Financial Analytics

Return :-

→ $R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$

R_t is simple return of asset with price series P_t. However, in statistical analysis of financial data, we consider log returns r_t, defined as:

[Generally used with high freq. data] ⇒ $r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(P_t) - \log(P_{t-1}) = \log[1+R_t]$

→ Simple and log returns are not one & same, although for small relative changes in price series P_t, they don't defer much.

→ We strongly assume a "normality in Financial Variables." [As assumed, that market can be beaten in long run] i.e follows a normal distribution.

[Mean or Avg. return = 0] → For Probability density func^n vs monthly/daily/yearly return.

↳ [graph centred at origin]

→ If histogram plot shifts rightwards of expectation, deviation of market is positive & if it shifts leftwards, deviation of market is negative.

→ If histograms becomes flat & doesn't reach the blue normal distribution, the market is indecisive as there is a confusion in market. [Risk too is high as graph is more scattered.]

[P.T.O]

Deviation from normalcy :-

- In basic statistics & probability theory, we almost exclusively deal with first & second central moment of a random variable, namely expectation & variance.

k^{th} moment of X : $m_k = E[X^k]$, eg:- expectation $\mu = E[$

k^{th} central moment of X : $\mu_k = E[(X-\mu)^k]$, eg:- variance

$Var(X) = E[(X-\mu$

- In statistical analysis of financial data, as better, in risk management, one is often also interested in the third & fourth central moments, which are the basis for "skewness" & "kurtosis".

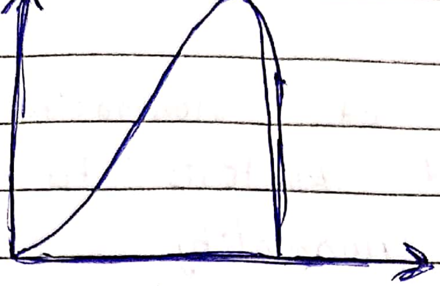
Skewness :-

- 3rd central moment of random variable

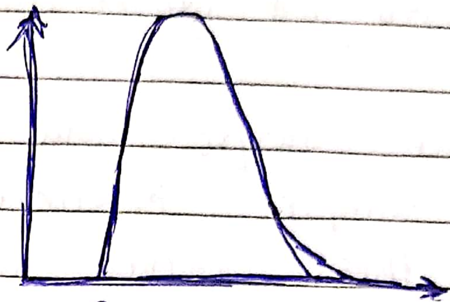
$$Skew = \frac{E[(X-\mu)^3]}{\sigma^3}$$

- Any random variable with symmetric dist. will have skew = 0.
- If skew > 0, shows positive skewness, i.e distributions have a heavy tail on right side
- If skew < 0, shows negative skewness i.e distributions have a heavy tail on left side

Sample Skewness : $Skew = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{x_i - \bar{x}}{\sigma} \right]^3$
(Unitless)



Negative Skewed



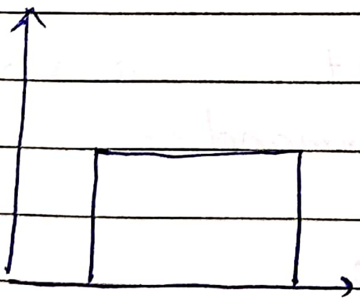
Positive Skewed

Kurtosis :-

- 4^{th} central moment
- Similar to variance it measures how ~~speed~~ spread out a distribution is, but it puts more weight on the tails.

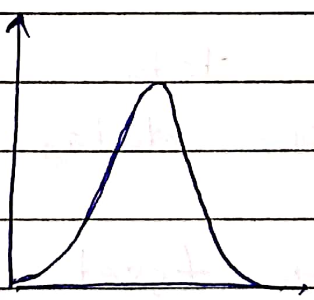
$$Kurt = \frac{E[(X-\mu)^4]}{\sigma^4}$$

Sample Kurt : $Kurt = \frac{1}{n}\sum_{i=1}^{n}\left[\frac{x_i - \bar{x}}{\sigma}\right]^4$



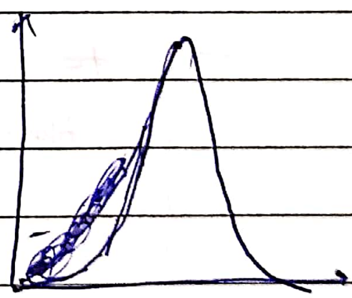
(–) Platykurtic

$Kurt < 3$



(0) mesokurtic (Normal)

$Kurt = 3$



(+) Leptokurtic

$Kurt > 3$

Tests of Normality :-

→ The Jarque - Bera test of normality compares the sample skewness and kurtosis to 0 & 3, their values under normality.

The test statistic is :

$$JB = \frac{n}{24}\left[4 \cdot \hat{Skew}^2 + \hat{Kurt}^2_{Ex}\right] \sim \chi^2_2$$

Note :-

→ 2 types of data :

- Time series data → Single variable on diff. times {eg. :- Nifty for long time
- Cross - section data → multiple variables on same time
  eg :- Nifty, Sensex, DJI on one day

Stata Commands :-

{generate var. {var. name start from value 1 to end

Eg :- gen time = _n

→ to tell stata that a variable is time series data, command is 'tsset'

eg → tsset time

(If this isn't done, stata will consider data as cross - section data).

→ 'tsline' to plot graph of time series data

$r_t = \ln\left(\frac{y_t}{y_{t-1}}\right) * 100$

$\underbrace{\left(\frac{y_t}{y_{t-1}}\right)}$

l. close_nifty means `lag value' of close_nifty

→ gen return = ln(close_nifty/l.close_nifty) * 100
(ln = log)

→ `Lag variable' only is for time series data.

→ `sum (return)' gives us `summary' of variable return. Gives major statistical features like no. of obs., mean, std. dev, min & max.

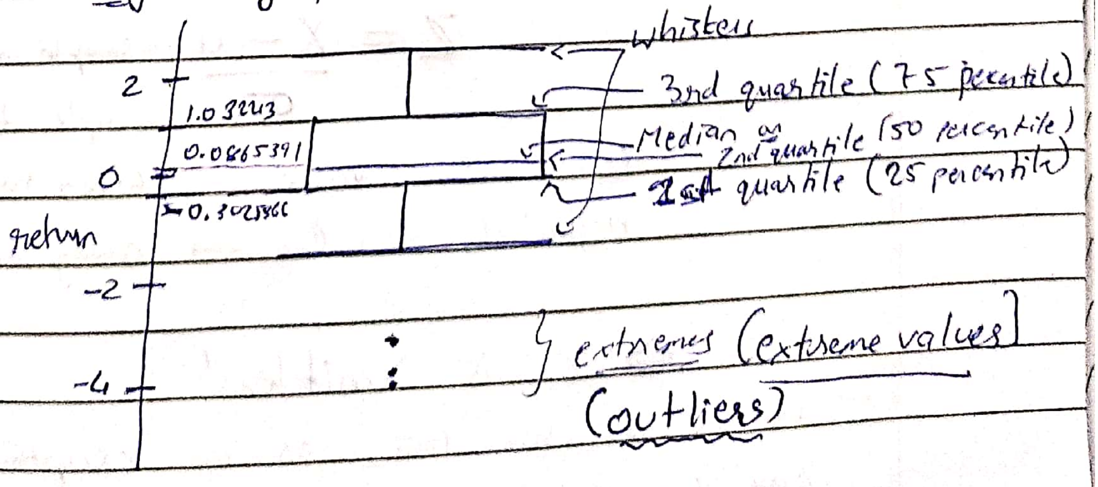
→ to see skewness, kurtosis, etc ~~extra~~ use command `detail' with sum & variable name.

eg:- sum return, detail
sum (return), detail

→ To ~~xxx~~ understand any command use `findit commandname' to open documentation & understand it. eg:- findit sum

→ `histogram variablename' to get histogram graph figure of ~~extra~~ respective variable.
eg:- histogram return

→ `graph box return' to get box plot of return variable.
eg:- graph box return



IMP →

- Borders of whiskers

$$\boxed{\begin{matrix} q_1 - 1.5(iqr) \\ q_3 + 1.5(iqr) \end{matrix}}$$

(iqr) → inter quartile range

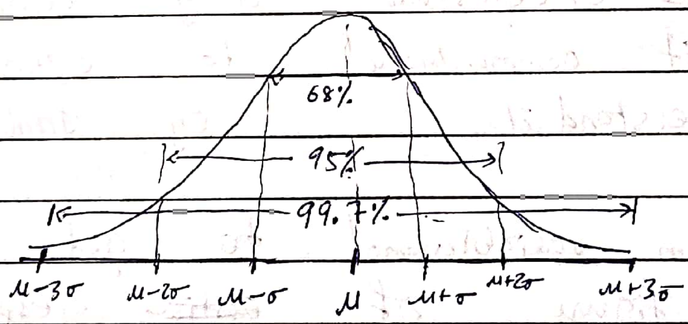
q_3 → 3^{rd} quartile (75 percentile) $[q_3 - q_1]$

q_1 → 1^{st} quartile (25 percentile)

⇛ stata uses

$$\boxed{\begin{matrix} q_1 - (iqr) \\ q_3 + (iqr) \end{matrix}}$$

Normal Distribution :-



mean → μ

std. dev → σ

Z - statistic (Population Based) [mean = 0, std.dev = 1] $\mu = 0$ $\sigma = 1$

$$Z = \frac{X - \mu}{\sigma}$$

X ← value of var.

μ ← sample mean

σ ← sample std

Lower abs. Z value → closer to mean

Higher abs. z value → away from mean

⇛ Z is 'unitless'

→ to see if var. lies as an exception or example depends on our results based on 'Hypothesis Testing'.

