Time Series & Survival Analysis

[Time Series]
Midsem → 40%
[Survival Analysis]
Endsem → 40%
Labtest → 10%

Class Test 1 [Time Series] → 5%
Class Test 2 [Survival Analysis] → 5%

Books →
- Time Series — Brockwell-Davis [Application]
- Time Series — Shumway Stoffer.

↓
For Time Series

Time Series

~~Observations~~ Errors are ~~co~~uncorrelated → This assumption of regression is violated in case of Time Series data.

[R-package → [tsmr]]

Any data varies over time is Time Series Data
↓
Mathematical Representation → $\{X_t, t \in T\}$
↓
It can give 4 types of Time Series Data

T – [Discrete, Continuous] X_t – [Discrete, Continuous]

Combination of these two will give 4 types.

We are going to learn how models are built on Discrete time and Continuous Variable.

By the Graphs Shown →
- Repeatative Pattern
- Variation [Slow, Fast]
- Irregular changing Pattern
- Variation are also increasing.

Time Series has 4 parts →

X_t

- m_t [Slow Change] → Trend
- S_t [Seasonal Change] → Seasonality [Period at most 12 months]
- C_t [Slower Repeatative] → Cyclical [Period is much more longer]
- I_t [Irregular] [Random]

In Seasonal Change → Amplitude Constant
In Cyclical Change → Amplitude Changes.

$\boxed{X_t = m_t + S_t + C_t + I_t}$ → It has limitations, as we cannot express the change of variation by this.

or $\boxed{X_t = m_t \cdot S_t \cdot C_t \cdot I_t}$ We can take log of the multiplicative model to build additive model.

Stochastic Processes → Collection of Random Variables indexed over ~~something for~~ time.

- **Trend** (m_t) : Smooth, regular, long-term movement of the time series data.

- **Seasonality** (S_t): A periodic movement, with period of movement less than one year. Eg → Winter Garment

- **Cyclic** (C_t) : An ~~oscill~~ oscillatory movement with period of oscillation more than one year. Eg → Earthquake

- **Random** (X_t) : Irregular component of time series.

We have to check $Y_t \to \left.\begin{matrix} m_t \\ S_t \\ X_t \end{matrix}\right\}$ These 3 components.

Non seasonal Model with Trend →

$$Y_t = m_t + X_t \quad , t = 1, \dots, n \qquad EX_t = 0$$

- Trend Estimation and the elimination
- Direct Trend Elimination

Trend Estimation & Elimination :→

(i) **Smoothing with a finite moving average filter** :

q → non-negative integer — considering two-sided moving average.

$$\boxed{\hat{m}_t = (2q+1)^{-1} \sum_{j=-q}^{q} Y_{t-j}} \quad \text{for } q+1 \le t \le n-q$$

$y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6 \quad y_7 \quad y_8$

Here $E(X_t) = 0$, So the Errors will be compensating with each other.

$\hat{m}_2$ → Average of y_1, y_2, y_3 will be taken.

$\hat{m}_3$ → Average of y_2, y_3, y_4 will be taken.

• Why $E(X_t)=0$? $\longrightarrow$ If x_t is +ve, it will be a part of the trend Part.

⊙ Exponential smoothing:

• For any fixed $\alpha \in (0,1)$, the one-sided moving averages ~~fns~~ defined by the recursions $\longrightarrow$ $\boxed{\hat{m}_t = \alpha Y_t + (1-\alpha)\hat{m}_{t-1}}$ for $t=2,\dots,n$.

$\therefore \hat{m}_1 = Y_1 \qquad \alpha = 1$

$$\hat{m}_t = \alpha Y_1 + (1-\alpha)\hat{m}_{t-1} = \alpha Y_1 + \alpha(1-\alpha) Y_{t-1} + (1-\alpha)^2 \hat{m}_{t-2}$$
$$= \alpha Y_1 + \alpha(1-\alpha) Y_{t-1} + (1-\alpha)^2 \alpha . Y_{t-2}$$
$$+ \dots + \alpha(1-\alpha)^{t-1} Y_1$$

More and more closer ~~mto~~ means

⊙ Polynomial fitting: $\longrightarrow$ Regression

$$\hat{m}_t = \hat{a}_0 + \hat{a}_1 t + \hat{a}_2 t^2$$

⊙ Once we estimate $\hat{m}_t$, we subtract it from Y_t to get the x_t (noise) $\boxed{X_t = Y_t - \hat{m}_t}$

■ Trend Elimination by Differencing: $\longrightarrow$

• Lag-1 difference Operator.

$$\nabla Y_t = Y_t - Y_{t-1} = (1-B) Y_t$$

B is called backshift operator $\rightarrow$ $BY_t = Y_{t-1}$

$\boxed{B^j(Y_t) = Y_{t-j}}$

Let, $\nabla^2 Y_t = (1-B)^2 Y_t = Y_t - 2Y_{t-1} + Y_{t-2}$

$\therefore \nabla Y_t = a_1 + (x_t - x_{t-1})$

$Y_t = \underbrace{a_0 + a_1 t}_{\text{This part is removed by } \nabla \text{ operator}} + x_t$

$$\nabla^j (Y_t) = \nabla(\nabla^{j-1} Y_t)$$

☐ Model with trend and seasonality: $\longrightarrow$

$$Y_t = m_t + s_t + X_t, \quad \text{for } t = 1,\dots,n. \qquad E(X_t) = 0$$

$Y_t \qquad \nabla_{12} Y_t \qquad \nabla\nabla_{12} Y_t = (1-B)(1-B^{12}) Y_t$

$$= Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13}$$

$\{Y_t\} \longrightarrow \boxed{} \longrightarrow m_t, \; s_t, \; x_t$

Is x_t random?
First we have to check whether x_t is i.i.d or not.

$\{x_t\} \longrightarrow \{x_1, x_2, \dots, x_n\} \longrightarrow$ Sequence of Random Variables.

$\hat{Y}_{n+k} = \hat{m}_{n+k} + \hat{s}_{n+k}$

- To check x_t is i.i.d. or not we need Hypothesis Testing.

$\{x_1, \dots, x_n\} \longrightarrow$ Sample.

$H_0 : \mu = 100$ hrs.

Since it is continuous distribution so, $\mu = \bar{x}_n$ has almost zero chance

$\bar{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i$

we will check whether $\bar{x}_n : \mu \approx 100$ (close to)

$\{x_1, \dots x_n\} \longrightarrow \mu, \; \sigma$

$\frac{1}{n} \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma} \sim N(0, \frac{1}{n})$

Central Limit Theorem is used here

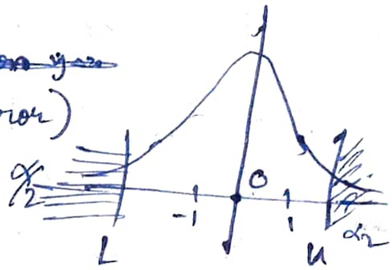
- Type 1 error $\longrightarrow$ Null Hypothesis is true

Let $\alpha = 0.05$ or 0.01

$=$ Prop (Type 1 error)

$P(t > u) = \frac{\alpha}{2}$

$P(t < l) = \frac{\alpha}{2}$



~~$H_0 : \mu = 100$ hrs~~

~~$H_a : \mu = 100$~~

$\{X_1, X_2, \dots, X_n\}$
$x_1, x_2, \dots x_n$

$H_0 : \mu = 100$ hrs

1. $H_A : \mu \neq 100$ hrs.
2. Construct the random variable (Test Statistics) [TS] as a fn of my samples.
   $TS = f(X_1, \dots, X_n)$
3. Any assumption
4. Find a distribution of Test Statistics under that assumption. $\longrightarrow$ $TS \sim$ Distr.
   ~~$TS \sim$ Distri.~~
5. Specify 'α' [$\alpha = P$(Type 1 error)]

6) Construction the rejection region. R.R = $(-\infty, L] \cup [U, \infty]$

7) Calculate the value of Test Statistics (ts)

$$ts = f(x_1, \ldots, x_n)$$

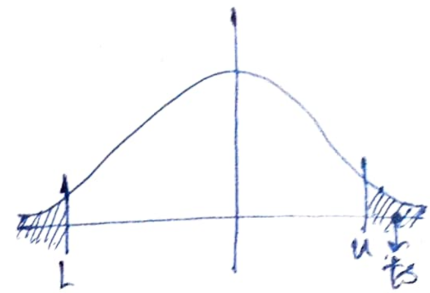
8) $ts \in R \Rightarrow$ reject H_0 in ~~factor~~ favour of H_1. (Alternative Hypothesis)

⊙ Rejection with p-value.

The first (1–5) steps will remain same.

$\boxed{2P(T_S > |t_s|)}$ → p-value

$2P(T_S$



(6) $ts = f(x_1, \ldots, x_n)$ ← Calculate ts.

(7) Calculate p-value.

(8) Reject depending upon p-value $>$ or $< \alpha$

⊙ Now we will check whether $\{X_t\}$ is i.i.d. or i.i.d. normal.

Auto Correlation function → Correlation b/w the sequence and ~~the~~ with its own but with lags (h)

For large n, the sample auto-correlations of an i.i.d. seq. $X_1, \ldots, X_n$ is defined as—

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \quad \text{for } -n < h < n.$$

where $\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-|h|} (x_{t+h} - \bar{x})(x_t - \bar{x})$. and $\bar{x} = n^{-1} \sum_{t=1}^{n} x_t$.

$$\begin{matrix} x_1, & x_2, & \cdots\cdots & x_n \\ & x_1, & \cdots\cdots & x_{n-1}, & x_n \end{matrix}$$

Taking these as pairs

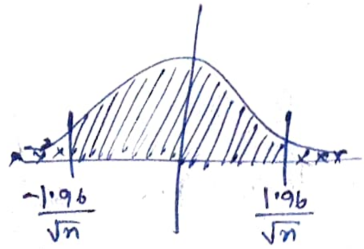
lag 1 [[illegible]]

$$\text{Cov.} = \frac{1}{n} \sum_{i=1}^{n-1} (x_i - \bar{x})(x_{i+1} - \bar{x}) \qquad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

For h lag → $\boxed{\frac{1}{n} \sum_{i=1}^{n-h} (x_i - \bar{x})(x_{i+h} - \bar{x}) = \hat{\gamma}(h)}$ → Auto Covariance.

$\gamma(-1) = \gamma(1)$ ∵ Symmetric $\qquad \hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$

As i.i.d. seq → Corr$(X_1, X_{1+h}) = 0$



(1)

$H_0 : \{X_t\}$ is i.i.d. with $Var.(X_t) < \infty$

$H_A : \{X_t\}$ is not i.i.d.

