Basic Properties of a Time Series (I)

A time series is simply a discrete time stochastic process $\{X_t, t \in \mathbb{Z}\}$. ($\mathbb{Z}$ is the set of integers $0, \pm 1, \pm 2, \ldots$)

We will assume that $\{X_t, t \in \mathbb{Z}\}$ is a <u>second-order</u> process, that is,

$$\mathbb{E}[X_t^2] < \infty$$

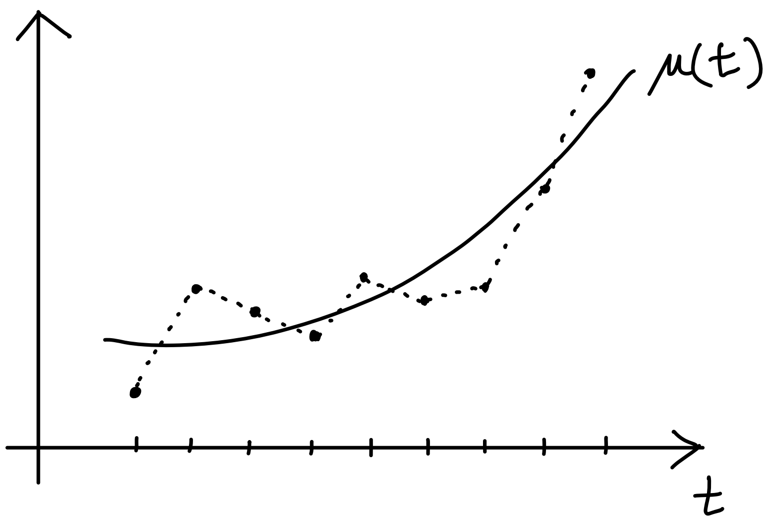
for all $t \in \mathbb{Z}$.

II

Mean Function

The mean function $\mu(t)$, $t \in \mathbb{Z}$ of a time series is defined by

$$\mu(t) = \mathbb{E}[X_t].$$



II

Covariance Function

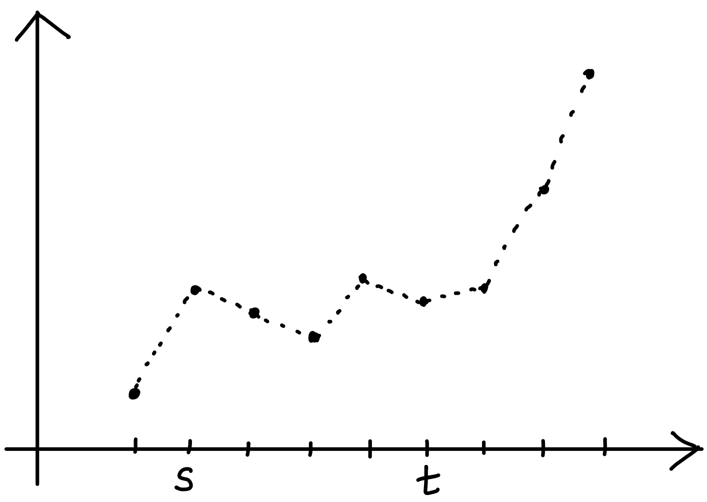
The (auto) covariance function of a time series $\{X_t, t \in \mathbb{Z}\}$ is

$$\gamma(s,t) = \mathrm{Cov}(X_s, X_t)$$

$$:= \mathbb{E}\left[(X_s - \mu(s))(X_t - \mu(t))\right]$$

$$= \mathbb{E}[X_s X_t] - \mu(s)\mu(t).$$

$s,t \in \mathbb{Z}$

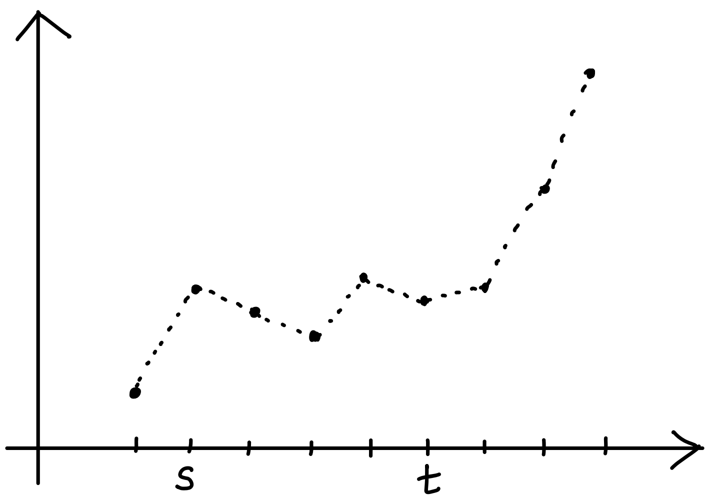


II

Covariance Function

Notice:

$$\mathrm{Var}(X_s) = \mathrm{Cov}(X_s, X_s)$$
$$=: \gamma(s, s)$$
$$=: \gamma(s), \qquad s \in \mathbb{Z}$$



