Time series analysis.

Time series is a collection of random variables $\{X_t \mid t \in T\}$ where T is an index set representing time.

Distribution of time series:

$X_0 = 0 \quad X_i \overset{iid}{\sim} N(0,1)$

$$z_i = \sum_{j=0}^{i} x_j$$

$$\text{Var}(Z_i) = \text{Var}\left(\sum_{j=0}^{i} x_j\right) = i$$

$\Rightarrow \text{Var}(Z_{60}) = 60$

and $\text{Var}(Z_{90}) = 90.$

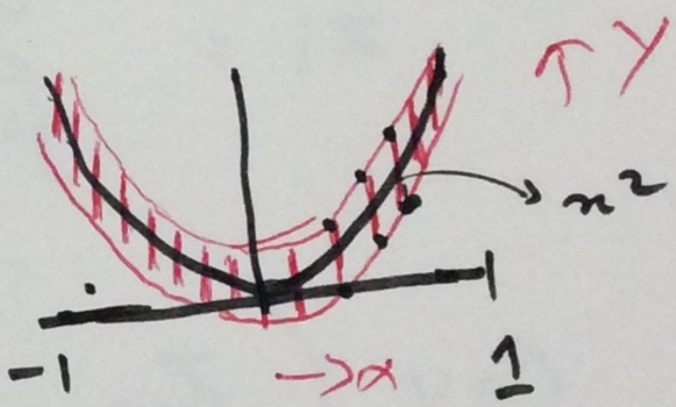
$$\begin{aligned}
&\text{Cov}(Z_{60}, Z_{90}) \\
&= \text{Cov}\left(Z_{60}, Z_{60} + \sum_{j=61}^{90} x_j\right) \\
&= \text{cov}(Z_{60}, Z_{60}) + \underbrace{\text{Cov}\left(Z_{60}, \sum_{j=61}^{90} x_j\right)}_{} \\
&= \text{Var}(Z_{60}) + 0 \\
&= \underline{60}
\end{aligned}$$

Note { Even if X_i's are uncorrelated then also this cov. is zero.

* The necessary condition for $\mathrm{cov}\left(Z_{60}, \sum_{j=61}^{90} X_j\right) = 0$ is that X_js are uncorrelated.

* Normally distributed random variables with correlation zero implies independence.

* Correlation or Covariance zero does not imply independence in general.



$X \sim U(-1, 1)$

$Y | X \sim U(x^2 - \epsilon, x^2 + \epsilon)$

$E(XY) = E(X^3) = 0$ $\quad E(X)E(Y) = 0$

Hence covariance $\underline{(X, Y) = 0}$.

$E(X_i) = 0 \quad V(X_i) = \underline{\underline{1}}$ and they are uncorrelated.

(iid) with finte 2nd order moment

$$Z_t = \frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} X_i \qquad t \in [0,1]$$

partial sum process.

what will be the distribution of Z_t when $n \uparrow \infty$??

$$Z_t = \frac{\sqrt{[nt]}}{\sqrt{n}} \left(\frac{1}{\sqrt{[nt]}} \sum_{i=1}^{[nt]} X_i \right)$$

when $n \uparrow \infty$

iid

$\{$ → indepenent but not idental
→ uncorrelated ✓
→ dependent – more restricting

$$\boxed{e^{\mu t + \frac{\sigma^2 t^2}{2}}}$$

$\downarrow$ $\sqrt{t}$

$\downarrow$ converge to $N(0,1)$ indistribut by CLT.

Z_t will converge in distribution to

$$\sqrt{t}\, N(0,1) \equiv N(0, \underline{\underline{t}}).$$

$Z_t \sim$ following Brownian motion.

$\{x_i\}$ series of numbers.

$$S_k = \sum_{i=1}^{k} x_i \qquad k \in \mathbb{N}.$$

→ partial sum.

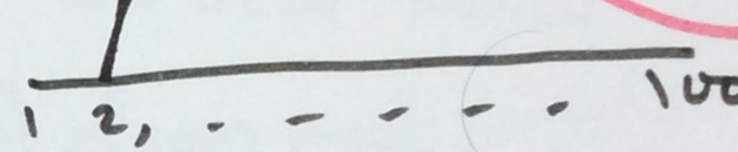
→ Cumulative sum.