Hypothesis Testing :-

- One 'ordinary' sample → Inference about population

- One 'properly chosen' Sample is a good representative → Inference about population

- Null & Alternative Hypothesis

- Null Hypothesis is a statement of the status quo or no effect. (H_0)
  eg:- Science & Arts students have no diff in IQ
- Alternative Hypothesis is one in which some diff. or effect is expected. (H_1)
  eg:- Science & Arts students have diff. in IQ

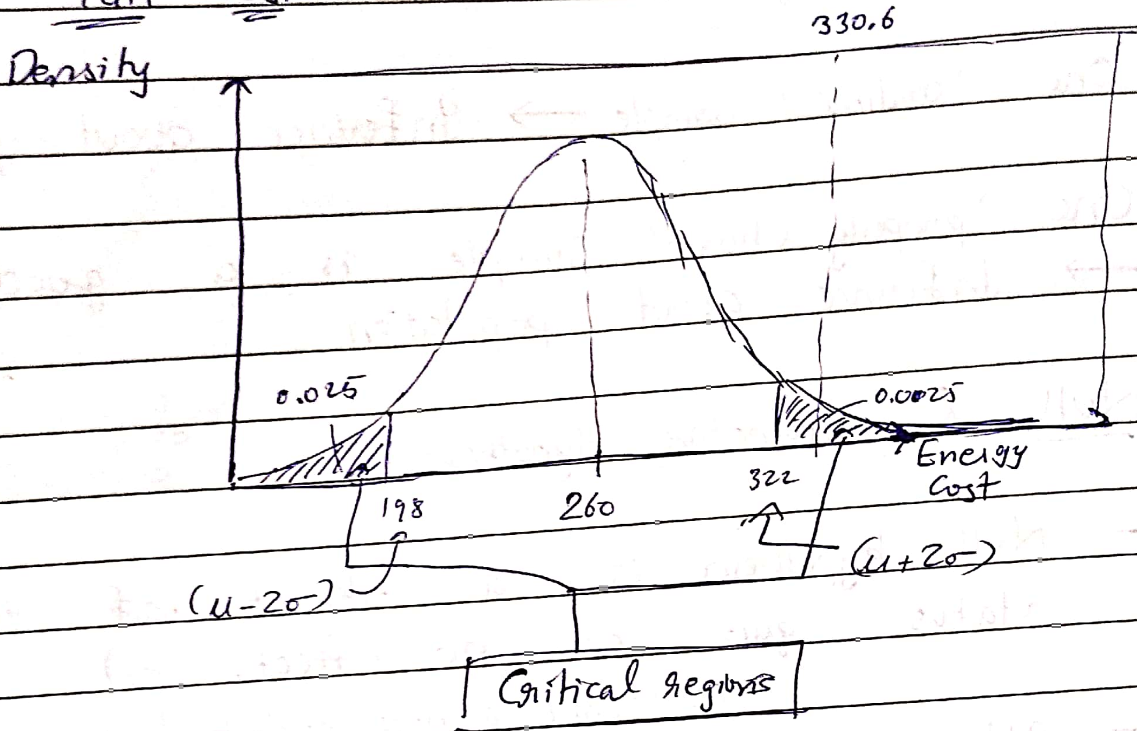
- Types of error

- Type - I error (You reject a true null hypothesis)
  Eg:- Punish a person who isn't guilty!

- Type - II error (You accept a false null hypothesis)
  Eg:- You release a truly guilty person!

- Power of a test = 1 - (Type - II test)
  {probability of avoiding a (type - II error)}

* **Two - Tail Test :-**



- Values outside critical regions will cause us to reject the null.
- Here null would be that the new Energy cost would not differ much in value from my avg. energy cost [260], calculated from past data.
- The 5% net critical region (2.5% on each side) is my 'Level of Significance'. (5% is general practive)

i.e z is in range of -2 to 2.

$|z| < 2 \rightarrow$ 5% Significance level

t-test statistics:- (Sample Based)

$$t = \frac{m - \mu}{s/\sqrt{n}}$$

obs. value ~~mean~~

theoretical value/mean

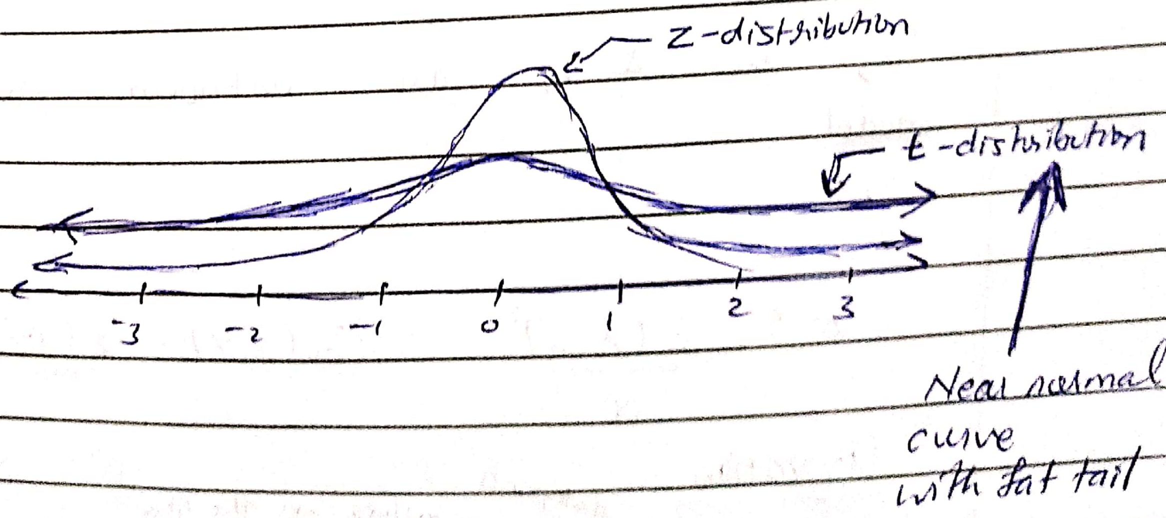
var. set size

std. dev.

- 2-t thumb of rule

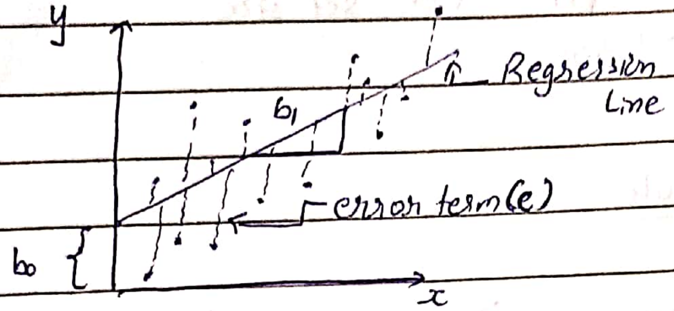
- If sample size is more than or equal to 20 (n>=20) & lvl of significance is 5% than an abs. value of 2 (|t|>=2) will suffice to reject the null hypothesis.

- t-distribution



* Notion of Linear Regression Model :-

$$Y = b_0 + b_1 x$$



RSS → sum of squares of errors

[Residual Sum of Squares]

→ We have datapoints for y & x but what we require is 'b_0' & 'b_1'.
We require b_0 & b_1 such that 'RSS' is minimum.
So, b_0 & b_1 are unknown parameters in my model.

$$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N} \qquad Cov(X,Y) = \frac{\Sigma(X_i - \bar{X})(Y_i - \bar{Y})}{n}$$

Derivation

datapoints, values on the line

$$Q = \sum_{i=1}^{n} (Y_i - \hat{Y})^2 = \sum_{i=1}^{n} (Y_i - a - bX_i)^2$$

$$\frac{\partial Q}{\partial a} = \sum_{i=1}^{n} -2(Y_i - a - bX_i) = 2\left[na + b\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} Y_i\right] = 0$$

$$a = \bar{Y} - b\bar{X}$$

$$\frac{\partial Q}{\partial b} = \sum_{i=1}^{n} -2X_i(Y_i - a - bX_i) = \sum_{i=1}^{n} -2(X_iY_i - aX_i - bX_i^2) = 0$$

$$\therefore \sum_{i=1}^{n} \left(x_i y_i - x_i \bar{y} + b x_i \bar{x} - b x_i^2 \right) = 0$$

$$b = \frac{\sum_{i=1}^{n} (x_i y_i - x_i \bar{y})}{\sum_{i=1}^{n} (x_i^2 - x_i \bar{x})} = \frac{\sum_{i=1}^{n} (x_i y_i) - n \bar{x} \bar{y}}{\sum_{i=1}^{n} x_i^2 - n \bar{x}^2}$$

$$\therefore b = \frac{\sum_{i=1}^{n} (x_i y_i - x_i \bar{y}) + \sum_{i=1}^{n} (\bar{x}\bar{y} - y_i \bar{x})}{\sum_{i=1}^{n} (x_i^2 - x_i \bar{x}) + \sum_{i=1}^{n} (\bar{x}^2 - x_i \bar{x})}$$

$$= \frac{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\therefore \boxed{b = \frac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)}}$$ IMP.

Regression in stata

```
regress close_nifty excangerate
predict niftyhat

tsline close_nifty niftyhat

scatter close_nifty excangerate || lfit close_nifty excangerate
```

[linear fit]

$\rightarrow$ Normal Distribution $(0, \sigma^2)$

$\Rightarrow$ $$y = \alpha + \beta x + \varepsilon$$

$$\frac{\Sigma \bar{y}}{n} = \bar{y} = \alpha + \beta \bar{x} + 0$$

$$(y-\bar{y})^2 = \cancel{\alpha} - \cancel{\alpha} + \beta(x-\bar{x})^2 + \varepsilon^2$$

$$E(y-\bar{y})^2 = E\left[\beta(x-\bar{x})^2\right] + \cancel{E(\varepsilon^2)}^{\,0}$$

$$TSS = ESS + RSS$$

Total Sum of Squares = Explained Sum of Squares + Residual Sum of Squares

$\rightarrow$ In ideal model, we expect ESS to be large.

Dividing by TSS,

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS}$$

$$\therefore 1 = R^2 + \frac{RSS}{TSS}$$

$\rightarrow$ So, 'R^2' is a measure of a model wherein we expect it to be higher & higher and closer to '1'.

$\Rightarrow$ So now making a new model,

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 x_t + \varepsilon$$

Now we add previous day data to improve model

l. → lag of → (T-1)
l2. → second lag of → (T-2)
l3. → 3rd lag of → (T-3)

→ Command.

```
regress close_nifty l.close_nifty exchangerate
predict niftyhat2
```

⇒ Root Mean Square Error

$$RMSE = \sqrt{\frac{\Sigma(y_i - \hat{y})^2}{n}}$$

should be as small as possible (desirable)

→ Command

```
gen error = (close_nifty - niftyhat)
gen sqerror = (error)^2
sum (sqerror)
gen rmse (21338)^0.5
```

⇒ Stationarity [3 properties]

→ Mean is constant
→ Var. is ~~finite~~ finite & time independent
→ Auto-covariance is constant & time independent

[Derivation ahead →]

Derivation of Stationarity

$\Rightarrow$ $y_t = \alpha + \beta y_{t-1} + \varepsilon_t$ (AR-1 Process)

$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t$ (AR-2 Process)

order of regression

$y_t = \alpha + \beta_1 y_{t-1} + \dots + \beta_k y_{t-k} + \varepsilon_t$ (AR(k) process)

From AR-1 process,

$$y_t = \alpha + \beta(\alpha + \beta y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

(Iteration)

$$y_t = \alpha + \beta\alpha + \beta^2 y_{t-2} + \beta\varepsilon_{t-1} + \varepsilon_t$$

$$= \alpha + \beta\alpha + \beta^2(\alpha + \beta y_{t-3} + \varepsilon_{t-2}) + \beta_{t-2} + \varepsilon_t$$

$$= \alpha + \beta\alpha + \beta^2\alpha + \beta^3 y_{t-3} + \beta^2\varepsilon_{t-2} + \beta\varepsilon_{t-2} + \varepsilon_t$$

$$= \alpha(1 + \beta + \beta^2 + \dots) + \beta^t y_0 + \sum_{i=0}^{t} \beta^i \varepsilon_{t-i}$$

If $|\beta| < 1$ then as $t \to \infty$

$$y_t = \underbrace{\frac{\alpha}{1-\beta} + \text{constant}}_{A} + \sum_{i=0}^{t} \beta^i \varepsilon_{t-i}$$

Mean of y_t will converge to a constant A

Because

Mean of y_t = $E(y_t) = E(A) + E\left(\sum_{i=0}^{t} \beta^i \varepsilon_{t-i}\right)$

$$= A + \beta^i E\left(\sum_{i=0}^{t} \varepsilon_{t-i}\right)$$

$$= A + 0$$

→ For a random var., expectation is synonymous to mean.

