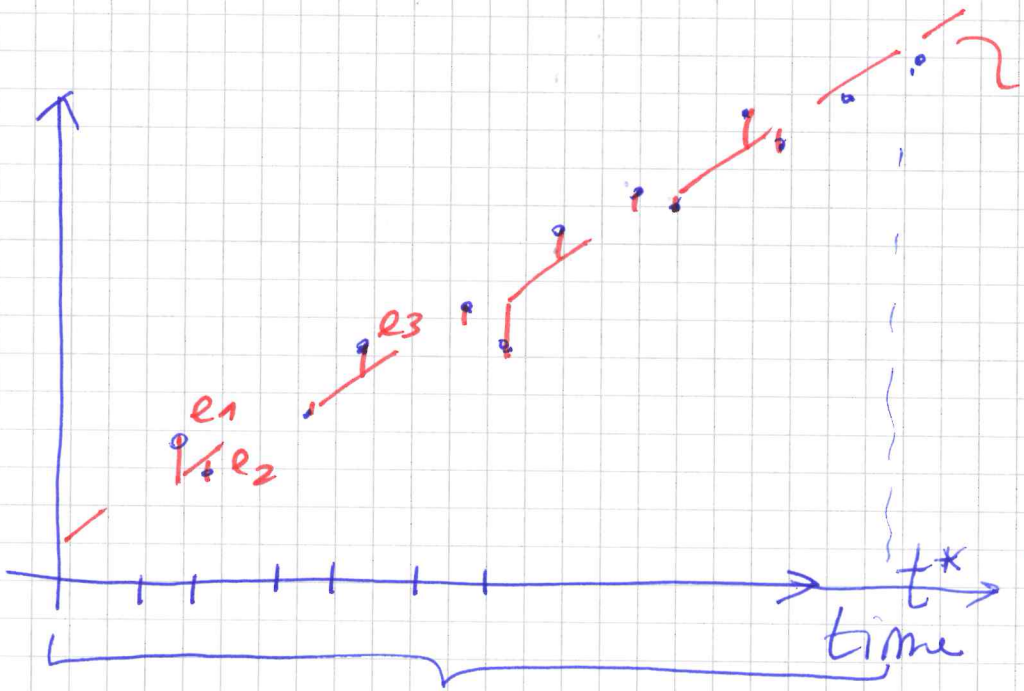
QDA 2023.05.17

SPC or SQM with NON iid data



$$\text{MODEL} = b_0 + b_1 t + \hat{\varepsilon}$$

$$\downarrow$$

$$e = \hat{\varepsilon} \overset{iid}{\sim} N(0, \hat{\sigma}_\varepsilon^2)$$

$$\uparrow$$

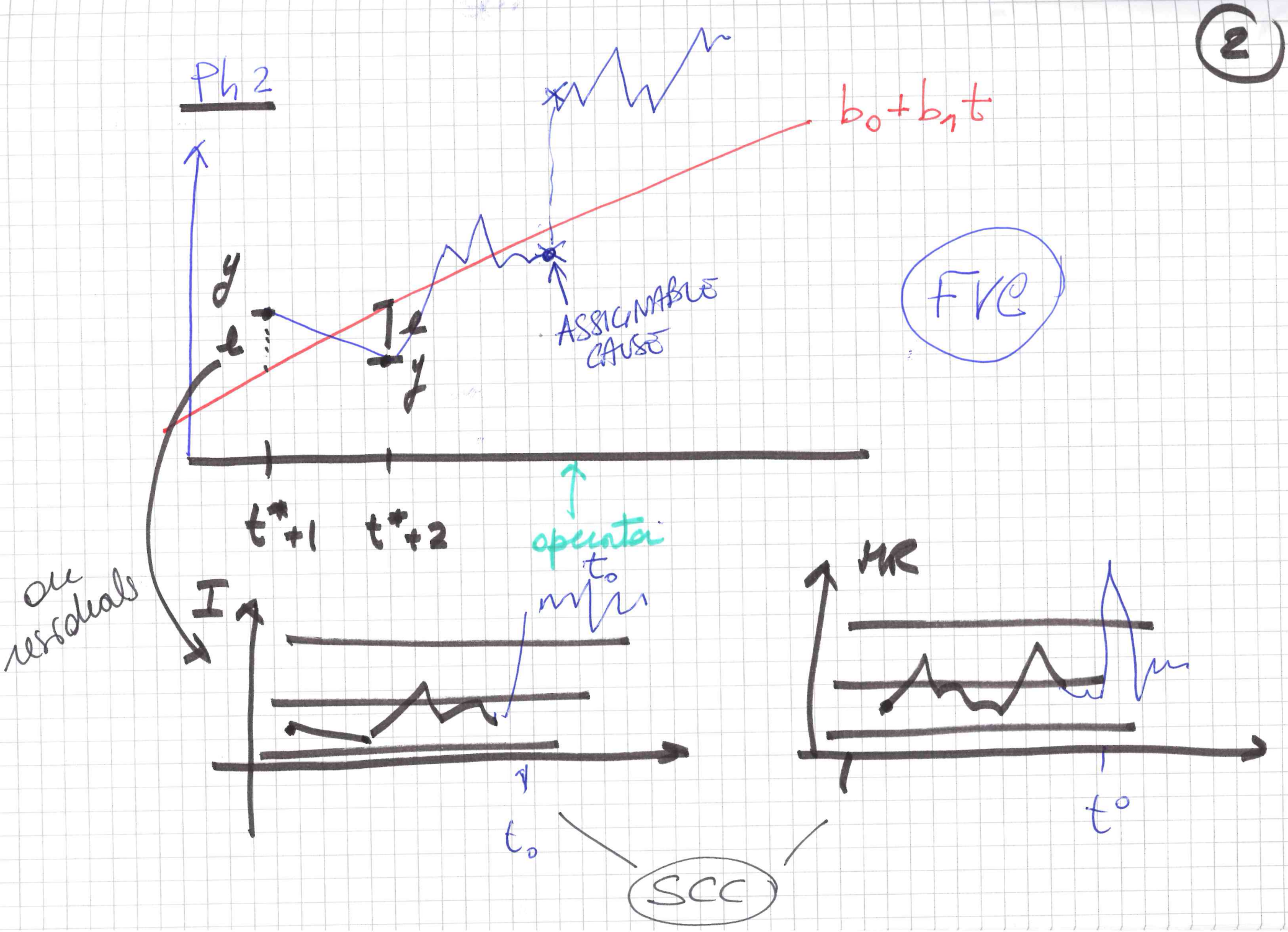
$$\hat{\sigma}_\varepsilon = \frac{\overline{MR}_e}{d_2(2)}$$

$e \rightarrow$ I-MR control chart

Ph 1

↓

MODEL (complete, with all the unknown parameters estimated) for the IN CONTROL DATA


Ph 2
$b_0 + b_1 t$
y
t
ASSIGNABLE CAUSE
FVC
t^*_{+1}
t^*_{+2}
operator
t_0
our residuals
I
MR
y
t_0
t°
SCC