$X_k \rightarrow k$

$S_k \rightarrow k.$

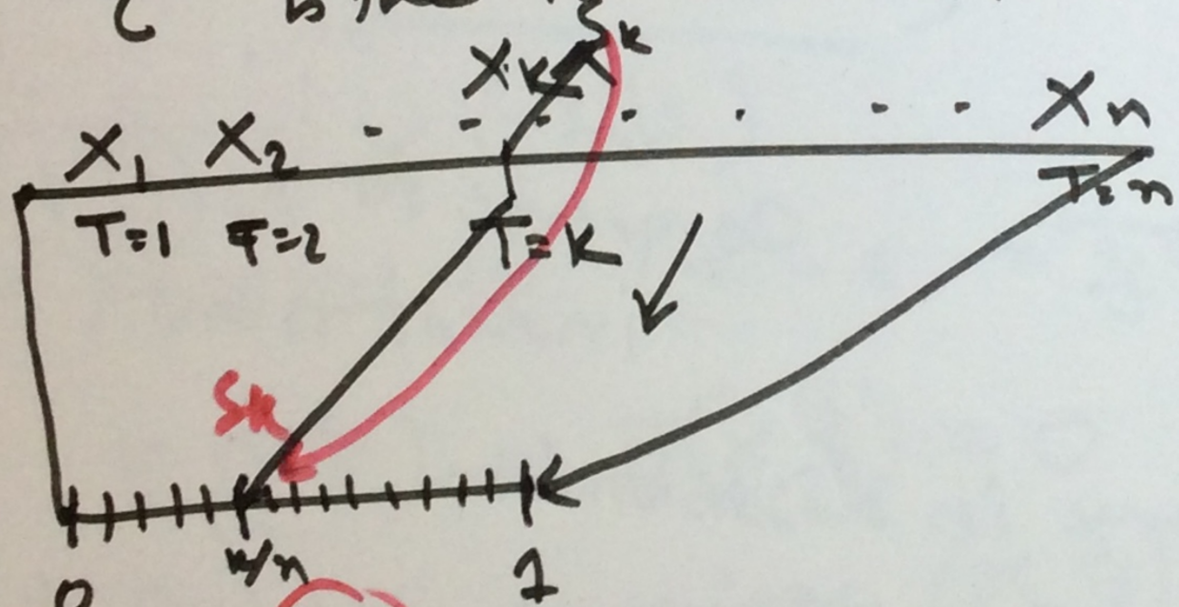
$X_k \rightarrow k/n$

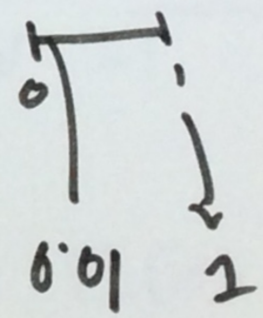
$S_k \rightarrow k/n$



n = sample size.

t is the time index.





$$t = \frac{k}{n} \longrightarrow X_k.$$

1. White noice:

A time series $\{W_t\}$ such that $E(W_t)=0$, $V(W_t)=\sigma_w^2$

$$W_t \sim WN(0, \sigma_w^2).$$

and they are uncorrelated.

Example: (1) $x_i \overset{iid}{\sim} N(0, \sigma^2)$

(2) $x_i \sim \begin{cases} N(0, \sigma^2) & \text{when } i \text{ is even} \\ \exp(1) - 1 & \text{when } i \text{ is odd} \end{cases}$ > independent.

→ (a) WN need not be normally distributed..

(b) WN need not be iid.

(c) iid sequence is always WN, with mean zero, and finite variance.

(d). WN is weakly stationary.

(e) WN with normally distributed r.v.s are strongly stationary.

(2) Binary process:

Consider a time series $\{X_t\}$ such that:

$$X_t = \begin{cases} +1 & \text{with prob. } p = 1/2 \\ -1 & \text{with prob. } 1-p = 1/2 \end{cases}$$

$E(X_t) = 0 \qquad V(X_t) = 1$

In general when $p \neq 1/2$

$E(X_t) = 2p - 1 \qquad V = 1 - (2p-1)^2$

$= \underline{4\,p(1-p)}$

$W \sim \text{bernoulli}(p).$

$X = \dfrac{\cancel{2}\,W - 0.5}{0.5}$

$X = a + bW$

$V(X) = b^2\, V(W)$

$= 4\, p(1-p)$

(3). Random walk:

Let $\{W_t\}$ be an iid sequen of random variables.

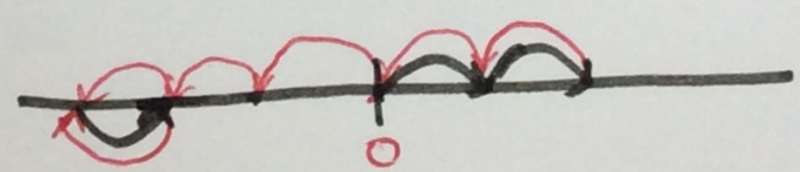
and define $X_0 = 0, \quad X_t = \sum_{i=1}^{t} W_i$ where.

$E(W_t) = 0 \quad Var(W_t) = \sigma_w^2.$

$\{W_t\} \sim$ binary process.