(2) Construct test statistics by auto-correlation fn.

$\{x_1, x_2, \ldots, x_n\}$ i.i.d.

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{E[(X_{t+h}-\mu)(X_t-\mu)]}{E(X_t-\mu)^2} = 0$$

Here $\frac{E[X_{t+h} X_t] - \mu^2}{\gamma(0)}$ $[E(X_{t+h}) = E(X_t) = \mu]$

(X_t) may not be i.i.d. but it ~~has to~~ should have common mean.

So, we get $\rightarrow$ $$\hat{\rho}(h) = \frac{\frac{1}{n}\sum_{t=1}^{n-|h|}(X_{t+h}-\bar{X})(X_t-\bar{X})}{\frac{1}{n}\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

Here $X_t \rightarrow$ Random variable.

This is our test statistics.

(3) $\hat{\rho}(h) \sim N(0, \frac{1}{n})$

$$E(\hat{\rho}(h)) = \frac{E\left(\sum_{t=1}^{n-|h|}(X_{t+h}-\bar{X})(X_t-\bar{X})\right)}{\sum_{t=1}^{n}(X_t-\bar{X})^2} = \frac{\sum_{t=1}^{n-|h|} E\left(X_{t+h}X_t - \bar{X}\cdot X_{t+h} - \bar{X}X_t + \bar{X}^2\right)}{\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

$$= \frac{\sum_{t=1}^{n} E(X_t)E(X_{t+h}) - E(\bar{X})\cdot E(X_{t+h}) - E(\bar{X})\cdot E(X_t) + E(\bar{X}^2)}{\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

$$= \frac{\sum_{t=1}^{n} \mu\cdot\mu - \mu\cdot\mu - \mu\cdot\mu + \mu\cdot\mu}{\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

$= 0$

(4) $\alpha = 0.05$

(5) $$T.S. = \frac{\sum_{t=1}^{n-|h|}(X_{t+h}-\bar{X})(X_t-\bar{X})}{\sum_{t=1}^{n}(X_t-\bar{X})^2}$$

(6) $-\frac{1.96}{\sqrt{n}} < t_s < \frac{1.96}{\sqrt{n}} \Rightarrow$ fail to reject H_0

O.W. $\Rightarrow$ Reject H_0.

$x_t \sim$ Bernoulli $\quad x_t' = \sum_{i=1}^{t} x_i$

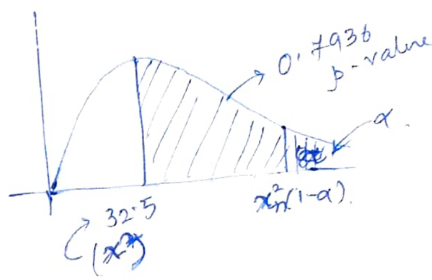
Even null hypothesis is true, output of auto-correlation functio maybe outside the confidence interval. If it is "frequently" outsid the interval, we reject it. "Frequent" can be measured with the help of α.

- The portmanteau test $\longrightarrow$

$$Q = n \sum_{j=1}^{h} \hat{\rho}^2(j)$$

- Ljung-Box Test $\longrightarrow$ $\quad Q_{LB} = n(n+2) \sum_{j=1}^{n} \hat{\rho}^2(j)/(n-j)$

Ljung Box Q value will ~~approximate~~ converge normal. more rapidly
So, it is preferred.



- Turning Point Test $\longrightarrow$

(1) If $x_1, \ldots, x_n$ is a sequence of observations, we say that there is a turning point at time $i, 1 < i < n$.
(a) if $x_{i-1} < x_i$ and $x_i > x_{i+1}$ or if $x_{i-1} > x_i$ and $x_i < x_{i+1}$

(2) If T is the number of turning points of an iid sequences of length n, then
(a) the probability that a point at time i is a turning point is $\frac{2}{3}$
(b) $\mu_T = E(T) = 2(n-2)/3$ [The last & first point is not taken so $(n-2)$ points having probability $\frac{2}{3}$ always $\therefore \frac{2}{3}(n-2)$]
(c) $\sigma_T^2 = Var.(T) = (16n-29)/90$

$$\frac{T - \mu_t}{\sigma_t}$$

$$\cancel{P(X_1 < X_2 > X_3)} = P(\cancel{X_1 < X_2}) + \cancel{P(X_2 > X_3)}$$
$$= P(\cancel{X_1 < X_2}) + \cancel{1 - P(X_3 < X_3)}$$

$$= \cancel{P(X_1 < X_2 \mid X_2 > X_3)}.$$

$$P(X_1 < X_2 > X_3) = \int_{x_2} P(X_1 < x_2 > X_3)\, f_{X_2}(x_2)\, dx_2$$

$$= \int_{x_2} [P(X_1 < x_2)]^2 f_{X_2}(x_2)\, dx_2 \quad [X_1 \,\&\, X_3 \text{ are identical}]$$

$$= \int_{x_2} F^2_{X_1}(x_2) \cdot f_{X_2}(x_2)\, dx_2 \quad [\text{c.d.f. of } X_1]$$

$$= \int_{x_2} F^2_{X_1}(x_2) \cdot f_{X_1}(x_2)\, dx_2 \quad [X_1, X_2 \to \text{Identical}]$$

$$= \frac{1}{3} F^3_{X_1}(x_2)\Big]_0^{\infty} = \frac{1}{3} F^3_{X_1}(\infty) = \frac{1}{3}$$

$$P(X_1 > X_2 < X_3) = \int_{x_2} P(X_1 > x_2 < X_3) f_{X_2}(x_2)\, dx_2$$

$$= \int_{x_2} [P(X_3 < x_2)]^2 f_{X_2}(x_2)\, dx_2 \quad [X_1, X_3 \to \text{identical}]$$

$$= \int_{x_2} F^2_{X_3}(x_2) \cdot f_{X_2}(x_2)\, dx_2 \quad [\text{c.d.f. of } X_3]$$

$$= \int_{x_2} F^2_{X_3}(x_2) \cdot f_{X_3}(x_2)\, dx_2 = \frac{1}{3} F^3_{X_3}(x_2)\Big]_0^{\infty}$$

$$= \frac{1}{3} F^3_{X_3}(\infty) = \frac{1}{3}$$

<u>OR</u>

$$P(X_1 > X_2 < X_3) = \int_{x_2} P(X_1 > x_2 < X_3)\, f_{X_2}(x_2)\, dx_2$$

$$= \int_{x_2} [P(X_1 > x_2)]^2 f_{X_2}(x_2)\, dx_2$$

$$= \int_{x_2} [1 - P(X_1 < x_2)]^2 \cdot f_{X_2}(x_2)\, dx_2$$

$$= \int_{x_2} f_{X_2}(x_2)\, dx_2 - 2\int_{x_2} P(X_1 < x_2)\, f_{X_2}(x_2)\, dx_2 + \int_{x_2} P(X_1 < x_2)^2 f_{X_2}(x_2)\, dx_2$$

$$= 1 - 2 \cdot \frac{1}{2} + \frac{1}{3} = \frac{1}{3}$$

Let us Define $\rightarrow$ $T = \sum_{i=2}^{n-1} I_{\{x_{i-1} > x_i < x_{i+1}\} \cup \{x_{i-1} < x_i > x_{i+1}\}}$ $\leftarrow$ Indicator fn

⊙ Mean

$$E(T) = \sum_{i=2}^{n-1} P[\{X_1 > X_2 < X_3\} \cup \{X_1 < X_2 > X_3\}]$$

$$= (n-2) \cdot \left[\frac{1}{3} + \frac{1}{3}\right] = \frac{2}{3}(n-2)$$

$\cancel{I_a = A}$

⊙ Variance

$\cancel{E(I_a) = 1}$

$$Var(T) = Var(I_2) + Var(I_3) + \cdots + Var(I_{n-1}) + 2\,Cov(I_2, I_3) + 2\,Cov(I_2, I_4) + \cancel{2\,Cov(I_2, I_5)} + \cdots + 2\,Cov(I_{n-2}, I_{n-1})$$

[$Cov(I_2, I_5) = 0$ as no association
$Cov(I_3, I_6) = 0$ as no association
& so on]

$$= (n-2)\,Var(I_2) + 2(n-3)\,Cov(I_2, I_3) + 2(n-4)\,Cov(I_2, I_4)$$

$$= \cancel{\frac{2}{9}(n-2)} + \frac{2}{9}(n-2) - \frac{n-3}{18} + \frac{(n-4)}{90} = \frac{20n - 40 - 5n + 15 + n - 4}{90} = \frac{16n - 29}{90}$$

$P(x_3 > x_4 > x_2)$
$P(x_4 > x_3 > x_2)$

$$Cov(I_2, I_3) = E(I_2 I_3) - E(I_2)E(I_3)$$

$$E(I_2 I_3) = P(X_1 < X_2 > X_3 < X_4) + P(X_1 > X_2 < X_3 > X_4)$$

$$= 2P(X_1 < X_2 > X_3 < X_4) = 2 \times [P(X_1 < X_2 > X_3) - P(X_1 < X_2 > X_3 > X_4)]$$

$$= 2 \times [P(X_1 < X_2 > X_3) - P(X_2 > X_3 > X_4) + P(X_1 > X_2 > X_3 > X_4)]$$

$$= 2 \times \left[\frac{1}{3} - \frac{1}{3!} + \frac{1}{4!}\right]$$

[As 6 cases appears so 3!. Similarly for 4!]

$$= 2 \times \left[\frac{1}{3} - \frac{1}{6} + \frac{1}{24}\right]$$

$$= \frac{1}{3} + \frac{1}{12} = \frac{5}{12}$$

$$Cov(I_2, I_3) = \frac{5}{12} - \frac{2}{3} \cdot \frac{2}{3} = \frac{5}{12} - \frac{4}{9} = \frac{15 - 16}{36} = -\frac{1}{36}$$

$$Cov(I_2, I_4) = E(I_2 I_4) - E(I_2)E(I_4)$$

$$E(I_2 I_4) = P(X_1 < X_2 > X_3 < X_4 > X_5) + P(X_1 < X_2 > X_3 > X_4 < X_5) + P(X_1 > X_2 < X_3 < X_4 > X_5) + P(X_1 > X_2 < X_3 > X_4 < X_5)$$

$$= 2P(X_1 > X_2 < X_3 < X_4 > X_5) + 2P(X_1 > X_2 < X_3 > X_4 < X_5)$$

$$= 2[P(X_2 < X_3 < X_4 > X_5) - P(X_1 < X_2 < X_3 < X_4 > X_5)] + 2 \times \frac{16}{5!}$$

$$= 2\left[\frac{1}{3!} - \frac{1}{4!} - P(X_1 < X_2 < X_3 < X_4) + P[X_1 < X_2 < X_3 < X_4 < X_5]\right] + \frac{32}{5!}$$

$$= 2 \times \left(\frac{1}{3!} - \frac{1}{4!} - \frac{1}{4!} + \frac{1}{5!}\right) + \frac{32}{5!} = 2 \times \frac{11}{5!} + \frac{32}{5!} = \frac{54}{5!} = \frac{54}{120} = \frac{9}{20}$$

$$Cov(I_2, I_4) = E[I_2 I_4] - E(I_2)E(I_4) = \frac{9}{20} - \frac{4}{9} = \frac{81 - 80}{180} = \frac{1}{180}$$

$$P[X_1 > X_2 < X_3 > X_4 < X_5] = 3! \times P[X_2 < X_4 < X_1 < X_3 < X_5] + 3! \times P[X_4 < X_2 < X_1 < X_3 < X_5] + 2 \times P[X_2 < X_1 < X_4 < X_3 < X_5] + 2 \times P[X_4 < X_5 < X_2 < X_1 < X_3]$$

$$= 16 \times \frac{1}{5!} = \frac{16}{5!}$$

(1) Why Portmanteau Test instead of ACF?