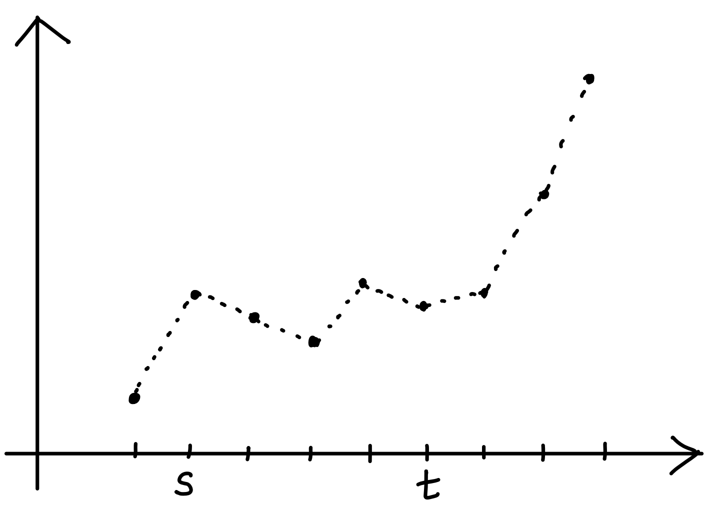
(II)

Correlation Function

The (auto) correlation function of a time series $\{X_t, t \in \mathbb{Z}\}$ is

$$\rho(s,t) = \operatorname{Corr}(X_s, X_t)$$

$$:= \gamma(s,t) \Big/ \sqrt{\gamma(s)\,\gamma(t)} \qquad s,t \in \mathbb{Z}$$



II

Example 1 (Coin Flips)

Suppose $Y_t, t \in \mathbb{Z}$ are independent and identically distributed with

$$Y_t = \begin{cases} -1 & \text{w.p. } 1-p \\ 1 & \text{w.p. } p. \end{cases}$$

(i) What is the mean function of $\{Y_t, t \in \mathbb{Z}\}$?

(ii) What is the correlation function of $\{Y_t, t \in \mathbb{Z}\}$?

II

Example 2 (Random Walk)

Let $Y_t,\ t \in \mathbb{Z}^+$ be as in the previous example. ($\mathbb{Z}^+ := 0, 1, 2, \dots$)

Define

$$X_t = X_{t-1} + Y_t, \quad t \in \{1, 2, \dots\}$$

$$X_0 = 0.$$

(i) What is the mean function of $\{X_t, t \in \mathbb{Z}\}$?

(ii) What is the correlation function of $\{X_t, t \in \mathbb{Z}\}$?

II

The mean function:

$$\underbrace{\mu(t)}_{\mathbb{E}[X_t]} = 0 \qquad t \in \mathbb{Z}^+.$$

Proof?

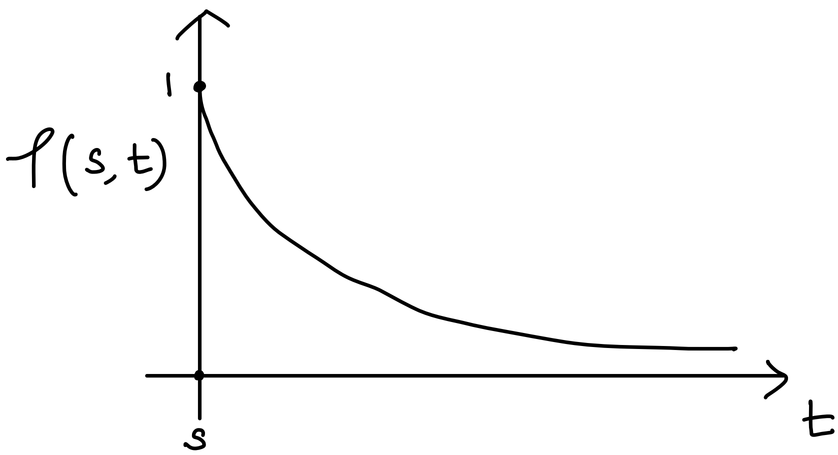
The autocorrelation function:

$$\rho(s,t) = \sqrt{\frac{\min(s,t)}{\max(s,t)}}, \quad s,t \in \{1,2,\dots\}$$

Proof?