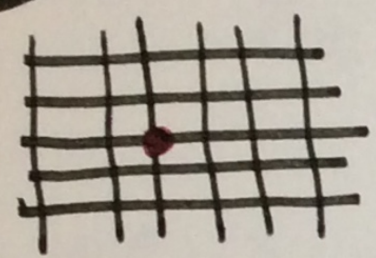
or $N(0, \sigma^2)$

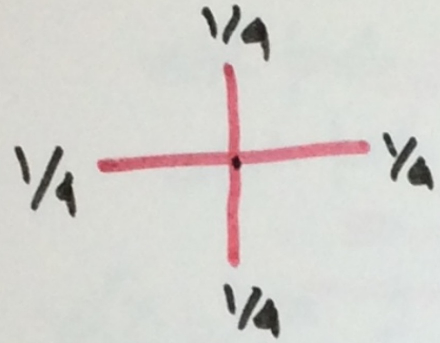
or. Laplace$(0, \lambda)$



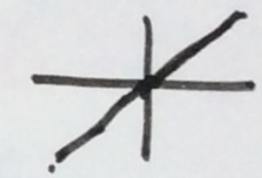
(a) $\mathbb{Z}$ $w_t \sim$ Binary process $(1/2)$. then it will eventually comeback to the place where it has started from.

(b) $\mathbb{Z} \times \mathbb{Z}$





The random varible eventually comeback to (0,0).



(c) $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

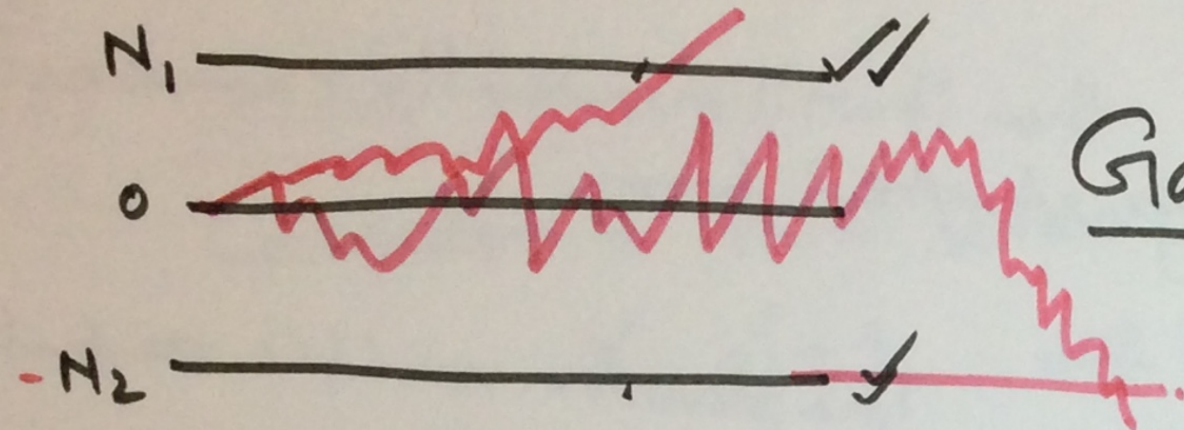
with equal probabity to each 6 ~~8~~ direction.
then the random variable may not come back to (0,0,0).

(4) Random walk with drift:

$$\begin{cases} W_i \overset{iid}{\sim} & E(W_i) = \delta \quad V(W) = \sigma^2 \\ X_t = \sum_{i=1}^{t} W_i \end{cases}$$

$$\begin{cases} X_t = \underline{t\delta} + \sum_{i=1}^{t} Z_i \\ Z_i \overset{iid}{\sim} & E(Z_i) = 0 \\ & V(Z_i) = \sigma_Z^2. \end{cases}$$

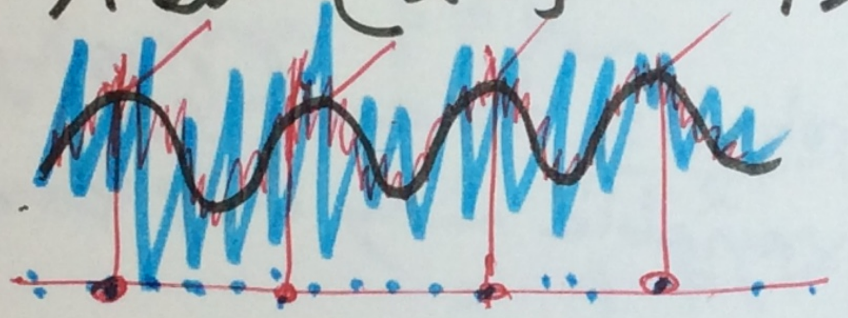
Mean of X_t is a function of t.



