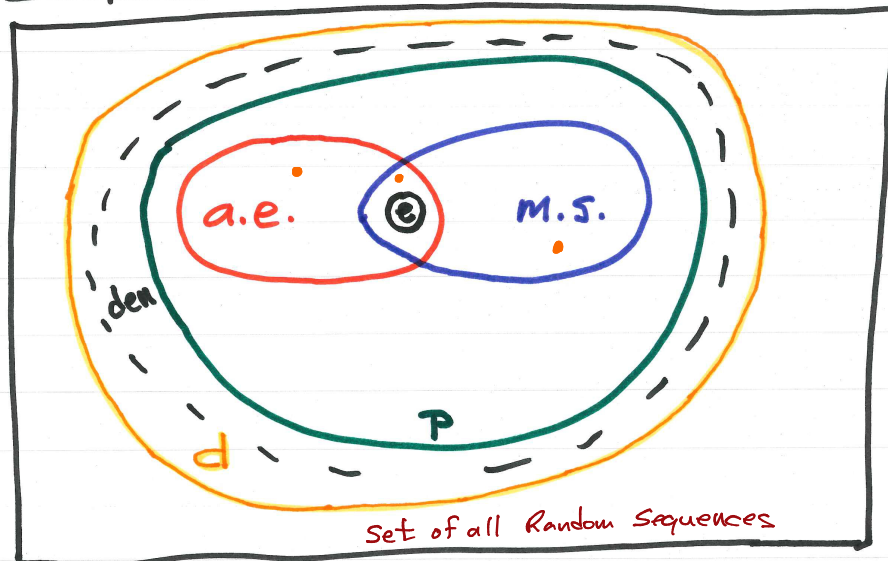


# Session 27

27.1

## Comparison of Modes of Convergence



1. M.S. convergence  $\Rightarrow$  convergence (p)

27.2

$$\text{n.b. } P(\{|X-\mu| \geq \varepsilon\}) \leq \frac{E[|X-\mu|^2]}{\varepsilon^2} = \frac{\text{Var}(X)}{\varepsilon^2}$$

$$\Rightarrow P(\{|X_n - X| \geq \varepsilon\}) \leq \frac{E[|X_n - X|^2]}{\varepsilon^2}$$

$$E[|X_n - X|^2] \xrightarrow{n \rightarrow \infty} 0 \quad (\text{m.s. convergence})$$

$$\Rightarrow P(\{|X_n - X| \geq \varepsilon\}) \xrightarrow{n \rightarrow \infty} 0 \quad (\text{convergence (p)})$$

2. convergence (a.e.)  $\Rightarrow$  convergence (p).

27.3

Follow from definitions

Converse is not true.

(See Papoulis).

3. Convergence (d) is "weaker" than convergence (a.e.), (m.s.) or (p).

$$(a.e.) \Rightarrow (d)$$

$$(m.s.) \Rightarrow (d)$$

$$(p) \Rightarrow (d).$$

4. (a.e.)  $\not\Rightarrow$  (m.s.)  
 (m.s.)  $\not\Rightarrow$  (a.e.)

27.4

n.b. The Chebyshev Inequality is an important tool when working with (m.s.) convergence.

Chebyshev Inequality: Let  $X$  be a R.V. with mean  $\mu$  and variance  $\sigma^2 < \infty$ . Then  $\forall \varepsilon > 0$ ,

$$P(\{|X - \mu| \geq \varepsilon\}) \leq \frac{\sigma^2}{\varepsilon^2}$$

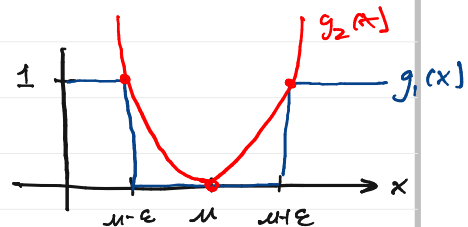
Proof: Let  $g_1(x) = \mathbb{1}_{\{x \in \mathbb{R} : |x - \mu| \geq \varepsilon\}}$

27.5

and  $g_2(x) = \frac{(x - \mu)^2}{\varepsilon^2}$

Note  $g_2(x) \geq g_1(x), \forall x \in \mathbb{R}$ .

Let  $\phi(x) \triangleq g_2(x) - g_1(x) \geq 0, \forall x \in \mathbb{R}$ .



$$\Rightarrow E[\phi(X)] \geq 0$$

$$\Rightarrow E[\phi(X)] = E[g_2(X) - g_1(X)] = E[g_2(X)] - E[g_1(X)] \geq 0$$

$$\Rightarrow E[g_1(X)] \leq E[g_2(X)]$$

$$\text{But } E[g_1(X)] = P(\{|X - \mu| \geq \varepsilon\})$$

$$E[g_2(X)] = \frac{\sigma^2}{\varepsilon^2}$$

$$\therefore P(\{|X - \mu| \geq \varepsilon\}) \leq \frac{\sigma^2}{\varepsilon^2}$$

## The Weak of Large Numbers

27.6

Let  $\{X_n\}$  be a sequence of i.i.d Random variables with mean  $\mu$  and variance  $\sigma^2$ .

