

Return:

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

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

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

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

→ We strongly assume a "normality in Financial Variables." [As assumed, that market can't be beaten in long run] i.e follows a normal distribution.

[Mean ~~is~~ ^{an} Avg. return = 0] → For Probability density func 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 or there is a confusion in market. [Risk too is high as graph is more scattered.]

Deviations from normality

→ 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 $X = m_k = E[X^k]$, e.g. expectation $\mu = E[X]$
 k^{th} central moment of $X: \mu_k = E[(X-\mu)^k]$, e.g. variance $\text{Var}(X) = E[(X-\mu)^2]$

→ 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

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

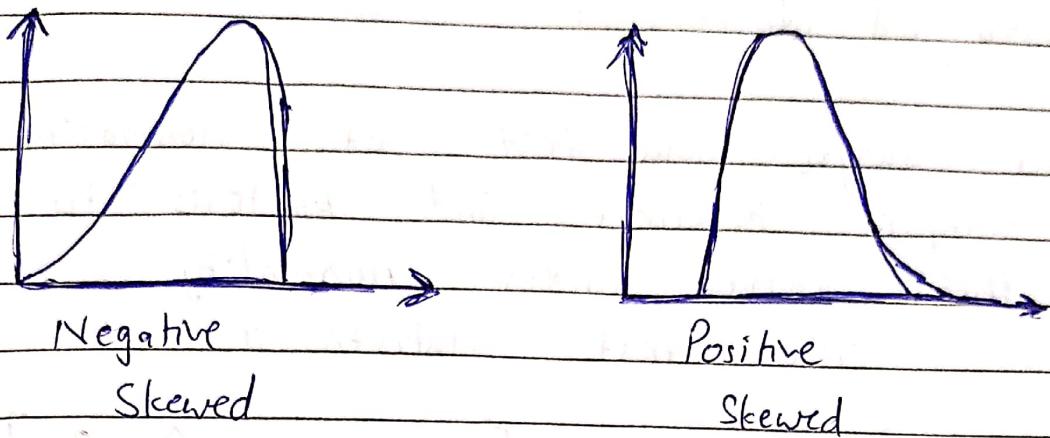
→ Any random variable with symmetric dist. will have skew = 0.

→ 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

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

(Unitless)

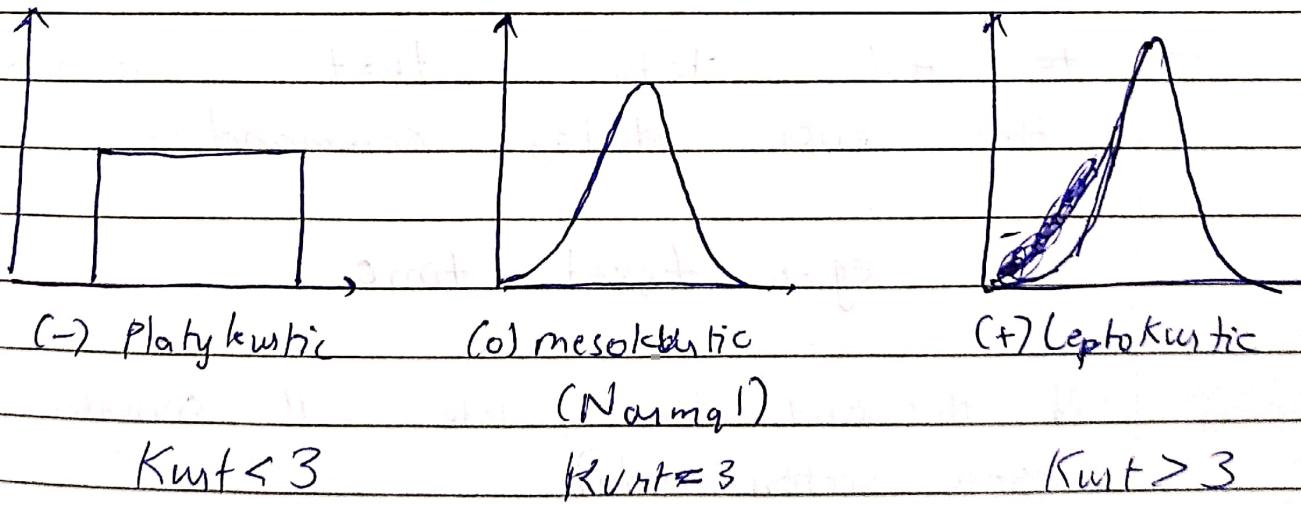


Kurtosis :-

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

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

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



Test 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:

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

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 :-

Eg:- `(gen) (time)=(-n)`

generate var. 5 var. name
start from value 1 to end

- 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

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

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

Q. close-nifty
means "lag
value" of
close-nifty

classmate

Date _____

Page _____

→ gen return = $\ln(\text{close-nifty}/\text{Q. close-nifty}) * 100$

→ '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. use command 'detail' with sum & variable name.

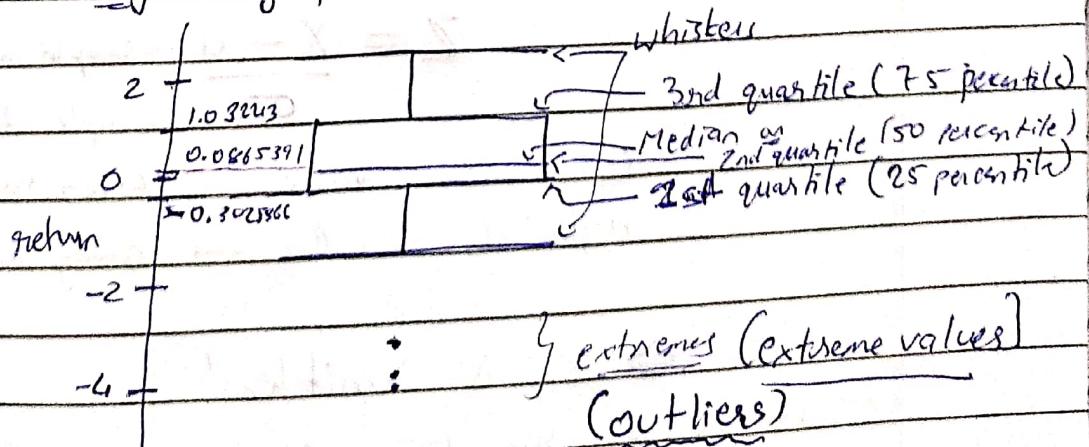
e.g.: sum return, detail
sum (return), detail

→ To ~~understand~~ understand any command use 'findit commandname' to open documentation & understand it.
e.g. findit sum

→ 'histogram varianlename' to get histogram graph figure of ~~the~~ respective variable.
e.g. histogram return

→ 'graph box return' to get box plot of return variable.

e.g.: graph box return



JMP →

Borders of whiskers

$$q_1 - 1.5(iqr)$$

$$q_3 + 1.5(iqr)$$

(iqr) → inter quartile range

q_3 → 3rd quartile (75 percentile)

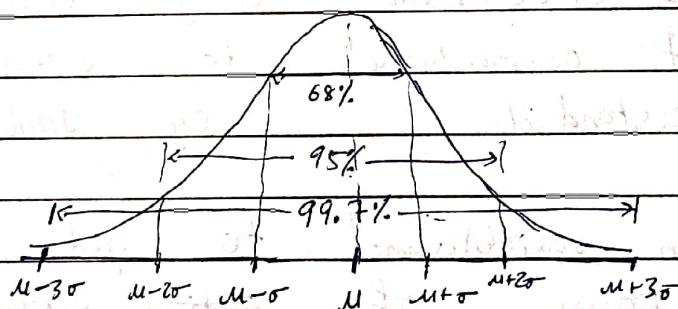
$$[q_3 - q_1]$$

q_1 → 1st quartile (25 percentile)

⇒ stata uses

$$\left(\begin{array}{l} q_1 - (iqr) \\ q_3 + (iqr) \end{array} \right)$$

Normal Distribution =



mean → μ

std. dev → σ

Z - statistic

(Population Based) [mean = 0, std.dev = 1]

value of var.