Gamblers Ruin problem.

5 Signal with noise.

$$X_t = A \sin(2\pi f t + \varphi) + W_t .$$



6 Moving average process: (order one).

$$W_t \sim WN(0, \sigma_w^2).$$

$$X_t = W_t + \underline{\theta} W_{t-1}.$$

$$X_t$$

$$\uparrow$$

$$(\theta_1 W_{t-1} + \theta_2 W_t + \theta_3 W_{t+1})$$

7 Auto regressive process. : (Order one).

$\{W_t\} \sim WN(0, \sigma_w^2)$.

$$X_t = \phi X_{t-1} + W_t. \quad |\phi| < 1, \; \phi \neq 0$$

$$Y = \beta X + \epsilon.$$

8 Wiener process / Brownian motion.

$\{X_t\}$ is said to follow wiener process ~~of~~ or BM. if.
($a=0 \Rightarrow$ standard BM).

(i) $X_0 = a$

(ii) $X_{t+u} - X_t$ is independently distributed to $X_s \; \forall s \leq t$. $u > 0$.

(iii) $X_{t+u} - X_t \sim N(0, u)$., $u > 0$

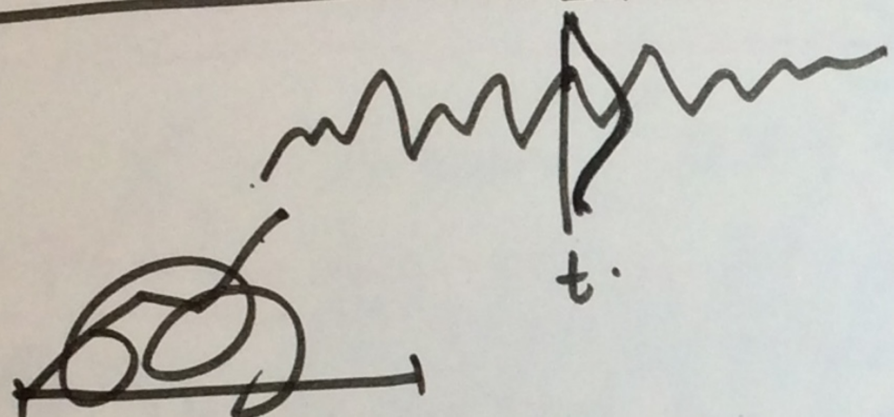
(iv) $\{X_t\}$ has continuous path on $t \in$ Time.
(but not differentiable)

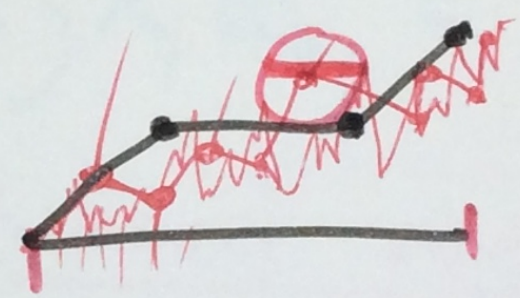
$$\begin{cases} W_k \overset{iid}{\sim} \quad E(W_k)=0 \quad V(W_k)=1. \\ \cancel{X_t} \quad Y_n(t) = \underbrace{\dfrac{1}{\sqrt{n}}\displaystyle\sum_{i=1}^{[nt]} W_i}_{\to N(0,t)} \longrightarrow X_t \sim BM. \end{cases}$$

For large n.

$t \in [0,1]$

Continuous time process.





3 Brownian Bridge.

Brownian motion with restriction.

$(t=0,\ X_t=0)$ and $(t=T,\ X_t=b)$.

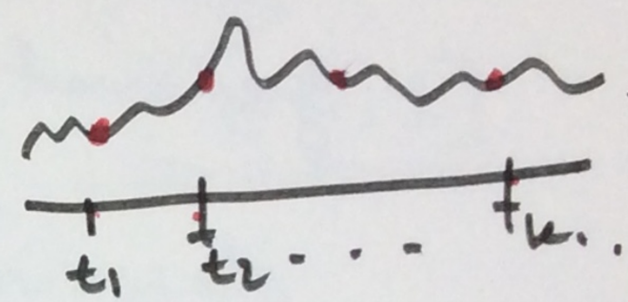
If $(T=1$ and $b=0) \Rightarrow$ standard Brownian bridge.

$$B_0(t) = \underset{N(0,t)}{X_t} - t\underset{N(0,1)}{X_1} \qquad X_t \sim BM \text{ on } [0,1].$$

$X_t \sim$ Standard BN.