→ Even if it is not iid ACF at different lag small but not enough so if we sum them up they become large so their randomness comes into question.

(2)

$$x_t = t/_2 + \cos\left(t/_{10} * pi/_{180}\right) + rnorm(200, 0, 5)$$

↓

Comparing with $2\pi T$ we get → $T = \frac{1}{3600}$

- Trend is a dominating component.
- If ACF function lies down very slowly then there is Trend in the data
- ACF ~~f~~ cannot be monotonically increasing as it cannot reach 1.

(3) IID Normal →

$\{x_t\}$

H_0: $\{x_t\}$ i.i.d. normal

H_A: not i.i.d. normal.

EDF → Empirical Distribution Function

↓

To estimate a distribution function.

CDF → $F(x) = P(X \le x)$

In case of EDF, we do not assume that the ~~sample~~ data follows any distribution. Here we directly calculate the distribution function.

Eg → $\{2, 5, 4, 5, 6\}$ ($x_1\ x_2\ x_3\ x_4\ x_5$) $\quad n = 5 \rightarrow \{2, 4, 5, 5, 6\}$ ($x_{(1)}, x_{(2)}, x_{(3)}, x_{(4)}, x_{(5)}$) ← Ascending order (1st step)

If there is any tie then the 1st one is taken and the rest are dropped.

I will calculate $F_x(x)$ for my available x

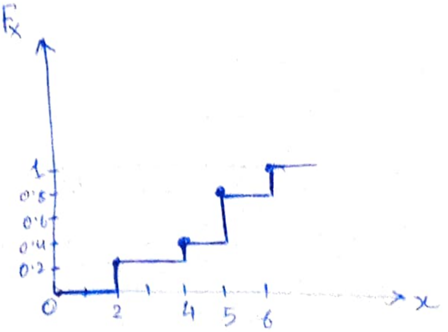
$[\because F_x(x) = P(X \le x)]$

$\therefore F_x(2) = \frac{1}{5}$

$F_x(4) = \frac{2}{5}$

$F_x(5) = \frac{4}{5}$

$F_x(6) = \frac{5}{5}$



This is estimated Distribution function for the given data.

General Definition $\rightarrow$ $F_X(x) = \dfrac{\text{\# data points lesser than equal to } x}{\text{\# total data points}}$

Here, $F_X(x) = \begin{cases} 0 & x < 2 \\ \frac{1}{5} & 2 \le x < 4 \\ \frac{2}{5} & 4 \le x < 5 \\ \frac{4}{5} & 5 \le x < 6 \\ 1 & x \ge 6 \end{cases}$

$$\boxed{F_Z(z) = \frac{j}{n} \quad ; Z_{(j)} \le Z < Z_{(j)+1}}$$

Here all datapoints are diff. [Continuous]

Sometimes, it is also written as –

$$\boxed{F_Z(z) = \frac{j - \frac{1}{2}}{n} \quad ; Z_{(j)} \le Z < Z_{(j)+1}}$$

This is more preferrable.

Q-Q Plot $\rightarrow$ Graphical check for normality

Steps $\rightarrow$ (i) Given an i.i.d. sequence $\{x_1, x_2, \dots, x_n\}$ turn it to standardized form $\{z_1, \dots, z_n\}$

(ii) Sort this to $\{z_{(1)}, \dots, z_{(n)}\}$

(iii) Corresponding to each $z_{(i)}$ calculate the associated quantile from standard normal. i.e. ~~$N_{\alpha/2(i)} = \frac{i - 0.5}{n}$~~ = [wrong]

$\boxed{P(Y \le N_{\alpha/2(i)}) = \frac{i - 0.5}{n}}$; $Y \sim N(0,1)$

(iv) Note: Empirical Distribution of $z_{(i)}$:

$$P(Z_{(i)} \le z_{(i)}) = \frac{\text{No. of points less than equal } z_i}{\text{Total Points}} = \frac{i}{n}$$

(iv) Plot the pair points $(N_{\alpha/2(i)}, z_{(i)})$ for $i = 1, \dots, n$

(v) If the sequence $\{x_1, x_2, \dots, x_n\}$ is coming from normal, the plot described above will be straight line passing through origin with slope 1.

- If Q-Q plot looks like the graph $y = x$ with slope 1, then we can say the data is coming from Normal Distribution.
- The maximum value of $F_Z(z) = \frac{j - \frac{1}{2}}{n}$ is when $j = n$.

  $\therefore F_Z(z) = \frac{n - \frac{1}{2}}{n} = 1 - \frac{1}{2n} \approx 1$ So Normal will have large (∞) value.

- We will check for the middle points of Q-Q plot and not the extremums.
- If we don't standardize the points then the slope and intercepts will change but the linearity holds.
- For any distribution the above process will work because the points will be checked for that distribution.

Shapiro Francia Test :→ Try to find the correlation b/w $N_{\alpha_{z(i)}}$ & $Z_{(i)}$

(i) Test Statistics, under normality assumption :

$$R^2 = \frac{\left[\sum_{i=1}^{n} N_{\alpha_{z(i)}} Z_{(i)}\right]^2}{\sum_{i=1}^{n}\left[N_{\alpha_{z(i)}}\right]^2 \sum_{i=1}^{n}\left[Z_{(i)}\right]^2}$$

All IID process is Stationary Process but not all Stationary process is IID process.

- IID process → x_t follows i.i.d. Random Variables.

If $\{x_t\}$ is stationary process →

$\{x_1, x_2, x_3, \ldots, x_n\}$ Let us take $x_{i_1}, x_{i_2}, x_{i_3}$

A joint distribution function with them is $F_{x_{i_1}, x_{i_2}, x_{i_3}}(a, b, c)$

$$= P(x_{i_1} \leq a, x_{i_2} \leq b, x_{i_3} \leq c) \quad \text{(i)}$$

Given for some integer value t, we take another 3 random variables

$x_{i_1+t}, x_{i_2+t}, x_{i_3+t}$. Joint Distribution is $F_{x_{i_1+t}, x_{i_2+t}, x_{i_3+t}}(a,b,c)$

$$= P(x_{i_1+t} \leq a, x_{i_2+t} \leq b, x_{i_3+t} \leq c) \quad \text{(ii)}$$

If for any n & t, (i) = (ii) then it will be called strictly stationary process.

All IID processes are strict stationary process. IID is special case of strict stationary process.

Stochastic Processes

⊙ Types of Stochastic Process

① $\{X_t\}$, X_t is i.i.d. and follows $N(0, \sigma^2)$ $\forall t = 1, 2 \ldots$

② $\{X_t\}$, X_t is i.i.d.

③ $\{X_t\}$ is ~~i.i.d.~~ stationary (strictly) $\Rightarrow$ $\{X_t\}$ is weakly stationary [Mathematically more tractable]

$$f_{X_{i_1}, X_{i_2}, \ldots, X_{i_n}}(x_1, x_2, \ldots, x_n)$$

$$= f_{X_{i_1+h}, X_{i_2+h}, \ldots, X_{i_n+h}}(x_1, x_2, \ldots, x_n) \quad \forall n, h > 0$$

All stationary processes are i.i.d. process.
We cannot reach the reality through any model.
Stationary Process $\longrightarrow$ Relaxing the assumptions of i.i.d. for reality.

■ $\{X_t\}$ weakly stationary Process $\rightarrow$

(i) $E(X_t^2) < \infty$ [i.e. 2nd Moment of X_t is finite]

(ii) $EX_t = \mu$ [Constant, independent of time]

(iii) Dependent Structure $\longrightarrow$ $\gamma_{X_{i_1}, X_{i_2}} = E[(X_{i_1} - \mu)(X_{i_2} - \mu)] = \gamma_{X_{i_1+h}, X_{i_2+h}}$

$$= E[(X_{i_1+h} - \mu)(X_{i_2+h} - \mu)]$$

⊙ It depends upon the ~~diff.~~ lag b/w i_1 & i_2
Not depends upon i_1 & i_2 values.

⊙ Correlation function
~~It~~ does not depend upon the ~~real~~ position of the two random variables rather depends upon the diff. position of the two random variables.
$\boxed{\gamma_{X_{i_1}, X_{i_2}} = f(|i_1 - i_2|)}$

$$E[(X_{i_1+h} - \mu)(X_{i_2+h} - \mu)] = \iint (x_1 - \mu)(x_2 - \mu) f_{X_{i_1+h}, X_{i_2+h}}(x_1, x_2)\, dx_1 dx_2$$

$$= \iint (x_1 - \mu)(x_2 - \mu) f_{X_{i_1}, X_{i_2}}(x_1, x_2)\, dx_1 dx_2 \quad [\because \text{Strictly Stationary}]$$

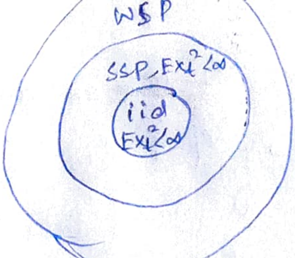
$$= E[(X_{i_1} - \mu)(X_{i_2} - \mu)]$$

Hence 3rd Condition Satisfied.
But 1st Condition will not be satisfied.

⊙ Cauchy Random Variable has not finite 2nd order moment and a Strictly Stationary Process.

⊙ A strict Stationary process with finite 2nd order moment, they are weak stationary process.

$\{X_t\}$, (weakly) Stationary Process,
iff the previous 3 conditions satisfied.



☒ White Noise : $\{X_t\} \sim WN(0,\sigma^2)$

- It's a sequence of uncorrelated random variables, each with zero mean and variance σ^2.

