^{(2)}(c) - g^{(2)}(c)$$

$$= 0$$

$$= \Psi_2(c),$$

i.e., $\Psi_2(\cdot)$ is continuous at point c. Consequently, using Theorem 3.2,

$$n^{2b}\big(g(X_n) - g(c)\big) = \frac{g^{(2)}(c)}{2}\left(n^b(X_n - c)\right)^2 + \frac{\left(n^b(X_n - c)\right)^2}{2}\Psi_2(X_n)$$

$$\xrightarrow{d} \frac{g^{(2)}(c)}{2} X^2,$$

since $\Psi_2(X_n) \xrightarrow{p} \Psi_2(c) = 0$ (as Ψ_2 is continuous at c and $X_n \xrightarrow{p} c$, as $n \to \infty$) and $\left(n^b(X_n - c)\right)^2 \xrightarrow{d} X^2$ (as $h(x) = x^2$ is a continuous function on $\mathbb{R}$ and $n^b(X_n - c) \xrightarrow{d} X$, as $n \to \infty$). ■

The following example demonstrates that the conclusion of Theorem 4.1 (The Delta-Method) may not hold if $b = 0$.

Example 4.1

Let $\{Z_n\}_{n \geq 1}$ be a sequence of random variables such that $Z_n \sim N(0,1), n = 1, 2, \ldots$. Then $n^0(Z_n - 0) = Z_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$. Let $g(x) = x^2$, $x \in \mathbb{R}$. Then

$$n^0\big(g(Z_n) - g(0)\big) = Z_n^2 \xrightarrow{d} Z_1^2 \sim \chi_1^2, \text{as } n \to \infty.$$

However $g^{(1)}(0)Z = 0 \times Z = 0$. ■

Corollary 4.1

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables, each having the mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in (0, \infty)$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, $n = 1, 2, \ldots$ and let $g: \mathbb{R} \to \mathbb{R}$ be a function that is differentiable at μ. Then

$$\sqrt{n}\big(g(\bar{X}_n) - g(\mu)\big) \xrightarrow{d} W \sim N\left(0, \left(g^{(1)}(\mu)\right)^2\sigma^2\right), \text{as } n \to \infty,$$

provided $g^{(1)}(\mu) \neq 0$. If $g^{(1)}(\mu) = 0$ then

$$\sqrt{n}\big(g(\bar{X}_n) - g(\mu)\big) \xrightarrow{p} 0, \text{as } n \to \infty \cdot$$

Proof. Let $Z \sim N(0,1)$ and let $V = \sigma Z$. Then by the CLT

$$\frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty$$

$$\Rightarrow \quad \sqrt{n}\,(\bar{X}_n - \mu) \xrightarrow{d} \sigma Z = V \sim N(0, \sigma^2\,), \text{as } n \to \infty$$

$$\Rightarrow \quad \sqrt{n}\,(g(\bar{X}_n) - g(\mu)) \xrightarrow{d} g^{(1)}(\mu)\,V, \text{as } n \to \infty \cdot$$

If $g^{(1)}(\mu) \neq 0$, then $W = g^{(1)}(\mu)\,V \sim N\big(0, (g^{(1)}(\mu))^2\sigma^2\,\big)$. However if $g^{(1)}(\mu) = 0$, then the random variable $g^{(1)}(\mu)\,V$ is degenerate at 0. Hence the result follows. ∎

Example 4.2

Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables such that $X_n \sim \chi^2_n, n = 1, 2, \ldots$. Show that

$$\sqrt{2}\,\big(\sqrt{X_n} - \sqrt{n}\big) \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty.$$

Solution. Let $Y_1, Y_2, \ldots$ be a sequence of i.i.d. χ^2_1 random variables. Then $E(Y_1) = 1, \mathrm{Var}(Y_1) = 2$ and $X_n \stackrel{d}{=} \sum_{i=1}^{n} Y_i = n\bar{Y}_n\,, n = 1, 2, \ldots$ (see Example 7.6 (i), Module 6). By the CLT

$$\frac{\sqrt{n}\,(\bar{Y}_n - 1)}{\sqrt{2}} \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty$$

$$\Rightarrow \sqrt{n}\,(\bar{Y}_n - 1) \xrightarrow{d} \sqrt{2}\,Z \sim N(0, 2), \text{as } n \to \infty \cdot$$

Since $g(x) = \sqrt{x}, x \in (0, \infty)$ is differentiable at $x = 1$, using the delta-method we have

$$\sqrt{n}\left(\sqrt{\bar{Y}_n} - 1\right) \xrightarrow{d} \frac{1}{2} \times \sqrt{2}\,Z = \frac{Z}{\sqrt{2}} \sim N\left(0, \frac{1}{2}\right), \text{as } n \to \infty.$$

$$\Rightarrow \quad \sqrt{2}\,\big(\sqrt{X_n} - \sqrt{n}\big) \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty. \blacksquare$$

Problems

1. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. $N(\mu, \sigma^2)$ random variables, where $\mu > 0$ and $0 < \sigma^2 < \infty$. Let $Z_n = \sum_{i=1}^{n} X_i$ and let $M_n = \sqrt{n}\,(\bar{X}_n - \mu)$, where $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i\,, n = 1, 2, \ldots$. Show that the sequence $\{Z_n\}_{n\geq 1}$ does not have a limiting distribution, however, the sequence $\{M_n\}_{n\geq 1}$ has a limiting distribution.

2. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables. Let $X_{1:n} = \min\{X_1, \ldots, X_n\}$ and let $Y_n = nX_{1:n}, n = 1, 2, \ldots$. Find the limiting distribution of $\{X_{1:n}\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ when

(i) $X_1 \sim \mathrm{U}(0, \theta), \theta > 0;$

(ii) $X_1 \sim \mathrm{Exp}(\theta), \theta > 0.$

3. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with mean μ and finite variance. Show that

(i) $\frac{2}{n(n+1)} \sum_{i=1}^{n} i\, X_i \xrightarrow{p} \mu, \text{as } n \to \infty;$

(ii) $\frac{6}{n(n+1)(2n+1)} \sum_{i=1}^{n} i^2\, X_i \xrightarrow{p} \mu, \text{as } n \to \infty \cdot$

4. Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables such that the p.m.f. of X_n is given by

$$f_n(x) = \begin{cases} \frac{1}{2}, & \text{if } x \in \left\{-n^{\frac{1}{4}}, n^{\frac{1}{4}}\right\} \\ 0, & \text{otherwise} \end{cases}.$$

Show that $\bar{X}_n \xrightarrow{p} 0$, as $n \to \infty$, where $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$, $n = 1, 2, \ldots$.

5. Let $X_n \sim \mathrm{NB}(n, p_n)$, where $p_n \in (0,1), n = 1, 2, \ldots$ and $\lim_{n\to\infty} n\,(1 - p_n) = \lambda > 0$. Show that $X_n \xrightarrow{d} X \sim \mathrm{P}(\lambda)$, the Poisson distribution with mean λ.

6. (i) Let $X_n \sim \mathrm{G}\left(n, \frac{1}{n}\right), n = 1, 2, \ldots$. Show that $X_n \xrightarrow{p} 1$, as $n \to \infty$.

(ii) Let $X_n \sim N\left(\frac{1}{n}, 1 - \frac{1}{n}\right), n = 1, 2, \ldots$. Show that $X_n \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$.

7. Consider a random sample of size 80 from the distribution having a p.d.f.

$$f(x) = \begin{cases} \frac{2}{x^3}, & \text{if } x > 1 \\ 0, & \text{otherwise} \end{cases}.$$

Compute, approximately, the probability that not more than 20 of the items of the random sample are greater than $\sqrt{6}$.

8. Let $X_1, X_2, \ldots, X_{200}$ be a random sample from $\mathrm{P}(2)$ distribution, and let $Y = \sum_{i=1}^{200} X_i$. Find, approximately, $P(\{420 \leq Y_{200} \leq 440\})$.

9. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables having a common p.d.f.

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + x^2}, -\infty < x < \infty.$$

Using the principle of mathematical induction, show that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \stackrel{d}{=} X_1, \forall\, n \in \{1, 2, \dots\}$. Hence show that $\{\bar{X}_n\}_{n\geq 1}$ does not converge to anything in probability (Note that $E(X_1)$ is not finite and therefore validity of WLLN is not guaranteed).

10. Let $X_n \sim \mathrm{P}(2n), Y_n = \frac{X_n}{n}$ and $Z_n = \frac{X_{n^2}}{n(2n+1)}, n = 1, 2, \dots$.
Show that
(i) $Y_n \xrightarrow{p} 2$ and $Z_n \xrightarrow{p} 1$, as $n \to \infty$;
(ii) $Y_n^2 + \sqrt{Z_n} \xrightarrow{p} 5$, as $n \to \infty$;
(iii) $\frac{n^2\, Y_n^2 + nY_n}{nY_n + n^2} \xrightarrow{p} 4$, as $n \to \infty$.

11. Let $\bar{X}_n$ be the sample mean based on a random sample of size n from a distribution having mean $\mu \in (-\infty, \infty)$ and variance $\sigma^2 \in (0, \infty)$. Let $Z_n = \frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma}, n = 1, 2, \dots$. If $\{Y_n\}_{n\geq 1}$ is a sequence of random variables such that $Y_n \xrightarrow{p} 2$, as $n \to \infty$, show that:
(i) $\frac{2Z_n}{Y_n} \xrightarrow{d} Z \sim \mathrm{N}(0,1)$, as $n \to \infty$;
(ii) $\frac{4Z_n^2}{Y_n^2} \xrightarrow{d} U \sim \chi_1^2$, as $n \to \infty$;
(iii) $\frac{(2n+Y_n)Z_n}{nY_n + Y_n^2} \xrightarrow{d} Z \sim N(0,1)$, as $n \to \infty$.