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

abs.

sample mean

lower abs. z value → closer to mean

High abs. z value → away from mean

⇒ Z is 'unitless'

→ to see if var. res 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)

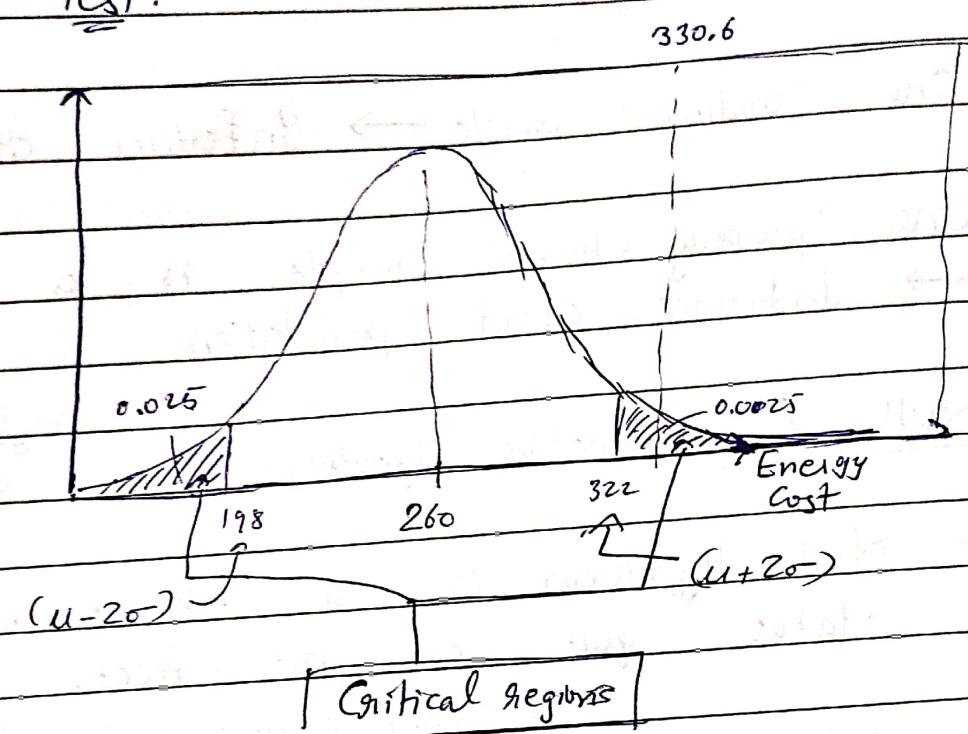
Eg:- You release a truly guilty person!

• Power of a test = $1 - (\text{Type-II error})$

probability of avoiding a (type-II error).

* Two - Tail Test :-

Density



- 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 practice)

$|z|$ is in range of -2 to 2.
 $|z| < 2 \rightarrow 5\% \text{ significance level}$

t-test statistics

(Sample Based)

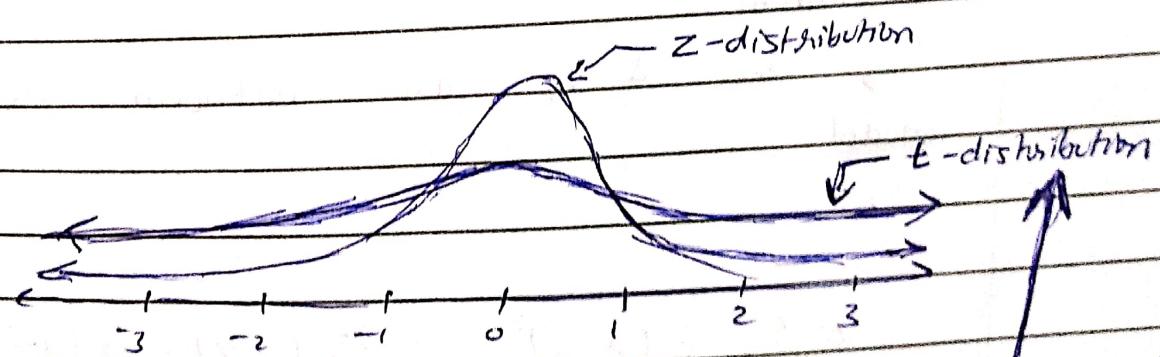
$$t = \frac{\text{obs. value} - \text{mean}}{\text{theoretical value / mean}}$$

s/\sqrt{n} - var. set size
std. dev.

2-t thumb of rule

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

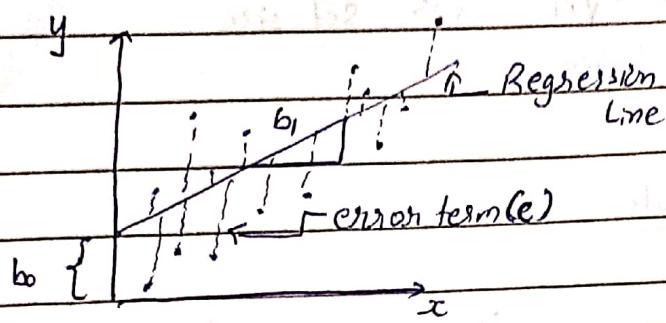
t-distribution



Normal
curve
with fat tail

* Notion of Linear Regression Model -

$$y = b_0 + b_1 x$$



RSS → sum of squares of errors

[Residual Sum of Squares]

→ We have data points for y & x but what we require is ' b_0 ' & ' b_1 '. We require b_0 & b_1 such that 'RSS' is minimum.

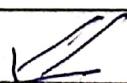
So, b_0 & b_1 are unknown parameters in my model.

$$\sigma^2 = \frac{\sum (x - \bar{x})^2}{n} \quad \text{Cov}(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n}$$

Derivation data points values on the line

$$Q = \sum_{i=1}^n (y_i - \hat{y})^2 = \sum_{i=1}^n (y_i - a - b x_i)^2$$

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



$$a = \bar{y} - b \bar{x}$$

$$\frac{\partial Q}{\partial b} = \sum_{i=1}^n -2 x_i (y_i - a - b x_i) = \sum_{i=1}^n -2 (x_i y_i - a x_i - b x_i^2) = 0$$

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

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

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

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

Regressum in stata

regress close_nifty excangrate
 predict niftyhat

tstat close_nifty niftyhat

scatter close_nifty excangrate || lfit close_nifty excangrate

$$\rightarrow y = \alpha + \beta x + (\epsilon)$$

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

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

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

$$E(y - \bar{y})^2 = E[\beta(x - \bar{x})^2] + E(\epsilon^2)$$

$$TSS = ESS + RSS$$

$$\text{Total Sum of Squares} = \text{Explained Sum of Squares} + \text{Residual Sum of Squares}$$

\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 & closer to $\underline{\underline{1}}$.

\Rightarrow So now making a new model,

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

\downarrow
Now we add previous day data to improve model

- l. \rightarrow lag of $\rightarrow (T-1)$
- l2. \rightarrow second lag of $\rightarrow (T-2)$
- l3. \rightarrow 3rd lag of $\rightarrow (T-3)$

classmate

Date _____
Page _____

\rightarrow Command.

regress close-nifty l.close-nifty exchangeRate
predict niftyhat2

\Rightarrow Root Mean Square Errors

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

should be as small as possible (desirable)

\rightarrow Command

gen error = (close-nifty - niftyhat)

gen sgerror = (error)^2

sum (sgerror)

gen rmse = (21338)^0.5

\Rightarrow Stationarity [3 properties]

- \rightarrow Mean is constant
- \rightarrow Var. is ~~finite~~ finite & time independent.
- \rightarrow Auto-covariance is constant & time independent
[Derivation ahead \rightarrow]

Derivation of Stationarity

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

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

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

↑ order of
regression

From AR-1 process,

$$y_t = \alpha + \beta(\alpha + \beta y_{t-2} + \epsilon_{t-1}) + \epsilon_t \quad (\text{Iteration})$$

$$y_t = \alpha + \beta\alpha + \beta^2 y_{t-2} + \beta \epsilon_{t-1} + \epsilon_t$$

$$= \alpha + \beta\alpha + \beta^2(\alpha + \beta y_{t-3} + \epsilon_{t-2}) + \beta \epsilon_{t-2} + \epsilon_t$$

$$= \alpha + \beta\alpha + \beta^2\alpha + \beta^3 y_{t-3} + \beta^2 \epsilon_{t-2} + \beta \epsilon_{t-2} + \epsilon_t$$

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

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

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

Mean of y_t will converge to a constant A
Because

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

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

$$= A + 0$$

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

Random Walk Model

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

If suppose $a_1 = 1$ then, [Strong Assumption]

$$y_t = a_0 + y_{t-1} + \epsilon_t \quad [\text{Now this Becomes Random Walk Model}]$$

$$y_t = a_0 + (a_0 + y_{t-2} + \epsilon_{t-1}) + \epsilon_t \quad (\text{iteration})$$

$$y_t = a_0 + a_0 + y_{t-2} + \epsilon_{t-1} + \epsilon_t$$

After 't' substitution,

Random walk Model $\Rightarrow y_t = \underbrace{t a_0 + y_0}_{\downarrow} + \sum_{i=0}^t \epsilon_{t-i} \rightarrow ①$

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

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

$$= A + 0 \quad [\text{By assumption}] \quad \begin{matrix} 1 \text{ to } t \\ \downarrow \\ \text{With prob 1} \end{matrix}$$

$$E(\epsilon_{t-i}) = 0 \quad \forall i, i = 1(1)t$$

Variance of y_t (when y_t is random variable)

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

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

From (1) we can write,

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

$$\left[\begin{array}{l} \text{var}(x+y) = \text{var}(x) + \text{var}(y) + 2 \text{cov}(x, y) \\ \text{cov}(x, y) = E(xy) - E(x)E(y), \\ \text{if } (x, y) \text{ are independent then, } \text{cov}(x, y) = 0 \end{array} \right]$$

Hence,

$$\begin{aligned} \text{var}(y_t) &= \text{var}(A) + \text{var}(\sum \varepsilon_{t-i}) \\ &= 0 + \sum \text{var}(\varepsilon_{t-i}) + \text{cov. terms} \end{aligned}$$

$$\text{var}(y_t) = t\sigma^2 \quad \left(\begin{array}{l} \text{et-i \& et-j are independent} \\ \text{for all } i \neq j \end{array} \right)$$

Second condⁿ of stationarity violated.

