# Financial Time Series

Lecture 3

Basic stochastic and time series models

# Basic models

- Models that give a rule for current or future observations based on past observations.
- We are concerned with a sequence of random variables that may be dependent.
- Our goal is to learn a set of possible models and make sensible model choices.
- Let’s start with very simple models and discuss their statistical properties.

# IID noise

- Simplest time series model
  - No trend
  - No seasonal variations
  - Independent observations from the same distribution (iid)
- Distributionally, ”iid-ness” implies

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2)\dots f(x_n)$$

- Limitation: Cannot be used for forecasting
- An i.i.d. mean zero Gaussian sequence is called Gaussian white noise: $w_1, w_2, \dots, w_t$.

# IID example

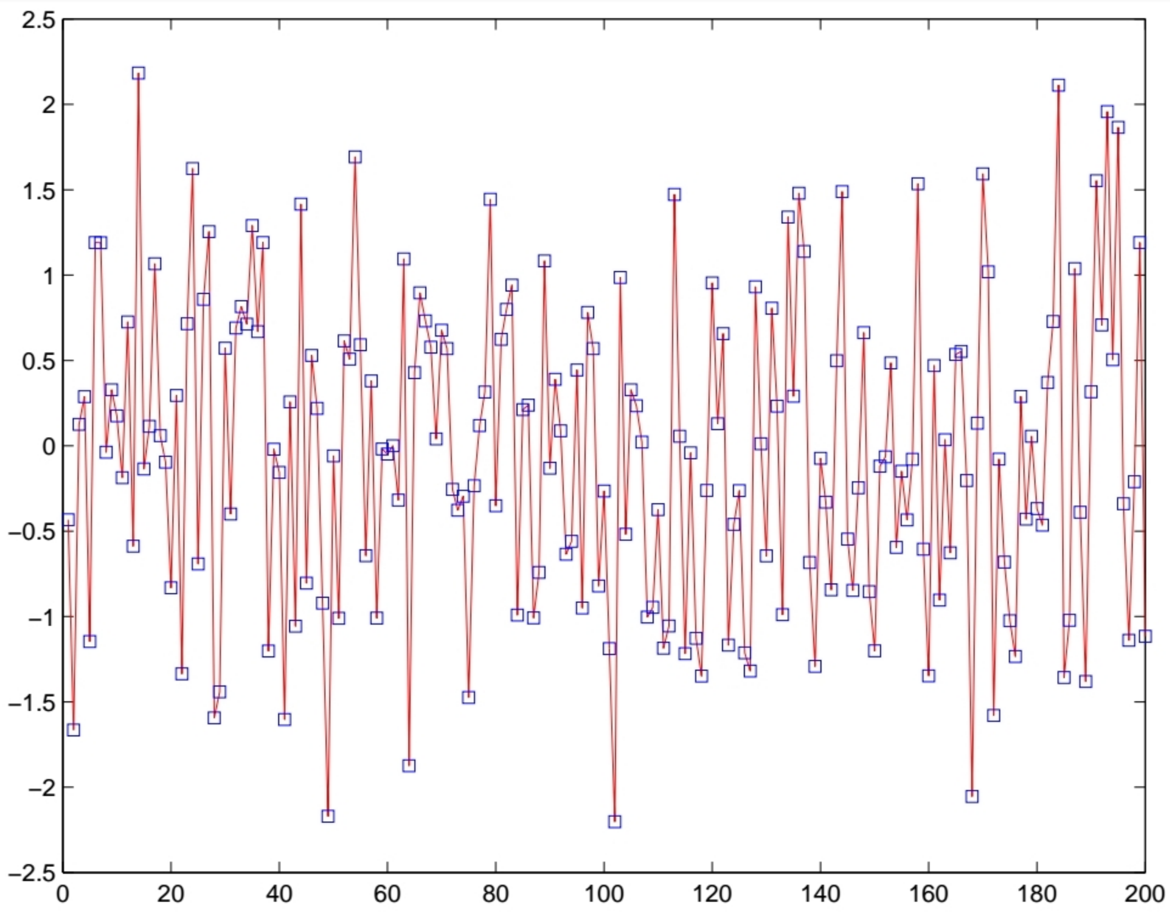



Figure: An iid sequence

# Random walk

- How would you model the position of a walk along a straight line?