II

For the (discrete) random walk, the (auto) correlation function looks like:



II

Generating a Random Walk

```
# initialize seed
  set.seed(420)

# generate coin flips
  Y ← rbinom(100, 1, 0.4)
# convert to 1s and -1s
  Y ← 2*Y - 1

# initialize the walk
  X ← rep(0, 100)
# create the walk
  for i in 2:100
  {  X[t] ← X[t-1] + Y[t]
  }
```

II

Example 3 (Moving Average)

Suppose the random variables in the discrete time stochastic process $\{X_t, t \in \mathbb{Z}^+\}$ are related as follows:

$$\boxed{\begin{array}{l} X_t = c_0 + \frac{1}{2}\left(\varepsilon_t + \varepsilon_{t-1}\right), \quad t = 1, 2, \ldots \\ X_0 = 0. \\ \varepsilon_t \text{ are iid}; \ \mathbb{E}[\varepsilon_t] = 0; \ \mathrm{Var}(\varepsilon_t) = \sigma_\varepsilon^2 \end{array}}$$

(The above "relationships" are sometimes collectively called a "model"; c_0 and σ_ε^2 are called model parameters.)

II

Let's calculate the mean and auto covariance and auto correlation functions.

$$\mu(t) = c_0, \quad t = 2, 3, \ldots$$

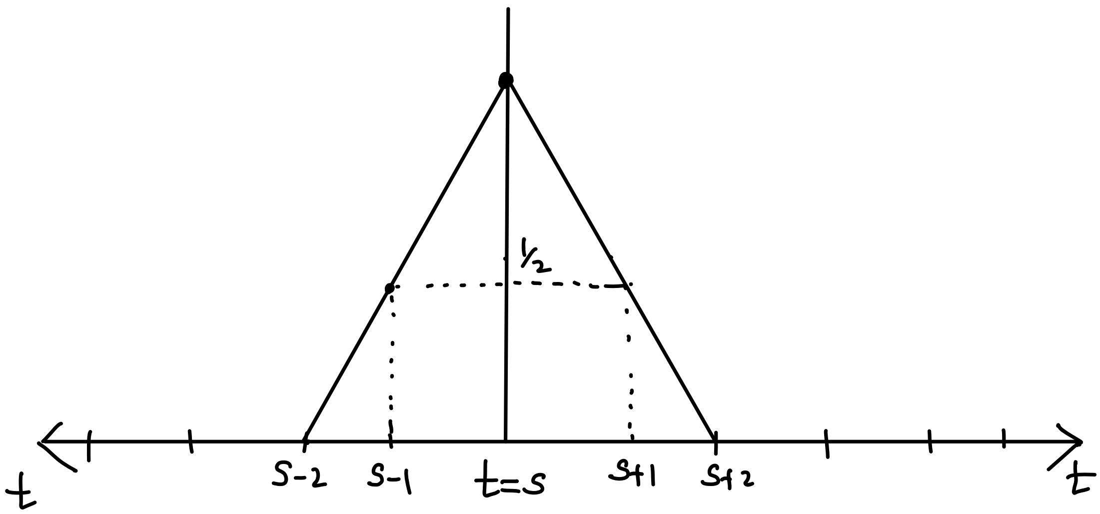
$$\gamma(s,t) = \begin{cases} \sigma_\varepsilon^2/2 & |s-t| = 0 \\ \sigma_\varepsilon^2/4 & |s-t| = 1 \\ 0 & |s-t| > 1 \end{cases}$$

$$\rho(s,t) = \begin{cases} 1 & |s-t| = 0 \\ 1/2 & |s-t| = 1 \\ 0 & |s-t| > 1 \end{cases}$$

Proof?

II

How does the autocorrelation function look?



More Intuition: Can we simulate (II) data from this process?

```
# generate the ε terms
  ε ← rnorm(100, 0, 2)

# initialize the series
  X ← rep(0, 100)
# choose constant
  c0 ← 1
# create the series
  for t in 2:100
{ X[t] ← c0 + 1/2 × (ε[t] + ε[t-1])
}
```

II

Let's make sure we understand some properties of the covariance function.