$t < s$

$$\begin{pmatrix} X_t \\ X_s \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} t & t \\ t & s \end{pmatrix}\right)$$



$B_0(t) \sim$ Standard BB.

$$\begin{pmatrix} B_0(t) \\ B_0(s) \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} t(1-t) & t(1-s) \\ t(1-s) & s(1-s) \end{pmatrix}\right)$$

Any $k \in \mathbb{N}$ and $\underline{t_1, t_2, t_3, \dots t_k} \in \mathbb{R}$.

As $(X_{t_1}, X_{t_2}, \dots, X_{t_k})$ or $(B_0(t_1), B_0(t_2), \dots B_0(t_k))$

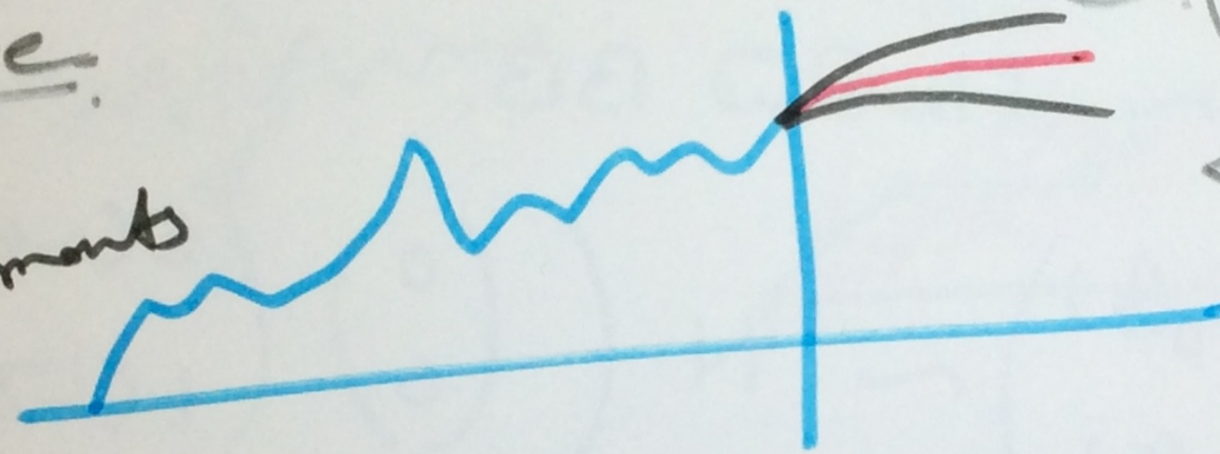
follows multivariate Normal.

X_t and $B_0(t)$ are known a <u>Gaussian process.</u>

10

$$X_t = (\delta_0 + \delta_1 t + \delta_2 t^2) + A \sin(2\pi f t) + W_t.$$

general features of a time series we might observe.

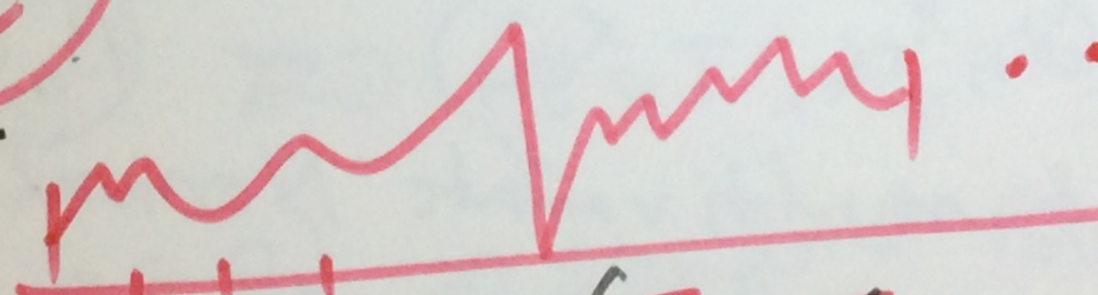


Even though all moments are same two random variables may have diffrt distn.

Weak Second order moments are same.

For all $k \in \mathbb{N}$. $t_1 t_2 \dots t_k \in \mathbb{R}$, $h \in \mathbb{R}$.

W_t.



Strong.

Joint distribution of $(Z_{t_1}, Z_{t_2} \dots Z_{t_k})$ is same as $(Z_{t_1+h}, Z_{t_2+h} \dots Z_{t_k+h})$.