→ ACVF

$$\gamma_X(h) = \begin{cases} \sigma^2 & \text{if } h=0 \\ 0 & \text{if } h\neq 0 \end{cases}$$

ACF

$$\rho_X(h) = \begin{cases} 1 & \text{if } h=0 \\ 0 & \text{if } h\neq 0 \end{cases}$$

All i.i.d. process are white noise process but the converse is not true.

☐ $X_1, X_2 \rightarrow$ independent
$\iff f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \quad \forall\, x_1, x_2$

☐ $X_1, X_2 \rightarrow$ uncorrelated
$\iff \iint x_1 x_2 f_{X_1,X_2}(x_1,x_2)\,dx_1\,dx_2 = \int x_1 f_{X_1}(x_1)\,dx_1 \cdot \int x_2 f_{X_2}(x_2)\,dx_2$

Eg →

$\{Z_t\}$ ~~follows~~ iid $\sim N(0,1)$

$$X_t = \begin{cases} Z_t & t = \text{even} \\ (Z_{t-1}^2 - 1)/\sqrt{2} & t = \text{odd} \end{cases}$$

☐ X_t is WN →

- $E(X_t) = \begin{cases} E(Z_t) \\ E\left[(Z_{t-1}^2-1)/\sqrt{2}\right] \end{cases} = \begin{cases} 0 & [\because N(0,1)] \\ \dfrac{1-0-1}{\sqrt{2}} & [\because N(0,1)] \end{cases} = \begin{cases} 0 \\ 0 \end{cases} = 0$

- $Var(X_t) = Var(Z_t) = 1 \quad [\because t = \text{even}]$

$Var(X_t) = Var\left((Z_{t-1}^2-1)/\sqrt{2}\right)$

$= E\left[\left((Z_{t-1}^2-1)/\sqrt{2}\right)^2\right] - E\left[(Z_{t-1}^2-1)/\sqrt{2}\right]^2$

$= E\left[\dfrac{Z_{t-1}^4 - 2Z_{t-1}^2 + 1}{2}\right] - 0$

$= \dfrac{1}{2}\left[E(Z_{t-1}^4) - E(Z_{t-1}^2) + 1\right]$

$\Rightarrow$

$Var(X_t) = Var\left((Z_t^2-1)/\sqrt{2}\right) = \dfrac{1}{2} Var(Z_{t-1}^2)$

$= \dfrac{1}{2}\left\{E[Z_{t-1}^4] - E^2[Z_{t-1}^2]\right\}$

$= \dfrac{1}{2}\{3-1\} = 1$

$E[E[Z_t^2-1]] = E(1) = 1$

$E[Z_{t-1}^4]$ = 4th order moment of Standard Normal = Kurtosis = 3

$\gamma_X(0) = 1$ [Variance]

Not independent but uncorrelated.

$X_3 = \frac{(Z_2^2 - 1)}{\sqrt{2}}$ $X_2 = Z_2$

$\gamma_X(1) = E[X_3 X_2]$
$= E\left[\frac{(Z_2^2-1)}{\sqrt{2}} \cdot Z_2\right]$
$= \frac{1}{\sqrt{2}} E\left[Z_2^3 \cdot Z_2\right]$ $= \frac{1}{\sqrt{2}}\left[E(Z_2^3) - E(Z_2)\right]$
$= \frac{1}{\sqrt{2}}\left[E(Z_2^3)\right] = \frac{1}{\sqrt{2}} \times 0 = 0$ $\left[E(Z_2^3) = 0\right.$ as symmetric$\left.\right]$

$\gamma_X(2) = E(X_4 X_2)$
$= E(Z_4 Z_2) = E(Z_4) \cdot E(Z_2) = 0$

So, $\{X_t\}$ is $WN(0,1)$ but not $iid(0,1)$

First order moving average or MA(1) process:

$X_t = Z_t + \theta Z_{t-1}$, $t = 0, \pm 1, \ldots$

where $\{Z_t\} \sim WN(0, \sigma^2)$ and θ is real valued constant.

- $E(X_t) = 0$
- ACVF $\gamma_X(h) = \begin{cases} \sigma^2(1+\theta^2) & \text{if } h = 0 \\ \sigma^2\theta & \text{if } h = \pm 1 \\ 0 & \text{if } |h| \geq 2 \end{cases}$

- ACF $\rho_X(h) = \begin{cases} 1 & \text{if } h = 0 \\ \theta/(1+\theta^2) & \text{if } h = \pm 1 \\ 0 & \text{if } |h| \geq 2 \end{cases}$

$\gamma_X(0) = \gamma(Z_t) + \gamma(\theta Z_{t-1})$
$= var.(Z_t) + var.(\theta Z_{t-1})$
$= \sigma^2 + \theta^2 \cdot \sigma^2 = \sigma^2(1+\theta^2)$

$\gamma_X(1) = E[X_3 X_2]$
$= E[(Z_3 + \theta Z_2)(Z_2 + \theta Z_1)]$
$= E[Z_3 Z_2 + \theta Z_1 Z_3 + \theta Z_2^2 + \theta^2 Z_1 Z_2]$
$= E(Z_3 Z_2) + \theta E(Z_1 Z_3) + \theta E(Z_2^2) + \theta^2 E(Z_1 Z_2)$ [$\because$ WN]
$= E(Z_3)E(Z_2) + \theta E(Z_1)E(Z_2) + \theta E(Z_2^2) + \theta^2 E(Z_1)E(Z_2)$ [$\because$ uncorrelated]
$= 0 \cdot 0 + \theta \cdot 0 \cdot 0 + \theta\sigma^2 + \theta^2 \cdot 0 \cdot 0 = \theta\sigma^2$

$\gamma_X(2) = E[X_4 X_2]$
$= E[(Z_4 + \theta Z_3)(Z_2 + \theta Z_1)]$ $= E(Z_4 Z_2) + E\,\theta(Z_2 Z_3) + E(\theta Z_1 Z_4) + E(\theta^2 Z_1 Z_3)$
$= 0$

First Order Autoregressive Process or AR(1) process :→

$$X_t = \phi X_{t-1} + Z_t \qquad t = 0, \pm 1, \ldots$$

where $\{Z_t\} \sim WN(0, \sigma^2)$, $|\phi| < 1$ and Z_t is uncorrelated with X_s for each $s < t$.

- $E(X_t) = 0$
- ACVF

$$\gamma(h) = \phi^{|h|} \cdot \frac{\sigma^2}{1-\phi^2}$$

- ACF

$$\rho(h) = \phi^{|h|}$$

$$\cancel{\gamma(0) = Var.(X_t) = Var.[\phi(X_{t-1})] + Var.(Z_t)}$$

$$X_t = \phi X_{t-1} + Z_t = \phi^2 X_{t-2} + \phi Z_{t-1} + Z_t$$
$$= Z_t + \phi Z_{t-1} + \phi^2 \cdot Z_{t-2} + \ldots$$

$\gamma_X(0) =$
$$Var.(X_t) = Var.(Z_t) + Var.(\phi(Z_{t-1})) + Var.(\phi^2 Z_{t-2}) + \ldots$$
$$= \sigma^2 + \phi^2\sigma^2 + \phi^4 \cdot \sigma^2 + \ldots = \sigma^2(1 + \phi^2 + \phi^4 + \ldots)$$
$$= \sigma^2 \cdot \frac{1}{1-\phi^2} = \frac{\sigma^2}{1-\phi^2}$$
$$= \phi^0 \cdot \frac{\sigma^2}{1-\phi^2}$$

$$\cancel{\gamma_X(1) = E(X_3 X_2) = E[(\phi X_2 + Z_3)(\phi X_1 + Z_2)]}$$
$$\cancel{= \phi^2 E(X_2 X_1) + \phi E(X_2 Z_2) + \phi E(X_1 Z_3) + E(Z_3 Z_2)}$$
$$\cancel{= \phi}$$

$$\gamma_X(1) = E(X_3 X_2) = E[(\phi X_2 + Z_3) X_2] = \phi E(X_2^2) + E(X_2 Z_3)$$
$$= \phi \gamma_X(0) + 0 \quad [X_2, Z_3 = \text{Uncorrelated}]$$
$$= \phi \cdot \frac{\sigma^2}{1-\phi^2}$$

$$\gamma_X(2) = E[(\phi X_3 + Z_4) X_2] = E[(\phi^2 X_2 + \phi Z_3 + Z_4) X_2]$$
$$= \phi^2 E(X_2^2) + \phi E(Z_3 X_2) + E(Z_4 X_2)$$
$$= \phi^2 \cdot \frac{\sigma^2}{1-\phi^2} + 0 + 0$$
$$= \phi^2 \cdot \frac{\sigma^2}{1-\phi^2}$$

First Order ARMA or ARMA(1,1) process :⟶

$$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1} \qquad t = 0, \pm 1, \dots$$

where $\{Z_t\} \sim WN(0, \sigma^2)$, $|\phi| < 1$, Z_t is uncorrelated with X_s for each $s < t$ and $\phi + \theta \neq 0$

- $E(X_t) = 0$
- ACVF

$$\gamma_x(h) = \begin{cases} \sigma^2\left[1 + \frac{(\theta+\phi)^2}{1-\phi^2}\right] & \text{if } h = 0 \\ \sigma^2\left[\theta + \phi + \frac{(\theta+\phi)^2\phi}{1-\phi^2}\right] & \text{if } h = \pm 1 \\ \phi^{|h|-1}\gamma_x(1) & \text{if } |h| \geq 2 \end{cases}$$

$\gamma_x(0) = var.(X_t) = \phi^2 var.(X_{t-1}) + var.(Z_t) + \theta^2 var.(Z_{t-1})$

$= \cancel{\phi^2[\ \phi^2 var.(X_{t-1}) + \sigma^2 + \theta^2\sigma^2}$

$= \phi^2[$

$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1} = \phi^2 X_{t-2} + \phi Z_{t-1} + Z_t + \theta Z_{t-1}$

$= Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \theta Z_{t-1}$

$\gamma_x(0) = var(X_t) = \sigma^2 + \phi^2\sigma^2 + \phi^4\sigma^2 + \theta^2\sigma^2$

$= \cancel{\sigma^2} \ \cancel{\sigma^2(1+\theta^2)} + \sigma^2(1+\phi^2+\phi^4+\dots) + \theta^2\sigma^2$