(Assumed)

$$\left(\begin{array}{l} \text{Homoscedasticity is a cond' in} \\ \text{which variance of residual} \\ \text{at every term, in a regression} \\ \text{model is constant} \end{array} \right)$$

Now,

$$\text{Cov}(x, y) = E(xy) - E(x)E(y)$$

If we want to see covariance b/w

y_t & $y_{t-\tau}$ [τ is some lag value and $\tau \neq 0$]

then,

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

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

$$\bar{y}_{t-\tau} = \bar{y}_t (= u) \quad \text{(Same)}$$

$$\text{Cov}(y_t, y_{t-\tau}) = E[(y_t - u)(y_{t-\tau} - u)]$$

$$= \gamma_\tau \quad \left\{ \begin{array}{l} \text{Autocovariance at } "T" \\ \text{lag} \end{array} \right\}$$

* Note:- $\gamma_0 (\tau=0) \Rightarrow$ Autocovariance at lag 0
is nothing but variance

So, from ①,

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

So, third condⁿ of stationarity violated.

\Rightarrow So, variance & auto covariance 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.

$$Y_t = \alpha_0 + \beta Y_{t-1} - \textcircled{1} \quad [\varepsilon \text{ is not part of eqn}]$$

\rightarrow If t is discrete i.e. $\pm 1, \pm 2, \pm 3, \dots$

then eq $\textcircled{1}$ is called difference equation.

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

then eq $\textcircled{1}$ becomes differential equation.

Here, we assume 't' is discrete.

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

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

$$\frac{a_1}{2} - \bar{y} = 1$$

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

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

$$1. Ab^t (b-5) = 0$$

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

$$\therefore y_t = PI + CF$$

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

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

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

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

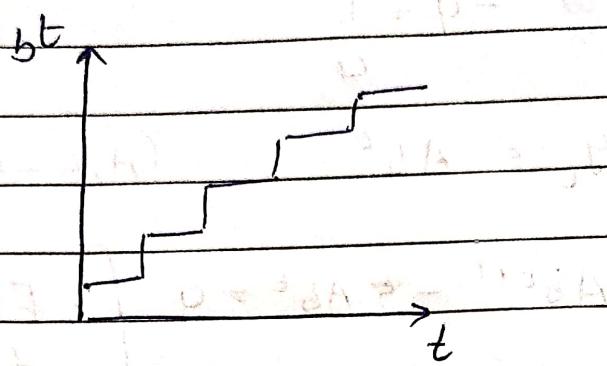
$$\therefore \boxed{A=2}$$

So solution is

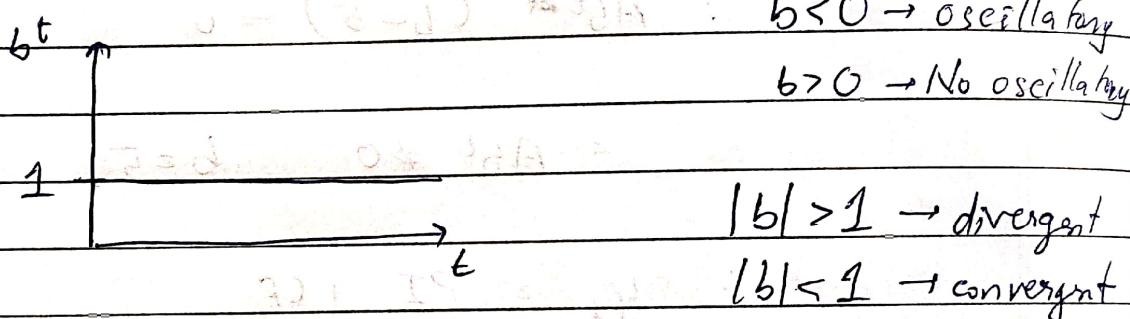
$$\therefore \boxed{y_t = \left(-\frac{1}{4}\right) + 2(5)^t}$$

$$\Rightarrow \text{In solution of } y_t = A b^t + \bar{y} \quad P.I$$

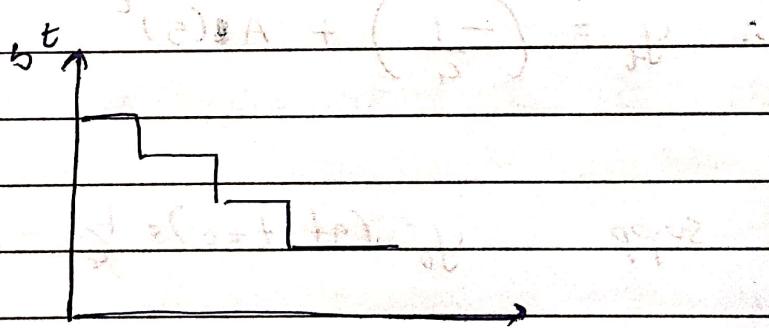
\rightarrow If $b > 1$,



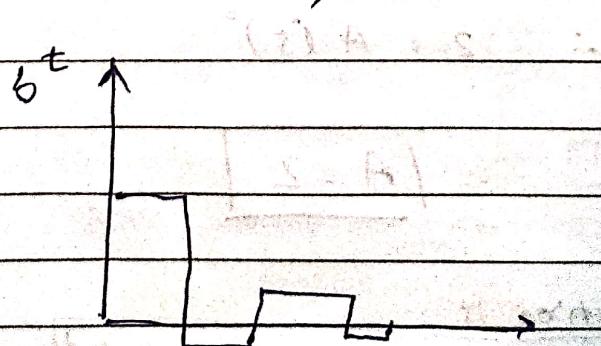
\rightarrow If $b = 1$,



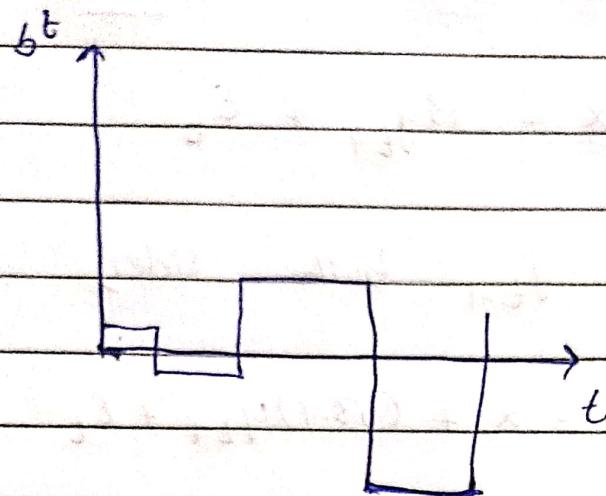
\rightarrow If $0 < b < 1$



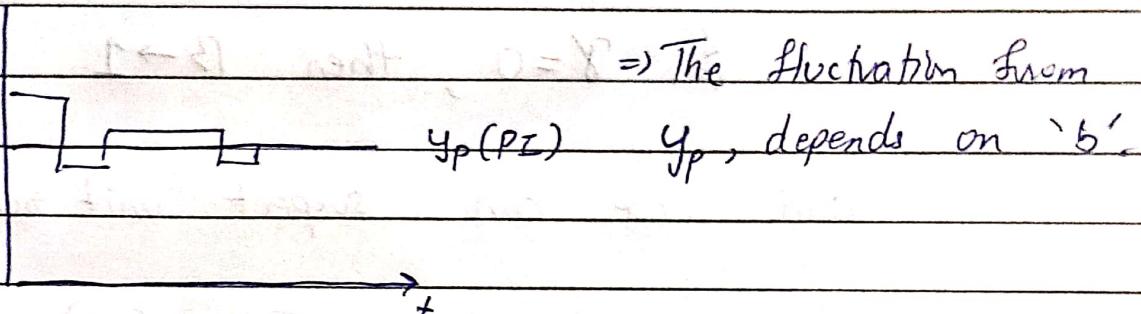
\rightarrow If $-1 < b < 0$



\rightarrow If $b < -1$



$$\Rightarrow \text{So, } y_t = A b^t + y_p(\text{PI})$$



\Rightarrow Stationarity of financial time series

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

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

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

If $|\beta| = 1$ then it is Random Walk Model, which is not ~~ever~~ convergent.

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

* Practical way to check stationarity:-

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

Subtracting y_{t-1} both sides,

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

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

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

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

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.exchange-rate l.exchange-rate

regress d.close-nifty l.close-nifty if t < 40

Concept of ACF (Auto correlation Funcⁿ) :-

→ Correlation defines linear association b/w two variables

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

ACF of order ① mean correlation shows y_t & y_{t-1}
 ACF " " " " " " " " y_t & y_{t-2}
 and so on.