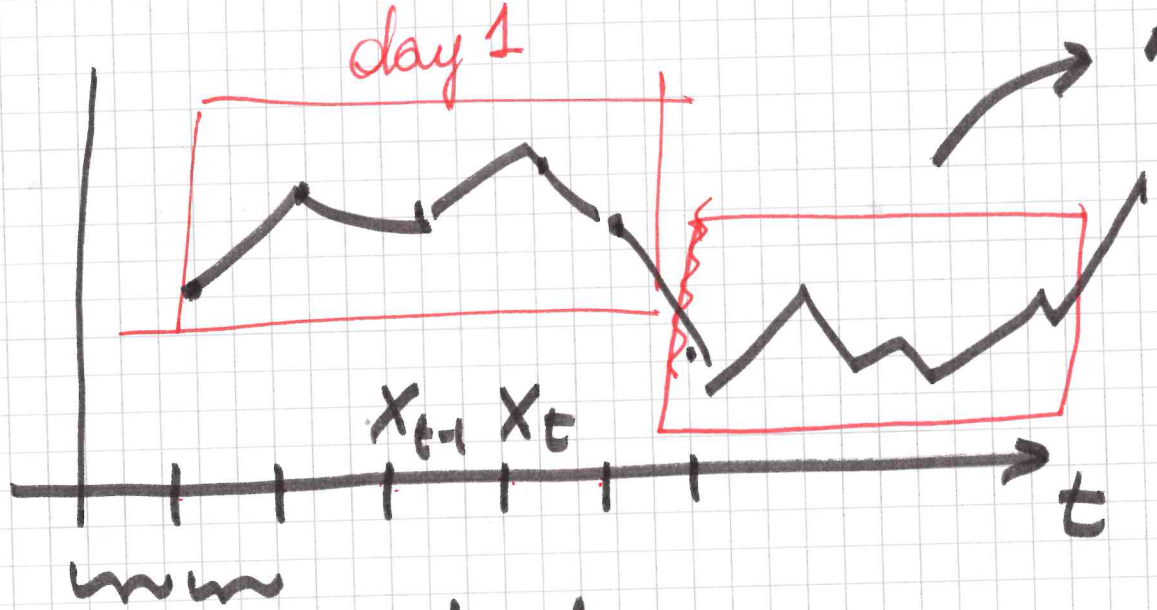
→ SAMPLING STRATEGY

(3)

$m > 1$

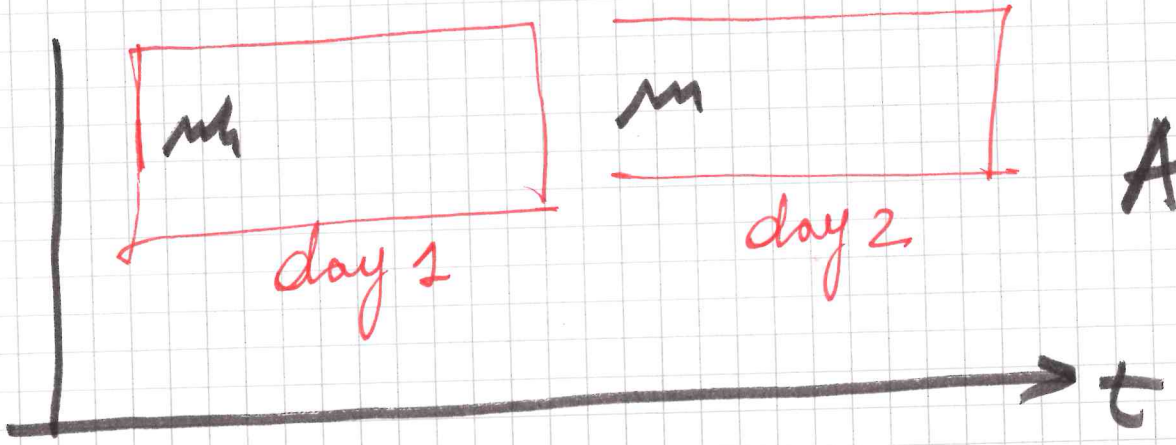
ex AR(1)

Ishikawa example

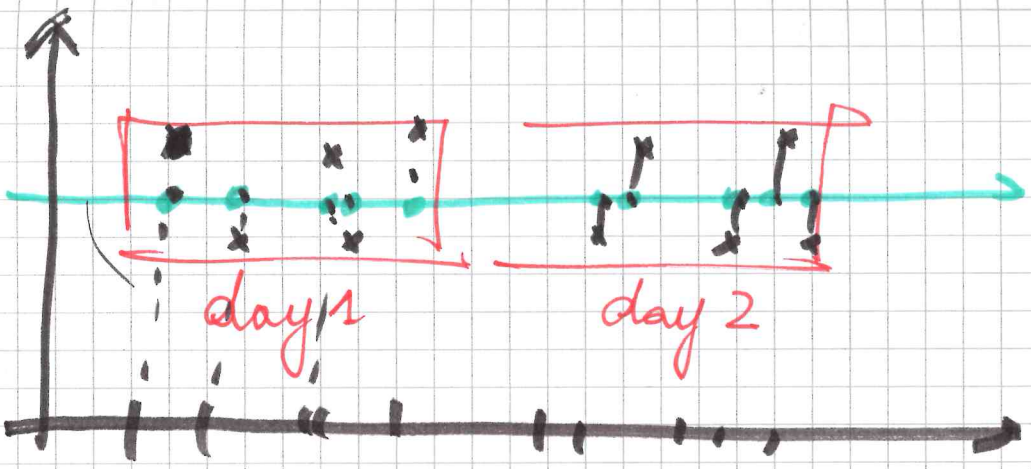


h constant

$$X_t = \beta_0 + \beta_1 X_{t-1} + \varepsilon_t$$



AR ?