$= \frac{\sigma^2}{1-\phi^2} + \theta^2\sigma^2$

$= \sigma^2 \frac{(1+\theta^2)}{1-\phi^2}$

$= \sigma^2\left[\frac{1-\phi^2+1+\theta^2 - 1 + \phi^2}{1-\phi^2}\right]$

$= \sigma^2\left[\frac{1}{1-\phi^2} + \theta^2\right]$

$= \sigma^2\left[\frac{1+\theta^2-\phi^2\theta^2}{1-\phi^2}\right]$

$= \sigma^2[\ \theta^2$

$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1} = \phi[\phi X_{t-2} + Z_{t-1} + \theta Z_{t-2}] + Z_t + \theta Z_{t-1}$

$= \cancel{\phi^2 X_{t-2} + \phi Z_{t-1} + \phi}$

$= \phi^2 X_{t-2} + \phi Z_{t-1} + \phi\theta Z_{t-2} + Z_t + \theta Z_{t-1}$

$= Z_t + \phi Z_{t-1} + \dots + \theta Z_{t-1} + \phi\theta Z_{t-2} + \dots$

$= Z_t + (\phi+\theta)Z_{t-1} + \dots$

$var.(X_t) \neq \sigma^2 + \phi^2\sigma^2 + \dots + \theta^2\sigma^2 + \phi^2\theta^2\sigma^2 + \dots$

$= \sigma^2 \cdot \frac{1}{1-\phi^2} + \theta^2\sigma^2 \cdot \frac{1}{1-\phi^2} = \cancel{\frac{\sigma^2}{1-\phi^2}[1+\theta^2]} = \sigma^2[$

$$\text{Var.}(X_t) = \sigma^2 + (\theta+\phi)^2\sigma^2 + \phi^2(\theta+\phi)^2\sigma^2 + \cdots$$
$$= \sigma^2\left[1+(\theta+\phi)^2\left(1+\phi^2+\cdots\right)\right]$$
$$= \sigma^2\left[1+\frac{(\theta+\phi)^2}{1-\phi^2}\right]$$

Properly →

$$X_t = \phi X_{t-1} + Z_t + \theta Z_{t-1} = Z_t + \theta Z_{t-1} + \phi^2 X_{t-2} + \phi Z_{t-1} + \phi\theta Z_{t-2}$$
$$\cancel{= Z_t + (\theta+\phi)Z_{t-1} + \phi(\theta+\phi)Z_{t-2}} +$$

$$= Z_t + \theta Z_{t-1} + \phi^2\left[\phi X_{t-3} + Z_{t-2} + \theta Z_{t-3}\right] + \phi Z_{t-1} + \phi\theta Z_{t-2}$$
$$= Z_t + \theta Z_{t-1} + \phi^3 X_{t-3} + \phi^2 Z_{t-2} + \phi^2\theta Z_{t-3} + \phi Z_{t-1} + \phi\theta Z_{t-2}$$
$$= Z_t + (\theta+\phi)Z_{t-1} + \phi(\theta+\phi)Z_{t-2} + \phi^2\theta Z_{t-3} + \phi^4 X_{t-4} + \phi^3 Z_{t-3} + \theta\phi^3 Z_{t-4}$$
$$= Z_t + (\theta+\phi)Z_{t-1} + \phi(\theta+\phi)Z_{t-2} + \phi^2[\theta+\phi]Z_{t-3} + \theta\phi^3 Z_{t-4} + \phi^4 X_{t-4}$$
$$= Z_t + (\theta+\phi)Z_{t-1} + \phi(\theta+\phi)Z_{t-2} + \phi^2[\theta+\phi]Z_{t-3} + \phi^3[\theta+\phi]Z_{t-4} + \phi^4[\theta+\phi]Z_{t-5} + \cdots$$

$$\gamma_X(0) = \text{Var.}(X_t) = \sigma^2 + (\theta+\phi)^2\sigma^2 + \phi^2(\theta+\phi)^2\sigma^2 + \phi^4(\theta+\phi)^2\sigma^2 + \cdots$$
$$= \sigma^2\left[1+\frac{(\theta+\phi)^2}{1-\phi^2}\right]$$

$$\gamma_X(1) = E(X_3X_2) = E\left[(\phi X_2 + Z_3 + \theta Z_2)X_2\right]$$
$$= E\left[\phi X_2^2 + X_2Z_3 + \theta X_2Z_2\right]$$
$$= \phi E(X_2^2) + E(X_2Z_3) + \theta E(X_2Z_2)$$
$$= \phi\gamma(0) + 0 + \theta E\left[(\phi X_1 + Z_2 + \theta Z_1)Z_2\right]$$
$$= \phi\cdot\sigma^2\left[1+\frac{(\theta+\phi)^2}{1-\phi^2}\right] + \theta\left[E(\phi X_1Z_2) + E(Z_2^2) + \theta E(Z_1Z_2)\right]$$
$$= \phi\sigma^2\left[1+\frac{(\theta+\phi)^2}{1-\phi^2}\right] + \theta\cdot\sigma^2$$
$$= \sigma^2\left[\theta+\phi+\frac{(\theta+\phi)^2\phi}{1-\phi^2}\right]$$

$\sigma^2 \; \sigma^2 \ldots$
$(n\sigma^2)$

$X_t = \phi X_{t-1} + Z_t \longrightarrow$ when $\phi < 1$ then it is stationary Process.

$X_t = X_{t-1} + Z_t \qquad Z_t \sim WN(0, \sigma^2)$

$\text{Cov.}(\cancel{X_m, X_n}) = \cancel{E(X_m)E(X_n)} \quad \cancel{\text{cov}(X_m, X_{n-1} + Z_n)} = \cancel{E(X_m}$

$= \cancel{E[(X_m - 0)(X_n - 0)]} \; : \; E[\cancel{X_m X_n}] =$

$$\text{Cov.}(X_m, X_n) = \text{Cov}(X_m, X_{n-1} + Z_n)$$
$$= \text{Cov}(X_m, (X_{n-2} + Z_{n-1} + Z_n))$$
$$= \text{Cov.}(X_m, X_{n-2}) + 0 + 0 = \text{Cov.}(X_m, X_{n-2})$$
$$= \text{Cov.}(X_m, X_{n-3} + Z_{n-1})$$
$$= \text{Cov.}(X_m, X_{n-3})$$

So, for $m < n$, $\text{Cov.}(X_m, X_n) = \text{Cov.}(X_m, X_m) = m \cdot \sigma^2$

& Similarly for $m > n$.

$\text{Cov.}(X_m, X_n) = \text{Cov.}(X_n, X_n) = n \cdot \sigma^2$

$\therefore \quad \text{Cov.}(X_m, X_n) = \min.(m, n)\sigma^2$

So, Cov. does not depend on the diff. Hence not a ~~weak~~ Stationary Process.

There are two kinds of Trends : $\longrightarrow$ (i) Deterministic trend
$(X_t = at + Z_t)$
(ii) Stochastic trend
$(X_t = X_{t-1} + Z_t)$

Diff b/w deterministic and stochastic Trend is the variance

For (i) Variance will be $t\sigma^2$
For (ii) Variance will be σ^2

Applying Method of Differencing, we get $\rightarrow$

(i) $\nabla X_t = Z_t$ $\qquad$ (ii) $\nabla X_t = a + (Z_t - Z_{t-1})$

Both ∇X_t is stationary Process

First one has lower variance.

Applying Estimation and Elimination, we get $\rightarrow$

(i) $\qquad$ (ii) $\nabla X_t = \hat{Z}_t$

For Stochastic Trend, Differencing operator is well but for Deterministic Trend Estimation and Elimination method is well.

q-order moving average or MA (q) process $\longrightarrow$

$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \qquad t = 0, \pm 1, \ldots$$

$\{Z_t\} \sim WN(0, \sigma^2)$ $\quad \theta_1, \ldots, \theta_q \rightarrow$ Real valued Constants.

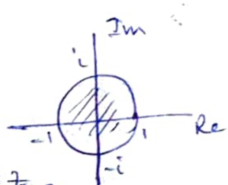
p-order autoregressive or AR (p) process $\longrightarrow$

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t, \quad t = 0, \pm 1, \ldots$$

$\{Z_t\} \sim WN(0, \sigma^2)$, $Z_t \rightarrow$ Uncorrelated with X_s for $s < t$.

and all roots of the polynomials $(1 - \phi_1 z - \phi_2 z^2 \cdots - \phi_p z^p)$ lie outside the unit circle.

$$1 - \phi_1 z - \cdots \quad \therefore \frac{1}{|\phi_1|} > 1 \Rightarrow |\phi_1| < 1$$



ARMA (p,q) process $\longrightarrow$

$$X_t = \phi X_{t-1} + \cdots + \phi_p X_{t-p} + Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q} \qquad t = 0, \pm 1, \ldots$$

$\{Z_t\} \sim WN(0, \sigma^2)$, with all roots of the polynomial $(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p)$ lie outside the unit circle and the polynomials $(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p)$ and $(1 + \theta_1 z + \cdots + \theta_q z^q)$ have no common factors.

- For ARMA (1,1) $\longrightarrow$ $1 - \phi_1 z \neq 1 + \theta_1 z \qquad z(\theta_1 + \phi_1) \neq 0$

Each Stationary Processes are linearly dependent on White Noise.

$$\underset{\text{(White Noise)}}{Z_t} \longrightarrow \boxed{\text{Linear Process}} \longrightarrow X_t$$

MA (1) $\longrightarrow$ $X_t = Z_t + \theta_1 Z_{t-1} = (1 + \theta_1 B) Z_t$ $\quad$ [B $\rightarrow$ Backshift Operator]

MA (2) $\longrightarrow$ $X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} = (1 + \theta_1 B + \theta_2 B^2) Z_t$

MA (q) $\longrightarrow$ $X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \cdots + \theta_q Z_{t-q}$
$= (1 + \theta_1 B + \cdots + \theta_q B^q) Z_t$

AR(1) $\longrightarrow$ $X_t = \phi X_{t-1} + Z_t = Z_t + \phi Z_{t-1} + \phi^2 Z_{t-2} + \cdots$
$= (1 + \phi B + \phi^2 B^2 + \cdots) Z_t$

Linear Process → The process $\{X_t\}$ is linear process if

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \quad \forall t$$

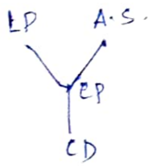
$\{Z_t\} \sim WN(0,\sigma^2)$ and $\{\psi_j\}$ is a sequence of constants with $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$

Alternate representation by backward shift operator: $X_t = \psi(B) Z_t$

$$\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$$

$\psi(B)$ is linear ~~operator~~ filter which is applied on white noise 'input' series $\{Z_t\}$ produces the 'output' $\{X_t\}$.

- The condition $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ (absolute summability) ensures that the infinite sum $X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$ converges (with probability one)



LP → LP Convergence
A.S. → Almost Sure Convergence
CP → Convergence in Probability
CD → Convergence in Distribution.

