Definition 5.1 (converge almost surely, TBp. 178)

A sequence of random variables $\{Z_n : \Omega \to \mathbb{R}\}$ is said to **converge almost surely** to a random variable $Z : \Omega \to \mathbb{R}$, and denoted as $Z_n \xrightarrow{\text{a.s.}} Z$, if for any $\epsilon > 0$,

$Z_n \xleftrightarrow{\text{alike}} Z$ as $n \to \infty$

converge pointwise

$$P\left(\left\{\omega \in \Omega : \lim_{n\to\infty} |Z_n(\omega) - Z(\omega)| < \epsilon\right\}\right) = 1.$$

A: $P(A)=1$, $P(A^c)=0$

$\lim_{n\to\infty} |Z_n(\omega)-Z(\omega)| = 0 \Rightarrow \lim_{n\to\infty} Z_n(\omega)=Z(\omega)$

3/5

Definition 5.2 (converge in probability, TBp. 178)

A sequence of random variables $\{Z_n : \Omega \to \mathbb{R}\}$ is said to **converge in probability** to a random variable $Z : \Omega \to \mathbb{R}$, and denoted as $Z_n \xrightarrow{P} Z$, if for any $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\{\omega \in \Omega : |Z_n(\omega) - Z(\omega)| < \epsilon\}\right) = 1.$$

A_n

$P(\cdot)$: $A_1, A_2, \ldots, A_n, \ldots \longrightarrow 1$

$\omega \quad \omega \quad \omega \quad \omega \quad \omega$

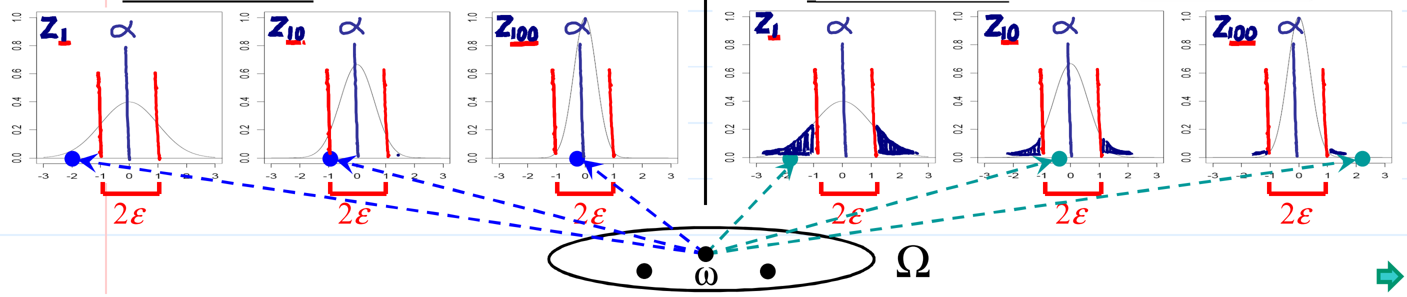
cont. r.v., $P(\{\omega\})=0$

cf.

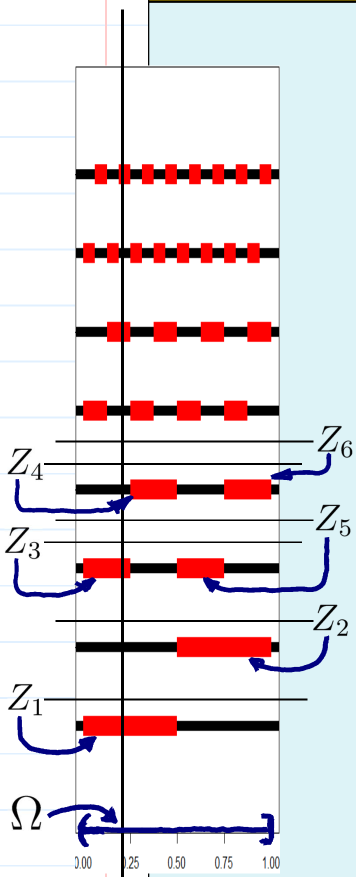
$Z=\alpha$

$Z_n \xrightarrow{a.s.} \alpha$, α : a constant. ($Z_n \xrightarrow{a.s.} Z \Leftrightarrow Y_n = Z_n - Z \xrightarrow{a.s.} 0$)