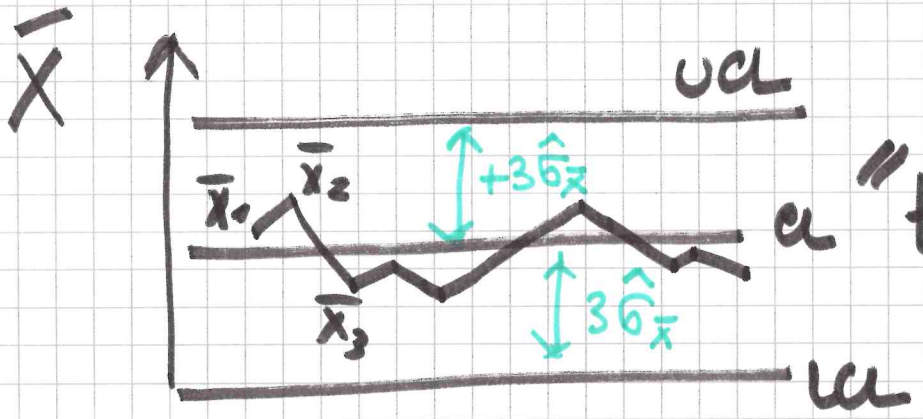


RND SAMPLING

~~iid~~ $X \sim N(\mu, \sigma^2)$ $n > 1$

VIOLATION of the IID assumption

batch $n = 5$ batch $n = 5$



a "HUGGING" or STRATIFICATION

it seems that the control chart on the $\bar{x}$ is overestimating the variability of $\bar{x}$

Remind that

$$X \overset{iid}{\sim} N(\mu, \sigma^2)$$

$$\Downarrow$$

$$\bar{X} = \frac{1}{n}\sum_i X_i \longrightarrow E(\bar{X}) = E\left(\frac{1}{n}\sum X_i\right) =$$

$$= \frac{1}{n}\sum \underbrace{E(X_i)}_{\mu} = \frac{1}{n}\, n\mu$$

$$\hat{\sigma}^2_{\bar{x}} \qquad = \mu \qquad \mu$$

ASSUMPTIONS IN $\bar{X}$ ee

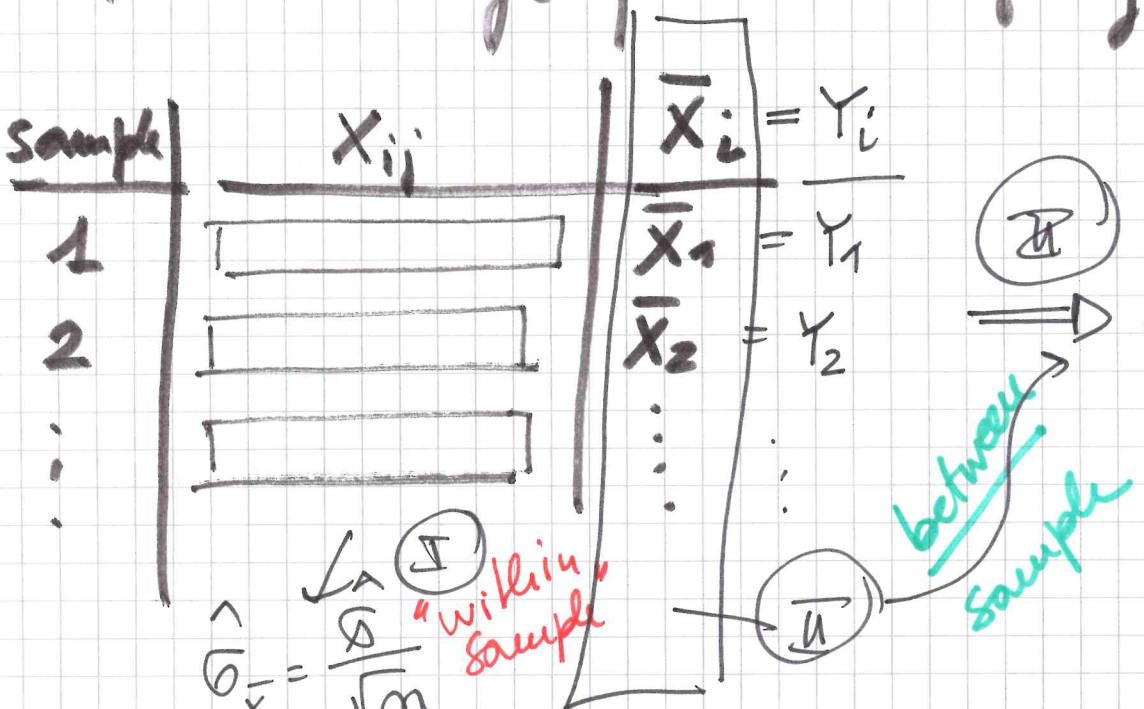
$$V(\bar{X}) = V\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2} V\left(\sum X_i\right)$$