⇒ Random Walk Model

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$ [AR-I Process]

If suppose $\boxed{a_1 = 1}$ then, [Strong Assumption]

$\Downarrow$

Now this [Becomes Random Walk Model]

$$y_t = a_0 + y_{t-1} + \varepsilon_t$$

$$y_t = a_0 + (a_0 + y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$ (iteration)

$$y_t = a_0 + a_0 + y_{t-2} + \varepsilon_{t-1} + \varepsilon_t$$

After 't' substitution,

Random walk Model ⟹ $$y_t = t a_0 + y_0 + \sum_{i=0}^{t} \varepsilon_{t-i} \rightarrow (1)$$

Hence, y_t = A [constant term) + Sum of Errors

$$\therefore E(y_t) = E(A) + E\left(\sum_{i=0}^{t} \varepsilon_{t-i}\right)$$

$$= A + 0$$ [By assumption] [1 to t with gap of 1]

$$E(\varepsilon_{t-i}) = 0 \ \forall \ i; \ i = 1(1)t$$

Variance of y_t (when y_t is random variable]

$$= E[y_t - E(y_t)]^2$$

Alt. Version: $Var(y_t) = E(y_t^2) - [E(y_t)]^2$

From (1) we can write,

$$\text{var}(y_t) = \text{var}(A) + \text{var}(\Sigma \varepsilon_{t-i}) + 2\,\text{cov}(A, \Sigma \varepsilon_{t-i})$$

$\left[\because \text{var}(X+Y) = \text{var}(X) + \text{var}(Y) + 2\text{cov}(X,Y) \right.$
$\text{cov}(X,Y) = E(XY) - E(X)E(Y)$,
$\left. \text{if } (X,Y) \text{ are independent then, } \text{cov}(X,Y) = 0 \right]$

Hence,

$$\text{var}(y_t) = \text{var}(A) + \text{var}(\Sigma \varepsilon_{t-i})$$
$$= 0 + \Sigma\,\text{var}(\varepsilon_{t-i}) + \text{cov. terms}$$

$$\underbrace{\text{var}(y_t) = t\sigma^2}_{\text{Second cond}^n\text{ of stationarity violated}}$$

$\because$ [Cond^{n}s Assumed]

[ε_{t-i} & ε_{t-j} are independent for all $i \neq j$]

[Homoscedasticity is a cond^n in which variance of residual or error term, in a regression model is constant]

Now,

$$\text{Cov}(X,Y) = E(X,Y) - E(X)E(Y)$$

If we want to see covariance b/w y_t & $y_{t-\tau}$ [τ is some lag value and $\tau \neq 0$]

then,

$$\text{cov}(y_t, y_{t-\tau}) = E\left[(y_t - \bar{y}_t)(y_{t-\tau} - \bar{y}_{t-\tau})\right]$$

For a long time series where $t >> 0$ and τ is very small then,

		$\tau=1$	$\tau=2$
t	y_t	y_{t-1}	y_{t-2}
1	10	—	—
2	12	10	—
3	11	12	10
4	21	11	12
5	19	21	11

$$\bar{y}_{t-\tau} = \bar{y}_t \ (= \mu)$$

(assume)

$$\therefore Cov(y_t, y_{t-\tau}) = E\left[(y_t - \mu)(y_{t-\tau} - \mu)\right]$$

$$= \gamma_\tau$$

[Autocovariance at "τ" lag]

* **Note:-** $\gamma_t(\tau=0) \Rightarrow$ Autocovariance at lag 0 is nothing but variance

So, from (1),

$Cov(y_t, y_{t-\tau})$ can be also derived as 'time' dependent [Try to derive]

So, third cond^n of stationarity violated.

$\Rightarrow$ So, variance & autocovariance both are time dependent.
Hence, model which follows random walk model is non-stationary.

[P.T.O]

Unit Root Test:-

$\Rightarrow$ $y_t = \alpha_0 + \beta y_{t-1} + \varepsilon_t$

If $\beta = 1$,

it becomes Random Walk Model.

$y_t = \alpha_0 + \beta y_{t-1}$ — (i) [ε is not part of eqn]

→ If t is discrete i.e $\pm 1, \pm 2, \pm 3, \ldots$

then eq (i) is called difference equation.

→ If t is continuous or fractional i.e $\pm 1, \pm \frac{1}{2}, \pm 0.0001$,

then eq (i) becomes differential equation.

Here, we assume 't' is discrete.

For PI, $y_{t+1} = y_t = \bar{y}$

$\therefore \bar{y} - 5\bar{y} = 1$

or $-\bar{y} = \frac{1}{4}$

For CF, $y_t = Ab^t$ ($A, b \rightarrow$ const.)

$\therefore Ab^{t+1} - 5Ab^t = 0$ [$\because$ Fluctuations are transitory]

$\therefore Ab^t (b-5) = 0$

$\because Ab^t \neq 0$, $b = 5$

$\therefore y_t = PI + CF$

$\therefore y_t = \left(-\frac{1}{4}\right) + A(5)^t$

If supp y_0 (at $t=0$) $= \frac{7}{4}$

$\therefore \frac{7}{4} = -\frac{1}{4} + A(5)^0$

$\therefore 2 = A(5)^0$

$\therefore \boxed{A = 2}$

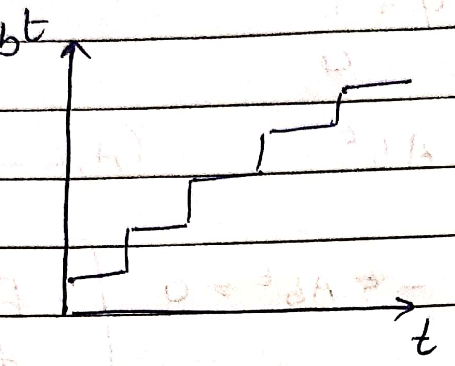
So, solution is,

$\therefore \boxed{y_t = \left(-\frac{1}{4}\right) + 2(5)^t}$ Ans

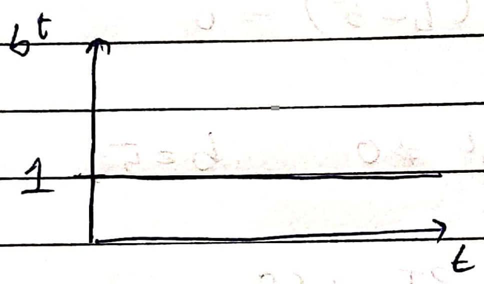
$\Rightarrow$ In solution of

$$y_t = \underbrace{Ab^t}_{CF} + \overbrace{\bar{y}}^{PI}$$

$\rightarrow$ If $b > 1$,



$\rightarrow$ If $b = 1$,



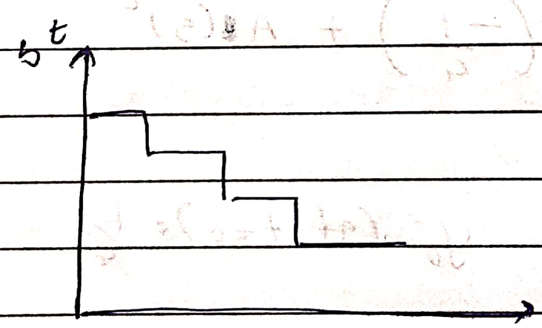
$b < 0 \rightarrow$ oscillatory

$b > 0 \rightarrow$ No oscillatory

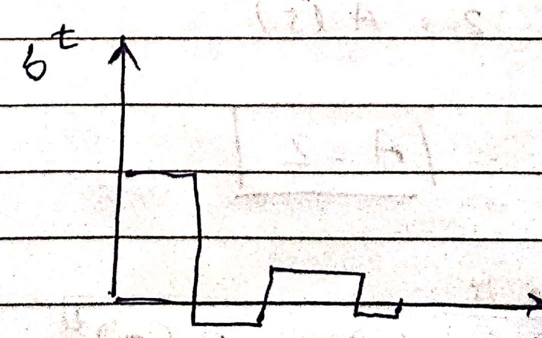
$|b| > 1 \rightarrow$ divergent

$|b| < 1 \rightarrow$ convergent

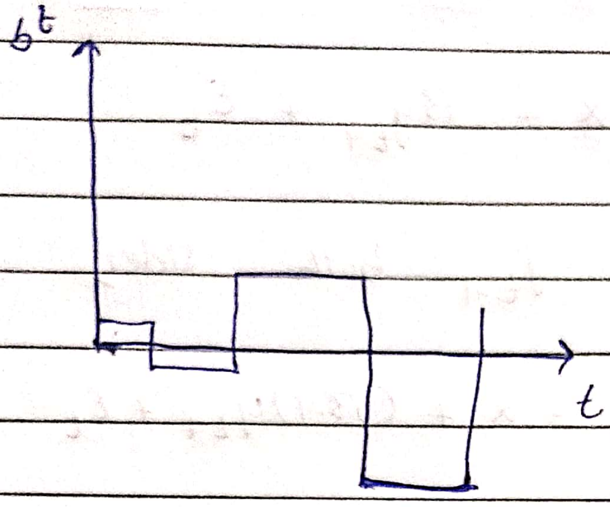
$\rightarrow$ If $0 < b < 1$



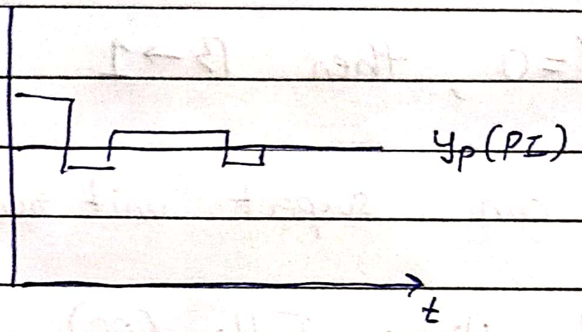
$\rightarrow$ If $-1 < b < 0$,



→ If $b < -1$



⇒ So, $y_t = Ab^t + y_p(PI)$



⇒ The fluctuation from y_p, depends on 'b'.