$$X_t = X_{t-1} + 1$$

- Now imagine a random walk, where you can take a step forward or backward with equal probability.

$$X_t = r_1 + r_2 + \cdots + r_t, \qquad t = 1,2,\ldots$$

where $r_t$ is i.i.d.

$$P(r_t = 1) = \frac{1}{2} \quad and \;\; P(r_t = -1) = \frac{1}{2}$$

- This is a simple symmetric random walk.
  - What are the key differences from the i.i.d. sequence
  - What if we take the differencing.

# Random walk example

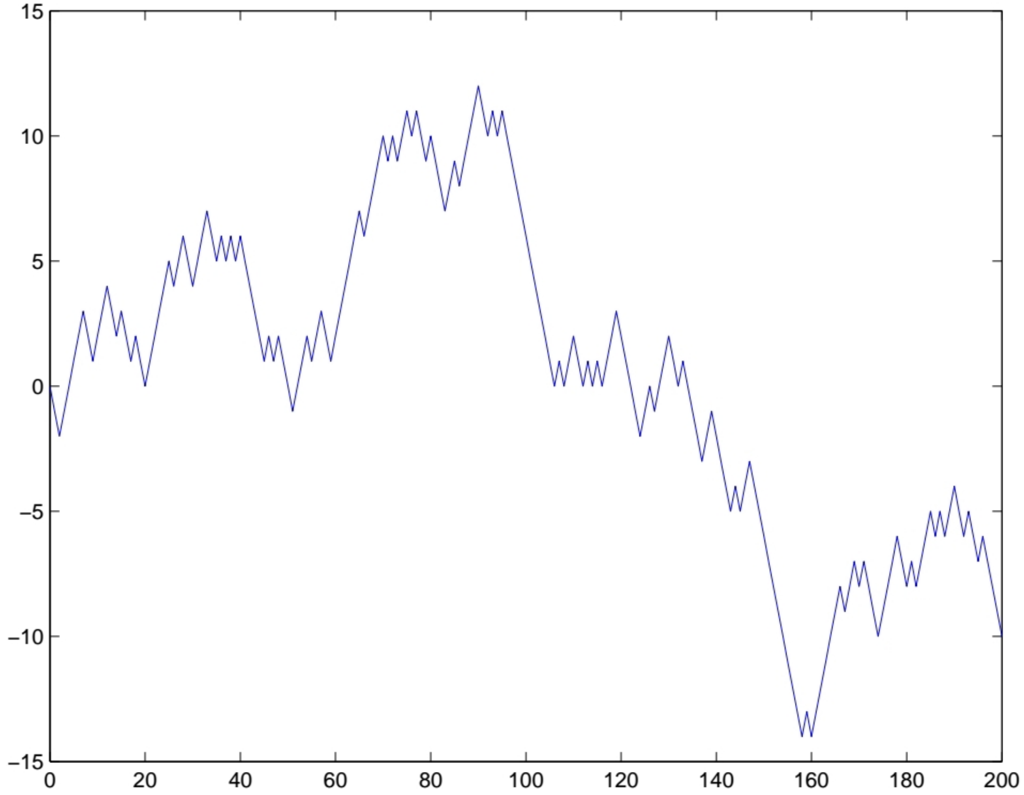



Figure: Simple symmetric random walk

# Random walk with a drift

- $X_t - X_{t-1} = r_t$
- $r_t$ can be any random variables, for instance a Gaussian white noise $w_t$.
- Adding a drift

$$X_t = \delta + X_{t-1} + w_t$$

- Assuming the starting point is zero, we have

$$X_t = \delta t + \sum_{i=1}^{t} w_i$$

- Random walk acts on the previous step, what if we want to generalize it to several steps before? Note the difference on dependency.

# Autoregressive and moving average model

- Autoregressive model (AR) is a class of models closely related to random walks.
- It is defined so that the current location is a linear combination of previous locations plus a random term

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + w_t$$

This is an AR(p) model.

- Another related set of models are the moving average models

$$X_t = w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$$

This is an MA(q) model. It takes sliding window and take a weighted average of the white noises within the window.

