Lec-5 Time Series-4

- ARIMA model family:
  - AR recap
  - MA
  - ARMA
  - ARIMA
- SARIMA model
- Ranged estimates - Confidence Interval

SARIMA

Auto Regression (AR)

idea: Past values in the TS can be used to forecast feature.

p = order of AR

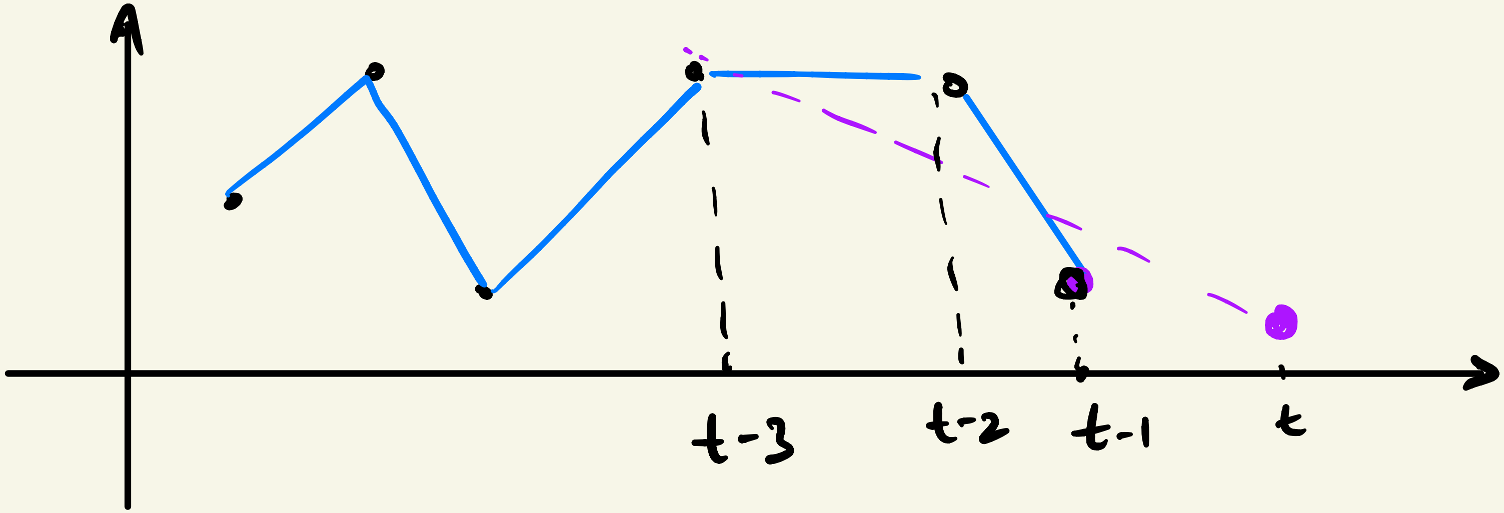
= no. of data points from past we are going to use

y_{t-3}	y_{t-2}	y_{t-1}	y_t	$\hat{y}_t$
			130	
			128	
			141	
130	128	141	150	144
128	141	150	161	159
			⋮	

Eg. $P = 3$

$$\hat{y}_t = \underbrace{f(y_{t-1}, y_{t-2}, y_{t-3})}_{\text{we use linear regression model}}$$

$$\hat{y}_t = a_0 + a_1 \, y_{t-1} + a_2 \, y_{t-2} + a_3 \, y_{t-3} \quad \text{for } P = 3$$