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

$$y_t = q_0 \sum_{i=0}^{t-1} q_i^i + a_i t y_0 + \sum_{i=0}^{t-1} q_i^i \epsilon_{t-i}$$

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

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

If $t \rightarrow \infty$ & $|a_i| < 1$ then

$$\lim_{T \rightarrow \infty} \bar{y}_T = \frac{a_0}{1-a_1} + \sum_{i=0}^{\infty} a_i^i \varepsilon_{t-i}$$

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

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

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

$$E(y_t) = E(y_{t-s}) = \frac{q_0}{1-q_1} = u$$

$$\text{Variance} \quad E(y_t - \mu)^2 = E[(\varepsilon_t + a_1 \varepsilon_{t-1} + a_2^2 \varepsilon_{t-2} + \dots)^2]$$

$$= \sigma^2 [1 + q_1 + q_1^2 + \dots]$$

$$\boxed{\text{Variance} = \frac{\sigma^2}{[1 - q_1^2]}} \quad \left\{ \because |q_1| < 1 \right\}$$

Similarly,

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

$$= E \left\{ [e_t + a_1 e_{t-1} + a_1^2 e_{t-2} + \dots] \right.$$

$$\left. [e_{t-s} + q_1 e_{t-s-1} + q_1^2 e_{t-s-2} + \dots] \right\}$$

$$= \sigma^2 (q_1)^s [1 + q_1^2 + (q_1)^4 + \dots]$$

$$= \frac{\sigma^2 q_1^s}{1 - q_1^2}$$

Now, if $s = 1$,

$$\text{Auto covariance} = \frac{\sigma^2 a_1}{1 - q_1^2}$$

If $s = 2$,

$$\text{''} = \frac{\sigma^2 a_1^2}{1 - q_1^2}$$

\Rightarrow

So,

$$\text{ACF} = \frac{\text{Autocovariance}}{SD_{y_t} SD_{y_{t-s}}}$$

[Auto Correlation Funcn]

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

van y_t

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

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

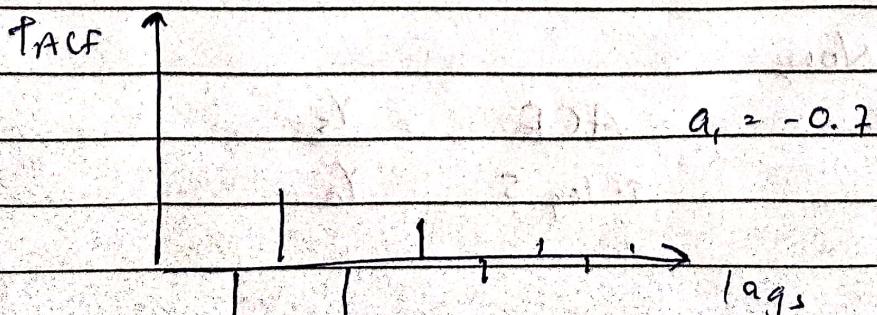
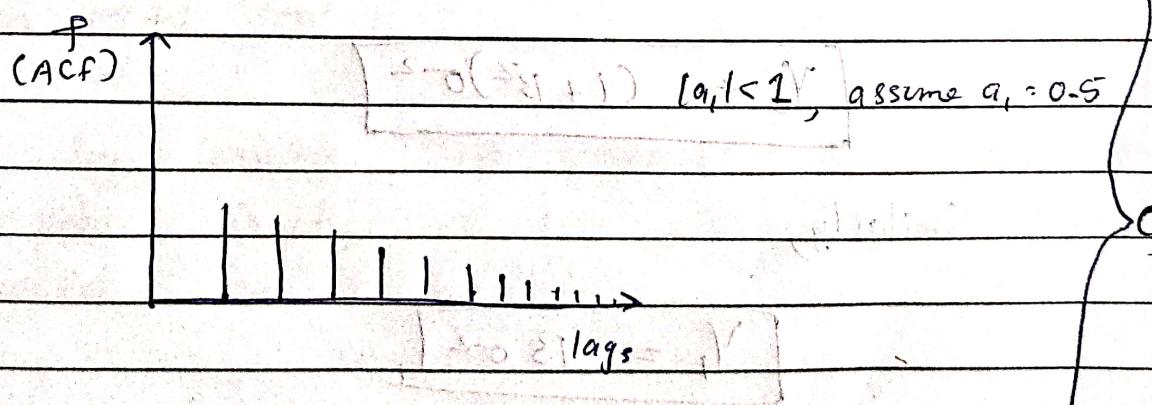
$$\boxed{\text{ACF at } l = \frac{\gamma_s}{\gamma_0} = \frac{\gamma_s}{\gamma_0} = a_1^s}$$

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

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

$$\boxed{\gamma_2 = \frac{\gamma_2}{\gamma_0} (= a_1^2)}$$

So, drawing graph,



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

classmate

Date _____
Page _____

Commands

ac closenisty

⇒ Derive the ACF of an AR(2) process,

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \varepsilon_t$$

Concept of MA [Moving Average]:-

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

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

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

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

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

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

Similarly,

$$\gamma_1 = \beta \sigma^2$$

Now,

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

$\therefore A(s=1)$

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

$$\left| \begin{array}{l} \gamma_1 = \frac{\gamma_1}{\gamma_0} = \beta \\ \gamma_0 = 1 + \beta^2 \end{array} \right.$$

$$\left| \begin{array}{l} \gamma_0 = \frac{\gamma_0}{\gamma_0} = 1 \\ \gamma_0 = 1 + \beta^2 \end{array} \right.$$

If we calculate γ_2 onwards then,

$$\left| \begin{array}{l} \gamma_i = 0 \quad \forall i \geq 2 \end{array} \right.$$

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

$$\gamma_0 = 1$$

$$\gamma_1 = -0.469$$

$$\gamma_2 (\text{onwards}) = 0$$

For MA-1 \Rightarrow It will get one spike in graph

for MA-2 \Rightarrow .. " two " \therefore think!

for MA-3 \Rightarrow .. " three " ..

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

Notes

→ 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.

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

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

$$Gy_t - y_t = \alpha + \beta y_{t-1} + \gamma y_{t-2} + \delta e_{t-1} + e_t$$

ARMA (2, 1)

[Since AR-2 & MA-1 comb]

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 Funcⁿs

→ If y_t follows an AR(1) structure then, y_t & y_{t-2} is related via y_{t-1} . So, when we calculate reln 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.

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

$$y_t^* = y_t - \bar{y}_t \quad | \quad \text{i.e. } y_t - \bar{y}_t = [y_{t-2} - \bar{y}_{t-2} (\approx \bar{y}_t)]$$

$$y_{t-2}^* = y_{t-2} - \bar{y}_t \quad | \quad (y_t - \bar{y}_t) = \alpha + \beta (y_{t-2} - \bar{y}_t) + \varepsilon_t$$

$$\bar{y}_2 \approx \bar{y}_t \quad | \quad \boxed{\text{de-meaned values}} \quad | \quad (y_t^*) = \alpha + \beta (y_{t-2}^*) + \varepsilon_t$$

⇒ ' β ' is PAC b/w y_t & y_{t-2}

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

$$\phi_{ss} = \frac{\rho_s - \sum \phi_{s-1,j} \rho_{s-j}}{1 - \sum_{s-1,j} \rho_j}; s=3,4,5, \dots$$

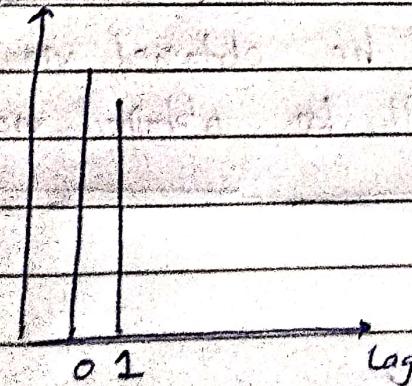
$$\phi_{11} = \rho_1$$

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

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

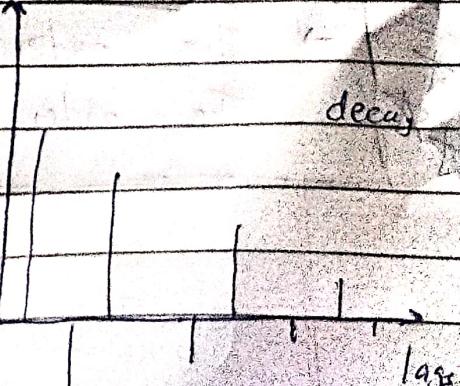
→ For AR(1) Process

PACF



→ For MA(2) process

PACF



→ The decay in MA(2) process could be exponential or sinusoidal depending upon the sign of PAC.

if neg. → sinusoidal & decay

if pos. → exponential decay

⇒ Calculation of PAC

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

This is an ARMA(2,1) process

we can calculate $\hat{\rho}_1$ as,

$$y_t = a_1 y_{t-1} + \varepsilon_t + b_1 \varepsilon_{t-1} \rightarrow \hat{\rho}_1 = \frac{(1 + a_1 + b_1)(a_1 + b_1)}{(1 + b_1^2 + 2a_1 b_1)}$$

∴ $\hat{\rho}_1$ for (1) is,

$$\hat{\rho}_1 = \frac{[1 + 0.49][(-0.7) - 0.7]}{[1 + 0.49 + 2 \times 0.49]} = -0.8445$$

$$\therefore \boxed{\phi_{11} = -0.8445 = \hat{\rho}_1}$$

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

V. IAP.

→ The order of AR will be obtained via PAC

The order of MA will be obtained via AC

Commands

poc return

The graph obtained with this will have a grey zone which shows our confidence interval.

CLASSMATE

Date _____
Page _____

→ So all spikes outside the cone, will have significant effect.

- 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 sign]
- 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,

$$y_t = \alpha y_{t-3} + \gamma \epsilon_{t-3} + \epsilon_t$$

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, i.e.

mean var.