# Mean and autocovariance function

- The mean function of $\{X_t\}$ is
$$\mu_X(t) = E(X_t)$$
- The variance function of $\{X_t\}$ is
$$\sigma_X^2(t) = V(X_t) = E\left[\left(X_t - \mu_X(t)\right)^2\right]$$
- The covariance function of $\{X_t\}$ is
$$\gamma_X(s,t) = cov(X_s, X_t) = E\left[\left(X_s - \mu_X(s)\right)\left(X_t - \mu_X(t)\right)\right]$$

# Mean and autocovariance function

- The mean function for random walk with drift is

$$\mu_t = E(X_t) = E\left(\delta t + \sum_{i=1}^{t} w_i\right) = E(\delta t) + E\left(\sum_{i=1}^{t} w_i\right)$$

$$= \delta t + \sum_{i=1}^{t} E(w_i) = \delta t$$

- How about the mean functions for AR(p) and MA(q) model?
- Exercise: variance and autocovariance function for an MA(2) model.
- Autocorrelation function

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(s,s)\gamma(t,t,)}}$$

# Stationarity

- A time series model for the observed data $\{x_t\}$ is a specification of the joint distributions (or possibly only the means and covariances) of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is a realization.
- In traditional statistics: Random sampling procedures enable us to obtain replicated observations under identical conditions. Besides, these observations are independent.
- For time series: We have only a single realization at each time point and also dependent over time. More precisely, it is a sample of size one.
- For any inference to be possible, we must recreate some notion of replicability.

# Stationarity

- A time series $\{X_t, t = 0, \pm 1, \pm 2, \cdots\}$ is said to be stationary if it has statistical properties similar to those of the "time-shifted" series $\{X_{t+h}, t = 0, \pm 1, \pm 2, \cdots\}$, for each integer h.
- Definition

A time series $\{X_t\}$ is said to be strongly or strictly stationary if the joint density functions depend only on the relative location of the observations, so that

$$f\left(x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\right) = f\left(x_{t_1}, x_{t_2}, \dots, x_{t_k}\right)$$

meaning that $(X_{t_1+h}, X_{t_2+h}, \cdots, X_{t_k+h})$ and $(X_{t_1}, X_{t_2}, \cdots, X_{t_k})$ have the same joint distributions for all h and for all time points $\{t_i\}$.

# Stationarity

- Strong stationarity is too strong to be practically useful. Besides, specifying the densities $f\ (x_{t_1}, x_{t_2}, \cdots, x_{t_k})$ is usually very complicated.
- Definition

A time series $\{X_t\}$ is weakly stationary if
(i) $\mu_X(t)$ is independent of t, i.e. $\mu_X(t) = \mu_X$ for all t and finite.
(ii) $\sigma_X^2(t)$ is finite.
(iii) $\gamma_X(t + h, t)$ is independent of t for each h.

- Weak stationarity is also referred to as second-order stationarity, or covariance stationarity. From now on, we shall say stationary to mean weakly stationary.
- Stationarity allows the re-creation of the notion of replicability that is crucial to statistical inference.

# Intuition and properties

- Location does not matter - ONLY distance. The sequence consists of identically distributed random variables.
- All strongly stationary time series are weakly stationary, but not vice versa.
- For Gaussian time series, these two concepts coincide. For any t,

$$\gamma(t+h,t) \ = \ E(X_{t+h}-\mu)(X_t-\mu)$$
$$= \ E(X_h-\mu)(X_0-\mu)=\gamma(h,0)$$

- Thus, for stationary processes $\gamma(s,t) \ = \ \gamma(s-t)$ and we write $\gamma(h)=E(X_{t+h}-\mu)(X_t-\mu)$

# Intuition and properties

- Thus for autocorrelation function, we have
$$\rho(h) = \frac{\gamma(h)}{\gamma(0)}$$
- Also when time series is stationary,
$$\gamma(h) = \gamma(-h)$$

and similarly
$$\rho(h) = \rho(-h)$$
- Without stationary, we have little hope of estimating the full $\gamma(s, t)$. With stationarity, we now have many observations that are h apart from one another (when h << T). Thus inference is possible for stationary processes.

