- Convergence in Probability

Convergence

$$\underset{Sequence}{a_n} \xrightarrow{n\to\infty} \underset{Constant}{a}$$

$$\lim_{n\to\infty} a_n = a$$

$$\underset{\underset{Convergence\ in\ Probability}{Random\ Variable}}{X_n \xrightarrow{n\to\infty} X}$$

$$\underset{Random\ Variable}{X_n} \xrightarrow{n\to\infty} \underset{Randon\ Variable}{X}$$

Definition

X_n : sequence of random variable

X : random variable

We say that "X_n convergence in probability to X". If

$$\lim_{n\to\infty} P[|X_n - X| \geq \varepsilon] = 0$$

$$\lim_{n\to\infty} P[|X_n - X| \leq \varepsilon] = 1, \quad \forall\varepsilon > 0.$$

If so, we write $X_n \xrightarrow{P} X$

- 약대수의 법칙Weak Law of Large Number

$$\bar{X}_n \xrightarrow{P} \mu \quad \Leftrightarrow \quad \lim_{n\to\infty} P[|\bar{X}_n - \mu| \geq \varepsilon] = 0$$

Chebyshe's ineguality

$X \sim (\mu, \sigma^2)$

$$P[|X - \mu| \geq a \cdot \sigma] \leq \frac{1}{a^2}$$

$$\Leftrightarrow \lim_{n\to\infty} P\left[|\bar{X}_n - \mu| \geq \left(\varepsilon \cdot \frac{\sqrt{n}}{\sigma}\right) \cdot \frac{\sigma}{\sqrt{n}}\right] \leq \frac{\sigma^2}{\varepsilon^2 n} \xrightarrow{n\to\infty} 0$$

- $E(X_i) = \mu$

$$\Rightarrow \frac{1}{n}\sum X_i \xrightarrow{P} E(X_i)$$

$$\frac{1}{n}\sum X_i^2 \xrightarrow{P} E(X_i^2)$$

$$\frac{1}{n}\sum X_i^n \xrightarrow{P} E(X_i^n)$$

- 확률 수렴 성질

  가정 : $X_n \xrightarrow{P} X, \quad Y_n \xrightarrow{P} Y$

  1. $X_n + Y_n \xrightarrow{P} X + Y$
  2. $aX_n \xrightarrow{P} aX$
  3. $X_n \xrightarrow{P} a \;\Rightarrow\; g(X_n) \xrightarrow{P} g(a), \quad$ g : a에서 연속
  4. $X_nY_n, \xrightarrow{P} XY$

$$\begin{aligned} X_nY_n &= \frac{1}{2}X_n^2 + \frac{1}{2}Y_n^2 - \frac{1}{2}(X_n - Y_n)^2 \\ &= \frac{1}{2}X^2 + \frac{1}{2}Y^2 - \frac{1}{2}(X - Y)^2 \\ &= XY \end{aligned}$$

- 일치추정량Consistency : 일치성

  $\hat{\theta}_n$: θ의 추정량

  $\hat{\theta}_n$: $X_1, \ldots, X_n$의 집합인 통계량

  If, $\hat{\theta}_n \xrightarrow{P} \theta$, then $\hat{\theta}_n$ is a consistent estimator of θ

$$\begin{aligned} &\Leftrightarrow \lim_{n\to\infty} P[|\hat{\theta}_n - \theta| \le \varepsilon] = 1 \\ &\Leftrightarrow \lim_{n\to\infty} P[\theta - \varepsilon \le \hat{\theta}_n \le \theta + \varepsilon] = 1 \\ &\Leftrightarrow \lim_{n\to\infty} [F(\theta + \varepsilon) - F(\theta - \varepsilon)] = 1 \end{aligned}$$

  example) sample mean $\bar{X}_n$ and sample variance S_n^2

$$\begin{aligned} \bar{X}_n &= \frac{1}{n}\sum X_i \xrightarrow{P} \mu \\ S_n^2 &= \frac{1}{n-1}\sum (X_i - \bar{X})^2 \xrightarrow{P} \sigma^2 \end{aligned}$$