⇒ <u>Stationarity of financial time series</u>

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t \quad -(1)$$

If we ignore the ε_t term and look at (1) as a difference eqn,

then $|\beta| < 1$ then solution y_t will have convergence to an equilibrium value (PI).

If $\beta = 1$ then it is Random Walk Model, which isn't ~~coverg~~ convergent.

So, a stationary process means it should converge as time extends.

* Practical way to check stationarity :-

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t$$

Substracting y_{t-1} both sides,

$$y_t - y_{t-1} = \alpha + (\beta - 1) y_{t-1} + \varepsilon_t$$

$$\therefore \Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t \quad [\gamma = \beta - 1]$$

If $\gamma = 0$, then $\beta \to 1$

and we can suspect unit root $(\beta \to 1)$

This is called Dickey Fuller (DF) unit root test.

Meaning,

regress Δy_t over y_{t-1} and see if co-eff. of y_{t-1} is closer to '0'. If true, then there is non-stationarity, else if false, it is stationary.

Command

calculates first difference

regress d.close_nifty l.close_nifty

regress d.excangerate l.excangerate

regress d.close_nifty l.close_nifty if t<40

Concept of ACF (Auto correlation Funcⁿ) :-

→ Correlation defines linear association b/w two variables

Auto correlation defines relation b/w y and its past values at different lag-length.

ACF of order (1) means correlation b/w y_t & y_{t-1}
ACF " " (2) " " " y_t & y_{t-2}
and so on.

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t \quad \left[\begin{array}{l} \varepsilon_t \text{ is normally distributed} \\ 0 \to \text{mean} \to \text{variance} \to \sigma^2 \text{ (const)} \end{array} \right]$$

$$y_t = a_0 \sum_{i=0}^{t-1} a_1^i \quad a_1^t y_0 + \sum_{i=0}^{t-1} a_1^i \varepsilon_{t-i}$$

$$E(y_t) = a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t y_0 + \{ E[\ ] = 0 \}$$

$$E[y_{t+s}] = a_0 \sum_{i=0}^{t+s-1} a_1^i + a_1^{t+s} y_0 \quad [\text{Subs. } t \text{ with } t+s]$$

If $t \to \infty$ & $|a_1| < 1$ then,

$$\lim y_t = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{t-i}$$

For $|a_1| < 1$, we can assume

$$\lim y_t \longrightarrow \frac{a_0}{1-a_1}$$

Since, t is sufficiently large and 's' is small,

$$E(y_t) = E(y_{t-s}) = \frac{a_0}{1-a_1} = \mu$$

$$\underbrace{E(y_t - \mu)^2}_{\text{variance}} = E\left[(\varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + \dots) \right]^2$$

$$= \sigma^2 [1 + a_1^2 + a_1^4 + \dots]$$

$$\boxed{\text{Variance} = \frac{\sigma^2}{[1-a_1^2]}} \quad [\because |a_1| < 1]$$

Similarly,

$$E[(y_t - \mu)(y_{t-s} - \mu)] = \text{Autocovariance b/w } y_t \text{ \& } y_{t-s}$$

$$= E\{[\varepsilon_t + a_1 \varepsilon_{t-1} + a_1^2 \varepsilon_{t-2} + \dots] [\varepsilon_{t-s} + a_1 \varepsilon_{t-s-1} + a_1^2 \varepsilon_{t-s-2} + \dots]\}$$

$$= \sigma^2 (a_1)^s [1 + a_1^2 + (a_1)^4 + \dots]$$

$$= \frac{\sigma^2 a_1^s}{1 - a_1^2}$$

Now, if $s = 1$,

$$\text{Auto covariance} = \frac{\sigma^2 a_1}{1 - a_1^2}$$

if $s = 2$,

$$'' = \frac{\sigma^2 a_1^2}{1 - a_1^2}$$

$\Rightarrow$ So,

$$\underset{\text{(Auto Correlation Func}^n)}{\text{ACF}} = \frac{\text{Autocovariance}}{SD_{y_t} \; SD_{y_{t-s}}}$$

$$= \frac{\sigma^2 a_1^s}{1-a_1^2} \quad (= \gamma_s)$$

$$\overline{\quad \text{var } y_t \quad}$$

$$= \frac{\left(\sigma^2 a_1^s\right) / \left(1-a_1^2\right)}{\left(\frac{\sigma^2}{1-a_1^2}\right)(= \gamma_0)}$$

$$\boxed{ACF = a_1^s \quad \forall \quad s = \pm 1, \pm 2, \ldots}$$

$$\Downarrow$$

$$\boxed{\text{ACF at lag } s = \rho_s = \frac{\gamma_s}{\gamma_0} = a_1^s}$$

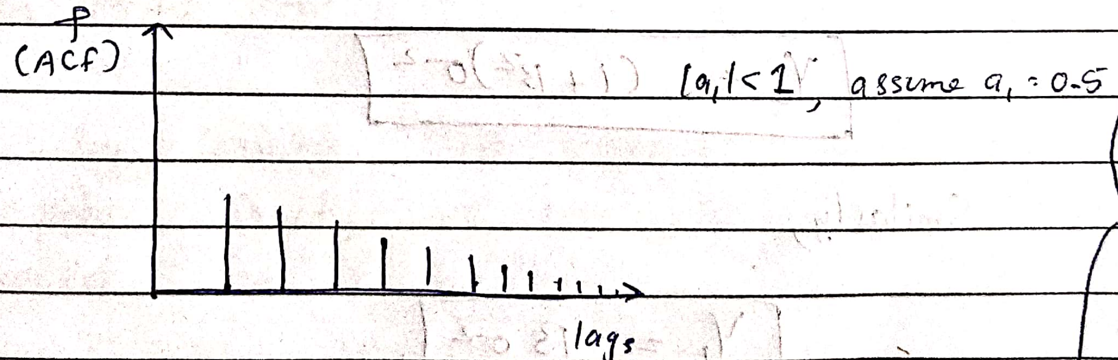
So,

$$\therefore \rho_0 = \frac{\gamma_0}{\gamma_0} = 1$$

$$\therefore \rho_1 = \frac{\gamma_1}{\gamma_0} = a_1$$

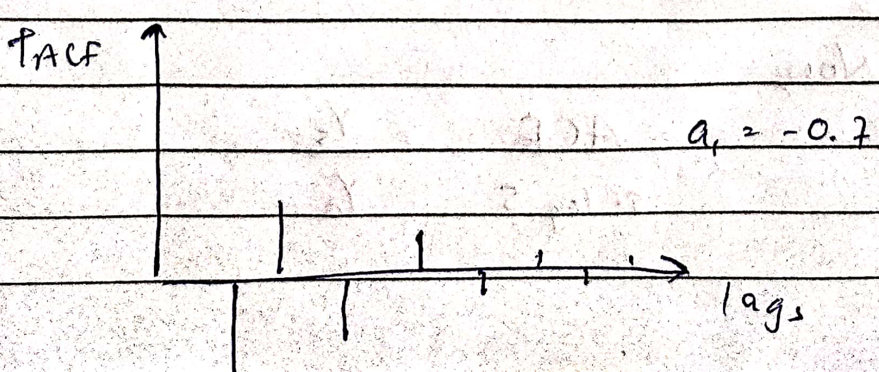
$$\therefore \rho_2 = \frac{\gamma_2}{\gamma_0} = a_1^2$$

So, drawing graph,



Correlograms

IMP.



Note: Just like correlation, auto-correlation also lies in b/w 1 & -1.

Commands

ac close_nifty

⇒ Derive the ACF of an AR(2) process,

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$

Concept of MA [Moving Average]:-

$$y_t = \varepsilon_t + \beta \varepsilon_{t-1} \quad \text{[Order-1 MA]}$$

[MA-1 process]

$$E(y_t \cdot y_{t-s}) = E\left[(\varepsilon_t + \beta \varepsilon_{t-1})(\varepsilon_{t-s} + \beta \varepsilon_{t-s-1})\right]$$

If $s=0$ then, $E(y_t \cdot y_{t-s}) = E(y_t y_t) = \gamma_0$

Autocovariance at lag 0

= variance

$$E(y_t y_t) = E\left[(\varepsilon_t + \beta \varepsilon_{t-1})(\varepsilon_t + \beta \varepsilon_{t-1})\right]$$

$$= E\left[\varepsilon_t^2 + \beta \varepsilon_t \varepsilon_{t-1} + \beta \varepsilon_{t-1} \varepsilon_t + \beta^2 \varepsilon_{t-1}^2\right]$$

$$\boxed{\gamma_0 = (1+\beta^2)\sigma^2}$$

Similarly,

$$\boxed{\gamma_1 = \beta \sigma^2},$$

Now,

$$\text{ACF at lag } s = \frac{\gamma_s}{\gamma_0}$$

$\therefore$ At $s=1$

$$ACF = \frac{\gamma_1}{\gamma_0} = \frac{\beta\sigma^2}{(1+\beta^2)\sigma^2} = \frac{\beta}{1+\beta^2}$$

$$\boxed{\rho_1 = \frac{\gamma_1}{\gamma_0} = \frac{\beta}{1+\beta^2}}$$

$$\boxed{\rho_0 = \frac{\gamma_0}{\gamma_0} = 1}$$

If we calculate ρ_2 onwards then,

$$\boxed{\rho_i = 0 \quad \forall\ i \geq 2}$$

Eg:- MA process $\Rightarrow$ $y_t = \varepsilon_t - 0.7\,\varepsilon_{t-1}$

$\rho_0 = 1$

$\rho_1 = -0.469$

ρ_2 (onwards) $= 0$

For MA-1 $\Rightarrow$ I will get one spike in graph
for MA2 $\Rightarrow$ " " two " " "
for MA-3 $\Rightarrow$ " " three " " "
think!

→ So by looking at graph itself, & seeing no of spikes in it, we can easily see what kind of process it is

Note:-

→ A series in reality could have a combo of both in one framework i.e A model could have components of both AR as well as MA process

Eg:- $y_t = \alpha + \beta y_{t-1} + \delta\,\varepsilon_{t-1} + \varepsilon_t$

Combo of MA-1 & AR-1 $\Rightarrow$ ARMA(1,1)

Eg:— $y_t = \alpha + \beta y_{t-1} + \gamma y_{t-2} + \delta \varepsilon_{t-1} + \varepsilon_t$

ARMA (2, 1)

[Since AR-2 & MA-1 combo]

All such combos of AR & MA could be seen. We can detect them easily.