$Z_n \xrightarrow{P} \alpha$, α : a constant. ($Z_n \xrightarrow{P} Z \Leftrightarrow Y_n = Z_n - Z \xrightarrow{P} 0$)



Example 5.1 (convergence in probability need not imply convergence a.s.)



$\omega = 0.0011\cdots$ (binary decimal value)

- $\Omega = (0, 1]$
- P: uniform probability measure on Ω (i.e., $0<a<b<1$, $P((a,b])=b-a$)
- For $k = 1, 2, \ldots$, divide $(0, 1]$ into 2^k subintervals of equal length. These intervals are given by
  $$I_{k,j} = \left(\frac{j-1}{2^k}, \frac{j}{2^k}\right] \rightarrow P(I_{k,j}) = \frac{1}{2^k}$$
  for $j = 1, 2, \ldots, 2^k$.
- Let $Z_1, Z_2, \cdots : \Omega \to \mathbb{R}$ be a sequence of r.v.'s defined as follows:
  $$Z_n(\omega) = \begin{cases} 1, & \text{if } \omega \in I_{k,j}, \\ 0, & \text{if } \omega \notin I_{k,j}, \end{cases}$$
  where $n = 2^k + j - 2$.
- $Z : \Omega \to \mathbb{R}$ such that $Z(\omega) = 0$ for $\omega \in \Omega$
- $Z_n \xrightarrow{P} Z$, because for $0 < \epsilon < 1$,
  $$P\left(\{\omega \in \Omega : |Z_n(\omega) - Z(\omega)| < \epsilon\}\right) = 1 - \frac{1}{2^k} \to 1.$$
- Z_n not converge to Z a.s. because
  $$P\left(\{\omega \in \Omega : Z_n(\omega) \to Z(\omega)\}\right) = P(\emptyset) = 0.$$

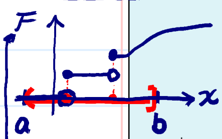
Definition 5.3 (converge in distribution, TBp. 181) ← why not mention Ω or ω? or $Z_n - Z$?

can be generalized to multivariate case: cdf → joint cdf, $\mathbf{z} \to \mathbf{Z}$

Let $Z_1, Z_2, \ldots$ be a sequence of random variables with cdf's $F_1, F_2, \ldots$ and let Z be a random variable with cdf F. Then Z_n **converges in distribution** to Z, denoted as $Z_n \xrightarrow{d} Z$, if

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

at every point z where F is continuous.



a.b: cont. pts of F: $P(Z_n \in (a,b]) \to P(Z \in (a,b])$

need no information about the sample space & how r.v.'s map to $\mathbb{R}$.

Recall $X = Y \underset{\not\Leftarrow}{\Rightarrow} X \stackrel{d}{=} Y$ $(F_X = F_Y)$

# of discontinuous pts of F: countably many

Theorem 5.1 (some properties about the 3 types of convergence)

1. $Z_n \xrightarrow{\text{a.s.}} Z \;\Rightarrow\; Z_n \xrightarrow{P} Z$ ← intuition
2. $Z_n \xrightarrow{P} Z \;\Rightarrow\; Z_n \xrightarrow{d} Z$ ← intuition
3. $Z_n \xrightarrow{d} c$, c: a constant $\Rightarrow Z_n \xrightarrow{P} c$ ← intuition

a.s. ⇒ in P. ⇒ in dist. ; if $Z = c$

converge in quadratic mean: $E(Z_n - Z)^2 \to 0$ (Roussas, 1997, CH8)

cdf: 1, 0; c; 2ε

$P(-\varepsilon < Z_n - c \le \varepsilon) = F_n(c+\varepsilon) - F_n(c-\varepsilon)$