(i) $\gamma(s, t) \overset{?}{=} \gamma(t, s)$

(ii) $\gamma(s, t) \overset{?}{=} \gamma(0, t-s)$

(iii) $\gamma(s, t) \overset{?}{=} \gamma(s+\tau, t+\tau)$

II

Second Order Stationary or Covariance Stationary Processes

A discrete-time stochastic process $\{X_t, t \in \mathbb{Z}\}$ is said to be covariance stationary if $\{X_{t+\tau}, t \in \mathbb{Z}\}$ and $\{X_t, t \in \mathbb{Z}\}$ have the same mean and covariance functions for all $\tau \in \mathbb{Z}$.

II

In other words, covariance stationary means:

I $\mu(t+\tau) = \mu(t) \quad \forall \tau \in \mathbb{Z}$

II $\gamma(s+\tau, t+\tau) = \gamma(s,t) \quad \forall \tau \in \mathbb{Z}$
(any fixed s,t)

The first condition is easy to interpret.

The second condition can be "tricky."

Best to interpret as "covariance between random variables is dependent only on the time lag $|t-s|$, but not on t, s."

II

For the "moving average" example recall the mean and the covariance functions that we calculated.

$$\mu(t) = c_0 , \quad t = 2, 3, \ldots$$

$$\gamma(s,t) = \begin{cases} \sigma_\varepsilon^2/2 & |s-t| = 0 \\ \sigma_\varepsilon^2/4 & |s-t| = 1 \\ 0 & |s-t| > 1 \end{cases}$$

$$\rho(s,t) = \begin{cases} 1 & |s-t| = 0 \\ 1/2 & |s-t| = 1 \\ 0 & |s-t| > 1 \end{cases}$$

II

In general if a process $\{X_t, t \in \mathbb{Z}\}$ is <u>covariance stationary</u>, then

(i) $\mu(t) \equiv \mu$

(ii) $\gamma(s,t) \equiv \gamma(0, t-s)$
$\equiv \gamma(h)$,

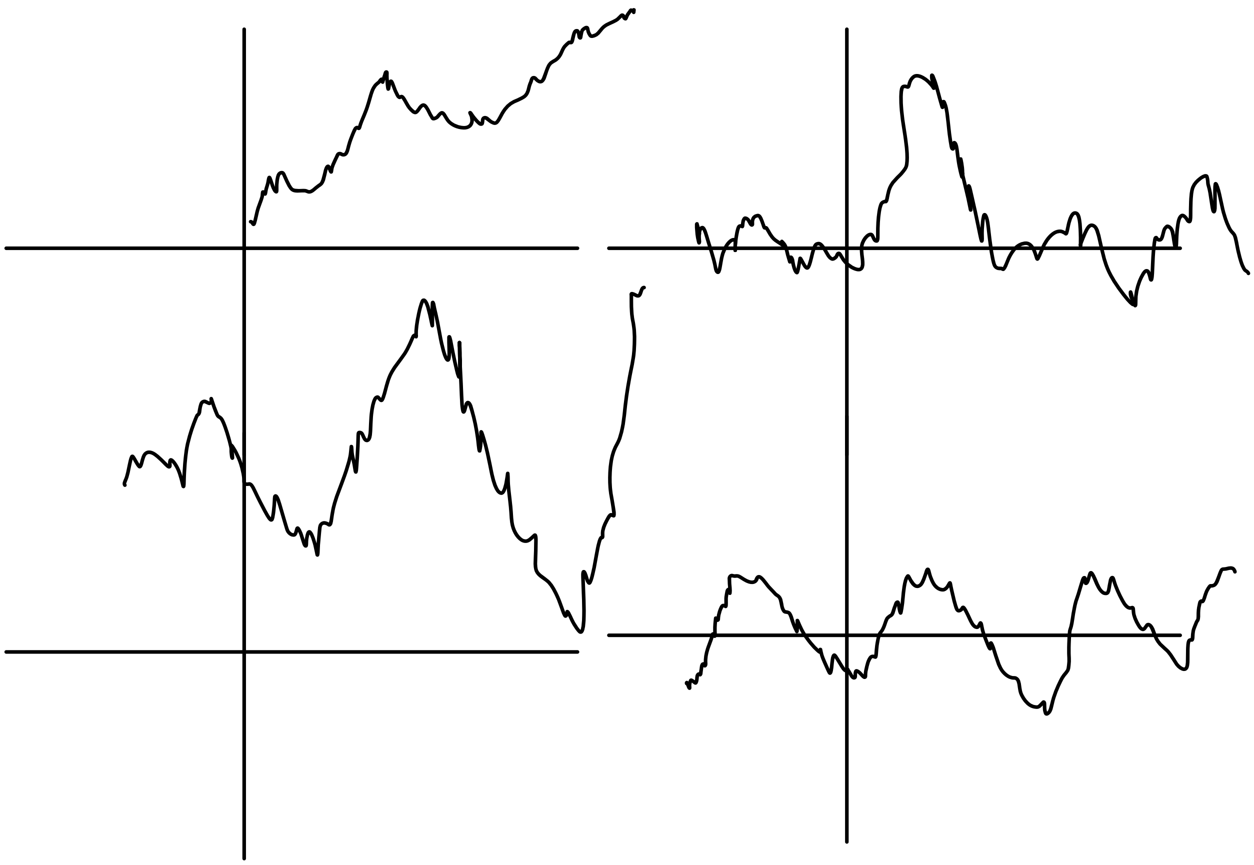
where h is called the "lag."