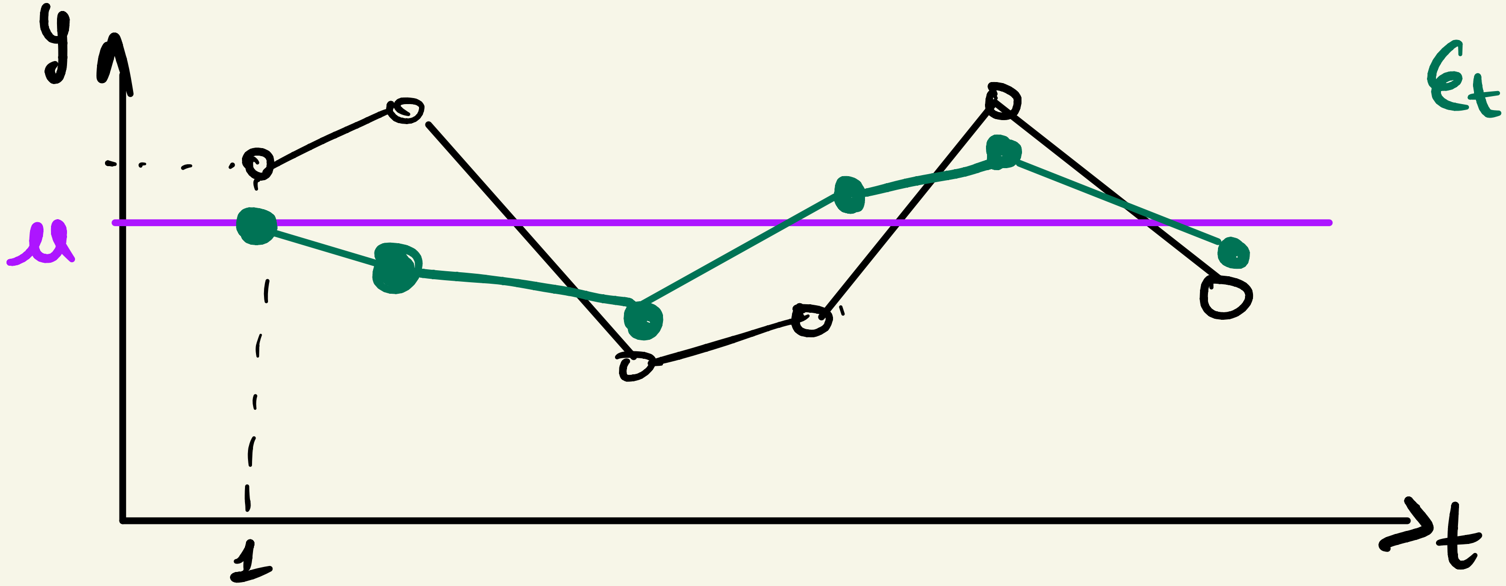
$P \rightarrow$ ACF

↳ Hyperparameter tuning

Moving Averages (MA model)

→ Nothing to do with the previously discussed moving/rolling average

Idea: Can we use the errors in the previous predictions to forecast the future values.

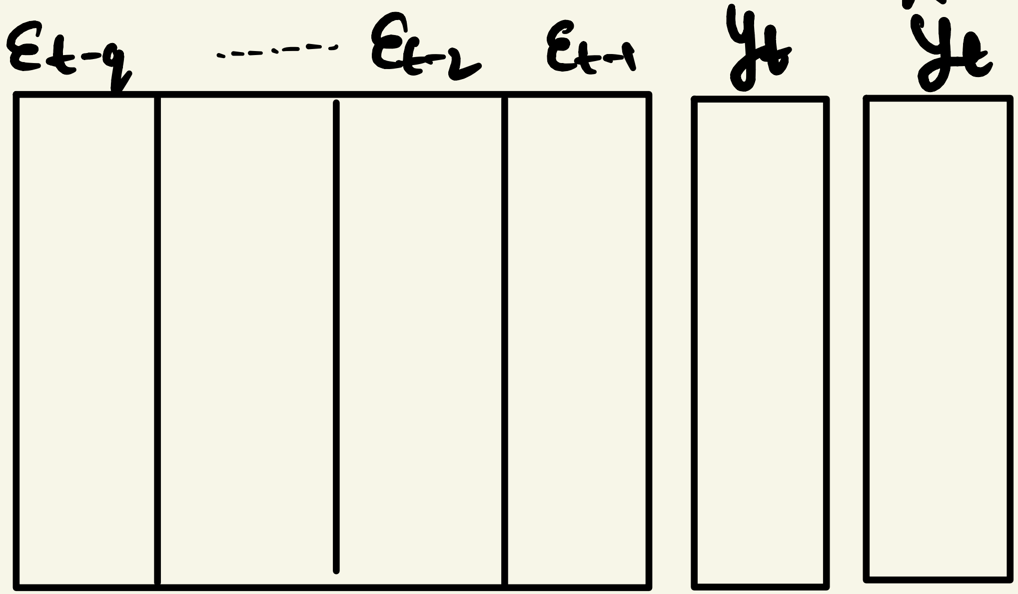


$$\epsilon_t = y_t - \hat{y}_t$$

$$\hat{y}_t = m_1 \varepsilon_{t-1} + m_2 \varepsilon_{t-2} + m_3 \varepsilon_{t-3} + \mu$$

q = order of MA = # of past errors term used to generate a forecast.

m_1, m_2, m_3 are weights



Eg. $q = 1$

$\hat{y}_1 = \mu$ (mean of TS)