ASSUMING NO COV UNCORRELATED

$$= \frac{1}{n^2}\sum V(X_i) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

If X_i are non iid $\Rightarrow$ $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$ in the $\bar{X}$ cc can be over- or under-estimating the real $\boxed{V(\bar{x})}$

If I suspect HUGGING or STRATIFICATION (or any other strange patterns of $\bar{X}_j$ in the Xbar cc)



$$\hat{\sigma}_{\bar{x}} = \frac{\bar{s}}{\sqrt{n}}$$

"within sample"

"within" sample

I Compare $\frac{\hat{\sigma}_i}{\sqrt{n}} = \hat{\sigma}_{\bar{x}}$ traditional

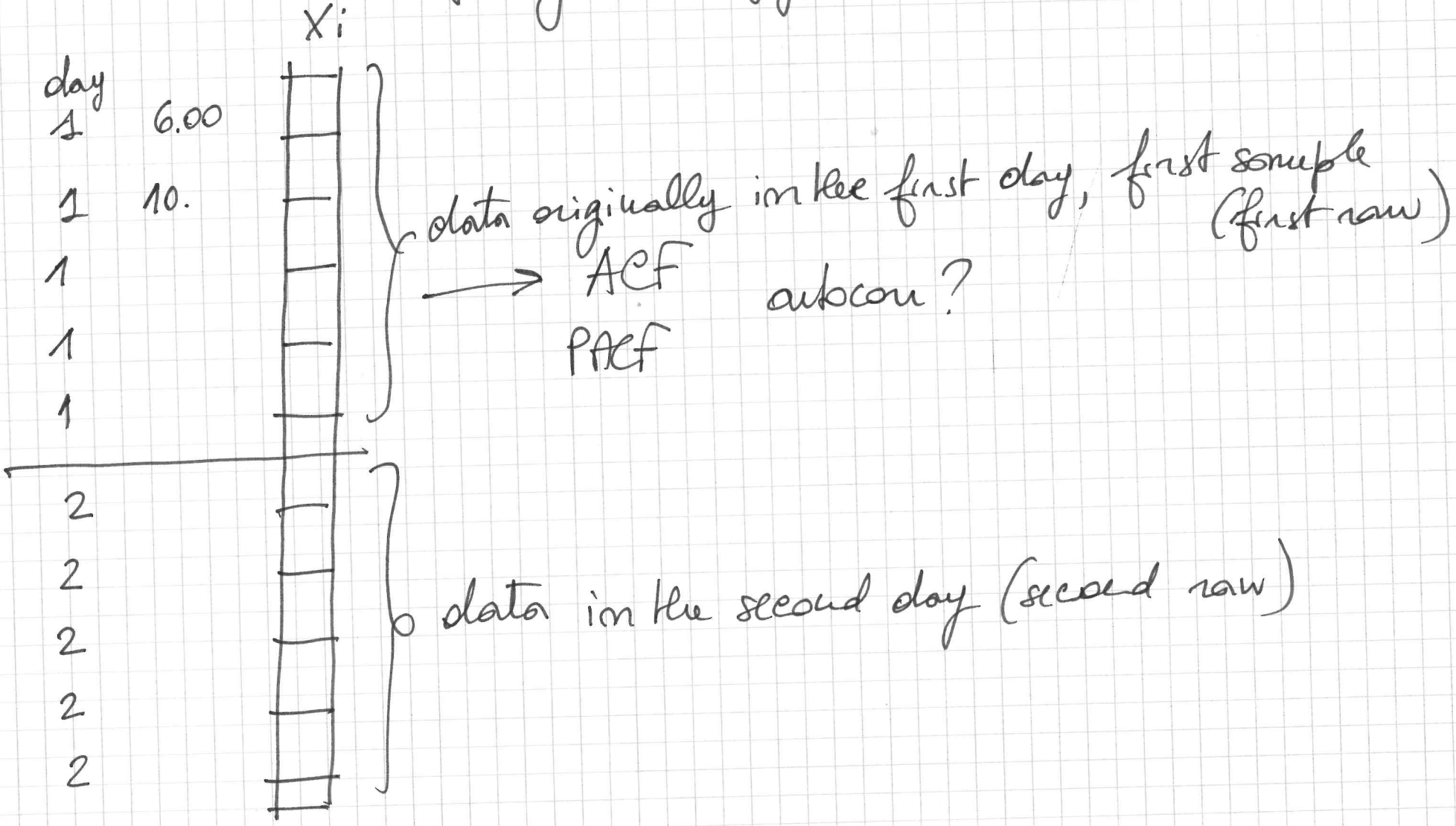
II $\Rightarrow$ Individual cc on Y_i (i.e. $\bar{X}_i$)