(iii) $\rho(s,t) \equiv \rho(h)$.

(Since $\gamma(h) = \gamma(-h)$, it is sufficient to consider $h \geq 0$.)

Ⅱ

How do time series generated from a covariance stationary process look?



II

Which of the following series are covariance stationary?

I. $X_t = 0.2 X_{t-1} + \varepsilon_t$

II. $X_t = 0.1t + 0.2 X_{t-1} + \varepsilon_t$

III. $X_t = X_{t-1} + \varepsilon_t$

IV. $X_t = 2\sin\left(2\pi\left(\frac{t}{6} + Y\right)\right)$

$Y \sim \text{Unif}(0,1)$.

(Assume $\varepsilon_t \overset{iid}{\sim} N(0,1)$, $t \in \mathbb{Z}^+$ and $X_0 = 0$.)

II

If the time series $\{X_t, t \in \mathbb{Z}\}$ is covariance stationary, then which of the following are true?

(i) $\mu(0) \overset{?}{=} \mu(t)$

(ii) $\rho(s, t) \overset{?}{=} \rho(-s, -t)$

(iii) $\rho(s, t) \overset{?}{=} \rho(-s, t)$

(iv) $\rho(0, h) \overset{?}{=} \rho(-h, 0)$

II

Example

Consider the time series

$$X_t = 2\cos\left(2\pi\left(\frac{t}{12}+U\right)\right),$$

$$t \in \mathbb{Z}$$

(i) Find the mean and autocorrelation functions

(ii) Is $\{X_t, t \in \mathbb{Z}\}$ cov. stationary?

II

A process $\{X_t, t \in \mathbb{Z}\}$ is said to be strictly stationary if $(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ and $(X_{t_1+\tau}, X_{t_2+\tau}, \ldots, X_{t_n+\tau})$ have the same joint distribution, for all $t_1, t_2, \ldots, t_n$.

Mind the Difference.

II

Covariance Stationary

$\{X_t, t \in \mathbb{Z}\}$ and $\{X_{t+\tau}, t \in \mathbb{Z}\}$

have the same mean and covariance functions. $\forall \tau \in \mathbb{Z}$.

Strictly Stationary

$(X_{t_1}, X_{t_2}, \ldots, X_{t_n})$ & $(X_{t_1+\tau}, X_{t_2+\tau}, \ldots, X_{t_n+\tau})$

have the same joint distribution.

II

Differencing and Backshift

"Differencing" is an operation on a time series, often in an attempt to stationarize it.

Differencing is denoted by ∇X_t:

$$\nabla X_t = X_t - X_{t-1}$$

$$\begin{aligned}\nabla^2 X_t &= \nabla(\nabla X_t) \\ &= \nabla(X_t - X_{t-1}) \\ &= X_t - 2X_{t-1} + X_{t-2}\end{aligned}$$

$\vdots$

II

II

Backshift is denoted by

BX_t :

$$BX_t = X_{t-1}$$

So,

$$\nabla X_t = X_t - X_{t-1} = (1-B)X_t$$

$$(1-B)^2 X_t = (1-2B+B^2)X_t$$
$$= \nabla^2 X_t .$$

II

Differencing m times is "multiplying" by $(1-B)^m$:

$$\nabla^m X_t = (1-B)^m X_t .$$

II

EXAMPLE

Find the mean and autocovariance functions of $Y_t := \nabla \varepsilon_t$, $t \in \mathbb{Z}$ where $\{\varepsilon_t, t \in \mathbb{Z}\}$ is white noise.