$\varepsilon_1 = y_1 - \hat{y}_1 = y_1 - \mu$

$\hat{y}_2 = \mu + m_1 \varepsilon_1 \Rightarrow \varepsilon_2 = y_2 - \hat{y}$

$\hat{y}_3 = \mu + m_1 \varepsilon_2$

ARMA

p = order of AR q = order of MA

$$\hat{y}_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_p y_{t-p} +$$
$$m_1 \varepsilon_{t-1} + m_2 \varepsilon_{t-2} + \dots + m_q \varepsilon_{t-q} + c$$

ARIMA model

I : integrated

$y \rightarrow y'(t) \longrightarrow y(t)$

to make the series stationary

$d = 1$

$d =$ order of integration

$=$ # of times differencing and integration

Eg.

$$y \rightarrow y' \rightarrow y'' \rightarrow y' \rightarrow y \quad d = 2$$

not Stationary (under y')

Stationary (under y'')

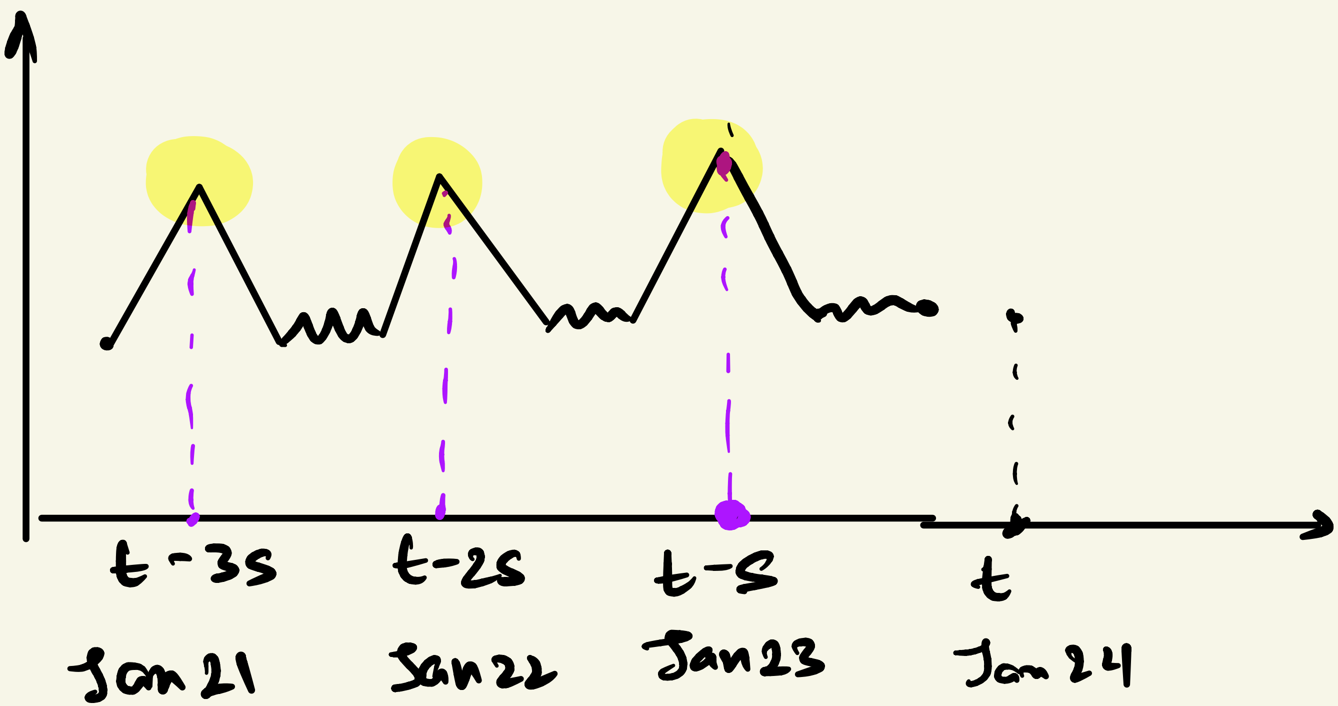
SARIMA model

idea: incorporate seasonality

*Seasonal model

s = period of seasonality

P = order of seasonal model



Eg: $p = 3$

$$\hat{y}_t = \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} + \gamma_3 y_{t-3s}$$