→ This is all a univariate model as with only one variable and its past values and lag values, we are trying to make predictions.

Partial Autocorrelation Func^n :-

→ If y_t follows an AR(1) structure then, y_t & y_{t-2} is related via y_{t-1}. So, when we calculate rel^n b/w y_t & y_{t-2}, intervention of y_{t-1} is inevitable.

If we can somehow remove the intervening influence of y_{t-1} then, y_t & y_{t-2} can have direct interaction.

PAC measures this very link / association b/w y_t & y_{t-2}, given that impact of y_{t-1} is removed.

meaning remove the mean

→ Simplest way to nullify the intervention of y_{t-1} is to ~~de-mean~~ 'de-mean' the data,

$y_t^* = y_t - \bar{y}_t$

$y_{t-2}^* = y_{t-2} - \bar{y}_t$

$\bar{y}_{t-2} \approx \bar{y}_t$ (long TS data)

i.e $y_t - \bar{y}_t = \left[y_{t-2} - \bar{y}_{t-2} (\approx \bar{y}_t) \right]$

$\therefore (y_t - \bar{y}_t) = \alpha + \beta (y_{t-2} - \bar{y}_t) + \varepsilon_t$

de-meaned values

$(y_t^*) = \alpha + \beta (y_{t-2}^*) + \varepsilon_t$

$\Rightarrow$ 'β' is PAC b/w y_t & y_{t-2}

It can be shown that PAC 'ϕ_{ss}' has a general expression.

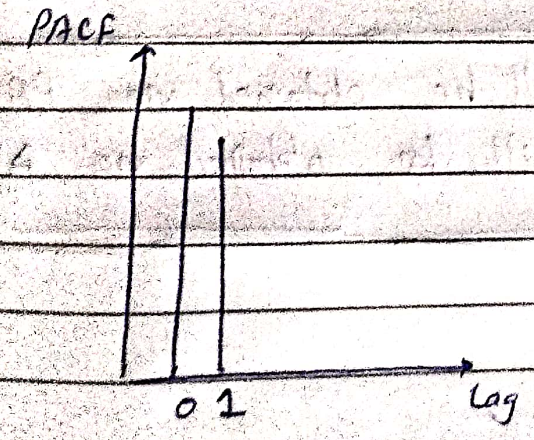
$$\phi_{ss} = \frac{\rho_s - \Sigma \phi_{s-1,j}\, \rho_{s-j}}{1 - \Sigma_{s-1,j}\, \rho_j} \quad ; s = 3, 4, 5, ---$$

$$\phi_{11} = \rho_1$$

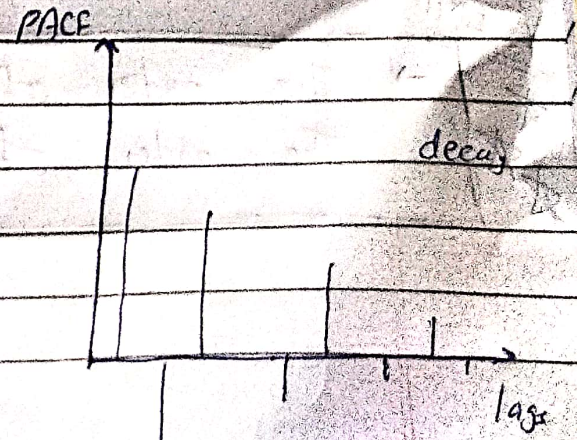
$$\phi_{22} = \frac{(\rho_2 - \rho_1^2)}{(1 - \rho_1^2)}$$

Note:- Yule-Walker equations are an alternate method ~~to~~ that could be used. Above formula has been derived using that very method.

→ For AR(1) Process



→ For MA(2) Process



→ The decay in MA(1) process could be exponential or sinusodial depending upon the sign of PAC.

if neg. → sinusodial decay

if pos. → exponential decay

⇒ Calculation of PAC

$$y_t = -0.7\,y_{t-1} + \varepsilon_t - 0.7\,\varepsilon_{t-1} \quad -(1)$$

This is an ARMA(1,1) process

we can calculate ρ_1 as,

$$y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} \rightarrow \rho_1 = \frac{(1 + a_1 \beta_1)(a_1 + \beta_1)}{(1 + \beta_1^2 + 2a_1\beta_1)}$$

∴ ρ_1 for (1) is,

$$\rho_1 = \frac{[1 + 0.49][-0.7 - 0.7]}{[1 + 0.49 + 2 \times 0.49]} = -0.8445$$

$$\therefore \boxed{\phi_{11} = -0.8445 = \rho_1}$$

$$\rho_{22} = \frac{0.591 - (-0.8445)^2}{1 - (-0.8445)^2} = -0.426$$

V. IMP.

★ → The order of AR will be obtained via PAC

★ The order of MA will be obtained via AC

The graph obtained with this will have a grey zone which shows our 'confidence interval'. → So all spikes outside the zone will have significant effect.

classmate

Date

Page

Commands

pac return

→ From graph, we infer that y_t is significantly effected from y_{t-3}, without intervention from y_t & y_{t-2}. [in data provided by sir]

→ So using PAC, we decided the order of AR i.e 3.

i.e y_t depends on y_{t-3}

→ Now for order of MA,

ac return

From graph, again the 3rd lag seems significantly effecting. So, order of MA is also 3.

So, we can say that our model that y_t series follows is,

$$y_t = \alpha y_{t-3} + \gamma \varepsilon_{t-3} + \varepsilon_t$$

i.e an ARMA (3, 3) Model.

So, using ACF & PAC, we can find the significant lag values that effect the series.

Note:-

→ For AR2 Model, ie

mean var.
$N(0, \sigma^2)$

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \varepsilon_t$$

$$\text{Variance} = \gamma_0 = \frac{\sigma^2}{1 - a_1 \rho_1 - a_2 \rho_2}$$

Autocorrelation

$$\rho_1 = \frac{a_1}{1 - a_2}$$

$$\rho_2 = a_1 \rho_1 + a_2 = \frac{a_1^2 + (1-a_2)a_2}{(1-a_2)}$$

$$\gamma_0 = \left[\frac{1 - a_2}{1 + a_2}\right] \frac{\sigma^2}{[(1 - a_2)^2 - a_1^2]}$$

$$\gamma_1 = \left[\frac{a_1}{1 - a_2}\right]\left[\frac{1 - a_2}{1 + a_2}\right] \frac{\sigma^2}{[(1 - a_2)^2 - a_1^2]} = \frac{a_1}{1 - a_2} \gamma_0$$

$$\gamma_\tau = a_1 \gamma_{\tau-1} + a_2 \gamma_{\tau-2}$$

$$\rho_\tau = a_1 \rho_{\tau-1} + a_2 \rho_{\tau-2}$$

Unit Root & Stationarity :-

$$y_t = a_0 + b_1 y_{t-1} + \varepsilon_t$$

Substracting y_{t-1} both sides,

$$y_t - y_{t-1} = a_0 + (b_1 - 1) y_{t-1} + \varepsilon_t$$

$$\therefore \Delta y_t = a_0 + y_{t-1}(b_1 - 1) + \varepsilon_t = a_0 + \gamma y_{t-1} + \varepsilon_t \longrightarrow (4)$$

If $b_1 \to 1$ then $\gamma \to 0$

⇒ It can be shown that Δy_t is stationary and for given data "y_{t-1}" we can run a ~~single~~ simple 'OLS' estimation on equation (4),

$b_1 \to 1$

Differencing variable $(\Delta y_t = y_t - y_{t-1})$ is crux of unit root test (Dickey Fuller Test)

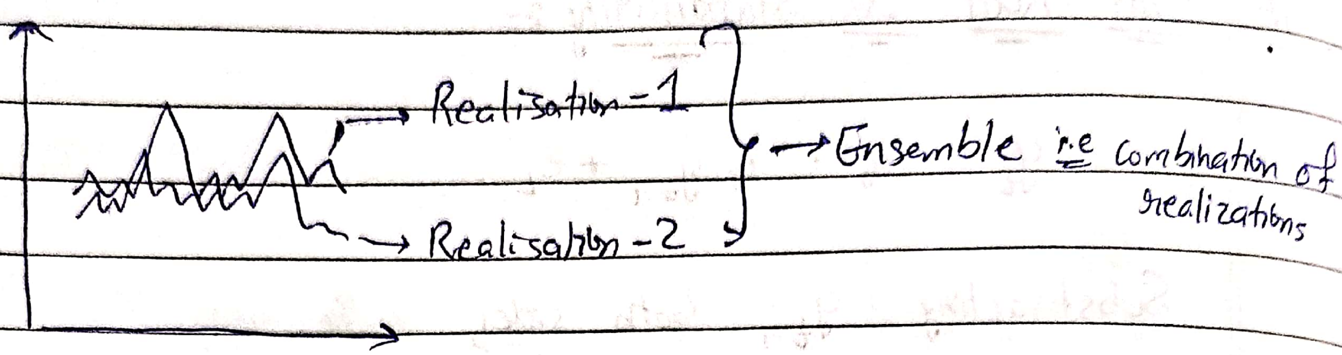
* Realization and Ensembles (Relooking at Stationarity)

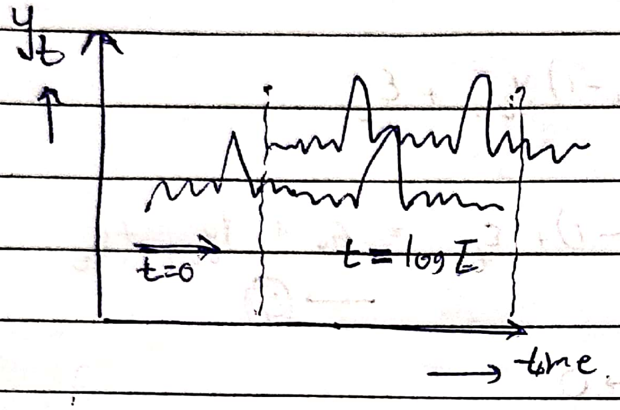
Mechanical Experiments in pure science
→ repeated experiments and generate multiple samples (altering parameters)
eg Law of gravitation

$$F = \frac{G m_1 m_2}{r^2}$$

→ But in Economics / Finance you get one sample
How work with

| Findings / Results → Realizations | Club of realisation → Ensembles |





Two "snapchats"

Steps to Fit An Univariate ARMA Model :-

'Box Jenkins Approach'

1. Identification (Plot ACF & PACF)
2. Estimation (Co-effs are estimated for AR & MA term)
3. Diagnostic Testing (residual is stationary or not)
4. Prediction

Other issues with time series

- Seasonality
- Structural break

IMP- Command → tsline excangerate (y_t)
tsline d.exchangerate $(y_t - y_{t-1})$
gen y = d.excangerate
pac y
ac y

// All on Stata //

Date ____
Page ____

→ In pac y, we get 10th line significantly outside, but we have only 60 obs. so instead I will take the 5th lag that's marginally outside

→ In ac y, no lag is outside the interval.