Chebyshev's Inequality → $P(|X - E(X)| > \epsilon) \leq \frac{Var.(X)}{\epsilon^2}$

Pdf ≥ 0 ⇒ Never Converges to −ve number

$\int f(x)dx = 1$ ⇒ Cannot Converge to any other value than 0.

- When $X_1, \ldots, X_n \xrightarrow{p} X$. then $\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$
- When $X_1, \ldots, X_n \xrightarrow{a.s.} X$ then $P(\{w : \lim_{n\to\infty} X_t^n(w) = X_t(w)\}) = 1$

Proof $\rightarrow$

$$X^n = \sum_{j=-n}^{n} \psi_j Z_{t-j}$$ for small $\varepsilon > 0$, define

$$A_n(\varepsilon) = \{|x_t^n - x_t| > \varepsilon\} = \{|\sum_{|j|>n} \psi_j Z_{t-j}| > \varepsilon\}$$

$$= \{w : |x_t^n(w) - x_t(w)| > \varepsilon\}$$ $$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}$$

By Chebyshev's Inequality $\rightarrow$ $= \{w : |\sum_{|j|>n} \psi_j Z_{t-j}| > \varepsilon\}$

$$P(A_n) \le E\left[\left|\sum_{|j|>n} \psi_j Z_{t-j}\right|^2\right] / \varepsilon^2$$

We try to show $\rightarrow$

$$\sum_{n=1}^{\infty} P(A_n) < \infty$$

$$\sum_{n=1}^{\infty} P(A_n) \le \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} E\left|\sum_{|j|>n} \psi_j Z_{t-j}\right|^2$$

$$\le \sum_{n=1}^{\infty} E\left|\sum_{|j|>n} \psi_j Z_{t-j}\right|^2 = \sum_{n=1}^{\infty} E\left|\sum_{|j|>n} \sum_{|k|>n} \psi_j \psi_k (Z_{t-j} Z_{t-k})\right|$$

$|ab| \le |a||b|$

$$\le \sum_{n=1}^{\infty}\left[\sum_{|j|>n} \sum_{|k|>n} |\psi_j||\psi_k| E|(Z_{t-j} Z_{t-k})|\right]$$

[Triangular Inequality]

$$\le \sum_{n=1}^{\infty}\left[\sum_{|j|>n} \sum_{|k|>n} |\psi_j||\psi_k| |E(Z_{t-j}^2)|^{1/2} |E|(Z_{t-k}^2)|^{1/2}|\right]$$

[Cauchy Schwarz Inequality]

$$\le \sigma^2 \sum_{n=1}^{\infty}\left[\sum_{|j|>n} \sum_{|k|>n} |\psi_j||\psi_k|\right]$$ Stationarity of WN

$$= \sigma^2 \sum_{n=1}^{\infty}\left[\sum_{|j|>n} |\psi_j|\right]^2 < \infty$$, absolute summability.

By Borel Cantelli Lemma $\rightarrow$

$$\sum_{n=1}^{\infty} P(A_n) < \infty \Rightarrow P(A^{(\varepsilon)}) = 0$$, where $A^{(\varepsilon)} = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$

$= \limsup_n A_n$

- $A^{(\varepsilon)}$ called lim sup event of the infinite sequence $\{A_n\}$
- $A^{(\varepsilon)}$ occurs iff $\forall n \ge 1$, $\exists$ a $m \ge n$ such that A_m occurs.
- Equivalently, $A^{(\varepsilon)}$ occurs iff infinitely many of the A_m occur.

By definition of limit. $\omega \in \underset{n \to n}{lt}\, X_n = X$. iff $\forall\, u \geq 1\ \exists\, n \geq 1$ such that for every $m \geq n$, $|X_m(\omega) - X(\omega)| \leq \frac{1}{u}$. Equivalently it holds iff

$$\omega \in \bigcap_{u=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcap_{m \geq n}^{\infty} \left[A_m\left(\tfrac{1}{u}\right)\right]^c = \left(\bigcup_{u}^{\infty} \bigcup_{n \geq 1}^{\infty} \limsup_{n} \left(\tfrac{1}{u}\right)\right)^c$$

Thus, $P\left(\omega : \lim_{n \to \infty} X_t^n(\omega) = X_t(\omega)\right) = P\left(\left(\bigcup_{n=1}^{\infty} \limsup_n A_n\left(\tfrac{1}{u}\right)\right)^c\right)$

$$= 1 - P\left(\bigcup_{n=1}^{\infty} \limsup_n A_n\left(\tfrac{1}{u}\right)\right)$$

$$\geq 1 - \sum_{u=1}^{\infty} P\Big(\underbrace{\limsup_n A_n\left(\tfrac{1}{u}\right)}_{0}\Big) = 1$$

So, almost convergence to 1

The condition $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ ensures that $\sum_{j=-\infty}^{\infty} \psi_j^2 < \infty$, therefore,

the series $\sum$

Lp ~~Norm~~ Convergence $\longrightarrow \underset{n \to \infty}{lt} E|X_n - X|^p = 0$

∴ L2 convergence / Mean square convergence $\longrightarrow \underset{n \to \infty}{lt} E|X_n - X|^2 = 0$

If we apply any stationary Process with mean 0 and covariance function γ_Y through Linear Filter it will give Stationary Process.

$$\xrightarrow{SP\ Y_t} \boxed{\text{Linear Filter}} \Longrightarrow \underset{SP}{X_t}$$

if $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$, then the process $X_t = \psi(B) Y_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j}$

is also stationary with mean 0 and autocovariance function as

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k E[Y_{t-j} Y_{t+h-k}]$$

$$= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k [h+j-k]$$

[Depends on difference]

Causal Process $\longrightarrow$ A linear process $\{X_t\}$ is called causal if the terms of current and past values $Z_s, s \le t$

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$$

Invertible $\longrightarrow$ A linear Process $\{X_t\}$ is invertible if Z_t can be expressed in terms of the current and past values

$X_t \to$ ARMA (p, q) process by Backshift operator.

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) X_t = (1 + \theta_1 B + \dots + \theta_q B^q) Z_t$$

$$\Rightarrow X_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)^{-1} (1 + \theta_1 B + \dots + \theta_q B^q) Z_t$$

AR(1) $\longrightarrow$ $(1 - \phi_1 B) X_t = Z_t \Rightarrow X_t = (1 - \phi_1 B)^{-1} Z_t$

$$= (1 + \phi_1 B + \phi_1^2 B^2 + \dots) Z_t$$

$|\phi_1| < 1 \leftarrow$ condition for this

Exercise $\to$ 3.1.b $\quad X_t + 1.9 X_{t-1} + 0.88 X_{t-2} = Z_t + 0.2 Z_{t-1} + 0.7 Z_{t-2}$

$$\therefore (1 + 1.9B + 0.88B^2) X_t = (1 + 0.2B + 0.7B^2) Z_t$$

Causality $\to$

$$(1 + 1.9z + 0.88z^2) = (1 + 0.8z)(1 + 1.1z)$$

$$\therefore z_1 = -\frac{1}{0.8}$$

$$z_1 = -\frac{1}{1.1} \qquad 1 > z_2 > -1$$

Inside Unit Disc So, Not causal.

Invertibility $\to$

$$1 + 0.2z + 0.7z^2$$

$$\therefore z = \frac{-0.2 \pm \sqrt{0.04 - 2.8}}{1.4} = \frac{-0.1 \pm \sqrt{-0.69}}{0.7}$$

$$= -0.14 \pm 1.19i$$

$\therefore |z| = 1.198$ $\longrightarrow$ Not Inside unit circle $\therefore$ invertible

Brockwell . Davis →

Ex – 2.2 $\quad X_t = A\cos(wt) + B\sin(wt) \qquad t = 0, \pm 1, \ldots$

$A, B \to$ uncorrelated $\qquad$ variance $= 1 \qquad$ mean $= 0$

$E(X_t) = \cos(wt)\cdot E(A) + \sin(wt)\cdot E(B) = 0 + 0 = 0$

$\mathrm{var}.(X_t) = \cos^2(wt)\,\mathrm{Var}.(A) + \sin^2(wt)\,\mathrm{Var}.(B)$

$= \cos^2 wt + \sin^2 wt = 1$

$E(X_t^2) = \mathrm{var}.(X_t) + E(X_t) = 1 + 0 = 1 < \infty$

$\gamma_X(h) = \mathrm{Cov}.(A\cos wt + B\sin wt,\ A\cos w(t+h) + B\sin w(t+h))$

$= \cos wt \cos w(t+h)\,\mathrm{Cov}(A,A) + \sin w(t+h)\cos wt\,\mathrm{Cov}(A,B)$

$+ \cos w(t+h)\sin wt\,\mathrm{Cov}(A,B)$

FORECASTING

$X_1, X_2, X_3, \ldots, X_n, X_{n+1}$?

$X_{n+1} = f(X_1, X_2, \ldots, X_n)$

Our main goal is to determine the value of f.

For AR(p) $\to f = \Phi_1 X_n + \Phi_2 X_{n-1} + \cdots + \Phi_p X_{n-p+1} + Z_{n+1}$

We can find the values of Φ_i by Regression or Curve Fitting

so for $p < n$

$X_{p+1} = \Phi_1 X_p + \Phi_2 X_{p-1} + \cdots + \Phi_p X_1$

$X_{p+2} = \Phi_1 X_{p+1} + \Phi_2 X_p + \Phi_3 X_{p-1} + \cdots + \Phi_p X_2$

Dependent Variable

$\vdots$

$X_n = \Phi_1 X_{n-1} + \Phi_2 X_{n-2} + \cdots + \Phi_p X_{n-p}$

$Y = AX$.