In general,

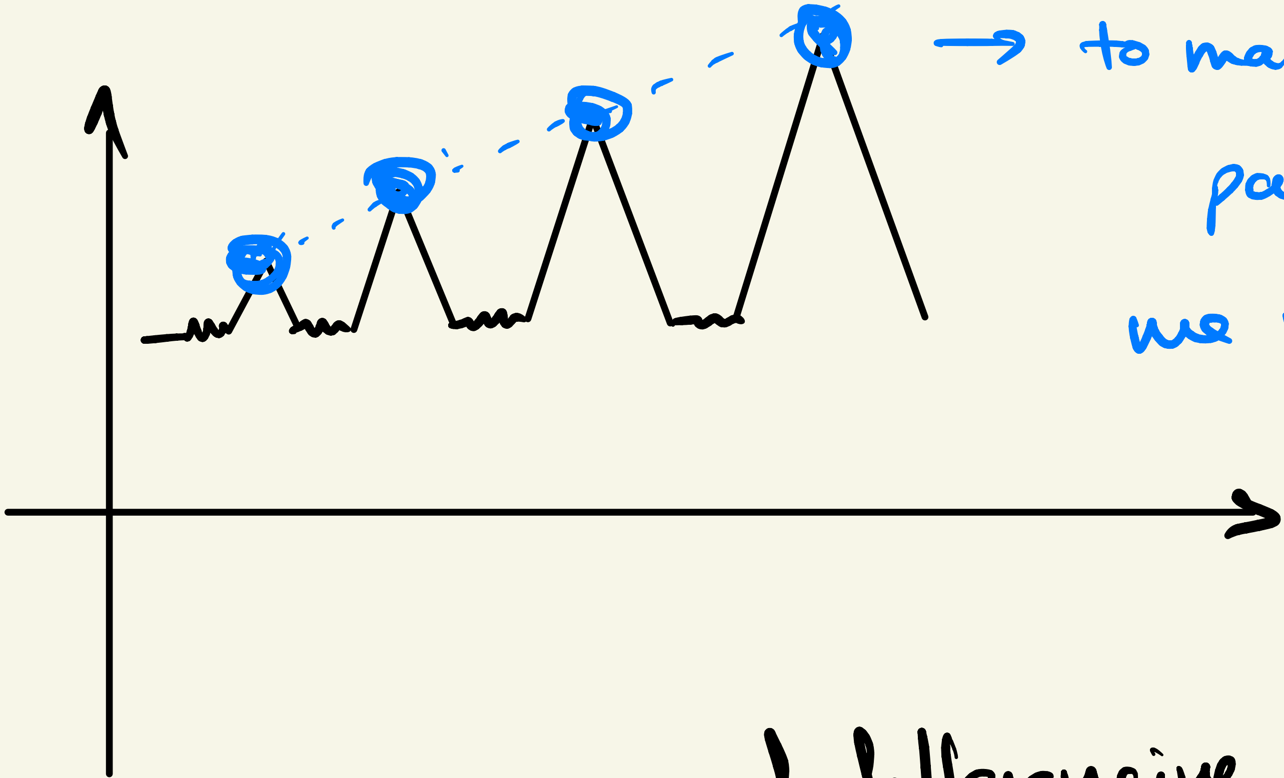
$$\hat{y}_t = \gamma_1 y_{t-s} + \gamma_2 y_{t-2s} + \cdots + \gamma_p y_{t-p \times s}$$

Seasonal Moving Average (SMA)

$$\hat{y}_t = \beta_1 \varepsilon_{t-s} + \beta_2 \varepsilon_{t-2s} + \beta_3 \varepsilon_{t-3s}$$
$$+ \cdots + \beta_Q \varepsilon_{t-Qs}$$

Q = order of SMA = # of past seasonal errors to consider

Eg. $Q = 3 \rightarrow$



→ to make the Seasonal pattern stationary we need to do differencing

D = Seasonal differencing

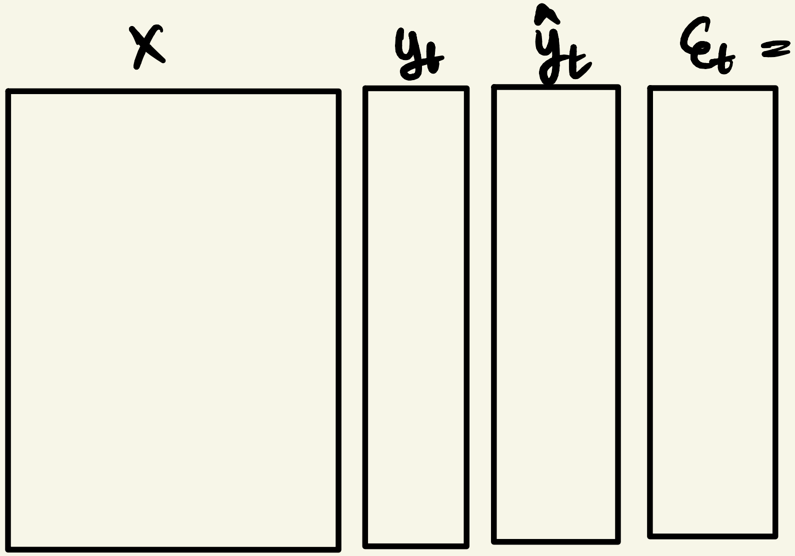
$$\underline{\text{SARIMA}} \rightarrow \underbrace{(p, d, q)}_{\text{order}} + \underbrace{(P, D, Q) + S}_{\text{seasonal order}}$$

