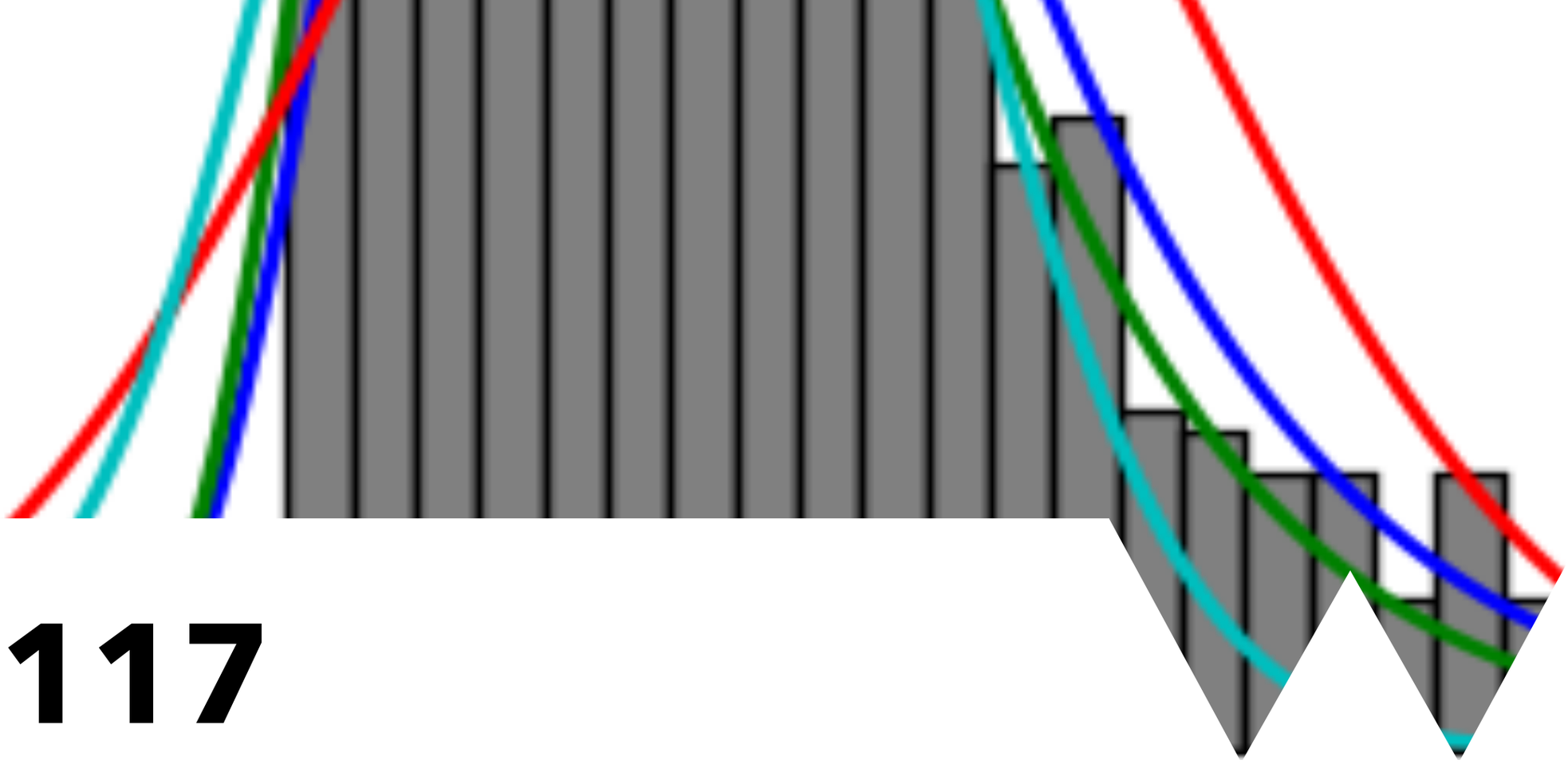

ST 117

3. Estimation

WARWICK

Lecture 10
(Week 4)

For this topic I will mostly use the visualiser or blackboard

Team work

SKILLS GAPS

Critical thinking

Managerial/supervisory

Communication/interpersonal

Leadership

Creativity/innovation

Proc. Improv./proj. managemt.

Managing/leading remotely

Technical skills

IT skills

---

https://www.stem.org.uk/resources/elibrary/resource/418157/top-ten-employability-skills

Team work in ST117: Pods!