Stationary timeseries: (Strong stationarity)

Let $\{X_t\}$ be a timeseries with joint distribution function of $(X_1, X_2, \dots, X_n)$ as $P(X_1 \le a_1, X_2 \le a_2, \dots, X_n \le a_n) = F_n(a_1, a_2, \dots, a_n)$

$$P(X_{k+1} \le a_1, X_{k+2} \le a_2 \dots X_{k+n} \le a_n) \qquad \forall n \in \mathbb{N}$$

$$= P(X_{k+h+1} \le a_1, X_{k+h+2} \le a_2 \dots X_{k+h+n} \le a_n) \qquad \forall k \in \mathbb{Z}, \ \forall h \in \mathbb{Z}$$

$$= F_n(a_1, a_2, \dots a_n) \qquad \forall a_i \in \mathbb{R} \quad i = 1, 2 \dots n.$$

Then $\{X_t\}$ is said to strongly stationary process / timeseries.

Weakly stationary timeseries

(I) $\mu = \mu_t = E(X_t)$ is free from t.

(II) $cov(X_t, X_{t+h}) = E(X_t - \mu_t)(X_{t+h} - \mu_{t+h})$

$= E[(X_t - \mu)(X_{t+h} - \mu)]$ is also free from t but may be a function of 'h'.

If (I) and (II) holds then $\{X_t\}$ is said to be weakly stationary.

⊛ $\mu_t = E(X_t)$ is a constant function of 't'.

⊛ $\gamma_X(h) = \mathrm{Cov}(X_t, X_{t+h})$ free from t. but dependent on lag 'h'.

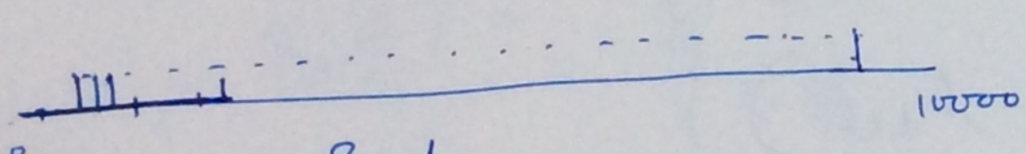
⊛ $\gamma_X(h) = \gamma_X(-h)$ because 'cov' is a symmetric function.

Auto correlation coefficient.

Auto correlation coefficient of an 'atleast' weakly stationary time series is defined as

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{\mathrm{Cov}(X_t, X_{t+h})}{\sqrt{V(X_t)\,V(X_{t+h})}}$$

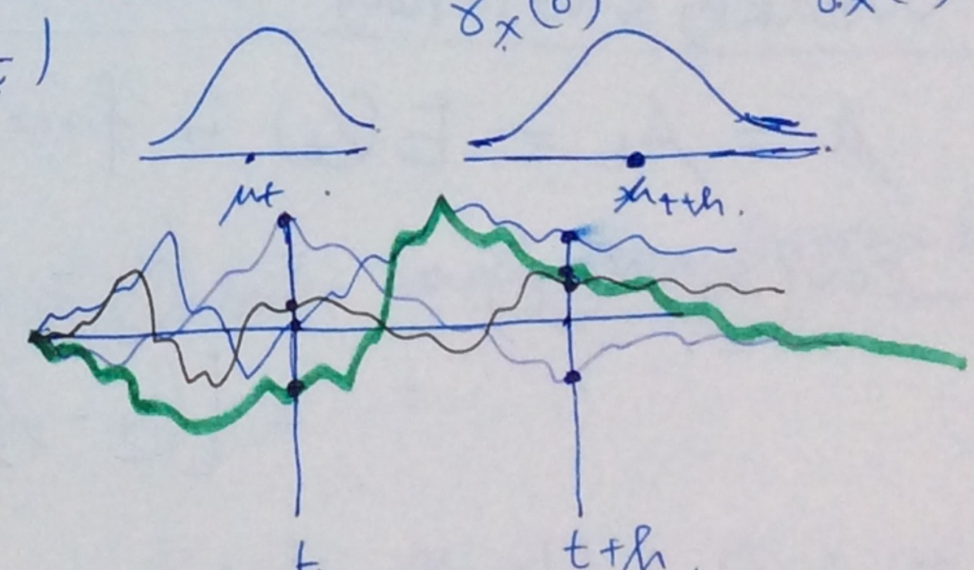
$$= \frac{\mathrm{cov}(X_t, X_{t+h})}{V(X_t)} = \frac{\gamma_X(h)}{\gamma_X(0)} = \frac{\gamma_X(-h)}{\gamma_X(0)}$$



$h = 1.$

$$x_1 \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \begin{pmatrix} x_3 \\ x_2 \end{pmatrix} \begin{pmatrix} x_4 \\ x_3 \end{pmatrix} \cdots \begin{pmatrix} x_n \\ x_{n-1} \end{pmatrix} x_n.$$

$$\hat{\gamma}_X(1)$$



* $\{X_t\} \overset{iid}{\sim} N(0, \sigma^2)$. strongly stationary $\Rightarrow$ weak stationary.

* $\{X_t\} \overset{iid}{\sim}$ Cauchy $(0, 1)$ strongly stationary but not with finite moment.