LONG TERM

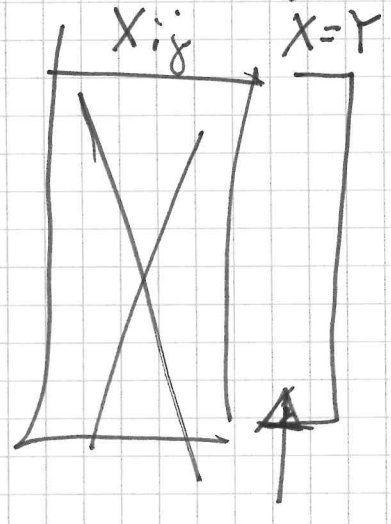
SHORT TERM

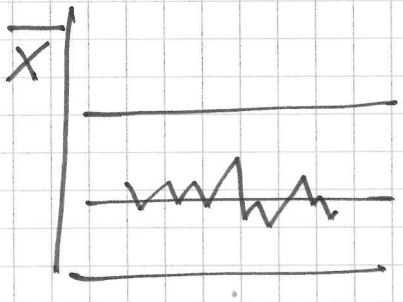
$$\bar{\bar{Y}} = \mu_Y \qquad \frac{\bar{s}}{c_4(m)} = \frac{\overline{MR}_Y}{d_2(2)} = \hat{\sigma}_Y$$

When the sampling strategy is structured

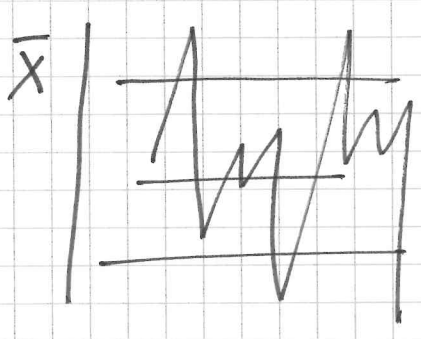


$n > 1$ → plots in $\bar{X}$ cc





HUGGING → suspect iid assumptions are violated



OVERDISPERSION

What can we do?

between sample — just take all the $\bar{X}$'s as a new sequence of quality indicator (i.e., a new random variable Y) and proceed with a I cc on this new Y's

pros: it is working?