So model is AR(5) only.

Command

arima y, ar(5) → insignificant z value
arima y, arima(5,0,0) → 4th lag only significant
arima y, ar(10) → insignificant z value
arima y, ar(4) → insignificant z value.

→ So we see when we take lags alone they are insignificant but when all 5 lags are taken using "arima y, arima(5,0,0)" command, 4th lag becomes significant
[shows interdependency of lags upon their significance]

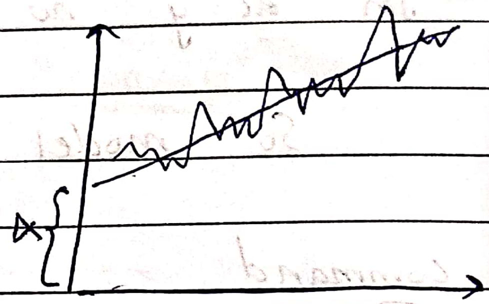
Command

- predict yhat
tsline y yhat
tsappend, add(5) (→ generating 5 extra values)
arima y, arima(5,0,0)
predict yhat2
tsline y yhat2

⇒ Any time series will have 4 major components:

- Trend
- Season
- Cycle
- Error/ Random Component

⇒ $y_t = \alpha + \beta t + \varepsilon_t$



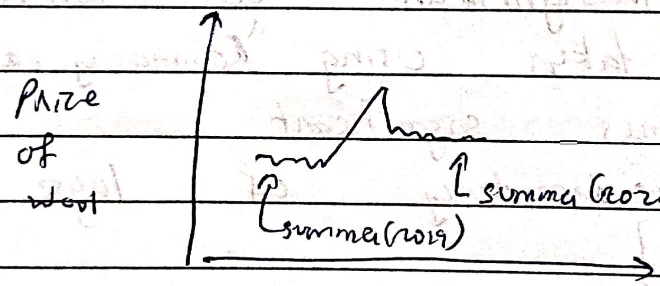
t = trend = 1, 2, ---n

Linear trend → represented by "t"

α = intercept

ε_t → contains (seasonal + cycle) Possibly!

⇒ Seasonality → sudden rise/fall at regular intervals [Not very smooth transition]



→ Moving Average method to remove seasonality from data

→ Cycle can be ~~filte~~ filtered by statistical filters

~~Hood~~ Hodrick – Prescott filter (HP)

TS ←

Ds

↓ → Trend

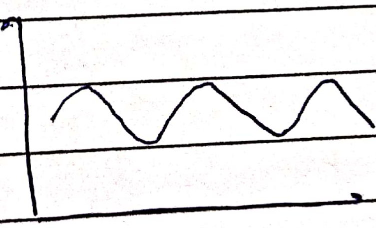
↓

Cycle + Errors

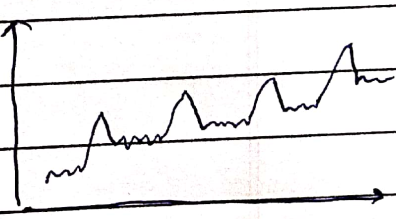
↓ ← HP filter

Error

→ In cycle, there would be sinusodial rise & flow of some kind, for it to be called a cycle.



Cycle



Seasonality

Deseasonalisation:

→ Elimination of seasonal variations from observed data of a time series.

→ Moving Average method used to achieve this.

IML → Seasonal ratio of a quarter is avg. of all quarter ratios
[eg:- seasonal value of quarter 2 is avg. of all q1 ratios]

- Quarter ratios are moving avg divided by actual value.
- Apply regression b/w variables and find the model. & predict the next variable value. [for next trend value].

⇒ The value obtained must be ~~divided~~ multiplied with respective seasonality ratio to incorporate effect of seasonality.

Eg:- In data sm sent,
we apply regression & get model as

$$Sale = 8 + 1.96 * trend$$

for trend = 13,

$$sale = 33.48 \quad (predicted)$$

Now to incorporate seasonality,

$$= 33.48 \times 0.874927$$

avg q1 seasonality ratio

~~Act~~. Sale Value = 29.4485
Predicted

Hodrick Prescott Filter

- Method to dissociate a time series into a non-linear trend & cycle has be developed by Hodrick & Prescott [1997].

$$y \begin{cases} \mu_t \rightarrow \text{time varying trend} \\ (y_t - \mu_t) \rightarrow \text{cyclic} \end{cases}$$

$$\frac{1}{T}(y_t - \mu_t)^2 + \frac{\lambda}{T} \sum_{t=2}^{T-1} [(\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1})]^2$$

λ = Parameters
$T \rightarrow$ No. of obs.

Minimize "z" w.r.t. parameters in such a way that sum of squares are minimum.

"λ" is cost as penalty parameters

If $\lambda = 0$ then z minimized when $y_t = \mu_t$

If $\lambda \to \infty$ then z minimized when $\Delta \mu_{t,t-1} = \Delta \mu_{t+1,t}$

$\therefore \mu_{t+1} - \mu_t = \mu_t - \mu_{t-1}$

i.e change in trend is constant.

$\therefore$ Linear trend

Smoothing parameter

Generally for
- quarterly data $\to \lambda = 1600$
- Monthly data $\to \lambda = 129600$ [$1600 \times 3 \times 4$]
- Annual data $\to \lambda = 6.25$
- Daily data $\to \lambda = 28,51,200$ [129600×22]

for there are 22 working days

Commands

```
tsfilter hp cycle = excangerate, smooth (129600)
gen trend = exchangerate - cycle
tsline trend
```

Multivariate Time Series Models

Vector - Autoregression [VAR Model]

↓

A matrix with a single column

$$y_t = b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{y_t}$$

$$z_t = b_{20} - b_{21} y_t + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{z_t}$$

→ ε_{y_t} & ε_{z_t} are white noise terms with constant variance σ_y^2 & σ_z^2

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{y_t} \\ \varepsilon_{z_t} \end{bmatrix}$$

$$B x_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$

or

$$x_t = A_0 + A_1 x_{t-1} + e_t \quad ; \quad \begin{cases} B^{-1} \Gamma_0 = A_0 \\ B^{-1} \Gamma_1 = A_1 \\ B^{-1} \varepsilon_t = e_t \end{cases}$$

$$y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{1t}$$

$$z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{2t}$$

where $e_{1t} = \dfrac{\varepsilon_{yt} - b_{12} \varepsilon_{zt}}{1 - b_{12} b_{21}}$

$$e_{2t} = \frac{\varepsilon_{z_t} - b_{21} \varepsilon_{y_t}}{1 - b_{12} b_{21}}$$

$$E(e_{1t}) = \frac{E(\varepsilon_{yt})}{1-b_{12}b_{21}} - \frac{b_{12}\,E(\varepsilon_{zt})}{1-b_{12}b_{12}} = 0$$

$$\therefore \quad var(e_{1t}) = (\sigma_y^2 + b_{12}^2\,\sigma_z^2]^2/(1-b_{12}b_{21})^2$$

Similarly $var(e_{2t})$ can be calculated.

$$\Rightarrow \quad cov(e_{1t}, e_{2t}) = E\left[(\varepsilon_{yt} - b_{12}\varepsilon_{zt})(\varepsilon_{zt} - b_{21}\varepsilon_{yt})\right]/(1-b_{12}b_{21})^2$$

$$= -(b_{21}\sigma_y^2 + b_{12}\sigma_z^2)/(1-b_{12}b_{21})^2$$

Hence,

Σ (variance covariance matrix of e_{1t} & e_{2t}) is denoted as,

$$\Sigma = \begin{bmatrix} var(e_{1t}) & cov(e_{1t}, e_{2t}) \\ cov(e_{2t}, e_{1t}) & var(e_{2t}) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

Concept of Impulse Response Function [IRF] :-

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{10} \\ a_{20} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

After substitution,

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix}$$

we know,

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{(1-b_{12}b_{21})} \begin{bmatrix} 1 & -b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$\therefore \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{(1-b_{12}b_{21})} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

Now let,

$$\phi_i = \frac{A_1^i}{(1-b_{12}b_{21})} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

Then,

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

$\phi_{11}(i)$, - - - are called Impulse Response Function (IRF)

$\phi_{jk}(0)$ are called instantaneous "impact multiplier"

$\phi_{12}(1)$ means (one std. dev. shock) on ε_{zt-1} will have ϕ_{12} amount change in "y_t" at period '1'.

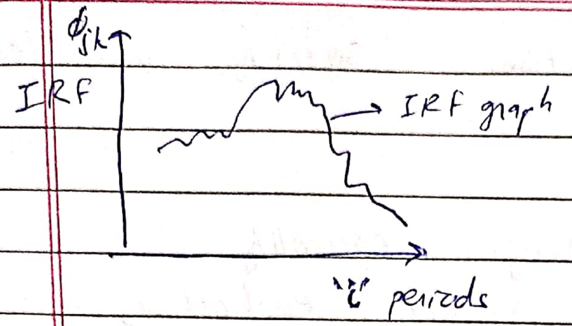
update by '1' period refers to

$\phi_{12}(1)$ as,

impact of ε_{zt} on y_{t+1} and so on ...

Meaning giving shock to ε_{zt} will show effect on y_{t+1}

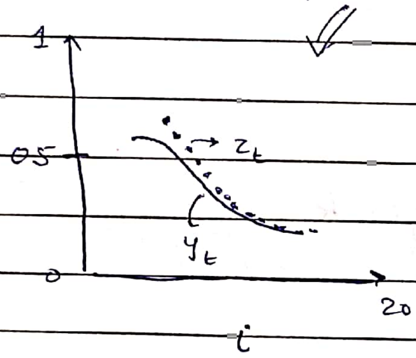
⇒ If IRF is not converging to zero, the shock has permanent effect. [Also shows instability in market].