- Forecasting AR Model →

$\hat{X}_{n+h} = a_0 + a_1 X_n + \cdots + a_n X_1 = X^n_{n+h}$ → Here a_0 can be ommitted when $\mu = 0$

Normal Equations →

$$E\left[X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right] = 0 \qquad \text{(i)}$$

$$E\left[\left(X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}\right) X_{n+1-j}\right] = 0 \quad \text{for } j = 1, \dots, n \qquad \text{(ii)}$$

(i) gives → $E(X_{n+h}) - E(a_0) - \sum_{i=1}^{n} a_i E(X_{n+1-i}) = 0$

$$\Rightarrow \mu - a_0 - \sum_{i=1}^{n} a_i \cdot \mu = 0$$

$$\Rightarrow a_0 = \mu\left[1 - \sum_{i=1}^{n} a_i\right]$$

(ii) gives → $\underbrace{E\left[X_{n+h} X_{n+1-j}\right]}_{K_1} - a_0 E\left[X_{n+1-j}\right] - \sum_{i=1}^{n} a_i \underbrace{E\left(X_{n+1-i} X_{n+1-j}\right)}_{K_2} = 0$

$$\Rightarrow \overset{K_1}{\cancel{\gamma_{n+h-n-1+j}}} - a_0 \mu - \sum_{i=1}^{n} a_i \overset{K_2}{\cancel{\gamma_{n+1-i-n-1+j}}} = 0$$

$$\Rightarrow \overset{K_1}{\cancel{\gamma_{h-1+j}}} - a_0 \mu - \sum_{i=1}^{n} a_i \overset{K_2}{\cancel{\gamma_{j-i}}} = 0$$

$K_1 = E\left[X_{n+h} X_{n+1-j}\right]$ $\quad K_2 = E\left[X_{n+1-i} X_{n+1-j}\right]$ (iii)

Solution is →

$$\Gamma_n a_n = \gamma_n(h)$$

Here

$$\Gamma_n = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & & & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{bmatrix} \qquad a_n = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$\gamma_n(h) = \left[\gamma(h) \quad \gamma(h+1) \quad \cdots \quad \gamma(h+n-1)\right]^t$$

∴ From (iii) → $j = 1$

$$E\left[X_{n+h} X_n\right] = a_0 \mu + \sum_{i=1}^{n} a_i E\left[X_{n+1-i} X_n\right]$$

$$= \mu\left[1 - \sum_{i=1}^{n} a_i\right]\mu + \sum_{i=1}^{n} a_i E\left[X_{n+1-i} X_n\right]$$

$$\Rightarrow E\left[X_{n+h} X_n\right] - \mu^2 = \left\{E\left[X_{n+1-i} X_n\right] - \mu^2\right\} \sum_{i=1}^{n} a_i$$

$$\Rightarrow \gamma_h = \sum_{i=1}^{n} a_i \gamma(1-i) = a_1 \gamma(0) + a_2 \gamma(1) + \cdots + a_n \gamma(n-1)$$

To solve the equations we have to do matrix inversion. But if the ~~var~~ variables are dependent then the eigen values will be close to zero. So, we cannot do this.

We will do the following process $\longrightarrow$

- Best Linear Unbiased Estimator $\longrightarrow$ $\underbrace{X^{n}_{n+h}}_{\text{Predicting } n+h \text{ by using previous } n \text{ values}} = \mu + a_n' (X_n - \mu 1_n)$

$$\bar{X}_n = [X_n, X_{n-1}, \dots X_1]^t \qquad 1_n = [\underbrace{1, \dots, 1}_{n \text{ times}}]^t$$

- Expected value of the prediction error [i.e. first normal equation]

$$E[X_{n+h} - X^{n}_{n+h}] = 0$$

~~[both are~~

- Mean Square Prediction Error $\longrightarrow$

$$E(X_{n+h} - X^{n}_{n+h})^2 = E\left[(X_{n+h} - \mu) - a_n^t(\bar{X}_n - \mu 1_n)\right]^2$$

$$= E[(X_{n+h}-\mu)^2] - 2a_n^t E[(X_{n+h}-\mu)(\bar{X}_n - \mu \cdot 1_n)]$$

$$- a_n^t E[(\bar{X}_n - \mu 1_n)(\bar{X}_n - \mu 1_n)^t] \times a_n \quad (*)$$

$$E[(X_{n+h}-\mu)(\bar{X}_n - \mu \cdot 1_n)] = E\left[(X_{n+h}-\mu)\begin{bmatrix} X_n - \mu \\ X_{n-1} - \mu \\ \vdots \\ X_1 - \mu \end{bmatrix}\right] = \begin{bmatrix} \gamma(h) \\ \gamma(h+1) \\ \vdots \\ \gamma(h+n-1) \end{bmatrix} = \gamma_n(h)$$

$$E[(X_{n+h} - \mu)^2] = \text{Var.}(X_{n+h}) = \gamma(0)$$

$$(*) \quad a_n^t E[(\bar{X}_n - \mu \cdot 1_n)(\bar{X}_n - \mu \cdot 1_n)^t] a_n$$

$$= a_n^t \begin{bmatrix} (X_n - \mu)(X_n - \mu) & \cdots \\ (X_{n-1}-\mu)(X_n - \mu) & \cdots \\ \vdots & \\ (X_1 - \mu)(X_n - \mu) & \cdots \end{bmatrix} a_n = a_n^t \Gamma_n(h) a_n$$

$(*)$ gives $\longrightarrow$

$$E(X_{n+h} - X^{n}_{n+h})^2 = \gamma(0) - 2a_n^t \gamma_n(h) + a_n^t \Gamma_n(h) a_n$$

$$= \gamma(0) - 2a_n^t \gamma_n(h) + a_n^t \gamma_n(h) \qquad [\Gamma_n a_n = \gamma_n(h)]$$

$$= \gamma(0) - a_n^t \gamma_n(h)$$

$$= \gamma(0) - \gamma_n^t(h) \cdot \Gamma_n^{-1} \gamma_n(h)$$

Eg → One step prediction of an AR(1) series

$$X_t = \phi X_{t-1} + Z_t \quad t = 0, \pm 1, \dots$$

where $|\phi| < 1$ and $Z_t \sim WN(0, \sigma^2)$

→ Here $Z_t \sim WN(0, \sigma^2)$ $\therefore \mu = 0$

$\therefore a_0 = 0$

$$X_{n+1}^n = a_n^t X_n$$

$X_n = [X_n, X_{n-1}, \dots, X_1]^t \quad a_n = [\phi, 0, \dots, 0]^t$

$$\begin{bmatrix} 1 & \phi & \phi^2 & \dots & \phi^{n-1} \\ \phi & 1 & \phi & \dots & \phi^{n-2} \\ \vdots & & \ddots & & \\ \phi^{n-1} & \phi^{n-2} & \phi^{n-3} & \dots & 1 \end{bmatrix} a_n = \begin{bmatrix} \phi \\ \phi^2 \\ \vdots \\ \phi^n \end{bmatrix}$$

Therefore the best linear predictor of X_{n+1} in terms of $\{X_1 \dots X_n\}$ is

$$X_{n+1}^n = a_n^t X_n = \phi X_n$$

Mean Square Error is $E(X_{n+1} - X_{n+1}^n)^2 = \cancel{\gamma(0) - a_n^t \gamma_n(1)}$
$= \cancel{\gamma(0)}$

$= \gamma(0) - a_n^t \gamma_n(1)$

$= \gamma(0) - \phi\gamma(1) = \gamma(0)[1 - \phi\rho(1)]$

$= \gamma(0)[1 - \phi^2] = \frac{\sigma^2}{1-\phi^2}(1-\phi^2)$

$= \sigma^2$ → Variance Unexplained

When we are predicting X_{n+1} by $E(X_{n+1})$ then the error $= X_{n+1} - E(X_{n+1})$. Because if we take $E(X_{n+1})$ as predictor then the $E(\text{error}) = 0$

Now, $E[(\text{error})^2] = \gamma(0) = \frac{\sigma^2}{1-\phi^2}$ → Total Variance

Total Variance $\left[\frac{\sigma^2}{1-\phi^2}\right]$ → Unexplained Variance $[\sigma^2]$

Total Variance $\left[\frac{\sigma^2}{1-\phi^2}\right]$ → Explained Variance $\left[\frac{\sigma^2\phi^2}{1-\phi^2}\right]$

Durbin - Levinson Algorithm :→

It is known as One Step Recursive Forecast $\phi_n = \begin{bmatrix} \phi_{n1} \\ \phi_{n2} \\ \vdots \\ \phi_{nn} \end{bmatrix}$

- Set a one step predicting equation based on a single (current) observation.

$$X^{n,n}_{n+1} = \phi_{11} X_n$$

- Compute ϕ_{11} and v_0 as follows →

$$\phi_{11} = \frac{\gamma(1)}{\gamma(0)} \quad \text{and} \quad v(0) = \gamma(0)$$

$$E[(X_{n+1} - \phi_{11} X_n)^2] = E[(X_{n+1} - \phi_{11} X_n) X_n]$$
$$= E[X_{n+1} \cdot X_n] - \phi_{11} E[X_n^2]$$
$$= \gamma(1) - \phi_{11} \gamma(0)$$

This $= 0 \Rightarrow \gamma(1) - \phi_{11}\gamma(0) = 0 \Rightarrow \phi_{11} = \frac{\gamma(1)}{\gamma(0)}$

$$v_0 = E[(X_{n+1} - \text{Mean})^2] = E(X_{n+1} - 0)^2$$
$$= E(X_{n+1})^2 = \gamma(0)$$

$\boxed{v_i \to \text{Error}}$

$$X^{n,n}_{n+1} = \phi_{11} X_n \to v_1 \qquad X^{n-1,n}_{n+1} = \phi_{21} X_n + \phi_{22} X_{n-1} \to v_2$$

$$X^{1,n}_{n+1} = \phi_{n,1} X_n + \phi_{n,2} X_{n-1} + \cdots\cdots + \phi_{n,n} X_1 \to v_n.$$

Recursively, set one step predicting equations based on (current) n observation.

$$X^n_{n+1} = X^{1,n}_{n+1} = \phi_{n1} X_n + \cdots\cdots + \phi_{nn} X_1$$

We can compute the coefficients $\phi_{n1}, \ldots, \phi_{nn}$

$$\phi_{nn} = \left[\gamma(n) - \sum_{j=1}^{n-1} \phi_{n-1,j}\, \gamma(n-j)\right] v^{-1}_{n-1}$$

Alternative Compact form →

$$\phi_{nn} = \left[\gamma(n) - \phi^{(r)t}_{n-1} \gamma_{n-1}\right] v^{-1}_{n-1} \quad \text{(i)}$$

$\phi^{(r)}_k = [\phi_{k,k}, \ldots \phi_{k,1}]^t$ → reverse

- $\begin{bmatrix} \phi_{n1} \\ \vdots \\ \phi_{n,n-1} \end{bmatrix} = \phi_{n-1} - \phi_{nn} \phi^{(r)}_{n-1}$ $\qquad v_n = v_{n-1}[1 - \phi_{nn}^2]$

Proof $\rightarrow$

$\Gamma_1 \Phi_1 = \gamma_1$ follows from $\gamma(0)\phi_1 = \gamma(1)$

Let us consider $\Gamma_n \Phi_n = \gamma_n$ be true for $n = k$.

$$\Gamma_{k+1}\Phi_{k+1} = \begin{bmatrix} \Gamma_k & \gamma_k^{(r)} \\ \gamma_k^{(r)t} & \gamma(0) \end{bmatrix} \begin{bmatrix} \Phi_k - \phi_{k+1,k+1}\Phi_k^{(r)} \\ \phi_{k+1,k+1} \end{bmatrix}$$

$$= \begin{bmatrix} \Gamma_k \Phi_k - \phi_{k+1,k+1}\Gamma_k \Phi_k^{(r)} + \phi_{k+1,k+1}\gamma_k^{(r)} \\ \gamma_k^{(r)t}\Phi_k - \phi_{k+1,k+1}\gamma_k^{(r)t}\Phi_k^{(r)} + \gamma(0)\phi_{k+1,k+1} \end{bmatrix}$$

$$= \begin{bmatrix} \gamma_k - \phi_{k+1,k+1}\gamma_k^{(r)} + \phi_{k+1,k+1}\gamma_k^{(r)} \\ \gamma_k^{(r)t}\Phi_k + \phi_{k+1,k+1}\left(\gamma(0) - \gamma_k^{(r)t}\Phi_k^{(r)}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \gamma_k \\ \gamma_k^{(r)t}\Phi_k + \phi_{k+1,k+1} v_k \end{bmatrix} \quad \text{[From (i)]}$$

$$= \begin{bmatrix} \gamma_k \\ \gamma(k+1) \end{bmatrix} = \gamma_{k+1}$$

So, it is true for all n.