$$\hat{y}_t = \underbrace{\text{AR terms}}_{p} + \underbrace{\text{MA terms}}_{q}$$

$$+ \underbrace{\text{AR-seasonal}}_{P} + \underbrace{\text{MA-seasonal}}_{Q}$$

$$+ \underbrace{\text{diff}}_{d} + \underbrace{\text{seasonal-diff}}_{D}$$

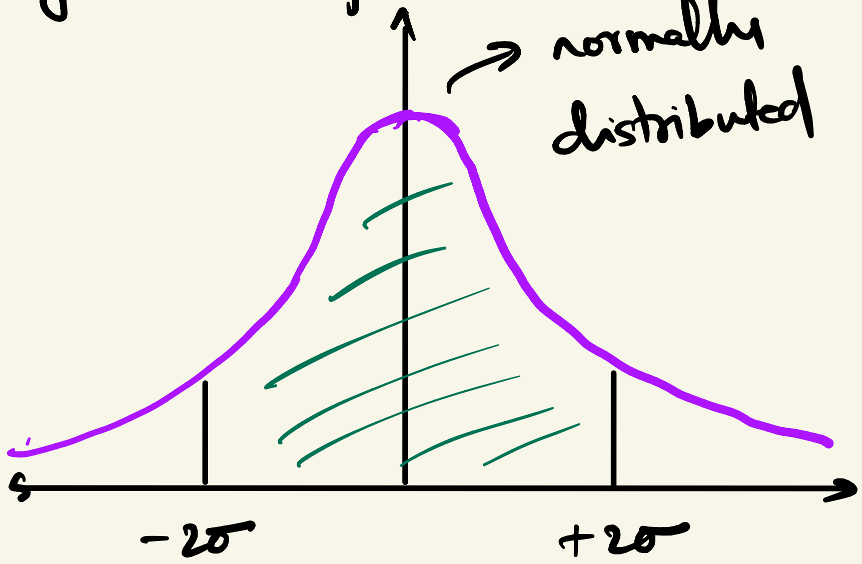
Confidence Intervals