4. (**convergence of transformation**) Let $g : \mathbb{R}^k \mapsto \mathbb{R}$ be a continuous function.

g; $Z_n(\omega) \to Z(\omega)$

exchange of g & lim. ← intuition

(a) $Z_n^{(j)} \xrightarrow{\text{a.s.}} Z^{(j)}$, $j = 1, \ldots, k \;\Rightarrow\; g(Z_n^{(1)}, \ldots, Z_n^{(k)}) \xrightarrow{\text{a.s.}} g(Z^{(1)}, \ldots, Z^{(k)})$.

(b) $Z_n^{(j)} \xrightarrow{P} Z^{(j)}$, $j = 1, \ldots, k \;\Rightarrow\; g(Z_n^{(1)}, \ldots, Z_n^{(k)}) \xrightarrow{P} g(Z^{(1)}, \ldots, Z^{(k)})$.

(c) $(Z_n^{(1)}, \ldots, Z_n^{(k)}) \xrightarrow{d} (Z^{(1)}, \ldots, Z^{(k)}) \Rightarrow g(Z_n^{(1)}, \ldots, Z_n^{(k)}) \xrightarrow{d} g(Z^{(1)}, \ldots, Z^{(k)})$.

joint cdf

continuous mapping Thm

By Thm 5.1, item 4, LNp.89

5. (**Slutsky's theorem**) If $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} a$, where a is a constant, then $\Omega \supset A_n = \{Y_n \le y\}$

① $y > a$. $A_n \to \Omega$
② $y < a$. $A_n \to \phi$

Then, $(X_n, Y_n) \xrightarrow{d} (X, a)$ (Ec) ← joint cdf

$P(X_n \le x, Y_n \le y)$

(a) $Y_n X_n \xrightarrow{d} aX$

(b) $X_n + Y_n \xrightarrow{d} X + a$

(c) $\dfrac{X_n}{Y_n} \xrightarrow{d} \dfrac{X}{a}$, provided that $P(Y_n \neq 0) = 1$ for all n and $a \neq 0$.

$X_n \xrightarrow{d} N(0,1)$
1. $Y_n = -X_n \xrightarrow{d} N(0,1)$, $X_n + Y_n = 0$
2. $Y_n = X_n \xrightarrow{d} N(0,1)$, $X_n + Y_n = 2X_n \xrightarrow{d} N(0,4)$

6. (**limit theorem for δ method**) Suppose

$$\text{cf. } \frac{X_n - \theta}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(X_n - \theta)}{\sigma} \xrightarrow{d} N(0,1).$$

standardization

$X_n \xrightarrow{P} \theta$; $E(X_n) \to \theta$; $Var(X_n)/\frac{\sigma^2}{n} \to 1$

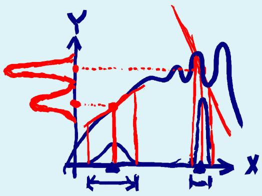
$Var(X_n) \approx \frac{\sigma^2}{n} (\to 0)$ when n is large.

For a given function g, suppose that $g'(\theta) \neq 0$ exists. Then

cf.

$$\frac{g(X_n) - g(\theta)}{(\sigma/\sqrt{n})\,|g'(\theta)|} = \frac{\sqrt{n}[g(X_n) - g(\theta)]}{\sigma |g'(\theta)|} \xrightarrow{d} N(0,1).$$

$\approx g(\theta)$; $\approx \frac{\sigma^2}{n} |g'(\theta)|^2$



By δ-method (LNp.56), $E[g(X_n)] \approx g(E(X_n))$; $Var[g(X_n)] \approx Var(X_n)[g'(EX_n)]^2$