  - sample variance

$$\begin{aligned} S_n^2 &= \frac{1}{n-1}\left[\sum X_i^2 - 2\bar{X}_n \sum X_i + n\bar{X}_n^2\right] \\ &= \frac{1}{n-1}\left[\sum X_i^2 - n\bar{X}_n^2\right] \\ &= \frac{n}{n-1}\left[\frac{1}{n}\sum X_i^2 - \bar{X}_n^2\right] \\ &\xleftrightarrow{n\to\infty} 1 \cdot \left[E(X_i^2) - \mu^2\right] \\ &= \sigma^2 \end{aligned}$$

$$S_n^2 \xrightarrow{P} \sigma^2$$

$\therefore$ 표본 분산 S_n^2은 모분산 σ^2의 일치추정량.

example) Suppose $X_1, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$. Let $Y = max\{X_1, \ldots, X_n\}$; order statistics.

1. cdf of Y_n

$$\begin{aligned} F_n(y) &= P(Y_n \le y) \\ &= P(X_1 \le y) \cdots P(X_n \le y) \\ &= [F_X(y)]^n \\ &= \begin{cases} \left(\frac{y}{\theta}\right)^n, & 0 \le y \le \theta \\ 0, & y < 0 \\ 1, & y \ge \theta \end{cases} \end{aligned}$$

2. pdf of Y_n

$$f_n(y) = F_n'(y) = n \left(\frac{y}{\theta}\right)^{n-1} \frac{1}{\theta} = \frac{n}{\theta^n} y^{n-1}$$

3. $E(Y_n)$

$$\begin{aligned} E(Y_n) &= \int_0^\theta y \cdot \frac{n}{\theta^n} y^{n-1} dy \\ &= \left[\frac{1}{n+1} y^{n-1} \frac{n}{\theta^n}\right]_0^\theta \\ &= \frac{n}{n+1} \frac{\theta^{n+1}}{\theta} \\ &= \frac{n}{n+1}\theta \end{aligned}$$

4. Unbiased estimator of Y_n

$$\begin{aligned} &\therefore\ E(Y_n) = \frac{n}{n+1}\theta \\ &\Leftrightarrow \frac{n+1}{n} E(Y_n) = \theta \\ &\Leftrightarrow E\left(\frac{n+1}{n} Y_n\right) = \theta \end{aligned}$$

즉, $\frac{n+1}{n} max(X_1, \ldots, X_n)$은 θ의 불편추정량

직관) n개를 균등하게 뽑는다. “평균적으로” 기대되는 $X_1, \ldots, X_n$은 등간격으로 뽑힘 , 간격 : $\frac{\theta}{n+1}$

5. Check consistency

$$Q\ :\ Y_n \xrightarrow{P} \theta \quad ?$$

$$
\begin{aligned}
\Leftrightarrow \lim_{n\to\infty} P[|Y_n - \theta| \le \varepsilon] = 1 \\
&= \lim_{n\to\infty} P[\theta - \varepsilon \le Y_n \le \theta + \varepsilon] \\
&= \lim_{n\to\infty} [F_n(\theta+\varepsilon) - F_n(\theta-\varepsilon)] \\
&= \lim_{n\to\infty} \left[1 - \left(\frac{\theta-\varepsilon}{\theta}\right)^n\right] \\
&= 1 - \lim_{n\to\infty} \left(\frac{\theta-\varepsilon}{\theta}\right)^n \\
&= 1
\end{aligned}
$$

$\therefore$ Y_n은 θ의 일치추정량이다. (θ의 불편추정량은 아니다.)

- 분포수렴Convergence in Distribution

$$X_n \xrightarrow{D} X$$

"확률변수의 분포"가 X의 "분포"로 수렴

Definition

$$X_n \to F_{X_n} : cdf\ of\ X_n$$

$$X \to F_X : cdf\ of\ X$$

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x), \quad \forall x \in C(F_n) : F_X\text{에서 연속인 점들}$$

$$\Rightarrow X_n \xrightarrow{D} X$$

확률수렴보다 약한 수렴이다.