Define

$$Y_n = \frac{1}{n} \sum_{k=1}^n X_k, \quad n=1, 2, 3, \dots$$

Then for any  $\varepsilon > 0$ ,

$$P\left(\left\{ \left| Y_n - \mu \right| \geq \varepsilon \right\}\right) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

$$\left( Y_n \xrightarrow{(P)} \mu \right).$$

Proof:  $E[Y_n] = E\left[\frac{1}{n} \sum_{k=1}^n X_k\right]$

27.7

$$= \frac{1}{n} \sum_{k=1}^n E[X_k] = \frac{1}{n} (n\mu) = \mu$$

and  $\text{var}(Y_n) = \dots = \frac{\sigma^2}{n}$  exercise (\*)

$$\begin{aligned} * \left[ \begin{aligned} E[Y_n^2] &= E[Y_n \cdot Y_n] = E\left[\left(\frac{1}{n} \sum_{j=1}^n X_j\right) \left(\frac{1}{n} \sum_{k=1}^n X_k\right)\right] \\ &= E\left[\frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n X_j X_k\right] = \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n E[X_j X_k] \\ &= \frac{1}{2} [n(\sigma^2 + \mu^2) + n(n-1)\mu^2] \\ &= \dots = \frac{\sigma^2}{n} + \mu^2 \\ \therefore \text{var}(Y_n) &= \frac{\sigma^2}{n} + \mu^2 - (\mu)^2 = \frac{\sigma^2}{n} \end{aligned} \right] \end{aligned}$$

Applying the Chebyshev Inequality  
to  $Y_n$ , we get

27.8

$$P(\{ |Y_n - \mu| \geq \varepsilon \}) \leq \frac{\text{var}(Y_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

$$\therefore P(\{ |Y_n - \mu| \geq \varepsilon \}) \leq \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0, \forall \varepsilon > 0$$

$$\Rightarrow Y_n \xrightarrow{(P)} \mu \text{ as } n \rightarrow \infty \blacksquare$$

27.8

Corollary: Suppose we repeat a simple  
experiment  $(\mathcal{S}_0, \mathcal{F}_0, P_0)$   
many times independently.

27.9

Let  $A \in \mathcal{F}_0$ , and define a RV

$$X_k \triangleq \mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

on the  $k$ -th repetition of the experiment

Note:  $X_1, \dots, X_n, \dots$  is an i.i.d. sequence  
of RVs with

$$E[X_k] = E[\mathbb{1}_A(\omega)] = P(A)$$

$$\text{var}(X_k) = \sigma^2 = P(A)(1 - P(A)).$$

27.9

Define

$$r_n(A) \triangleq \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$= \frac{1}{n} \sum_{k=1}^n X_k \quad \left. \vphantom{\sum_{k=1}^n X_k} \right\} \begin{array}{l} \text{relative frequency} \\ \text{with which } A \\ \text{occurs in } n \text{ trials} \end{array}$$

$$\begin{aligned} E[r_n(A)] &= E\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right] \\ &= \frac{1}{n} \underbrace{(P(A) + P(A) + \dots + P(A))}_{n \text{ times}} = P(A) \end{aligned}$$

Applying WLLN to  $r_n(A)$ , we get

$$P(\{ |r_n(A) - P(A)| \geq \varepsilon \}) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

for all  $\varepsilon > 0$  -  $(P\{|r_n(A) - P(A)| \geq \varepsilon\} \leq \frac{P(A)(1-P(A))}{n\varepsilon^2} \rightarrow 0 \text{ as } n \rightarrow \infty)$

(i.e.  $\text{var}(r_n(A)) = \frac{P(A)(1-P(A))}{n}$ .)

So this says that the relative frequency

$$r_n(A) = \frac{\left( \begin{array}{l} \text{number of times } A \text{ occurs} \\ \text{in } n \text{ independent trials} \end{array} \right)}{n}$$

converges in probability ( $P$ ) to the probability  $P(A)$ .

$$r_n(A) \xrightarrow{(P)} P(A) \text{ as } n \rightarrow \infty.$$

## Weak Law of Large Numbers

27.12

$$\bar{Y}_n \xrightarrow{(p)} \mu \text{ as } n \rightarrow \infty.$$

(You can show this even if the variance doesn't exist (i.e., is unbounded))  
- proof is harder.