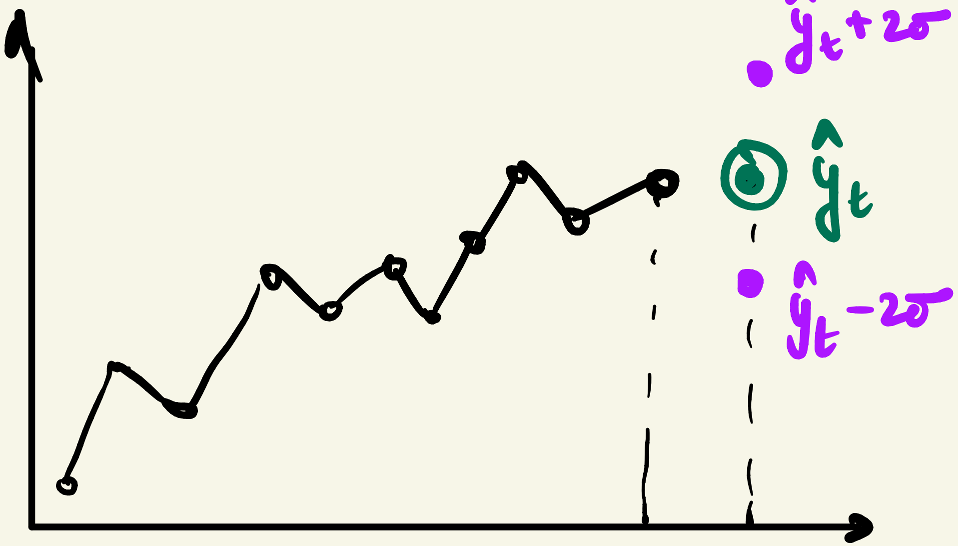


residuals

histogram

normally distributed



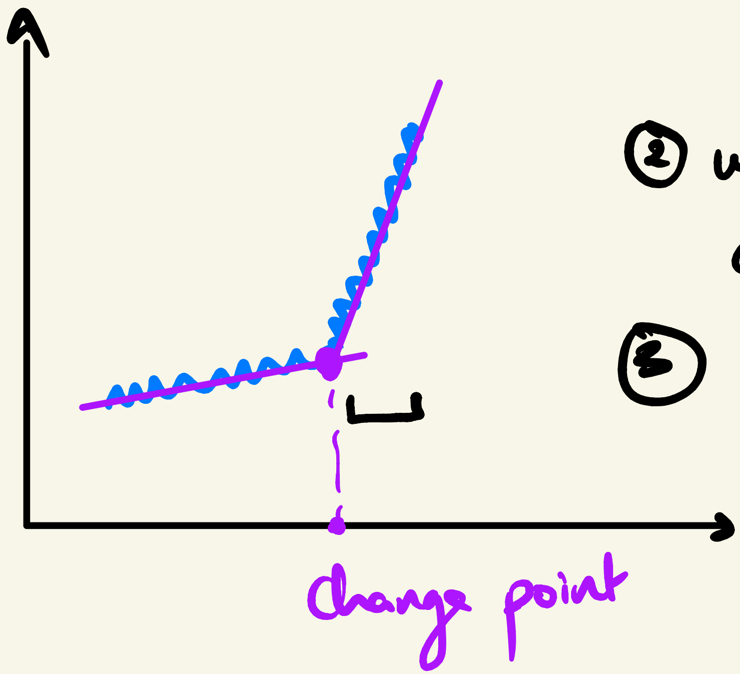


95% CI

$$\longrightarrow [\hat{y}_t - 2\sigma, \hat{y}_t + 2\sigma]$$

Change point

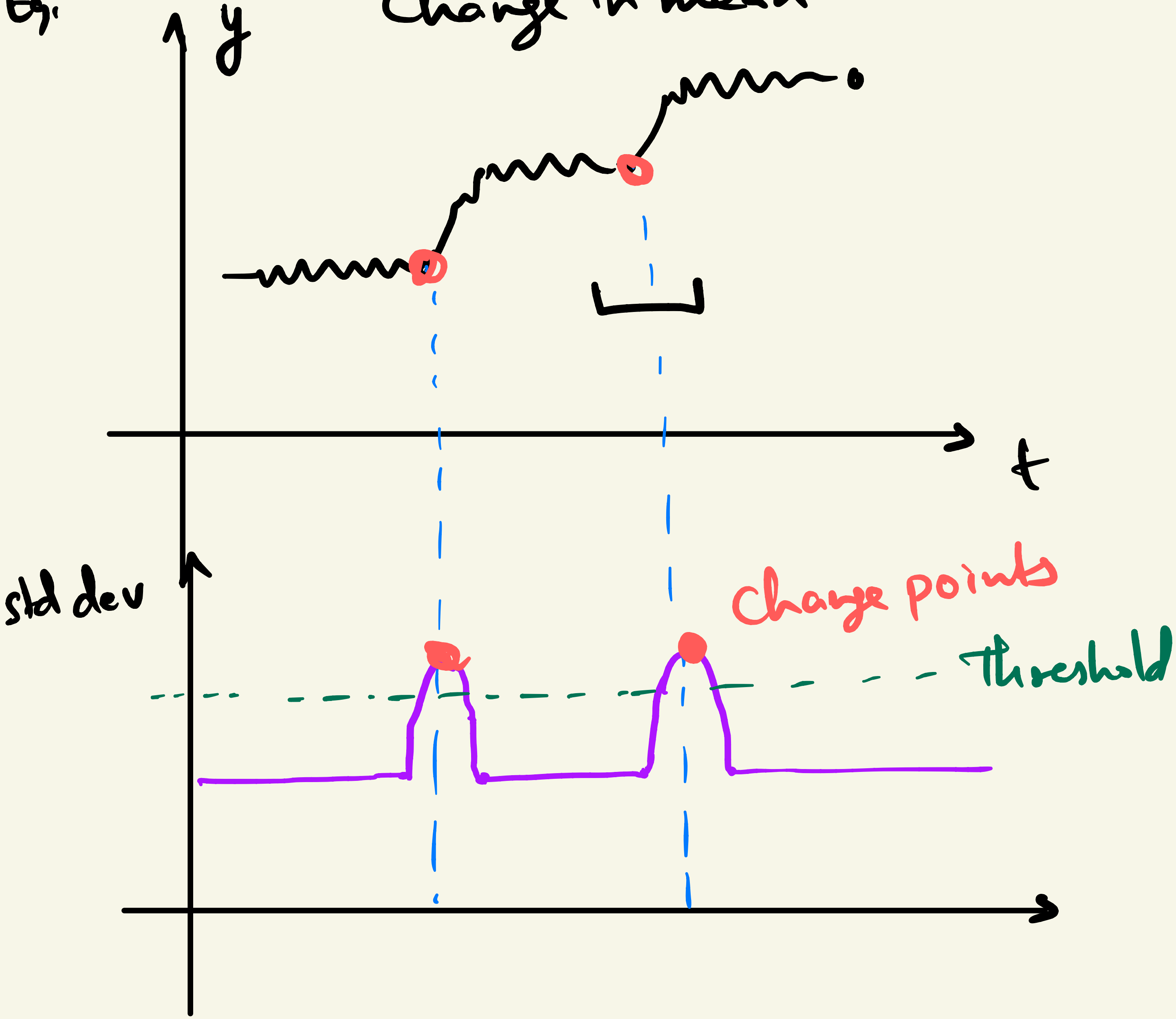
idea: How do you detect change in trend?