 $N(0, \sigma^2)$

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

$$\text{Variance} = \gamma_0 = \sigma^2$$

$$1 - a_1 \rho_1 - a_2 \rho_2$$

$$\text{Autocorrelation } \left\{ \begin{array}{l} \rho_1 = \frac{a_1}{1 - a_2} \end{array} \right.$$

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

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

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

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

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

Unit Root & Stationarity:

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

Subtracting y_{t-1} both sides,

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

$$\therefore \Delta y_t = a_0 + y_{t-1}(b_1 - 1) + \epsilon_t = a_0 + \gamma y_{t-1} + \epsilon_t \quad \rightarrow (4)$$

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

⇒ 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),

$$b_1 \rightarrow 2$$

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)

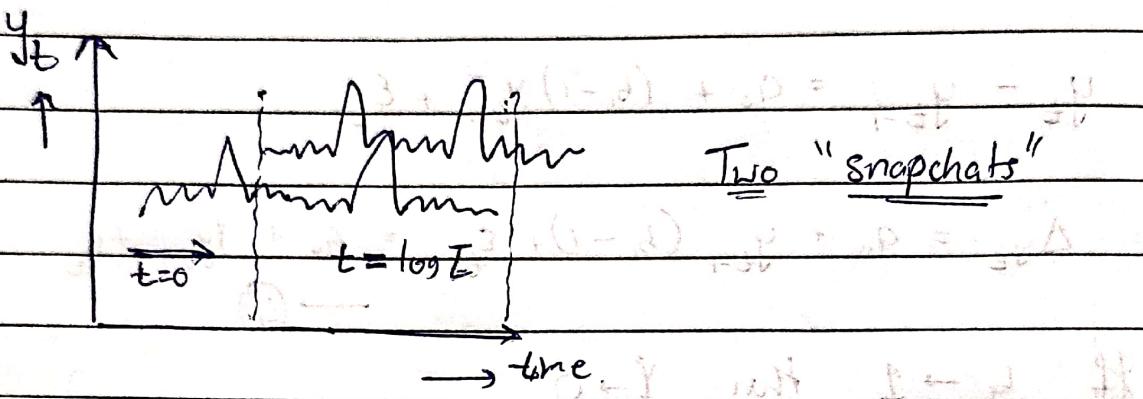
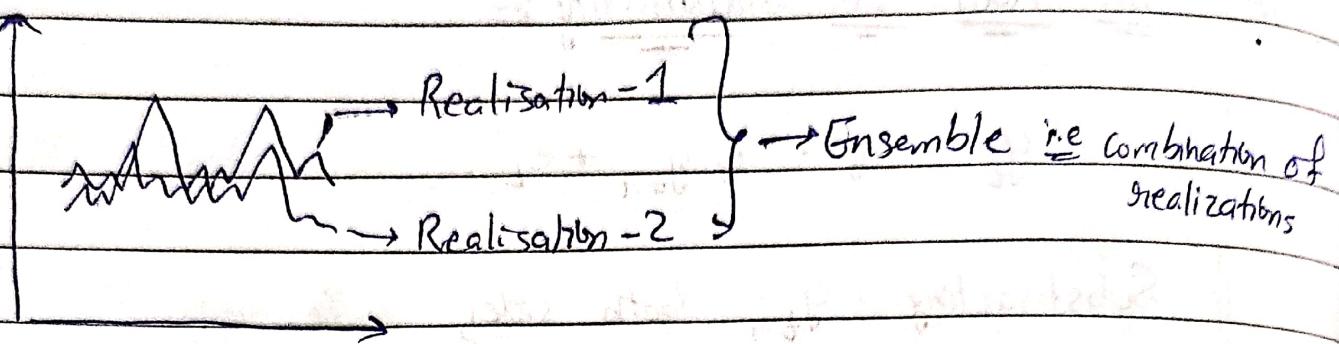
e.g. Law of gravitation

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

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

Findings / Results → Realizations

Club of realisation → Ensembles



* 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

Command → tsline excangerate (y_t)
tsline d.excangerate ($y_t - y_{t-1}$)
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 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)) → 6th 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 6th 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

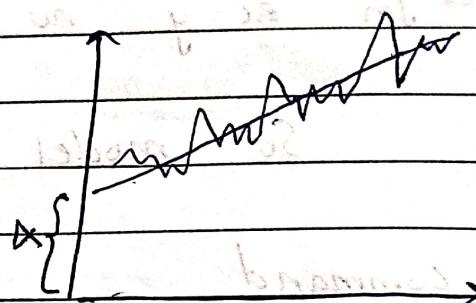
tsline y yhat2

⇒ Any time series will have 4 major components:

- Trend
- Season
- Cycle
- Error / Random Component

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

$$t = \text{trend} = 1, 2, \dots, n$$



Linear trend → represented by "t"

α = intercept

ε_t → contains (seasonal + cycle)

Possibly!

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

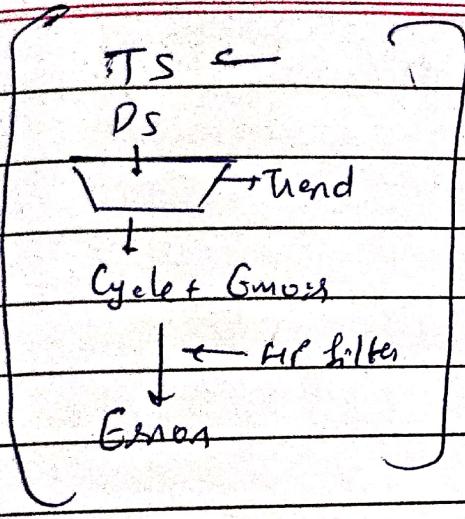
Price
of
West

↑
summer (2020)
summer (2019)

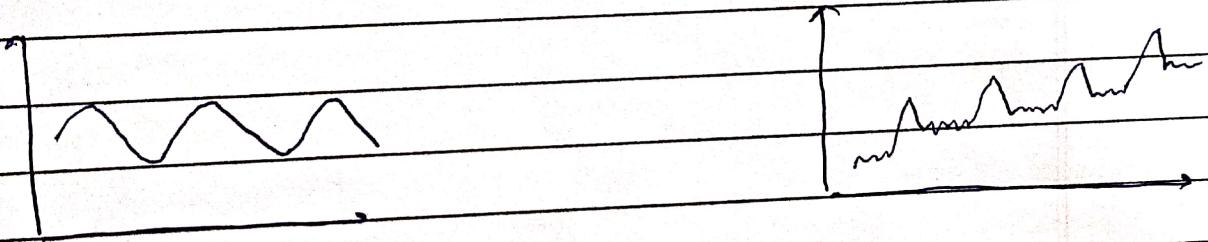
→ Moving Average method to remove seasonality from data

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

~~Hodrick - Prescott~~ Hodrick - Prescott Filter (HP)



→ In cycle, there would be sinusoidal rise & fall 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
[e.g. seasonal value of quarter 1 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 ~~multiplied~~ multiplied with respective seasonality ratio to incorporate effect of seasonality.

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

$$\text{Sale} = 8 + 1.06 \times \text{trend}$$

say trend = 13,

$$\text{Sale} = 33.48 \quad (\text{predicted})$$

Now to incorporate seasonality,

$$= 33.48 \times (0.8762)$$

\approx avg q1 seasonality ratio

Act. Sale Value 29.4455
Predicted

Hodrick-Prescott Filter

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

$$y \xrightarrow{\mu_t} \text{time varying trend}$$

$$\xrightarrow{(y_t - \mu_t)} \text{cyclic}$$

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

λ = Parameters

T → No. of obs.

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

" λ " is cost or penalty parameter

If $\lambda = 0$ then χ^2 minimized when $y_t = u_t$

If $\lambda \rightarrow \infty$ then χ^2 minimized when $\Delta u_{t+1} = \Delta u_{t+2}$

$$\therefore u_{t+1} - u_t = u_t - u_{t-1}$$

i.e. change in trend is constant.

