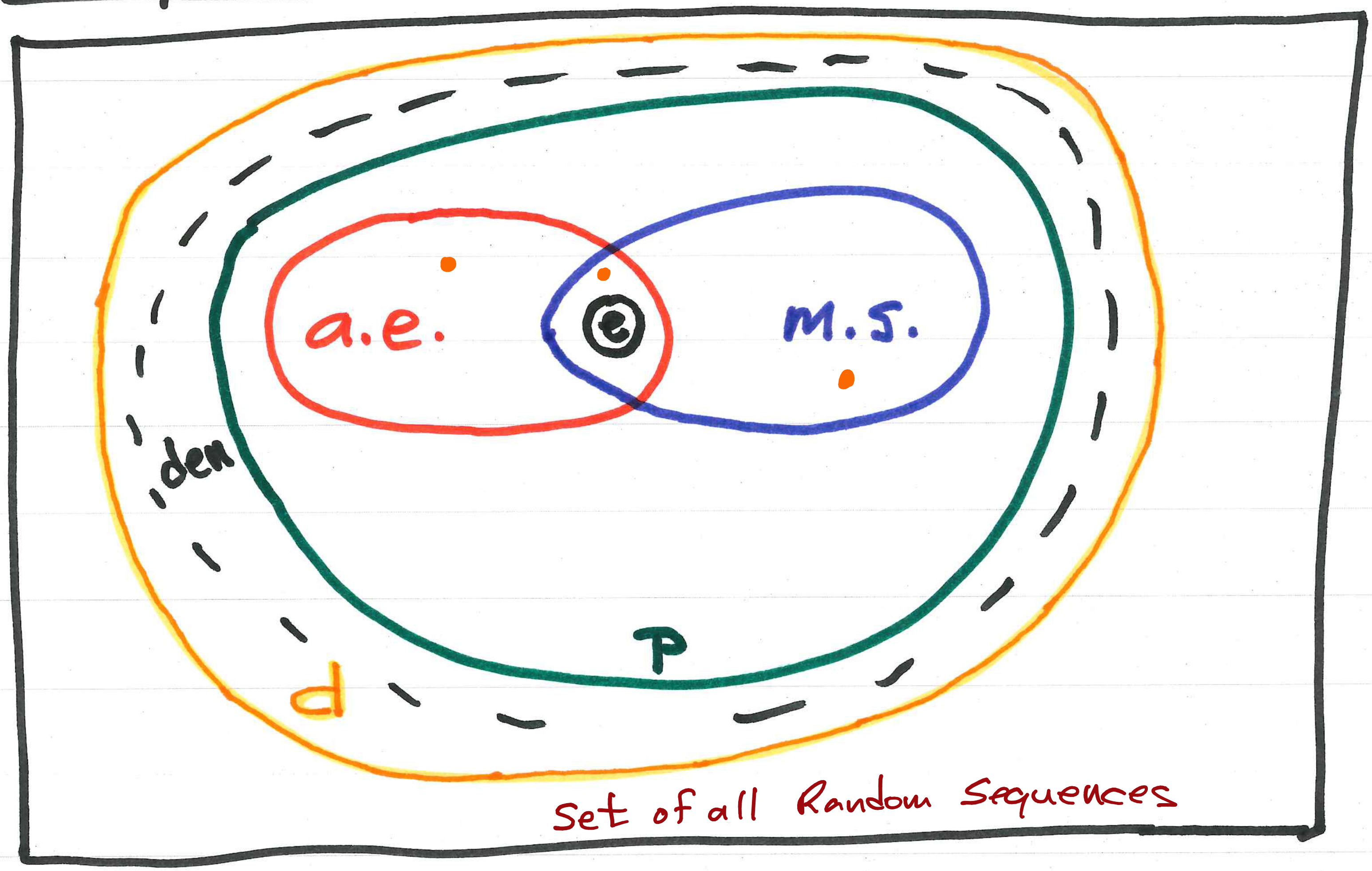
Session 27

Comparison of Modes of Convergence



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

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

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

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

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

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

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

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

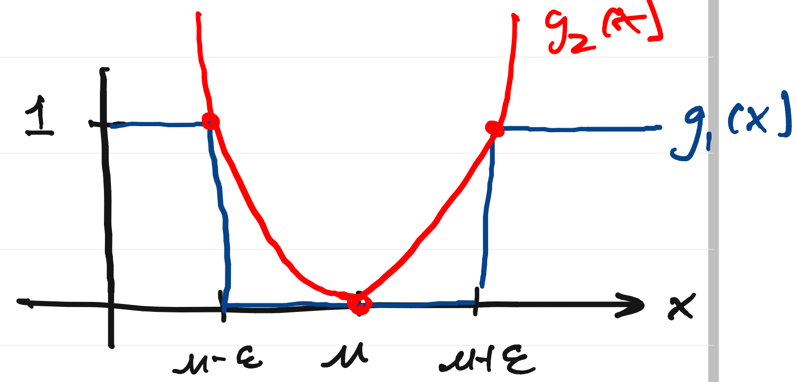
n.b. The byshev Inequality is an important tool when working with (m.s.) convergence.