There are also stronger forms of the Law of Large numbers

Weak: convergence (p)

Strong: convergence (a.e.)

## Strong Law of Large Numbers (Borel)

27.13

Let  $\{X_n\}$  be a sequence of identically distributed RVs with mean  $\mu$  and variance  $\sigma^2$ , and assume the RVs are uncorrelated:

$$E[(X_j - \mu)(X_k - \mu)] = 0, \quad j \neq k$$

(So the  $X_j$  are not necessarily independent.)

Then

$$\bar{Y}_n = \frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{(a.e.)} \mu \text{ as } n \rightarrow \infty$$

Proof: Beyond this course.

## The Central Limit Theorem

27.14

Let  $\{X_n\}$  be a sequence of i.i.d. RVs with mean  $\mu$  and variance  $\sigma^2$ .

Define  $Z_n \triangleq \frac{(X_1 + X_2 + \dots + X_n) - n\mu}{\sigma\sqrt{n}}$ ,  $n=1, 2, \dots$

Then  $\{Z_n\}$  converges in distribution to a RV  $Z$  that is Gaussian with mean 0 and variance 1.

i.e.,  $F_{Z_n}(z) \rightarrow \Phi(z)$  as  $n \rightarrow \infty$   
 $\forall z \in \mathbb{R}$ .

Sketch of proof: We will show that

27.15

$$\underbrace{\Phi_{Z_n}(w)}_{\text{char. fn. of } Z_n} \xrightarrow{n \rightarrow \infty} \underbrace{e^{-\frac{1}{2}w^2}}_{\text{char. fn. of } Z}, \quad \forall w \in \mathbb{R}$$

$$\begin{aligned} \Phi_{Z_n}(w) &= E[e^{iwZ_n}] \\ &= E\left[\exp\left\{\frac{iw}{\sigma\sqrt{n}} \sum_{k=1}^n (X_k - \mu)\right\}\right] \\ &= E\left[\prod_{k=1}^n e^{iw(X_k - \mu)/\sigma\sqrt{n}}\right] \\ &= \prod_{k=1}^n E\left[e^{iw(X_k - \mu)/\sigma\sqrt{n}}\right] = \dots \end{aligned}$$



$$= \left( E \left[ e^{i\omega(X-\mu)/\sigma\sqrt{n}} \right] \right)^n$$

27.16

15

We can expand this exponential as a power series (in  $\omega$  about  $\omega=0$ )



$$E \left[ e^{i\omega(X-\mu)/\sigma\sqrt{n}} \right]$$

$$= E \left[ 1 + \frac{i\omega(X-\mu)}{\sigma\sqrt{n}} + \frac{(i\omega)^2}{2n\sigma^2} (X-\mu)^2 + \underbrace{R(\omega)}_{\text{remainder}} \right]$$

$$= 1 + \frac{i\omega}{\sigma\sqrt{n}} E[X-\mu] + \frac{(i\omega)^2}{2n\sigma^2} E \left[ \frac{(X-\mu)^2}{\sigma^2} \right] + E[R(\omega)]$$

It can be shown that

$$\frac{E[R(\omega)]}{\omega^2/2n} \rightarrow 0 \text{ as } n \rightarrow \infty, \forall \omega \in \mathbb{R}$$

(It's a bit of work!)

Thus we have

27.17

16

$$E \left[ e^{i\omega(X-\mu)/\sigma\sqrt{n}} \right] = 1 - \frac{\omega^2}{2n} + E[R(\omega)]$$

$$\approx 1 - \frac{\omega^2}{2n}, \text{ as } n \rightarrow \infty.$$

$$\therefore \Phi_{Z_n}(\omega) \approx \left[ 1 - \frac{\omega^2}{2n} \right]^n \rightarrow e^{-\omega^2/2} \text{ as } n \rightarrow \infty$$

(Recall:  $(1 + \frac{x}{n})^n \rightarrow e^x$  as  $n \rightarrow \infty$ .)

$$\therefore \Phi_{Z_n}(\omega) \rightarrow e^{-\omega^2/2}, \forall \omega \in \mathbb{R} \quad \left[ 1 - \frac{\omega^2}{2n} \right]^n$$

$$\Rightarrow F_{Z_n}(z) \rightarrow \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-x^2/2} dx$$

as  $n \rightarrow \infty$ .

## General Forms of CLT:

We have assumed that  $\{X_n\}$  are i.i.d. for our proof, but a CLT will hold even if the  $\{X_n\}$  are not i.i.d. (i.e., they can be correlated and come from different distributions)

Very general form of CLT:

Lindeberg-Feller Central Limit Theorem

(See: P. Billingsley, Probability and Measure  
for conditions and proof. (Math/stat)  
538)