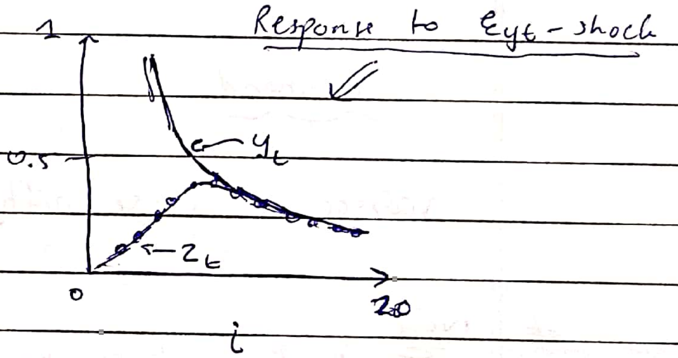


Eg:-

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$$

Response to ε_{zt} - shock





Granger Causality :- [Extension of VAR Model]

→ If all variables are stationary then granger causality is applicable.
[This makes IRF converging]

integrated of order '0' [Shows stationarity]

y_t and $z_t \sim I(0)$

$$y_t = \alpha_1 + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{i=1}^{q} \gamma_i z_{t-i} + \varepsilon_{t_1}$$

$$z_t = \alpha_2 + \sum_{i=1}^{m} \delta_i y_{t-i} + \sum_{i=1}^{n} \xi_i z_{t-i} + \varepsilon_{t_2}$$

→ If $\gamma_i = 0 \quad \forall \; i = 1(1)p$

then $z_t \not\rightarrow y_t$ [z_t doesn't cause y_t]

→ If $\delta_i = 0 \quad \forall \; i = 1(1)m$

then $y_t \not\rightarrow z_t$ [y_t doesn't cause z_t]

If above 2 conditions satisfy, variables are independent.
But, if not,
it is a granger causality.
If both y_t & z_t impact eachother,
it is Bidirectional Causality (Feedback Method)
If only one variable impacts the other,
it is unidirectional Causality

Command

varsoc close_nifty excangerate

Note:-

→ rely for small samples

Statistical Criteria:
- AIC → Akaike Information Criteria
- SBIC →
- HQIC →
- FPE →

→ Stas marked values shows optimal lags to be chosen as per criteria

Command

order of variable not important

(IRF Function) ⇒ varbasic close-nifty exchangerate, lags(4)

Varbasic close-nifty excangerate, lags(1/4)

to include all 4 lags into the model

var → only estimation

Varbasic → estimation + IRF graph (impulse response func^n)

Command

varsoc close_nifty excangerate

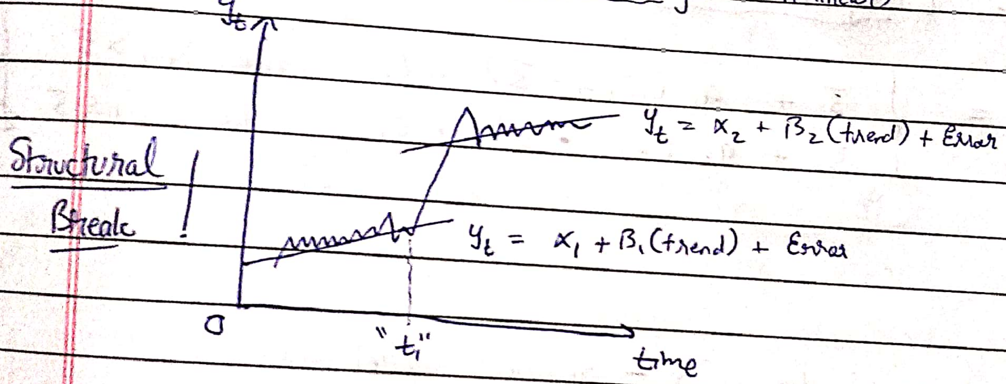
var close_nifty excangerate, lags(1/4)

(forecast) fcast compute m1_, step(4)

Non-Linear Time Series Models:-

[Non-linearity doesn't necessarily mean a curve]

Type - I [Piecewise - linearity] ← (Non-linear)



$\alpha_1 \neq \alpha_2$ [parameters changes after a certain time period, "t_i"]

$\beta_1 \neq \beta_2$ [Slope coefficient may also vary]

Type - II [Threshold Type (Model)] [Piecewise]

$y_t = 0$ is the threshold

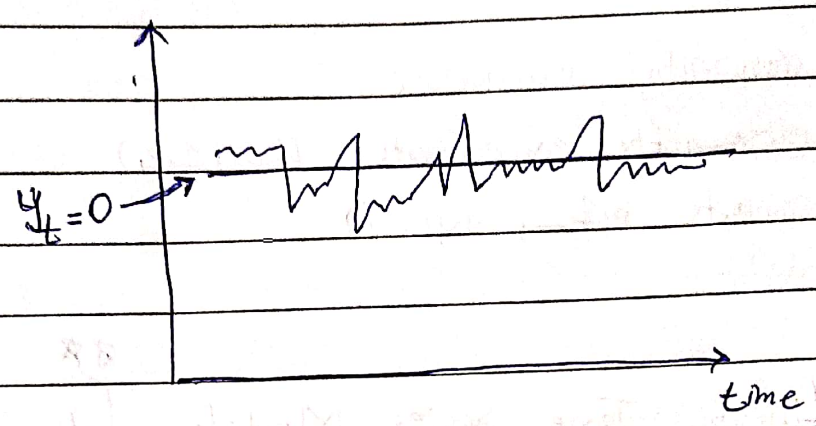
$$\begin{cases} y_t = \alpha_1 y_{t-1} + \varepsilon_{1t} \rightarrow \text{if } y_{t-1} > 0 \\ y_t = \alpha_2 y_{t-1} + \varepsilon_{2t} \rightarrow \text{if } y_{t-1} \leq 0 \end{cases}$$

$y_t = \alpha_1 I y_{t-1} + \alpha_2 (1 - I) y_{t-1} + \varepsilon_t$ → Combined

$$\left.\begin{cases} \text{if } & y_{t-1} > 0 \Rightarrow I = 1 \\ \& & y_{t-1} \leq 0 \Rightarrow I = 0 \end{cases}\right\} \text{Dummy variable.}$$

[P.T.O]

Threshold Model



Type – III [Regime Model] [Again piecewise]

2-regime Model

$$S_t = \begin{cases} \bar{S} + a_1 (S_{t-1} - \bar{S}) + \varepsilon_t & \text{when } S_{t-1} > \bar{S} \\ \bar{S} + a_2 (S_{t-1} - \bar{S}) + \varepsilon_t & \text{when } S_{t-1} \leq \bar{S} \end{cases}$$

→ There can be 3 – regime model :-

$$S_t = \bar{S} + a_1 (S_{t-1} - \bar{S}) + \varepsilon_t \quad \text{when } S_{t-1} > \bar{S} + c$$

$$S_t = S_{t-1} + \varepsilon_t \quad \text{when } \bar{S} - c < S_{t-1} \leq \bar{S} + c$$

$$S_t = \bar{S} + a_2 (S_{t-1} - \bar{S}) + \varepsilon_t \quad \text{when } S_{t-1} \leq \bar{S} - c$$

where, 'c' is some specific constant value

⇒ Here all three models are overall non-linear [not curves but piecewise functions]

Detecting Auto correlation func" of non-linear model:-

→ Assuming 'chaos' model,

$$y_t = 4y_{t-1}[1-y_{t-1}] \quad \text{for} \quad 0 < y_t < 1 \quad -(1)$$

If we get ACF of y_t, we will have small & insignificant values, but ACF of $\{y_t^2\}$ will be large.

If $y_1 = 0.7$, then next 99 values can be generated using (1) & autocorrelations will be :-

r_1	r_2	r_3	r_4	r_5	r_6
-0.074	-0.072	0.008	0.032	-0.016	-0.030

(Univariate)

But ACF of y_t^2 & y_{t-1}^2 is -0.281

y_t^3 and y_{t-1}^3 is ~~0.035~~ -0.386

So, looking at square & cubic values ACF, we can have non-linear ~~&~~ relation.

Ljung - Box Test :-

→ This test helps us to detect any non-linearity in model through the error terms behaviour.

(Univariate)

Let, P_i denotes correlation b/w $\hat{e}_t^2$ & $\hat{e}_{t-i}^2$ then Ljung Box test statistics is denoted by,

$$Q = T[T+2] \sum_{i=1}^{n} P_i/(T-i) \quad \}$$

$T \rightarrow$ No. of observation

$n \rightarrow$ no. of lags chosen

Alternatively,

$$e_t^2 = \alpha_0 + \alpha_1 \hat{e}_{t-1}^2 + \alpha_2 \hat{e}_{t-2}^2 + \cdots v_t$$

If $\alpha_i = 0 \quad \forall \quad i=(1)(1)n$

then there is no non-linearity in the error-square terms, hence no non-linearity in y_t too.

Command [Important]

wntestq (excangerate), lags(2) ← [significant if Prob > chi2(2) = 0.500]

we get lag 1 [AR 1] → pac excangerate ← [we see first lag significant]

wntestq (excangerate), lags (1) ← [significant as prob > chi2(1) : 0.00]

ac excangerate ← [we get lag 1, 2 & 3 as significant] [MA 1, 2 & 3]

arima excangerate, arima (1, 0, 3] [Considering all 3 lags]

predict exhat

tsline exhat excangerate

gen error = excangerate - exhat

tsline error ← [we can see that volatility of error is increasing]

gen vol = error^2

tsline vol

[Now we regress vol on its past values]

pac vol → [No lags significant] [Because growth in vol is very smoothly & not sharply Hence, with higher variation it ain't significant.]

Smooth Transition Autoregressive Model - [STAR Model]

→ This model allows autoregressive parameters to change slowly.

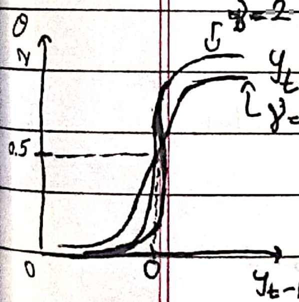
↳ if this becomes zero, it is normal AR-I process.

$$y_t = a_0 + a_1 y_{t-1} + \left(\beta_1 y_{t-1} f(y_{t-1})\right) + \varepsilon_t$$

→ $f(\cdot)$ is a smooth continuous funcⁿ.

→ AR parameter $(a_1 + \beta_1)$ will change smoothly along with y_{t-1}

⟹ Example 1



LSTAR Model [Logistic - STAR Model]

$$y_t = a_0 + a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p}$$

$$+ \theta\left[\beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p}\right]$$

$$+ \varepsilon_t \quad -(i)$$

$$\theta = 1 + \exp\left[-\gamma\left(y_{t-1} - c\right)\right]^{-1} \; ; \gamma > 0$$

[γ → Shows responsiveness based on past data]

⟹ If $\gamma \to 0$ or ∞ then,

(i) becomes an AR(p) model

⟹ If $0 < \gamma < \infty$ then,

y_t's decay depends on y_{t-1}

⟹ As $y_{t-1} \to -\infty$, $\theta \to 0$ then

y_t becomes, $a_0 + a_1 y_{t-1} + \cdots a_p y_{t-p} + \varepsilon_t$

→ As $y_{t-1} \to \infty$ and $\theta \to 1$ then,

$$y_t = (a_0 + \beta_0) + (a_1 + \beta_1) y_{t-1} + \cdots + \varepsilon_t$$

$\boxed{a_0 \to (a_0+\beta_0)}$ $\boxed{a_1 \to (a_1+\beta_1)}$

Thus, intercept and co-efficients smoothly change b/w $\underbrace{\text{two values}}_{\text{extreme}}$ as y_{t-1} changes.