# Estimation

- Let's first note that with stationarity, the (true) mean is constant. We can therefore estimate the mean using the sample mean

$$\bar{x} = \frac{\sum_{t=1}^{n} x_t}{n}$$

This converges to $\mu$ regardless of the dependency structure (up to a point).

- Now, let's look at the sample autocovariance function

$$\hat{\gamma}(h) = \frac{\sum_{t=1}^{n-1}(x_{t+h} - \bar{x})(x_t - \bar{x})}{n} \quad for\ h = 0,1,\ldots,n-1$$

# Estimation

- For fixed h, all the random variables $y_t = (x_{t+h} - \bar{x})(x_t - \bar{x})$ have the same distribution. Therefore, $\frac{\sum_{t=1}^{n} y_t}{n}$ converges to $\gamma(h) = E(x_h - \bar{x})(x_0 - \bar{x})$. To get the sample autocorrelation we simply scale by the variance

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

- An important property for the sample autocorrelation function is that when the true model is white noise $\hat{\rho}(h)$, (h=1,2,…,H) is approximately normally distributed with zero mean and standard deviation of $1/\sqrt{n}$. Since by central limit theorem,

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

is approximately normal for large n.

# Examples of stationary processes

- The iid noise process

If $\{X_t\}$ is an iid noise process, we write $\{X_t\} \sim IID(0, \sigma^2)$

We also have,

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

- The white noise process

Let $\{X_t\}$ be a sequence of

- Uncorrelated random variables, i.e. $\gamma_X(h) = 0\ for\ h \neq 0$.
- Each variable having zero mean, i.e. $E(X_t) = 0$.
- Each variable having finite variance, i.e. $V(X_t) = \sigma^2 < \infty$.

Such a sequence is referred to as white noise (with mean 0 and variance $\sigma^2$), and indicated by $\{X_t\} \sim WN(0, \sigma^2)$.

# Examples of stationary processes

- $\{X_t\} \sim WN(0, \sigma^2)$ is clearly a stationary process.
- The covariance function of $\{X_t\} \sim WN(0, \sigma^2)$ is the same as that of $IID(0, \sigma^2)$, namely

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2, \ if \ h = 0 \\ 0, \ if \ h \neq 0 \end{cases}$$

- The random walk process

$$X_t = w_1 + w_2 + \cdots + w_t$$

where $w_t \sim WN(0, \sigma^2)$.

  - $E[X_t]$ =? and $V[X_t]$ =?
  - If s > t then $cov(X_s, X_t)$ =? $\gamma_X$(t + h, t) =?
  - Is the series $\{X_t\}$ stationary?

# First order AR process

- A series $\{X_t\}$ is a first-order autoregressive or AR(1) process if

$$X_t = \phi X_{t-1} + w_t, \qquad t = 0, \pm 1, \pm 2, \dots$$

where

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

- It is easy to show that the autocovariance function (ACVF) is

$$\gamma_X(h) = \sigma^2 \frac{\phi^{|h|}}{1-\phi^2}, \quad h = 0, \pm 1, \pm 2, \dots$$