cons: loosing information

try to model X_{ij}'s

batching / gapping *

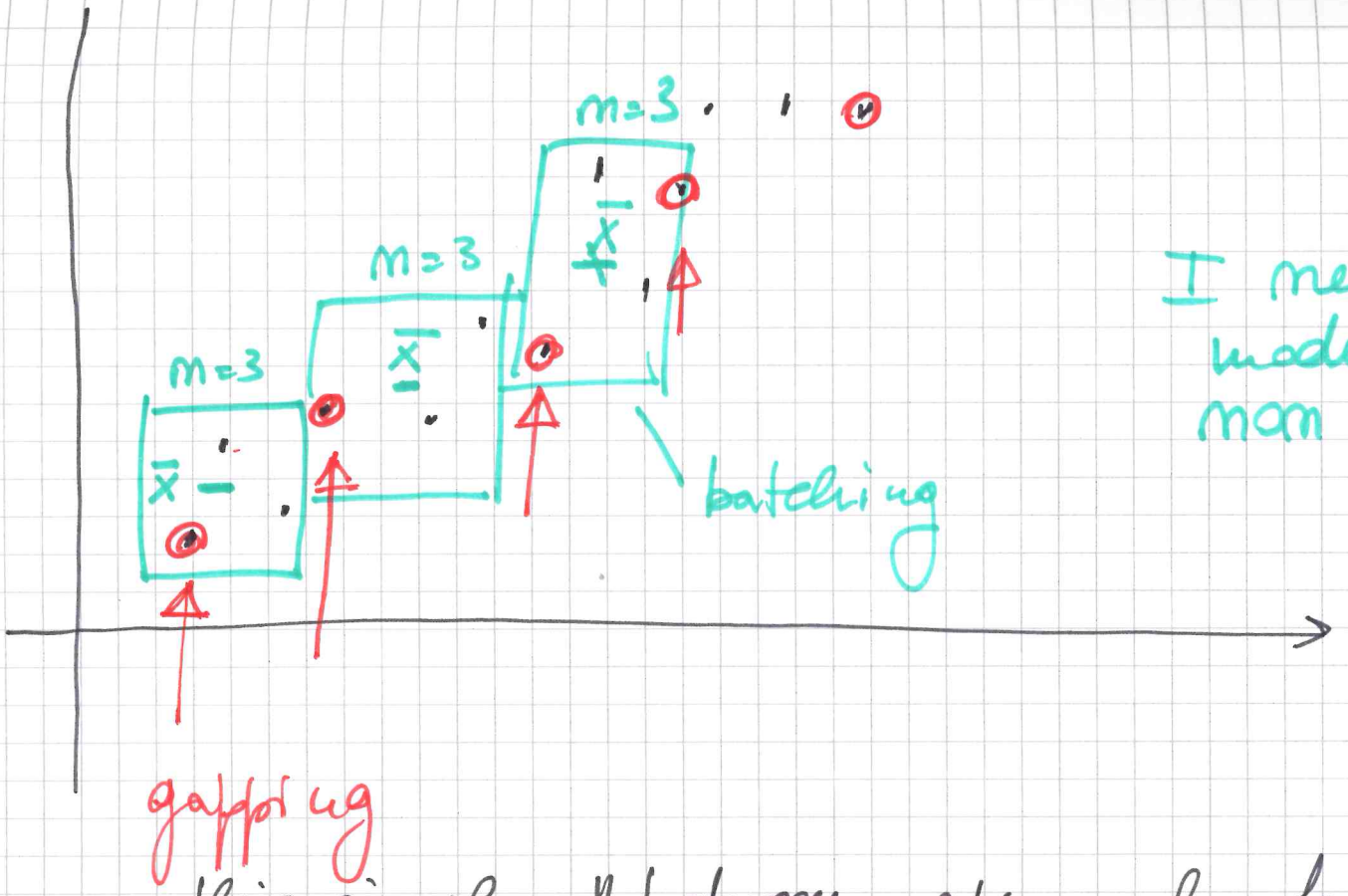
best model $X_{ij} = f(\overbrace{i, j, \text{dummy}, \text{trend}}^{\text{ARIMA}}) + \varepsilon$ ⇓ FYC + SCC

pros: better model for the data, without loosing info

cons: finding the right model can be a complex task

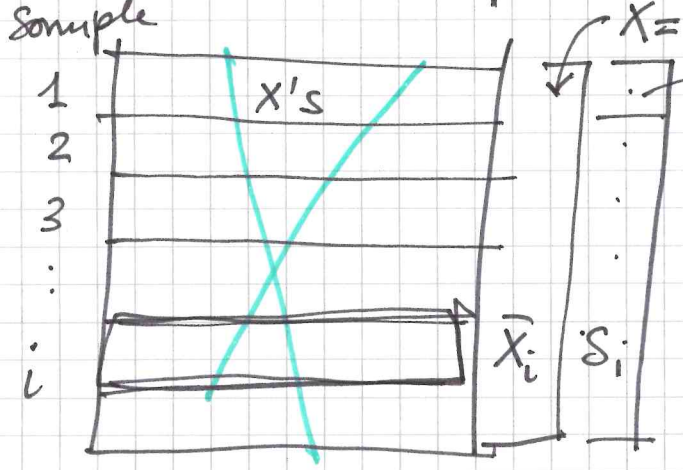
* a note: gapping or batching can work only for stationary processes

ex: trend = non stationary process



I need to model the non stationarity!

You can use this similar "between" approach for any statistic taken at each sample → Individual ce on Y



→ Individual ce on W
(Box Cox ?)