⟹ Example 2

ESTAR Model [Exponential STAR Model]

$$\theta = 1 - \exp[-\gamma (y_{t-1} - c)^2] \; ; \; \gamma > 0$$

$$y_t = a_0 + a_1 y_{t-1} + \dots + a_p y_{t-p} + \theta\,[\beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p}] + \varepsilon_t$$

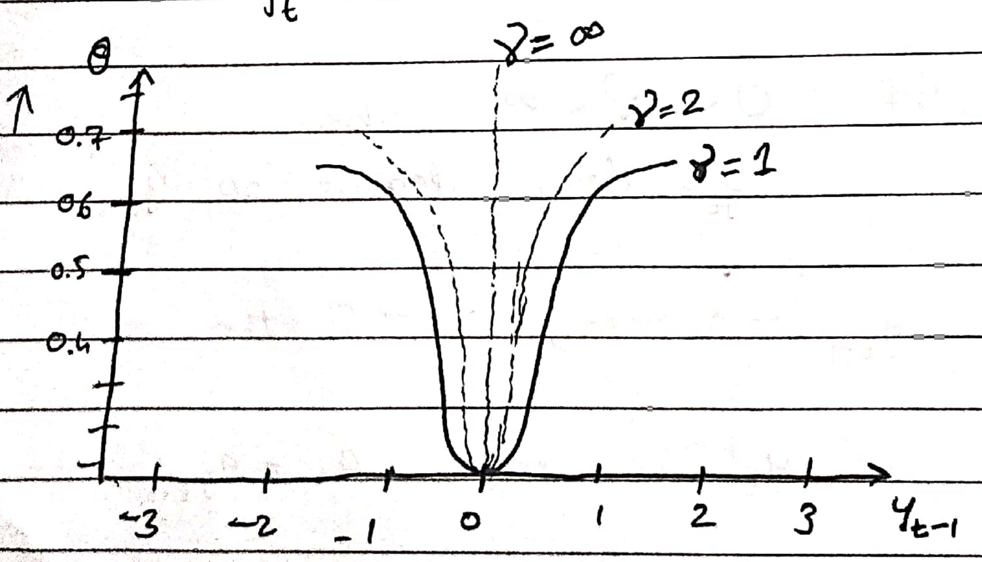
→ when, $y_{t-1} \to c$ then

$\theta \to 0 \Rightarrow y_t$ becomes AR(p) Model.

→ As y_{t-1} moves away from c,

$\theta \to 1$ and

$$y_t = (a_0 + \beta_0) + (a_1+\beta_1) y_{t-1} + \dots + \varepsilon_t$$



(Raj Rajaratnam & Rajat Gupta)

☆☆ [Goldman Sachs Insider Trading Case]

classmate
Date ____
Page ____

Markov Regime Switching Model :- [Transitions are swift & random [contrary to regime model]]

→ In general threshold models, regimes switch from one to another on the dynamics of y_{t-d}.
If y_{t-d} exceeds any specific threshold then system identifies itself with one regime.

⇒ EMH (Efficient Market Hypothesis)

→ Market assimilates all past records / experiences (or all available information) into its current price.

→ Directly implies that it is impossible to beat the market.

⇒

$$y_t = a_{10} + a_1 y_{t-1} + \varepsilon_{1t} \quad \text{if in regime-1}$$

$$y_t = a_{20} + a_2 y_{t-1} + \varepsilon_{2t} \quad \text{if in regime-2}$$

If p_{11} = probability that y_t will be in regime-1 when it's past value y_{t-1} is also in regime-1

So, $(1-p_{11})$ = " " y_t " " " regime-2
" " " " y_{t-1} is in regime 1.

Similarly for p_{22},

p_{22} → y_t in regime-2 & y_{t-1} in regime-2

$(1-p_{22})$ → y_t in regime-1 & y_{t-1} in regime-2

⇒ This model is generally used to examine conditional volatility of models

[P.T.O]

[Atheoretical Model]
↳ Difficult to justify a theory based on model

Artificial Neural Network Model :-

→ ANN can be useful non-linear processes that have an unknown funcn form -

A simple ANN model is -

$$y_t = a_0 + a_1 y_{t-1} + \sum_{i=1}^{n} \alpha_i f_i(y_{t-1}) + \varepsilon_t$$

$f_i(y_{t-1})$ → cummulative distribution or a logistic funcn such as -

LSTAR Model

For the logistic funcn -

$$y_t = a_0 + a_1 y_{t-1} + \Sigma \alpha_i \left[1 + \exp(-\gamma_i (y_{t-1} - c)) \right]^{-1} + \varepsilon_t$$

Though ANN & LSTAR model are similar but there are subtle differences

① ANN allows only the intercept to be time varying; but AR process considers a_1 as constant.

② ANN use 'n' different logistic funcn (called nodes)

This leads to complex and over parametrized model.

Volatility Models:-

→ Concept of conditional and unconditional volatility in stock market.

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

Conditional : $E(y_t) = E(a_0) + a_1 E(y_{t-1}) + \underbrace{E(\varepsilon_t)}_{(=0)}$

Variance

[Old values donot impact future value] ← (Only yesterday's value impacts)

$\therefore var(y_t) = E[y_t - E(y_t)] = \sigma^2$

Unconditional : $y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$

Variance

[Old values have significant effect] $y_t = a_0 + a_1(a_0 + a_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$

$$= a_0(1 + a_1 + \cdots) + a_1^t y_0 + \sum_{i=1}^{t} a_1^i \varepsilon_t$$

$$\therefore E(y_t) = \frac{a_0}{1-a_1} + a_1^t y_0 + 0$$

$$\therefore var(y_t) = \frac{\sigma^2}{1-a_1^2}$$

$|a_1| < 1$

$\therefore$ Unconditional Variance > Conditional Variance.

ARCH [Auto regressive Conditional Heteroskedasticity] Model:-

$y_{t+1} = \varepsilon_{t+1} x_t$ (written in product form)

If we take n-log or ln

$$\ln y_t = \ln \varepsilon_{t+1} + \ln x_t$$

$var(y_{t+1} | x_t) = x_t^2 a^2$ for given value of x_t

'Conditional variance of y_t for a given x_t'

Heteroskedasticity → square of error term is not constant.
Homoskedasticity → square of error terms are constant.

⇒ If $x_t \uparrow$ then $var(y_{t+1} | x_t)$ will also increase.

→ Now, assuming an ARMA process,

$$y_t = a_0 + a_1 y_{t-1} + \varepsilon_t$$

$$y_{t+1} = a_0 + a_1 y_t + \varepsilon_{t+1}$$

$$var(y_{t+1} | y_t) = E_t(\varepsilon_{t+1})^2 = \alpha^2 \text{ [under assumption]}$$

If suppose,

$E(\varepsilon_t^2)$ is not α^2 (constant) & varies over time

then,

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + V_t$$

V_t = White noise term.

$$\therefore E_t(\hat{\varepsilon}_{t+1}^2) = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \dots + \alpha_v \hat{\varepsilon}_{t-v}^2$$

⇒ Engle (1982) proposed first form of ARCH model

$$\varepsilon_t = V_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$$

V_t → white noise term
where $E(V_t) = 0$ & $var(V_t) = 1$

$$E(\varepsilon_t) = E(V_t \sqrt{\alpha_0 + \alpha_1 \varepsilon_{t-1}^2})$$

$$= E(V_t)\, E(\alpha_0 + \alpha_1 \varepsilon_{t-1}^2)^{1/2} = 0$$

$$E(\varepsilon_t \varepsilon_{t-i}) = 0 \quad \forall i \neq 0$$

$$E(\varepsilon_t^2) = E(v_t^2] \; E[\alpha_0 + \alpha_1 \varepsilon_{t-1}^2]$$
$$= 1 \cdot \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \quad [\text{for given } \varepsilon_{t-1}^2]$$

$$\therefore E[\varepsilon_t^2] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Precisely,

$$E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2$$

Conditional volatility ε_t^2 will increase if past value (ε_{t-1}^2) is high & so on.
This means that 'α' terms are high & strong. This phenomenon is called 'Volatility Persistance.'

⇒ A Way to fit ARCH Model

① ε is an OLS estimation ~~on the~~ model
② Extract errors
③ Square the errors
④ Regress

$$\hat{\varepsilon}_t^2 = \alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \cdots + \alpha_p \hat{\varepsilon}_{t-p}^2$$

If $\alpha_1 = \alpha_2 = \cdots = 0$
then there is no ARCH Effect.

i.e Past errors values are not significantly effecting the present errors value.

MLE → Maximum Likelihood Estimation

↳ Generate likelihood function ← maximize this function

GARCH Model – [Generalized ARCH Model]

Bollerslev (1986) extended Engle's model –

$$\varepsilon_t = \nu_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i h_{t-i}$$

ν_t → white noise term

This is referred to as GARCH (p, q) model.

** → If $q=1$ & $p=0$ then GARCH (p, q) becomes ARCH Model.

⇒ Fitting GARCH Model

① Estimate y_t using best fit ARMA model

Extract residual from model $\{\hat{\varepsilon}_t^2\}$

Calculate

$$\hat{\sigma}^2 = \frac{\sum \varepsilon_t^2}{T} \leftarrow \text{no. of obs.}$$

② Calculate ACF of $\hat{\varepsilon}_t^2$

$$\rho_i = \frac{\sum\left[(\hat{\varepsilon}_t^2 - \sigma^2)(\hat{\varepsilon}_{t-i}^2 - \sigma^2)\right]}{\sum(\hat{\varepsilon}_t^2 - \hat{\sigma}^2)^2}$$

at lag i

3. In large sample, standard dev. of r_i can be approximated as $(1/T^{0.5})$

Q statistic $\rightarrow T(T+2) \sum_{i=1}^{n} r_i^2/(T-i)$

→ If Q – statistic is significantly different from '0' then reject the null.

Null hypothesis is $r_i \; \forall i$ are insignificant [i.e No GARCH effect]

So, we can detect presence or absence of GARCH effect.

i.e ε_t^2 and ε_{t-i}^2 at different lags or 'i' values not strongly connected.

i.e There is no GARCH effect.