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$$

$$X_n \xrightarrow{D} X \not\Rightarrow X_n \xrightarrow{P} X$$

분포수렴 : $X_n \xrightarrow{D} N(0,1)$ 위의 표현과 같다.

퇴화분포 *degenerate distribution*

X가 가질수 있는 값이 하나뿐일 때 (상수)

$\bar{X}_n$: random variable

$$F(\bar{x}) = \int_{-\infty}^{\bar{x}} \frac{1}{\sqrt{\frac{1}{n}}\sqrt{2\pi}} e^{-\frac{nw^2}{2}} dw$$

적분 안의 식 : $N\left(0, \frac{1}{n}\right)$의 pdf

$v = \sqrt{n}w$

$$F_n(\bar{x}) = \int_{-\infty}^{\sqrt{n}\bar{x}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv$$

적분 안의 식 : $N(0, 1)$의 pdf

$$\lim_{n\to\infty} F_n(\bar{x}) = \begin{cases} 0, & \bar{x} < 0 \ (\sqrt{n}\bar{x} \to -\infty) \\ \frac{1}{2} & \bar{x} = 0 \\ 1, & \bar{x} > 0 \end{cases}$$

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

$\Rightarrow$ $\lim_{n\to\infty} F_n(x) = F(x)$, $\bar{x} = 0$을 제외

$\Rightarrow$ $\bar{X}_n \xrightarrow{D} F$

$\star$ X : 이산형인 경우

$X_n \overset{D}{\to} X$ 라고 해서

$$\lim_{n\to\infty} P_n(x) \neq P(x)$$

example) pmf of X_n

$$P_n(x) = \begin{cases} 1, & x = 2 + \frac{1}{n} \\ 0, & elsewhere \end{cases}$$

1.

$$\lim_{n\to\infty} P_n(x) = \begin{cases} 0, & x = 2 + \frac{1}{\infty} \\ 0, & elsewhere \end{cases}$$

$$= 0, \quad \forall x$$

2.

$$F_n(x) = \begin{cases} 0, & x < 2 + \frac{1}{n} \\ 1, & x \geq 2 + \frac{1}{n} \end{cases} \quad \text{퇴화분포}$$

3.

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0, & x \le 2 \\ 1, & x > 2 \end{cases} \quad \text{cdf 가 아님}$$

4.

$$F(x) = \begin{cases} 0, & x < 2 \\ 1, & x \ge 2 \end{cases} \quad \text{cdf 를 만족}$$

$\lim_{n\to\infty} F_n(x) = F(x)$, $\quad x = 2$를 제외! $\Rightarrow$ $X_n \xrightarrow{D} F$, $\quad$ 즉, $x = 2$에서 퇴화분포

X : 연속형

웬만한 경우 분포수렴을 보일 때

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

대신에 $\lim_{n\to\infty} f_n(x) = f(x)$로 확인가능.

LDCT[Lebesgue Dominated Convergence Theorem]를 만족할 때,

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \int_{-\infty}^{x} f_n(y)dy = \int_{-\infty}^{x} \lim_{n\to\infty} f_n(y)dy$$

$\Rightarrow$ pdf로 분포수렴 확인 가능

Example)

Let T_n have a t-distribution with n degrees of freedom, $n = 1, 2, 3, \ldots$. Thus its cdf is

$$F_n(t) = \int_{-\infty}^{t} \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\Gamma(n/2)} \frac{1}{(1+y^2/n)^{(n+1)/2}} dy,$$

where the integrand is the pdf $f_n(y)$ of T_n. Find an asymptotic distribution of T_n.