The autocorrelation function (ACF) of an AR(1) is given by

$$\rho_X(h) = \phi^{|h|}, \qquad h = 0, \pm 1, \pm 2, \dots$$

# First order AR process

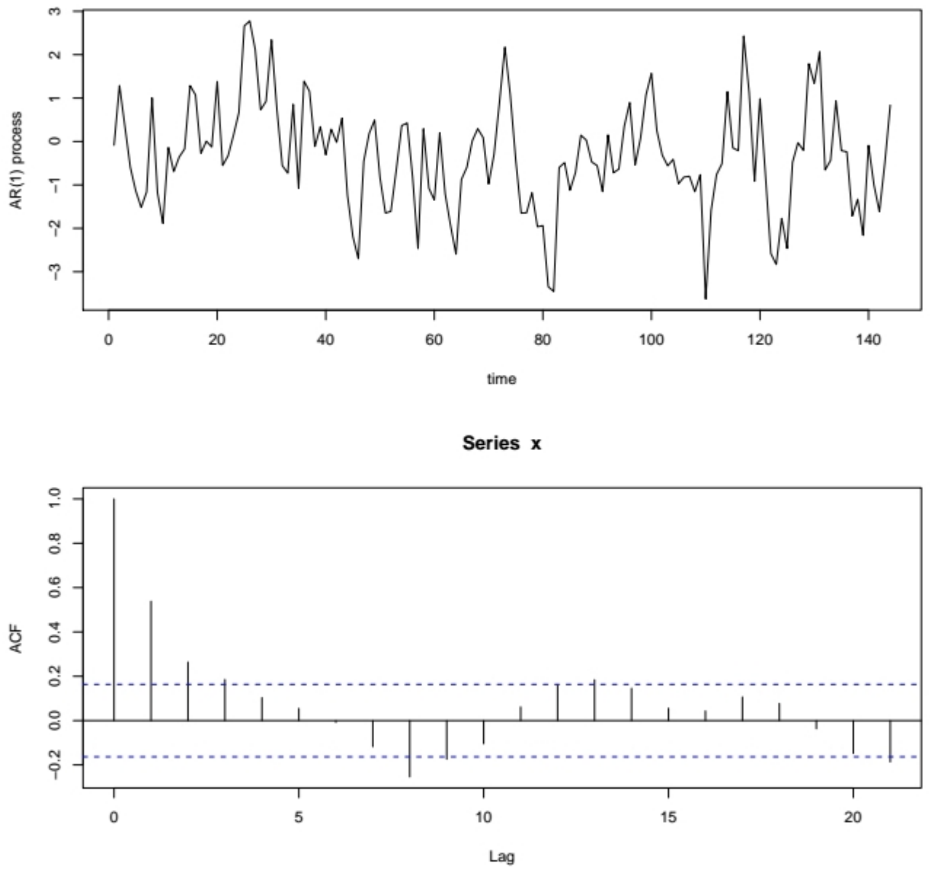



Figure: AR(1) with $\phi = 0.5$.

# First order MA process

- A series $\{X_t\}$ is a first-order moving average or MA(1) process if

$$X_t = w_t + \theta w_{t-1}, \qquad t = 0, \pm 1, \pm 2, \dots$$

where

  - $\{w_t\} \sim WN(0, \sigma^2)$
  - $\theta$ is a real-valued constant

- It is easy to show that the autocovariance function (ACVF) is

$$\gamma_X(t+h, t) = \begin{cases} \sigma^2(1+\theta^2), \ if \ h = 0 \\ \sigma^2\theta, \ if \ h = \pm 1 \\ 0, \ if \ |h| > 1 \end{cases}$$

- Clearly, an MA(1) process is a stationary process
- Also, the autocorrelation function (ACF) of an MA(1) is given by

$$\rho_X(t+h, t) = \rho_X(h) = \begin{cases} 1, & if \ h = 0 \\ \theta/(1+\theta^2), & if \ h = \pm 1 \\ 0, & if \ |h| > 1 \end{cases}$$