\therefore Linear trend

Generally for quarterly data $\rightarrow \lambda = 1600$

Monthly data $\rightarrow \lambda = 129600$ [1600 $\times 3 \times 4$]

Annual data $\rightarrow \lambda = 6.25$

Daily data $\rightarrow \lambda = 28,51,200$ [129600 $\times (22)$]

?
for
there are
22 working
days

Commands

tsfilter hp cycle = exchangenate, smooth(129600)

gen trend = exchangenate - cycle

tstime trend

Multivariate Time Series Models

Vector - Autoregression [VAR Model]



A matrix with a single column

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

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

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

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

$\uparrow \quad \uparrow \quad \uparrow \quad \uparrow$

$$B x_t = T_0 + T_1 x_{t-1} + \epsilon_t$$

or

$$x_t = A_0 + A_1 x_{t-1} + \epsilon_t ; \quad \left\{ \begin{array}{l} B^{-1} T_0 = A_0 \\ B^{-1} T_1 = A_1 \\ B^{-1} \epsilon_t = \epsilon_t \end{array} \right.$$

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

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

where $e_{1t} = \frac{\epsilon_{y_t} - b_{12} \epsilon_{z_t}}{1 - b_{12} b_{21}}$

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

$$E(\epsilon_{1t}) = \frac{E(\epsilon_{y_t}) - b_{12} E(\epsilon_{z_t})}{1 - b_{12} b_{21}} = 0$$

$$\therefore \text{Var}(\epsilon_{1t}) = (\sigma_y^2 + b_{12}^2 \sigma_z^2)^2 / (1 - b_{12} b_{21})^2$$

Similarly ~~if~~ $\text{Var}(\epsilon_{2t})$ can be calculated.

$$\Rightarrow \text{Cov}(\epsilon_{1t}, \epsilon_{2t}) = E[(\epsilon_{y_t} - b_{12} \epsilon_{z_t})(\epsilon_{z_t} - b_{21} \epsilon_{y_t})] / (1 - b_{12} b_{21})^2$$

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

Hence,

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

$$\Sigma = \begin{bmatrix} \text{Var}(\epsilon_{1t}) & \text{Cov}(\epsilon_{1t}, \epsilon_{2t}) \\ \text{Cov}(\epsilon_{2t}, \epsilon_{1t}) & \text{Var}(\epsilon_{2t}) \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_y^2 & \sigma_{12} \\ \sigma_{21} & \sigma_z^2 \end{bmatrix}$$

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,

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{(1 - b_{12} b_{21})} \begin{bmatrix} 1 & -b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} \epsilon_{y_t} \\ \epsilon_{z_t} \end{bmatrix}$$

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

Now let,

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

Then,

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

$\phi_{11}(i)$, $\phi_{12}(i)$, $\phi_{21}(i)$, $\phi_{22}(i)$ are called Impulse Response function (IRF)

$\phi_{jk}(i)$ are called instantaneous "impact multipliers"

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

update by '1' period refers to

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

Impact of ϵ_{zt-1} on y_{t+1} and so on

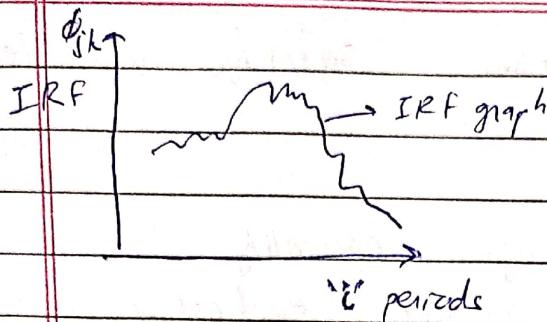
Meaning giving shock to ϵ_{zt} will show effect on y_{t+1}

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

classmate

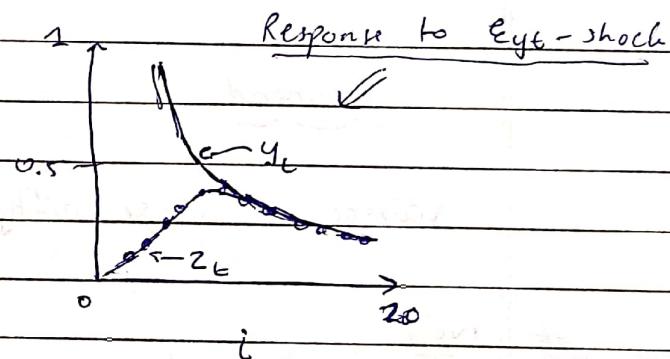
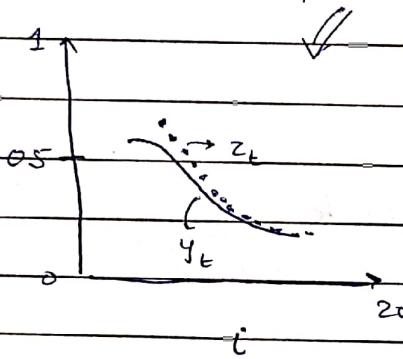
Date _____

Page _____



$$\text{Eg:- } \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} e_{yt} \\ e_{zt} \end{bmatrix}$$

Response to e_{zt} - shock



Granger Causality :- [Extension of VAR Model]

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

$$y_t \text{ and } z_t \sim I(0)$$

integrated of order '0' [shows stationarity]

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

$$z_t = \alpha_2 + \sum_{i=1}^m \delta_i y_{t-i} + \sum_{i=1}^n \tau_i z_{t-i} + \epsilon_t$$

→ If $\gamma_i = 0 \neq i=1(1)p$

then $z_t \not\rightarrow y_t$ (z_t doesn't cause y_t)

→ If $\delta_i = 0 \neq i=1(1)m$

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 stronger causality.

If both y_t & z_t impact each other,

it is Bidirectional Causality (Feedback Method)

If only one variable impacts the other,
it is unidirectional causality.

Command

~~varsoc~~ close-nifty exchangerate

Note:

→ rely for small samples

Statistical
Criteria

(AIC → Akaike Information Criteria)

SBC →

HQIC →

FPE →

→ Stat 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 exchangerate, lags(1/4)

to include all 4 lags
into the model

var → only estimation

Varbasic → estimation + Impulse response funcⁿ IRF graph

CLASSMATE

Date _____
Page _____

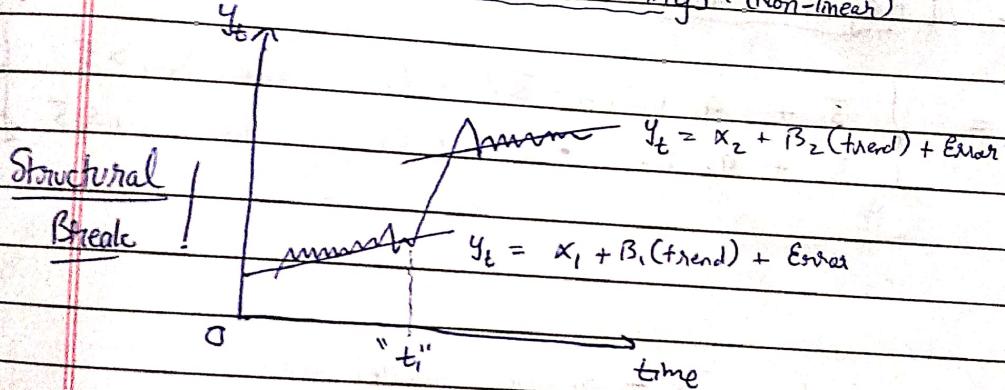
Command

VARSOC close nifty exaggerate

(Forecast) VFG close nifty exaggerate, lags(1/4)
~~EFcast~~ Compute m_1 , step (4)

Non - Linear Time Series Models ★

Type - I [Piecewise - (linearity) \leftarrow (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]

$$y_0 = 0$$

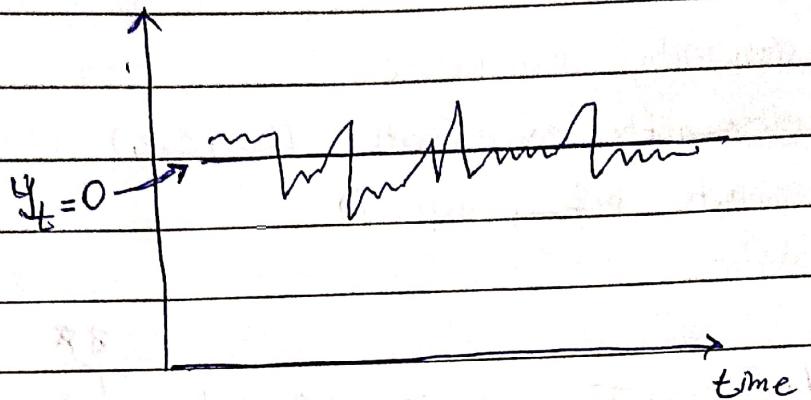
$$\text{is the threshold} \quad \begin{cases} y_t = \alpha_1 y_{t-1} + \epsilon_{1t} & \rightarrow \text{if } y_{t-1} > 0 \\ y_t = \alpha_2 y_{t-1} + \epsilon_{2t} & \rightarrow \text{if } y_{t-1} \leq 0 \end{cases}$$

$$y_t = \alpha_1 I y_{t-1} + \alpha_2 (1-I) y_{t-1} + \epsilon_t \rightarrow \text{Combined}$$

$$\begin{cases} \text{if } y_{t-1} > 0 \Rightarrow I = 1 \\ \text{& } y_{t-1} \leq 0 \Rightarrow I = 0 \end{cases} \rightarrow \begin{matrix} \text{dummy} \\ \text{variable.} \end{matrix}$$

[P.T.O]

Threshold Model



Type - III [Regime Model] [Again piecewise]

2-regime Mode

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

→ 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 function of non-linear models

→ Assuming 'chaos' model,

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

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 \textcircled{1} & autocorrelations will be -

P_1	P_2	P_3	P_4	P_5	P_6
-0.074	-0.072	0.008	0.032	-0.016	-0.030

But ACF of y_t^2 & y_{t-1}^2 is -0.281
 y_t^3 and y_{t-1}^3 is ~~-0.038~~ -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.