* Strongly stationary time series with finite second order moment will be also weakly stationary.

* Normally distributed weakly stationary time series ~~ost~~ will be strongly stationary if $\gamma_X(h) =$ constant for all $h \neq 0$.

$E(X_t) = \mu$ given, $V(X_t) = \sigma^2$ given.,

$$\begin{pmatrix} X_1 \\ X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{pmatrix} \cancel{\rho} & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ & & \ddots & \\ \rho & \rho & \cdots & 1 \end{pmatrix}\sigma^2\right)$$

if iid. $\Rightarrow$ strong stationarity.

<u>Ex 1.</u> $\{X_t\} \overset{iid}{\sim}$ $E(X_t) = 0$ Var. $(X_t) = \sigma^2 < \infty$.

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

Even if $\{X_t\}$ are WN. then also the above moment condition <u>holds</u>.

Ex. 2 $\quad S_t = \sum_{i=1}^{t} w_i \qquad w_i \overset{iid}{\sim} N(0, \sigma^2)$.

$$\left.\begin{array}{l} E(S_t) = 0 \\ V(S_t) = t\sigma^2 \end{array}\right\}$$

NOT strongly stationary.
NOT even weakly stationary.

$\text{Cov}(S_t, S_{t+h}) \qquad \underline{\underline{h \geq 0}}$.

$= \text{Cov}\left(S_t, S_t + \sum_{t+1}^{t+h} w_i\right) = \text{Cov}(S_t, S_t) + \text{Cov}\left(S_t, \sum_{i=t+1}^{t+h} w_i\right)$

$= V(S_t) + 0 = \sigma^2 t$ function of t.

HW. $\quad Z_t = \frac{1}{\sqrt{t}} S_t = \frac{1}{\sqrt{t}} \sum_{i=1}^{t} w_i$

$E(Z_t) =$

$V(Z_t) =$

$\gamma_Z(h) = \text{Cov}(Z_t, Z_{t+h}) =$

weak stationary?
strong stationary?

Ex 3. Moving average process of order one. (MA(1)).

$$Z_t = W_t + \theta W_{t-1} \qquad W_t \overset{iid}{\sim} \underline{N(0, \sigma^2)}$$

(need not be normal always. WN will be enough.)

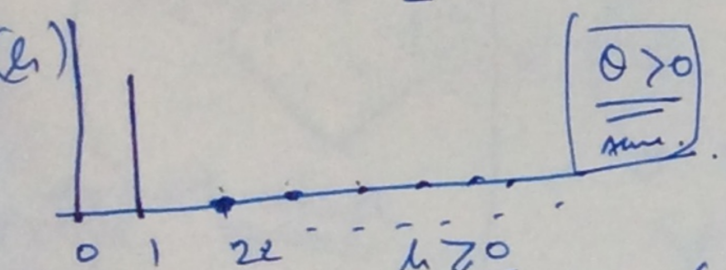
$$E(Z_t) = 0 \qquad \forall\, t \in \mathbb{Z} \qquad h \in \mathbb{Z}$$

$$V(Z_t) = (1+\theta^2)\sigma^2 . \qquad \forall\, t \in \mathbb{Z}$$

Z_t are normally distributed in this example.

MA(1)

$$\gamma_Z(h) = \begin{cases} (1+\theta^2)\sigma^2 & h = 0 \\ \theta\sigma^2 & h = \pm 1 \\ 0 & |h| > 1 \end{cases}$$



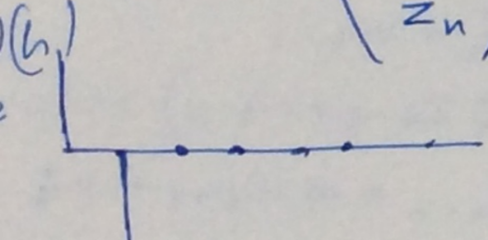
$$Z_t = W_t + \theta W_{t-1}$$
$$Z_{t+1} = W_{t+1} + \theta W_t$$
$$Z_{t+2} = W_{t+2} + \theta W_{t+1}$$

Remark 1 : $\{Z_t\}$ is weakly stationary.

Remark 2 : $\{Z_t\}$ is strongly stationary only because $W_t \overset{iid}{\sim} N(0, \sigma^2)$
[In general it may not be true.]