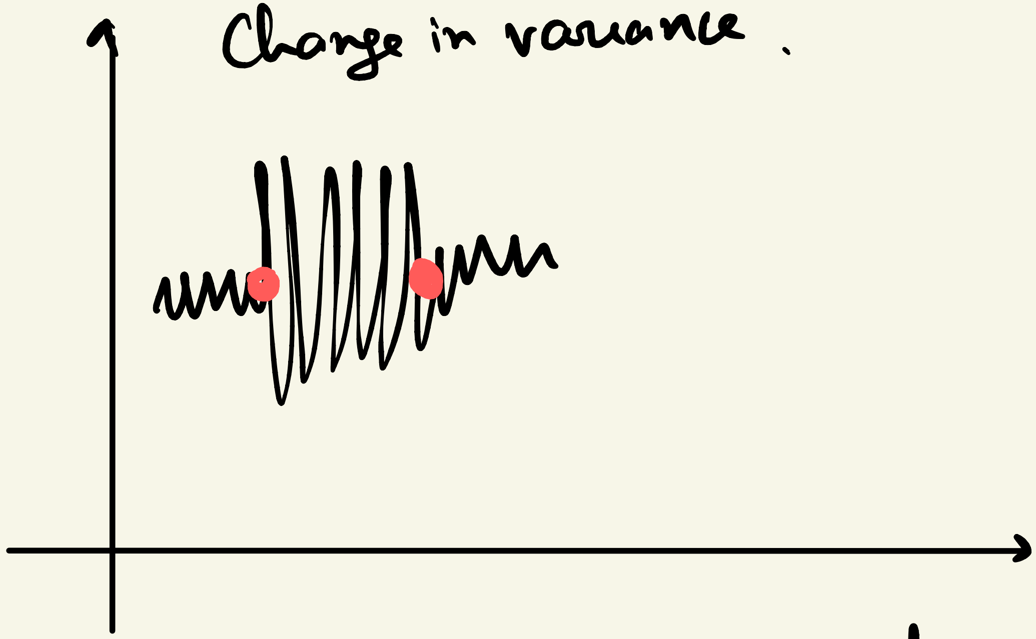
1. use a cost function:
   eg. change in stddev
2. use rolling window to calculate the cost function
3. wherever you observe a major change in the cost function, classify that point as a change point.

Eg.