Mean Square Error $\rightarrow$

Proof $\rightarrow$

Let $v_n = v_{n-1}\left[1 - \phi_{nn}^2\right]$ be true for $n = k$.

$$v_{k+1} = \gamma(0) - \Phi_{k+1}^t \gamma_{k+1}$$

$$= \gamma(0) - [\phi_{k+1,1} \dots, \phi_{k+1,k}]\gamma_k - \phi_{k+1,k+1}\gamma(k+1)$$

$$= \gamma(0) - \left(\Phi_k^t - \phi_{k+1,k+1}\Phi_k^{(r)t}\right)\gamma_k - \phi_{k+1,k+1}\gamma(k+1) \quad \text{[From (ii)]}$$

$$= v_k - \phi_{k+1,k+1}\left(\gamma(k+1) - \Phi_k^{(r)t}\gamma_k\right) \quad \text{[By assumption]}$$

$$= v_k - \phi_{k+1,k+1}\left(\phi_{k+1,k+1} v_k\right) \quad \text{[From (i)]}$$

$$= v_k\left(1 - \phi_{k+1,k+1}^2\right)$$

So, it is true for all n.

$$\hat{X}_{n+1} = X_{n+1}^{1,n} = \theta_{n,1}(X_n - \hat{X}_n) + \theta_{n,2}(X_{n-1} - \hat{X}_{n-1}) + \cdots + \theta_{n,n}(X_1 - \hat{X}_1)$$

Here, we consider the problem of predicting the values X_{n+h}, $h>0$, of a stationary time series with known mean and auto covariance function in terms of the values of successive differences prediction $\{X_n - \hat{X}_n\}$ and so on.

- We follow MA process here.

As started with X_1 · $\hat{X}_1 = 0$

- Here $\theta(X_n - \hat{X}_n)$ is proxy of Z_n.

The Innovations Algorithm :→

Set the predicting equation of time series at time $(n+1)$

$$X_{n+1}^{n} = \sum_{j=1}^{n} \theta_{nj}\left(X_{n+1-j} - X_{n+1-j}^{n-j}\right) \qquad n = 1, 2, \ldots \qquad X_1^0 = 0$$

Compute the coefficients $\theta_{n1}, \ldots, \theta_{nn}$ from the following eqn. →

$$v_0 = \gamma(0)$$

$$\theta_{n,n-k} = v_k^{-1}\left(\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j}\,\theta_{n,n-j}\, v_j\right) \qquad 0 \le k < n$$

$$v_n = \gamma(0) - \sum_{j=0}^{n-1} \theta^2_{n,n-j}\, v_j$$

$$X_2^1 = \theta_{11} X_1 \,;\, v_1 \qquad X_1^0 = 0$$

$$X_3^2 = \theta_{21}(X_2 - X_2^1) + \theta_{22}(X_1 - X_1^0) \,;\, v_2$$

$$X_4^3 = \theta_{31}(X_3 - X_3^2) + \theta_{32}(X_2 - X_2^1) + \theta_{33}(X_1 - X_1^0) \,;\, v_3$$

$$\vdots$$

$$X_{n+1}^{n} = \theta_{n1}(X_n - X_n^{n-1}) + \cdots + \theta_{nn}(X_1) \,;\, v_n$$

- Remarks →

(i) The one step prediction error, $U_n = X_n - X_n^{n-1}$ is named as innovation at time n.

(ii) Innovations $U_1, U_2, \ldots, U_n$ are uncorrelated.

[U_i's are uncorrelated due to the normal eqn. In case of regression the error terms are uncorrelated with predictors.]

Regression →

$Y = a_1 X_1 + \cdots + a_n X_n$

$Y - \hat{a}_1 X_1 - \cdots - \hat{a}_n X_n \longleftrightarrow$ Uncorrelated $X_1, \ldots, X_n$.

$$\hat{X}_{n+1} = \sum_{j=1}^{n} \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j}) \qquad \hat{X}_{k+1} = \sum_{j=0}^{k-1} \theta_{k,j+1} (X_{k-j} - \hat{X}_{k-j})$$

$$\theta_{n,n-k} = v_k^{-1} \left[\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j} \langle X_{n+1}, X_{j+1} - \hat{X}_{j+1} \rangle \right] \qquad \sum_{j=0}^{k-1} \theta_{k,k-j} (X_{j+1} - \hat{X}_{j+1})$$

$$= v_k^{-1} \left[\gamma(n-k) - \sum_{j=0}^{k-1} \theta_{k,k-j} \theta_{n,n-j} v_j \right]$$

h-Step Recursive Forecast :→

The predicting equation of time series at time $n+h$ depending on the n ~~obs~~ observations is as follows —

$$X_{n+h}^{n} = \sum_{j=h}^{n+h-1} \theta_{n+h-1,j} (X_{n+h-j} - X_{n+h-j}^{n+h-1-j})$$

$n = 1, 2, \ldots \qquad X_1^0 = 0$

Corresponding Mean Squared Error →

$$E(X_{n+h} - X_{n+h}^{n})^2 = \gamma(0) - \sum_{j=h}^{n+h-1} \theta_{n+h-1,j}^2 \, v_{n+h-1-j}$$

$x_1, x_2, \ldots\ldots, x_n$

$$\hat{x}_{n+1} = \phi_{n1} x_n + \phi_{n2} x_{n-1} + \ldots\ldots + \phi_{np} x_{n-p+1} + \ldots\ldots + \phi_{nn} x_1$$
$\hookrightarrow$ AR Process

$$\hat{x}_{n+1} = \theta_{n1}(x_n - \hat{x}_n) + \theta_{n2}(x_{n-1} - \hat{x}_{n-1}) + \ldots\ldots + \theta_{nn}(x_1 - \hat{x}_1)$$
$\hookrightarrow$ MA Process

$\hat{x}_1 = 0$

If it is only given that it is a stationary process then the following questions appear—

Q1. Check whether AR/MA/ARMA?

Q2. Order of the process?

Q3. Parameters?

A1, A2 $\rightarrow$ Given by two plots. (ACF plot, PACF plot)

$$\text{ACF} \rightarrow \rho(h) = \operatorname{cor}(x_{n+h}, x_n) = \frac{E(x_{n+h}, x_n)}{\sqrt{E[x_{n+h}^2}}$$

$$\hat{\gamma}(h) = \frac{1}{n}\sum_{i=1}^{n-|h|}(x_{i+h} - \bar{x})(x_i - \bar{x}) \qquad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

σ^2 is estimated by Sample Residuals.

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$$

Q. Gaussian Process $\rightarrow$ Every finite dimensional sequence of process follows Multivariate Normal Distribution.

$$X_1 \sim N(0, \sigma^2) \qquad L(\sigma^2) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2}} e^{-\frac{1}{2} x_1 \sigma^{-2} x_1}$$

$$\vec{X}_n = [X_1, \ldots\ldots, X_n] \sim N(\vec{0}_{n\times 1}; \Gamma_n)$$

$$\Gamma_n = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(n-2) \\ \vdots & & & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{bmatrix}$$

$$x_n = [X_1, \ldots X_n]^T$$

$$L(\Gamma_n) = \frac{1}{(2\pi)^{n/2} |\Gamma_n|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2} X_n^T (\Gamma_n)^{-1} X_n}$$

Matrix Inversion (Γ_n^{-1}) can be avoided by the use of following identity

$$\vec{z}_n = C_n(\vec{x}_n - \hat{\vec{x}}_n)$$

$$\vec{x}_n = (x_1, \ldots\ldots, x_n)$$

$$C_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \theta_{11} & 1 & 0 & & 0 \\ \theta_{22} & \theta_{21} & 1 & & 0 \\ \vdots & & & \ddots & \vdots \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \cdots & & 1 \end{bmatrix}$$

All elements of C_n matrix are uncorrelated.

$$\vec{x}_1 = \vec{x}_1 - \hat{\vec{x}}_1$$

$$x_2 = \theta_{11}(x_1 - \hat{x}_1) + (x_2 - \hat{x}_2) \Rightarrow \hat{x}_2 = \theta_{11} x_1$$

$$x_3 = \theta_{21}(x_2 - \hat{x}_2) + \theta_{22}(x_1 - \hat{x}_1) + (x_3 - \hat{x}_3)$$

$$[x_n - \hat{x}_n] = \begin{bmatrix} x_1 - \hat{x}_1 \\ \vdots \\ \vdots \\ x_n - \hat{x}_n \end{bmatrix}$$

$$L(\phi, \theta, \sigma^2) = \frac{1}{\sqrt{(2\pi\sigma^2)^n r_0 r_1 \cdots r_{n-1}}} \cdot e^{\left\{-\frac{1}{2\sigma^2} \sum_{j=1}^{n} (x_j - \hat{x}_j)^2 / r_{j-1}\right\}}$$

$\cancel{\text{log}} \cdot L \propto$

$$-l = \frac{n}{2} \log \sigma^2 + \frac{1}{2\sigma^2} S$$

$$\Rightarrow \frac{\partial l}{\partial \sigma^2} = \frac{n}{2\sigma^2} - \frac{S}{2\sigma^4} = 0 \Rightarrow n = \frac{S}{\hat{\sigma}^2} \Rightarrow \boxed{\hat{\sigma}^2 = \frac{S}{n}.}$$

AIC → Akai Information Criteria. ~~[Better~~

AICC → Akai Information Criteria Corrected [Better Fit for Trend]

BIC → Bayesian Information Criteria. [For Robust Model]
[Better Fit for Test]

Lower the AIC, AICC and BIC values, the model will be better.