Let, P_i denotes correlation \hat{e}_t

\hat{e}_t^2 & \hat{e}_{t-i}^2 then Ljung Box test statistics is denoted by,

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

$T \rightarrow$ No. of observation
 $n \rightarrow$ no. of lags chosen

Alternatively,

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

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

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 $\text{prob} > \chi^2(2) = 0.000$]

[we get lag 1]
[AR(2)]
→ pac excangerate ← [we see first lag significant]

wntestq (excangerate), lags(1) ← [significant as $\text{prob} > \chi^2(1) = 0.00$]

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

arima excangerate, arima (1, 0, 3)

predict exhat Considering all 3 lags

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 isn't significant.

Smooth Transition Auto regressive Model [STAR Model] ↗

→ This model allows autoregressive parameters to change slowly.

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

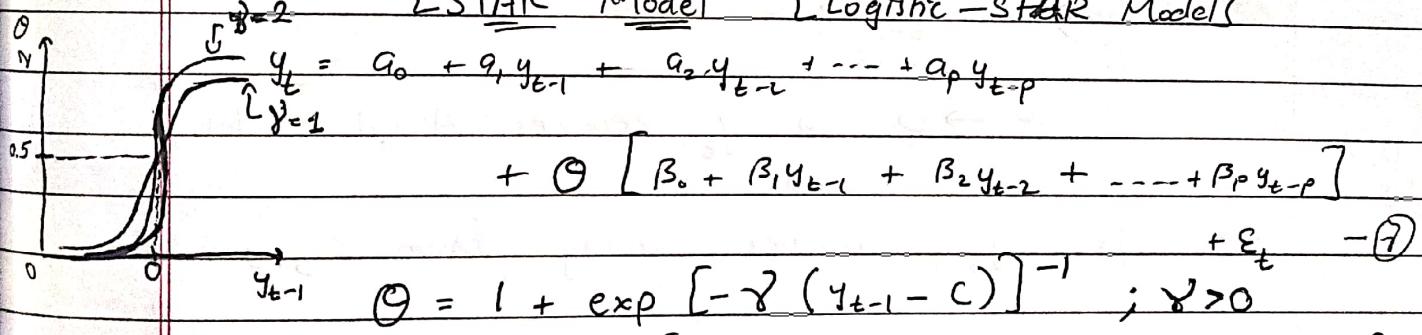
$$y_t = \alpha_0 + \alpha_1 y_{t-1} + (\beta_1 y_{t-1} f(y_{t-1})) + \epsilon_t$$

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

→ AR parameter $(\alpha_1 + \beta_1)$ will change smoothly along with y_{t-1} .

→ Example 1

LSTAR Model [Logistic STAR Model]



[$\gamma \rightarrow$ Shows responsiveness based on past data]

⇒ If $\gamma \rightarrow 0$ or ∞ then,

(i) becomes an AR(p) model

⇒ If $0 < \gamma < \infty$ then,

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

⇒ As $y_{t-1} \rightarrow -\infty$, $\theta \rightarrow 0$ then

y_t becomes, $\alpha_0 + \alpha_1 y_{t-1} + \dots + \beta_1 y_{t-p} + \epsilon_t$

⇒ As $y_{t-1} \rightarrow \infty$ and $\theta \rightarrow 1$ then,

$$y_t = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1) y_{t-1} + \dots + \epsilon_t$$

$$\alpha_0 \rightarrow (\alpha_0 + \beta_0)$$

$$(\alpha_1 \rightarrow (\alpha_1 + \beta_1))$$

Thus, intercept and co-efficients smoothly change b/w two values as y_{t-1} changes.

extreme

⇒ Example 2

ESTAR Model [Exponential STAR Model]

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

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} \\ + \theta [\beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p}] + \epsilon_t$$

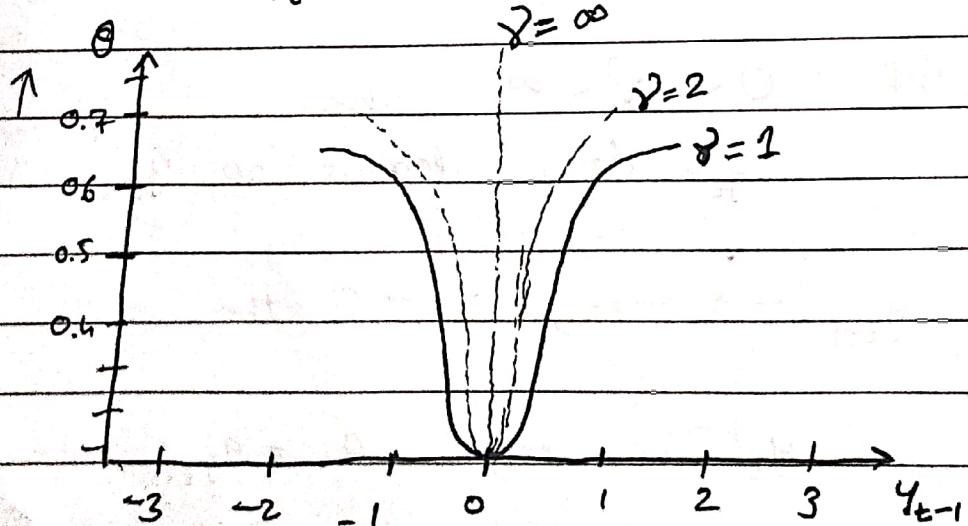
→ when, $y_{t-1} \rightarrow c$ then

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

→ As y_{t-1} moves away from c ,

$\theta \rightarrow 1$ and

$$y_t = (\alpha_0 + \beta_0) + (\alpha_1 + \beta_1)y_{t-1} + \dots + \epsilon_t$$



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) (or all available information)

→ Market assimilates all past records / experiences into its current price.
 → Directly implies that it is impossible to beat the market.

$$\Rightarrow y_t = a_0 + a_1 y_{t-1} + \epsilon_{tL} \text{ if in regime -1}$$

$$y_t = a_0 + a_2 y_{t-1} + \epsilon_{tL} \text{ if in regime -2}$$

If p_{11} = probability that y_t will be in regime -1 when its 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}~~

$p_{12} \rightarrow y_t$ in regime -2 & y_{t-1} in regime -2

$(1-p_{12}) \rightarrow y_t$ 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 func form -

A simple ANN model is -

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

$f_i(y_{t-1}) \rightarrow$ cumulative distribution or a logistic func such as -

LSTAR Model

For the logistic func -

$$y_t = a_0 + a_1 y_{t-1} + \sum x_i \left[1 + \exp(-\gamma_i (y_{t-1} - c)) \right]^{-1} + \epsilon_t$$

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_i as constant.

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

This leads to complex and overparametrized model.

Volatility Models

→ Concept of conditional and unconditional volatility in stock market.

$$y_t = a_0 + a_1 y_{t-1} + \epsilon_t \quad \rightarrow (\approx 0)$$

Conditional : $E(y_t) = f(a_0) + a_1 E(y_{t-1}) + E(\epsilon_t)$

Variance

[Old values don't impact future value] $\therefore \text{var}(y_t) = E[y_t - E(y_t)]^2 = \sigma^2$
 (Only yesterday's value impacts)

Unconditional, $y_t = a_0 + a_1 y_{t-1} + \epsilon_t$

Variance

[Old values have significant effect] $y_t = a_0 + a_1 (a_0 + a_1 y_{t-1} + \epsilon_{t-1}) + \epsilon_t$
 $= a_0 (1 + a_1 + \dots) + a_1 y_0 + \sum_{i=1}^t a_1^i \epsilon_i$

$$\therefore E(y_t) = \frac{a_0}{1-a_1} + a_1^t y_0 + 0 \quad \left. \begin{array}{l} \\ \end{array} \right\} |a_1| < 1$$

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

\therefore Unconditional Variance $>$ Conditional Variance.

ARCH [Auto regressive Conditional Heteroskedasticity] Model:-

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

If we take n -log or \ln

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

$$\text{var}(y_{t+1}|x_t) = x_t^2 \alpha^2 \quad \text{for given value of } x_t$$

Conditional variance of y_t for a given x_t

Heteroskedasticity \rightarrow square of error term is ^{not} constant

Homoskedasticity \rightarrow square of error terms are constant

classmate

Date _____
Page _____

\Rightarrow If $\sigma_e \uparrow$ then $\text{var}(y_{t+1} | y_t)$ will also increase.

\rightarrow Now, assuming an ARMA process,

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

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

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

If suppose,

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

$$\varepsilon_t^2 = \kappa_0 + \kappa_1 \hat{\varepsilon}_{t-1}^2 + \kappa_2 \hat{\varepsilon}_{t-2}^2 + \dots + \nu_t$$

ν_t white noise term.

$$\therefore E_t(\hat{\varepsilon}_{t+1}^2) = \kappa_0 + \kappa_1 \hat{\varepsilon}_{t-1}^2 + \kappa_2 \hat{\varepsilon}_{t-2}^2 + \dots + \kappa_v \hat{\varepsilon}_{t-v}^2$$

\Rightarrow Engle (1982) proposed first form of ARCH model

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

$v_t \rightarrow$ white noise term

where $E(v_t) = 0$ & $\text{var}(v_t) = 1$

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

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

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

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

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

Precisely,

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

Conditional volatility ε_t^2 will increase if past value (ε_{t-1}^2) is high & so on.

This means that 'x' terms are high & strong. This phenomenon is called 'Volatility Persistence.'