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

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

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

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



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

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

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

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

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

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

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

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

The Weak of Large Numbers

27.6

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

Define

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

Then for any $\varepsilon > 0$,

$$P(\{|Y_n - \mu| \geq \varepsilon\}) \longrightarrow 0 \quad \text{as } n \to \infty .$$

$$(Y_n \xrightarrow{(p)} \mu).$$

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

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

and

$$\operatorname{var}(\bar{Y}_n) = \overset{\text{exercise}^{(*)}}{\cdots} = \frac{\sigma^2}{n}$$

* $E[\bar{Y}_n^2] = E[\bar{Y}_n \cdot \bar{Y}_n] = E\left[\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right)\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)\right]$

$$= E\left[\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} X_j X_k\right] = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} E[X_j \cdot X_k]$$

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

$$= \cdots = \frac{\sigma^2}{n} + \mu^2$$

$$\therefore \operatorname{Var}(\bar{Y}_n) = \frac{\sigma^2}{n} + \mu^2 - (\mu)^2 = \frac{\sigma^2}{n}$$

Applying the Chebyshev Inequality to $\bar{Y}_n$, we get

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

$$\therefore \quad P(\{|\bar{Y}_n - \mu| \geq \varepsilon\}) \leq \frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0, \quad \forall \varepsilon > 0$$

$$\Rightarrow \quad \bar{Y}_n \xrightarrow{(p)} \mu \quad \text{as} \quad n \to \infty \quad \blacksquare$$

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

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

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

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

Define

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

$$= \frac{1}{n}\sum_{k=1}^{n} X_k$$ } relative frequency with which A occurs in n trials

$$E[r_n(A)] = E\left[\frac{X_1 + X_2 + \cdots X_n}{n}\right]$$

$$= \frac{1}{n}(\underbrace{P(A) + P(A) + \cdots P(A)}_{n \text{ times}}) = P(A)$$

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

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

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

$\left(\text{n.b. } var(r_n(A)) = \frac{P(A)(1-P(A))}{n}.\right)$

So this says that the relative frequency

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

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

$$r_n(A) \xrightarrow{(p)} P(A) \quad \text{as } n \to \infty.$$

Weak Law of Large Numbers

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

(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)

Let $\{X_n\}$ be a sequence of identically distributed RVs with mean μ and variance σ^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

$$Y_n=\frac{1}{n}\sum_{k=1}^{n} X_k \xrightarrow{(a.e)} \mu \text{ as } n \to \infty$$

Proof: Beyond this course.

The Central Limit Theorem

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

Define

$$Z_n \triangleq \frac{(X_1 + X_2 + \cdots + X_n) - n\mu}{\sigma\sqrt{n}}, \quad n = 1, 2, \ldots .$$

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

i.e.,

$$F_{Z_n}(z) \longrightarrow \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx \quad \text{as } n \to \infty \quad \forall z \in \mathbb{R}.$$

Sketch of proof: We will show that

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

$$\Phi_{Z_n}(w) = E\left[e^{iwZ_n}\right]$$

$$= E\left[\exp\left\{\frac{iw}{\sigma\sqrt{n}}\sum_{k=1}^{n}(X_k - \mu)\right\}\right]$$

$$= E\left[\prod_{k=1}^{n} e^{iw(X_k-\mu)/\sigma\sqrt{n}}\right]$$

$$= \prod_{k=1}^{n} E\left[e^{iw(X_k-\mu)/\sigma\sqrt{n}}\right] = \cdots$$

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

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

$$\Downarrow$$

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

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

$$= 1 + \frac{i\omega}{\sigma\sqrt{n}}\cancelto{0}{E[X-\mu]} + \frac{(i\omega)^2}{2n\sigma^2}\underbrace{E[(X-\mu)^2]}_{\sigma^2} + E[R(\omega)]$$

It can be shown that

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

(It's a bit of work!)

Thus we have

$$E\left[e^{iw(\bar{X}-\mu)/\sigma\sqrt{n}}\right] = 1 - \frac{w^2}{2n\cancel{\sigma^2}}\cdot\cancel{\sigma^2} + E[R(w)]$$

$$\cong 1 - \frac{w^2}{2n}, \text{ as } n \to \infty.$$

$$\therefore \Phi_{Z_n}(w) \cong \left[1 - \frac{w^2}{2n}\right]^n \longrightarrow e^{-w^2/2} \text{ as } n \to \infty$$

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

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

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

as $n \to \infty$.

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)