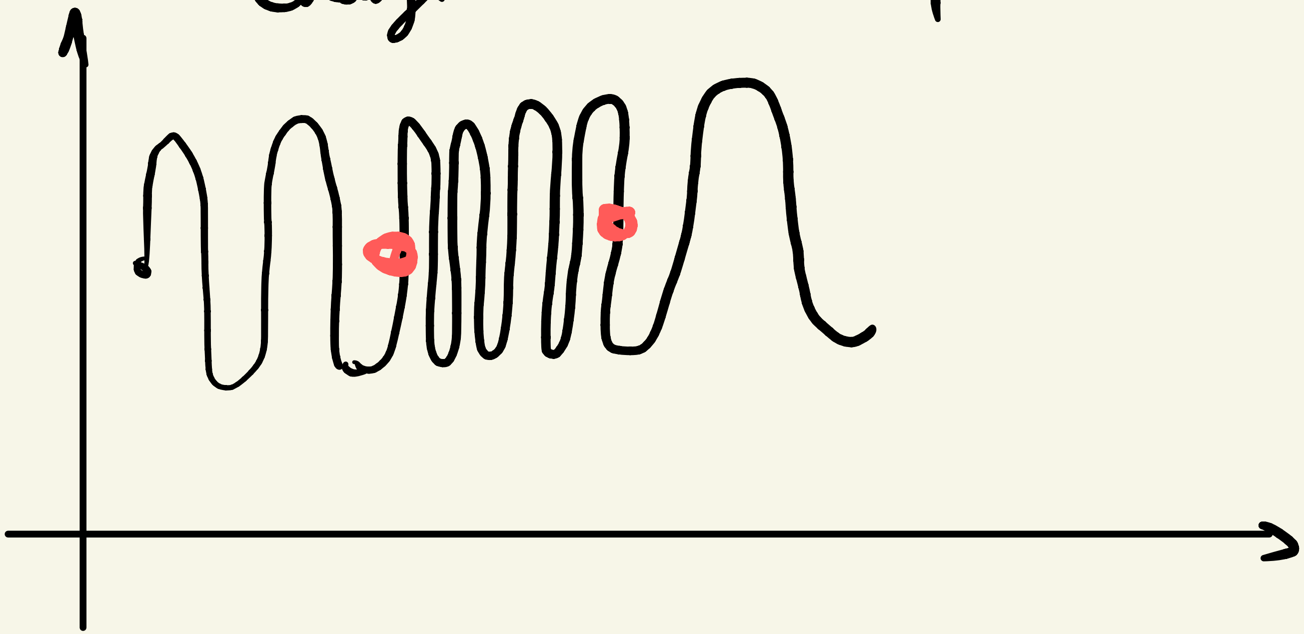
Change in mean

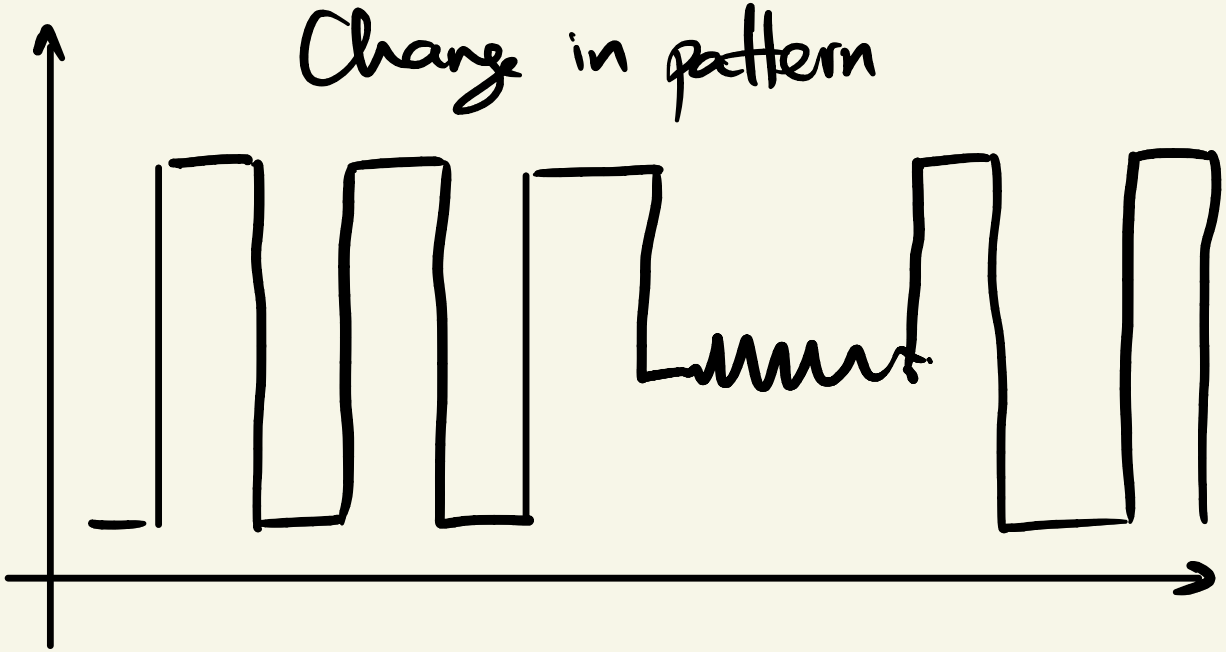


Eg. Change in variance.



Eg. Change in Seasonality




Change in pattern