\Rightarrow A Way to fit ARCH Model

- ① E is an OLS estimation ~~the model~~
- ② Extract errors
- ③ Square the errors
- ④ Regress

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

$$\text{If } \alpha_1 = \alpha_2 = \dots = 0$$

then there is no ARCH Effect,

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

MLE → Maximum Likelihood Estimation

↳ Generate likelihood function $\xrightarrow{\text{maximize this function}}$

classmate

Date _____

Page _____

GARCH Model [Generalized ARCH Model]

Bollerslev (1986) extended Engle's model -

$$\hat{\epsilon}_t = v_t \sqrt{h_t}$$

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

v_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 $\{\hat{\epsilon}_t^2\}$

Calculate

$$\hat{\sigma}^2 \rightarrow \frac{\sum \hat{\epsilon}_t^2}{T} \quad T \leftarrow \text{no. of obs.}$$

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

$$T_i = \frac{\sum [(\hat{\epsilon}_t^2 - \bar{\sigma}^2)(\hat{\epsilon}_{t-i}^2 - \bar{\sigma}^2)]}{\sum (\hat{\epsilon}_t^2 - \bar{\sigma}^2)^2}$$

(at lag i)

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

$$Q \text{ statistic} \rightarrow T(T+2) \sum_{i=1}^T \hat{\epsilon}_t^2 / (T-i)$$

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

Null hypothesis is $\hat{\epsilon}_t \nparallel \hat{\epsilon}_i$ are insignificant
[i.e. No GARCH effect]

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

i.e. $\hat{\epsilon}_t^2$ and $\hat{\epsilon}_i^2$ at different lags at 'i' values not strongly connected.

i.e. There is no GARCH effect.

ARCH - M Model:- [ARCH in Mean type Model]

→ Engle, Lilien & Robins (1987) extended the basic ARCH framework to allow the mean of sequence to depend on its own conditional variance.

$$y_t = \mu_t + \varepsilon_t$$

$$y_t = \text{excess return } [r_{t+1} - \text{T bill rate}]$$

μ_t = risk premium necessary to induce the risk-averse investor to hold this risky asset over bond.

$$\varepsilon_t = \text{error / stocks}$$

$$\mu_t = \beta + \delta h_t \quad \delta > 0$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad \text{If } \alpha_i = 0 \quad \text{then model boils down to,}$$

$$y_t = \beta + \kappa_0 + \varepsilon_t \leftarrow \text{constant risk premium market.}$$

iGARCH Model :- [integrated GARCH Model]

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

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

$$E(\varepsilon_t^2) = E(\sigma_t^2) E(h_t)$$

If $h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$

and if $\alpha_1 + \beta_1 = 1$ then

$$h_t = \alpha_0 + [\alpha_1 - \beta_1] \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

implies

If α_1 is high (say 0.9), it means lag value of ε_t has significant impact on present conditional vol. & also since β_1 is low, lag value of conditional vol. has little impact.

TARCH (Threshold ARCH Model) :-

Notion of leverage effect \rightarrow

\rightarrow Tendency for volatility to decline when return rises

\rightarrow and " " " " increase a " falls

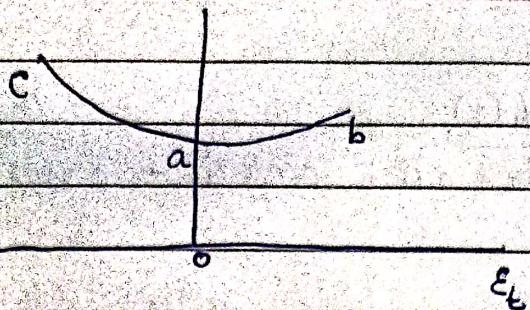
\rightarrow If $\varepsilon_t = 0$ then $E_t(h_{t+1}) = \text{distance } \underline{\alpha}$

\rightarrow Any news increases volatility.

\Rightarrow If news is good ($\varepsilon_t > 0$) volatility increases $\widehat{\alpha b}$.

\Rightarrow If news is bad ($\varepsilon_t < 0$) volatility increases $\widehat{\alpha c}$.

$E_t(h_{t+1})$



New information / shock

$$r_t = \alpha + \beta r_{t-1} + \varepsilon_t$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \lambda D \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

If $\varepsilon_{t-1} < 0$ then $D = 1$ [Dummy]

If $\varepsilon_{t-1} \geq 0$ then $D = 0$

∴ when $\varepsilon_{t-1} < 0$ [bad news]

$$h_t = \alpha_0 + (\alpha_1 + \lambda) \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

when $\varepsilon_{t-1} \geq 0$ (good / neutral news)

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Command

gen t = -n

tset t

gen return = ln (close/l.close)*100

tline return deaths

tline close deaths

gen dg = ln (deaths/l.deaths)*100

tline return dg

regress return l.return

estat archlm, lags(3) \leftrightarrow (Prob > chi2 < 0.0000) (significant)

arch return, arch(1/2)

if z value < 2,
No ARCH effect

Else there is ARCH effect.

If this \Rightarrow is < 0.05
then we reject the null
at 5% level of significance

So there is ARCH effect

predict ex_cy, variance \leftarrow predicting conditional vol. of above

conditional
vol.

model

tsline cv dg

arch return, arch(1) gauch(1)

$$L \ln = 0.65 + 0.954 \hat{e}_{t-1}^2 + 0.8598 h_{t-1}$$

predict cv1, variance

#

Co-integration [Market=Comovement]

If X & y are two stocks / indices and both X & y are non-stationary then,

a linear combination of X and y can generate a stationary variable.

i.e

$$aX + bY = Z$$

↑ ↑ ↓
non-stationary stationary

[where Z is a linear combination of X & y]

If above condition takes place, X & y are co-integrated.

It means there is a long run, stable, equilibrium relationship b/w X and y .

Steps : ① Unit root test of X and y

② Regress y over X $\rightarrow Y = A + BX + \epsilon$

Equivalent to $X\hat{X} + BY$

Since Y has co-efficient 1

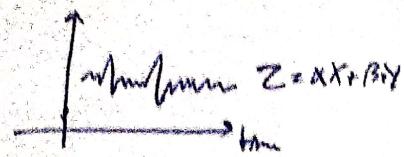
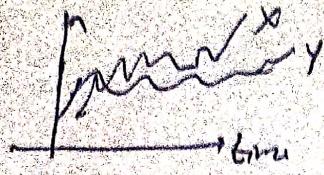
and X has $\approx B$

[Ignore A here]

$$\textcircled{3} \quad Y = XB + (\epsilon + A)$$

$$Y - XB = \epsilon'$$

$$\textcircled{3} \quad Y + (-B)(X) = \epsilon' \quad (\epsilon' = \epsilon + A)$$



If ϵ' is stationary then X and Y are cointegrated.

$[1, -\beta]$ is referred as cointegrated vector.

A Error Correction Model :-

y_t and z_t two variable [non-stationary]

$$y_t = \beta_0 + \beta_1 z_t + \epsilon_t \quad (1)$$

$$\Delta y_t = \alpha_1 \hat{\epsilon}_{t-1} + \epsilon_t \quad (\epsilon_t \text{ is white noise term})$$

$$\begin{aligned} \Delta y_t &= \alpha_1 + \alpha_2 [y_{t-1} - \beta_1 z_{t-1}] + \sum_{i=1}^n \alpha_{12}(i) \Delta y_{t-i} \\ &\quad + \sum_{i=1}^m \alpha_{12}(i) \Delta z_{t-i} + \epsilon_{yt} \end{aligned}$$

α_2 is the co-eff. of $[y_{t-1} - \beta_1 z_{t-1} = \hat{\epsilon}_{t-1}]$
past error of regression model (1).

If $|\alpha_2| = 1$ then old error impact change in y
and Δy_t changes 100% for unit change in ϵ_{t-1}
It means adjustments happens to max. possible limits

If $\alpha_2 = 0$ then response in y (Δy) hasn't been impacted by old error (ϵ_{t-1})

Ideally, $0 \leq |\alpha_2| \leq 1$

and it denotes the degree of adjustment

Sign of α_2 should be (-ve) because if past
error increases then ~~decrese~~ deviation in equilibrium
relation is high and to bring back this eqs.
change in y should drop or fall.

Hence α_2 (denoting response of Δy to unit
change in ϵ_{t-1}) should bear a negative sign.

This method is referred as Engle-Granger cointegration
and ECR (Error Correction Model)

If $\alpha_y = -0.9$ then there is 90% adjustment in chn
(or correction is happening by 90%) in current period
(vis-a-vis previous period)

Command

check stationarity

```
regress d.gold l.gold ← [coeff closer to zero] ⇒ (Non-stationary)
regress d.oil l.oil
regress gold oil
predict goldhat
gen error = gold - goldhat
```

{ observe → non-stationary

↳ hence, gold & oil are not
cointegrated.

```
regress gold oil if time < 200
```

```
predict goldhat1
```

```
gen error1 = gold - goldhat1
```

```
regress gold oil if time > 200
```

```
predict goldhat2
```

```
gen error2 = gold - goldhat2
```

fixed term

```
regress gold oil (time) if time < 200
```

```
predict yhat
```

```
gen e = gold - yhat if time < 200
```

tseries e

varsoc d.gold d.oil → taking SBIC stat mank which is at lag 1

```
regress d.gold l.e l-d.gold. l.oil
```