$$\begin{aligned}
\lim_{n\to\infty} f_n(y) &= \lim_{n\to\infty} \frac{\Gamma[(n+1)/2]}{\sqrt{\pi n}\Gamma(n/2)} \frac{1}{(1+y^2/n)^{(n+1)/2}} \\
&= \underbrace{\lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \left(1+\frac{y^2}{n}\right)^{-\frac{n}{2}}}_{(1)} \underbrace{\lim_{n\to\infty} \left(1+\frac{y^2}{n}\right)^{-\frac{1}{2}}}_{(2)} \underbrace{\lim_{n\to\infty} \frac{\Gamma[(n+1)/2]}{\sqrt{n/2}\Gamma(n/2)}}_{(3)}
\end{aligned}$$

(1)

$$\lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \left(1+\frac{y^2}{n}\right)^{-\frac{n}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad \text{e의 정의}$$

(2)

$$\lim_{n\to\infty} \left(1+\frac{y^2}{n}\right)^{-\frac{1}{2}} = 1^{-\frac{1}{2}} = 1$$

(3)

Stirlings Formula

$$\Gamma(k+1) \approx \sqrt{2\pi} k^{k+1/2} e^{-k}$$

$$\begin{aligned}
\lim_{n\to\infty} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} &= \lim_{n\to\infty} \frac{\Gamma\left(\frac{n-1}{2}+1\right)}{\frac{n}{2}^{\frac{1}{2}}\Gamma\left(\frac{n-2}{2}+1\right)} \\
&\approx \lim_{n\to\infty} \left(\frac{n}{2}\right)^{-\frac{1}{2}} \frac{\sqrt{2\pi}\left(\frac{n-1}{2}\right)^{\frac{n-1}{2}+\frac{1}{2}} e^{-\frac{n-1}{2}}}{\sqrt{2\pi}\left(\frac{n-2}{2}\right)^{\frac{n-2}{2}+\frac{1}{2}} e^{-\frac{n-2}{2}}} \\
&= \lim_{n\to\infty} \left(\frac{n}{2}\right)^{-\frac{1}{2}} \frac{\left(\frac{n-1}{2}\right)^{\frac{n}{2}} e^{-\frac{n-1}{2}+\frac{n-2}{2}}}{\left(\frac{n-2}{2}\right)^{\frac{n-2}{2}}\left(\frac{n-2}{2}\right)^{-\frac{1}{2}}} \\
&= \lim_{n\to\infty} \frac{\left(\frac{n}{2}\right)^{-\frac{1}{2}}}{\left(\frac{n-2}{2}\right)^{-\frac{1}{2}}} \left(\frac{n-1}{n-2}\right)^{\frac{n}{2}} e^{-\frac{1}{2}} \\
&= \lim_{n\to\infty} \left(\frac{n}{n-2}\right)^{-\frac{1}{2}} \left(1+\frac{1}{n-2}\right)^{\frac{n}{2}} e^{-\frac{1}{2}} \\
&= 1 \times e^{\frac{1}{2}} \times e^{-\frac{1}{2}} \\
&= 1
\end{aligned}$$

Example)

Let $X_1, \ldots, X_n \overset{iid}{\sim} U(0,\theta)$, and let $Y_n = max(X_1, \ldots, X_n)$. Find the limiting distribution of $Z_n = n(\theta - Y_n)$.

$$\begin{aligned}
F_n(z) &= P(Z_n \le z) = P(n(\theta - Y_n) \le z) \\
&= P(Y_n \ge \theta - \frac{z}{n}) = 1 - P(Y_n \le \theta - \frac{z}{n}) \\
&= 1 - \frac{\left(\theta - \frac{z}{n}\right)^n}{\theta^n} = 1 - (1 - \frac{z/\theta}{n})^n \\
\lim_{n\to\infty} F_n(z) &= 1 - e^{-\frac{z}{\theta}}, \quad z \ge 0
\end{aligned}$$

$\Rightarrow$ 지수분포의 cdf ($Exp\left(\frac{1}{\theta}\right) = \Gamma(1,\theta)$)