MA(1)

$\theta < 0$



$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{bmatrix} (1+\theta^2)\sigma^2 & \theta\sigma^2 & & 0 \\ \theta\sigma^2 & \ddots & \ddots & \\ & \ddots & \ddots & \theta\sigma^2 \\ 0 & & \theta\sigma^2 & (1+\theta^2)\sigma^2 \end{bmatrix} \right) \qquad \begin{pmatrix} Z_{1+h} \\ Z_{2+h} \\ \vdots \\ Z_{n+h} \end{pmatrix} \sim$$

Ex 4. Autoregressive process of order one (AR(1)).

$$Z_t = \varphi Z_{t-1} + W_t. \quad \left\{ \begin{array}{l} W_t \overset{iid}{\sim} N(0, \sigma^2) \qquad t \in \mathbb{Z} \\ |\varphi| < 1, \quad \varphi \neq 0. \\ \text{given that } Z_t \text{ is weakly stationary.} \end{array} \right.$$

Find the value of $\gamma_Z(h)$, $\rho_Z(h)$.

$Y = f(x) + \epsilon$.

$Z_t = \varphi(Z_{t-1}) + W_t$.

Auto Regressive process.

$$\left.\begin{array}{l} E(Z_t) = \varphi E(Z_{t-1}) + E(W_t) \\ E(Z_t) = \varphi E(Z_{t-1}) + 0 \end{array}\right\} \Rightarrow E(Z_t) = 0.$$

$$\Rightarrow E(Z_t) = 0 \quad \forall t.$$

$$V(Z_t) = V(\varphi Z_{t-1} + W_t).$$

$$\Rightarrow E(Z_t^2) = E\left[(\varphi Z_{t-1} + W_t)^2\right]$$

$$\Rightarrow E(Z_t^2) = E(\varphi^2 Z_{t-1}^2 + W_t^2 + 2\varphi Z_{t-1} W_t)$$

$W_t \overset{iid}{\sim} N(0, \sigma^2)$

Z_{t-1} and W_t are independent because Z_{t-1} does not involve W_t.

$$\Rightarrow E(Z_t^2) = \varphi^2 E(Z_{t-1}^2) + \sigma^2 + 0$$

use weak stationarity

$$\Rightarrow E(Z_t^2) = V(Z_t) = \frac{\sigma^2}{1-\varphi^2} > \sigma^2 = Var(W_t).$$

$$\begin{aligned} Z_{t-1} &= \varphi Z_{t-2} + W_{t-1} \\ &= \varphi(\varphi Z_{t-3} + W_{t-2}) + W_{t-1} \\ &= \varphi^2 Z_{t-3} + \varphi W_{t-2} + W_{t-1} \end{aligned}$$

$$\gamma_Z(h) = \operatorname{cov}(Z_{t+h}, Z_t) = \operatorname{cov}(Z_{t-h}, Z_t)$$

$h > 0$ AR(1)

$$= \operatorname{cov}(\varphi Z_{t+h-1} + W_{t+h}, Z_t)$$

$$= \varphi \operatorname{cov}(Z_{t+h-1}, Z_t)$$

$$= \varphi^2 \operatorname{cov}(Z_{t+h-2}, Z_t)$$

$$\vdots$$

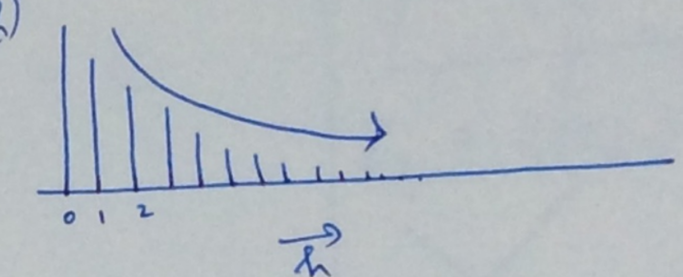
$$= \varphi^h \operatorname{cov}(Z_t, Z_t)$$

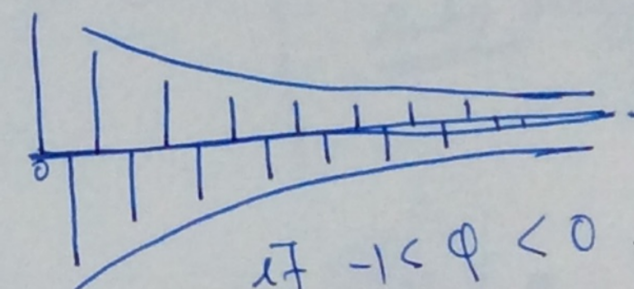
$$= \varphi^h \gamma_Z(0)$$

$$= \varphi^h \operatorname{Var}(Z_t)$$

$$= \varphi^h \left(\frac{\sigma^2}{1-\varphi^2}\right) = \varphi^{|h|}\left(\frac{\sigma^2}{1-\varphi^2}\right)$$

if $0 < \varphi < 1$





if $-1 < \varphi < 0$

HW

$$\rho_Z(h) = \varphi^{|h|}$$

Auto correlation function is a useful tool to identify AR and MA process.