EXAMPLE

II

Suppose

$$Y_t = \beta_0 + \beta_1 t + X_t, \quad t \in \mathbb{Z}$$

where $\{X_t, t \in \mathbb{Z}\}$ is mean-zero and cov. stat. with lag-k autocov $\gamma(k)$.

1. Show that $\{Y_t, t \in \mathbb{Z}\}$ is <u>not cov. stat</u>.
2. Show that ∇Y_t is <u>cov. stat</u>

II

In the previous problem, suppose

$$Y_t = f(t) + X_t .$$

and $f(t) = a_0 t^m + a_1 t^{m-1} + \cdots + a_m$.

Then, $\nabla^n Y_t$ is <u>cov. stat</u>

if $n \geq m$ and <u>not cov. stat</u>

if $n < m$.

Sample Estimators

II

If $\{X_t, t \in \mathbb{Z}\}$ is covariance stationary, then

(i) $\mu(t) = \mu \quad \forall t \in \mathbb{Z}$

(ii) $\gamma(s,t) = \gamma(0, |s-t|)$

$\equiv \gamma(\underbrace{h}_{\text{"lag"}}), \; h = 0, 1, \ldots$

(iii) $\rho(s,t) = \rho(0, |s-t|)$

$\equiv \rho(h), \; h = 0, 1, \ldots$

Can μ, $\gamma(\cdot)$, & $\rho(\cdot)$ be estimated?

II

Suppose $\{X_t,\ t = 1, 2, \dots\}$ is covariance stationary and $X_1, X_2, \dots, X_n$ are the first n observations.

We want to estimate μ, $\gamma(h)$, $\rho(h)$:

$$\mu = \mathbb{E}[X_t]$$

$$\gamma(h) = \mathrm{Cov}(X_t, X_{t+h})$$

$$\rho(h) = \gamma(h)/\gamma(0)$$

II

$$\hat{\mu}(t) = \hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} X_j$$

$$\hat{\gamma}(h) = \frac{1}{n-h}\sum_{j=1}^{n-h}(X_j-\hat{\mu})(X_{j+h}-\hat{\mu})$$

$$h = 0, 1, 2, \ldots, n-1.$$

$$\hat{\rho}(h) = \hat{\gamma}(h) \big/ \hat{\gamma}(0), \quad h = 0, 1, \ldots, n-1.$$

Beware of Dependence During Estimation

II

Let's take an example.

Suppose:

$$Y_t = \mu + X_t, \quad t \in \mathbb{Z}$$

where $\{X_t, t \in \mathbb{Z}\}$ is a covariance stationary time series with autocovariance function $\gamma(h), h \geq 0$, and zero mean.

Beware of Dependence During Estimation

II

We do not know μ and $\gamma(h)$, $h \geq 0$ and so we want to estimate these from "data".

Suppose we have observed data:

$$Y_1, Y_2, \ldots, Y_n$$

Beware of Dependence During Estimation

II

Notice that the mean function

$$\mu(t) = \mathbb{E}[Y_t] = \mu + 0$$

$$\Rightarrow \quad \mu(t) = \mu \quad \forall\, t \in \mathbb{Z}.$$

Hence,

$$\hat{\mu} = \frac{1}{n} \sum_{j=1}^{n} Y_j$$

Beware of Dependence During Estimation

II

Let's see how sure we can be about $\hat{\mu}$.

$$\mathrm{Var}(\hat{\mu}) = \mathrm{Var}\left(\frac{1}{n}\sum_{j=1}^{n} Y_j\right)$$

$$= \frac{1}{n^2}\mathrm{Var}\left(\sum_{j=1}^{n} Y_j\right)$$

Beware of Dependence During Estimation

II

$$Var(\hat{\mu}) = \frac{\gamma(0)}{n}\left(1 + 2\sum_{h=1}^{n-1}\left(1 - \frac{h}{n}\right)\rho(h)\right)$$

(Make sure you understand how I got the above expression.)

Beware of Dependence During Estimation

II

When n is large, and

$$\sum_{h=1}^{\infty} |\rho(h)| < \infty,$$

$$\text{Var}(\hat{\mu}) \approx \frac{\gamma(0)}{n} \sum_{h=-\infty}^{\infty} \rho(h).$$

e.g., if $\rho(h) = \phi^h$, $h \geq 0$,

then

$$\text{Var}(\hat{\mu}) \approx \frac{\gamma(0)}{n} \left(\frac{1+\phi}{1-\phi}\right).$$