12. Let $X_1, X_2, \dots$ be a sequence of i.i.d. $\mathrm{U}(0,1)$ random variables. Let $G_n = (X_1\, X_2 \cdots X_n)^{\frac{1}{n}}, n = 1, 2, \dots$ be the sequence of geometric means. Show that, as $n \to \infty$,
(i) $G_n \xrightarrow{p} \frac{1}{e}$;
(ii) $n^b \left(G_n^2 - \frac{1}{e^2}\right) \xrightarrow{d} N(0, \sigma^2)$, for some $b > 0$ and $\sigma^2 > 0$. Find the values of b and σ^2.

13. Let $\{(X_{1n},\, X_{2n})\}_{n\geq 1}$ be a sequence of i.i.d. bivariate random vectors such that $E(X_{11}) = \mu_1 \in \mathbb{R}, E\,(X_{21}) = \mu_2 \in \mathbb{R}, \mathrm{Var}(X_{11}) = \sigma_1^2 > 0$, $\mathrm{Var}\,(X_{21}) = \sigma_2^2 > 0$, and $\mathrm{Corr}(X_{11}, X_{21}) = \rho \in (-1, 1)$. Let $\bar{X}_{1n} = \frac{1}{n}\sum_{i=1}^{n} X_{1i}\,, \bar{X}_{2n} = \frac{1}{n}\sum_{i=1}^{n} X_{2i}\,,$ $C_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_{1i} - \bar{X}_{1n})\,(X_{2i} - \bar{X}_{2n}), S_{1n}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{1i} - \bar{X}_{1n})^2\,, S_{2n}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_{2i} - \bar{X}_{2n})^2$ and $R_n = \frac{C_n}{S_{1n}S_{2n}}, n = 2, 3, \dots$. Show that, as $n \to \infty$,
(i) $C_n \xrightarrow{p} \rho\sigma_1\sigma_2$ and $R_n \xrightarrow{p} \rho$;
(ii) $\sqrt{n}\,(C_n - \rho\,\sigma_1\,\sigma_2) \xrightarrow{d} N(0, (\theta - \rho^2)\sigma_1^2\,\sigma_2^2)$, where $\theta = \frac{E\left((X_{11}-\mu_1)^2(X_{21}-\mu_2)^2\right)}{\sigma_1^2\sigma_2^2}$.

14.

(i) Let $X_n \sim \text{Bin}\,(n, p_n)$, where $p_n \in (0,1), n = 1, 2, \ldots$ and $\lim_{n\to\infty} p_n = p \in (0,1)$. Show that

$$\frac{X_n - np_n}{\sqrt{np_n(1-p_n)}} \xrightarrow{d} Z \sim N(0,1), \text{as } n \to \infty;$$

(ii) Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables of absolutely continuous type. Let $F(\cdot)$ and $f(\cdot)$, respectively, denote the d.f. and the p.d.f. of X_1 and let θ be the median of $F\left(\text{i.e.}, F(\theta) = \frac{1}{2}\right)$. Suppose that $f(\theta) > 0$. Let $M_n = X_{n+1:2n+1}, n = 1, 2, \ldots$, be the middle observation (called the sample median) based on random sample $X_1, X_2, \ldots, X_{2n+1}$. Show that, as $n \to \infty$,

(a) $\sqrt{n}\,(M_n - \theta) \xrightarrow{d} N\left(0, \frac{1}{4f^2(\theta)}\right)$;

(b) $M_n \xrightarrow{p} \theta$.

15. Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables such that, for real constants μ and $\sigma > 0, \sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2)$, as $n \to \infty$. Find the limiting distributions of

(i) $S_n = \sqrt{n}\,(X_n^2 - \mu^2), n = 1, 2, \ldots,$

(ii) $T_n = n(X_n - \mu)^2, n = 1, 2, \ldots,$

(iii) $U_n = \sqrt{n}\,(\ln X_n - \ln \mu), n = 1, 2, \ldots$, where $\mu > 0$.

16. Let $X_n \sim \text{Bin}\,(n, p), n = 1, 2, \ldots$. Find the limiting distribution of $Z_n = \sqrt{n}\,\left(\frac{X}{n}\left(1 - \frac{X}{n}\right) - p(1-p)\right), n = 1, 2, \ldots$. Find the limiting distribution (non degenerate) of a normalized version of $Y_n = \frac{X}{n}\left(1 - \frac{X}{n}\right)$ when $p = \frac{1}{2}$.