- Homework pods and report pods
- Every student is expected to engage with **all question of each of the exercises sets**
- You meet to **present** your approaches for solutions to your pod fellows
- You **compare approaches** (critical appraisal)
- Plan **together which approaches to submit**
- **Each pod member needs to have engaged and understood** (at least at a minimal level) **all submitted work**
- **Practicalities: Has your pod communicated?**

Moodle page

- Please keep visiting as we add material
- Please watch for announcements
- Please bear in mind IT can fail
- New: timeline for first half of the lecture

	Due	Lectures	Lab	Posted	Individual tasks	Pod tasks
Term 2						
	A/E/WR Tue 1pm *Q* Wed	W1 Thu 11-12pm & 1-2pm & Fri 11-12pm W2+ Mon 1-2pm & Tue 1-3pm	W2+ Tue-Fri	A/E/W Tue Q Wed		*Hw Pod* = Homework Pod
1. Introduction & 2. R basics						
1		Teaching team, module assessment & org., tasters, syllabus R basics (covering material on cheat sheet): getting started, data types & structures, predefined functions,		*A0*	*A0* Install R, R Studio Small exercises from lecture	
2	*A0* indiv.	data input & output, tables, graphics, constructing functions, controlling flow, wrapping up and looking ahead, finding resources	Practice R with your *A0* datasets		Small exercises from lecture Practice for *Q1*	
3. Estimation						
3	*Q1 (Wed)* indiv.	Review R for quiz, normal distribution, normal approximation, joint distributions	Practicing data processing in R with gapminder data set, Fibonacci sequence	*Q1* *Hw Pod (Wed)* *E1 (Thu)*	Practice for *Q1* *Start working on E1*	Meet your *Hw Pod*
4	*E1 (Fri)* *Hw Pod*	Sequences of random variables, survey sampling, introduction to estimation	Birthday problems (probability theory) in R	*A1*	*Continue working on E1* Prepare submission *E1* *Start A1*	Discuss and compare indiv. approaches to *E1* Prepare submission *E1*
5	*A1* indiv.	[tentative: MLE, other estimators, properties of estimators, model fit]	TBA	*E2*	*Submit AI* *Start working on E2*	Discuss and compare indiv. approaches to *E2*

Timeline

	Due	Lectures	Lab	Posted	Individual tasks	Pod tasks
Term 2						
	A/E/WR Tue 1pm *Q* Wed	W1 Thu 11-12pm & 1-2pm & Fri 11-12pm W2+ Mon 1-2pm & Tue 1-3pm	W2+ Tue-Fri	A/E/W Tue Q Wed		*Hw Pod* = Homework Pod
1. Introduction & 2. R basics						
1		Teaching team, module assessment & org., tasters, syllabus R basics (covering material on cheat sheet): getting started, data types & structures, predefined functions,		*A0*	*A0* Install R, R Studio Small exercises from lecture	
2	*A0* indiv.	data input & output, tables, graphics, constructing functions, controlling flow, wrapping up and looking ahead, finding resources	Practice R with your *A0* datasets		Small exercises from lecture Practice for *Q1*	
3. Estimation						
3	*Q1 (Wed)* indiv.	Review R for quiz, normal distribution, normal approximation, joint distributions	Practicing data processing in R with gapminder data set, Fibonacci sequence	*Q1* *Hw Pod (Wed)* *E1 (Thu)*	Practice for *Q1* *Start working on E1*	Meet your *Hw Pod*
4	*E1 (Fri)* *Hw Pod*	Sequences of random variables, survey sampling, introduction to estimation	Birthday problems (probability theory) in R	*A1*	*Continue working on E1* Prepare submission *E1* *Start A1*	Discuss and compare indiv. approaches to *E1* Prepare submission *E1*
5	*A1* indiv.	[tentative: MLE, other estimators, properties of estimators, model fit]	TBA	*E2*	*Submit AI* *Start working on E2*	Discuss and compare indiv. approaches to *E2*

The two key questions for Part 3

① What can we observe from (repetitions) of the (same) (random process) over and over again?

$X_1, X_2, X_3, \ldots$ RVs on prob. space Ω

Sequence $\sim P$

② Can we use the (observations generated) by a (random process) to guess some aspects about how this process operates? And how can we (guess)

all X_i $(i \in \mathbb{N})$ $\sim$ $X \sim P \in (P_\vartheta)_{\vartheta \in (\Theta)}$

→ (infer which $\vartheta \in (\Theta)$ is optimal)

Example: Rolling a die

Rolling a die, observe

2 6 5 4 6 1 6 5 1 ⋯

4

4.33

$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ Sample mean

$X_1, X_2, \ldots, X_n$ Sample

Asymptotics, deriving law of large numbers

What happens for n large to $\bar{X}_n$?

$$\bar{X}_n \longrightarrow E[X] = \mu$$

$$E[\bar{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[X_i]}_{\mu} = \mu$$

$$Var(\bar{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right)$$

independent assumption

$$= \frac{\sigma^2}{n}$$

$$P(|\bar{X}_n - \mu| > \varepsilon) < \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n \cdot \varepsilon^2} \longrightarrow 0$$