# First order MA process

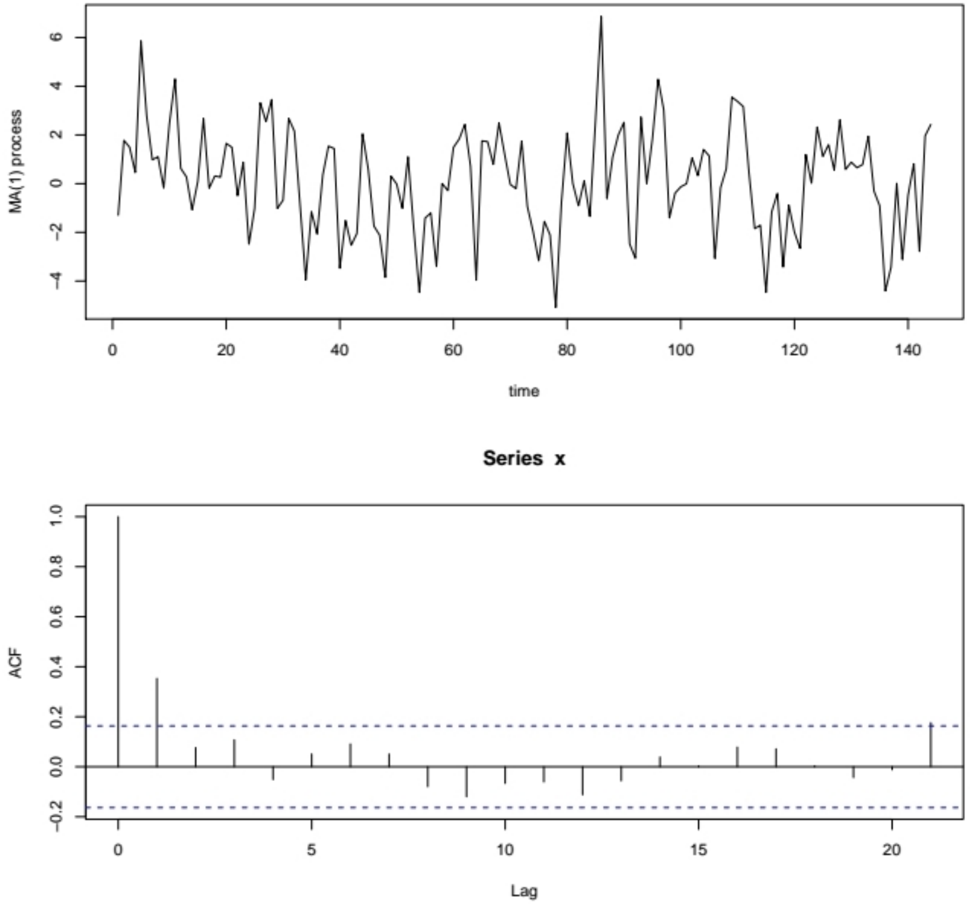


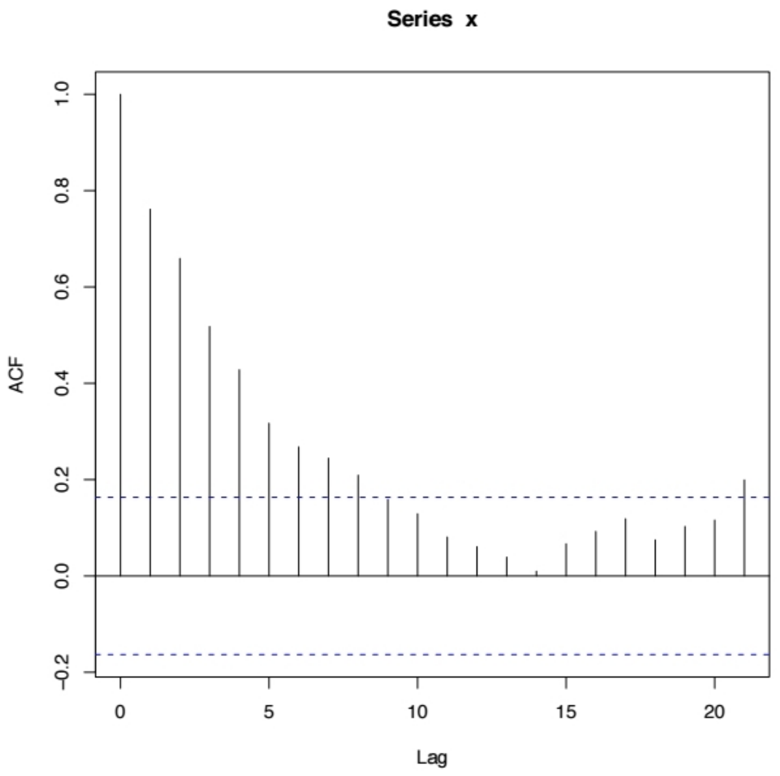

Figure: MA(1) with $\theta = 2$.

# A test for IID noise using the sample (ACF)

- For iid noise with finite variance, we have, for $h \neq 0$,
$$\hat{\rho}(h) \sim N(0, 1/n)$$
- Steps for the diagnostic for iid noise
  - Plot the lag h versus $\hat{\rho}(h)$
  - Draw two horizontal lines at $\pm 1.96/\sqrt{n}$. These two lines are drawn automatically in R
  - You should have about 95% of the values of $\{\hat{\rho}(h): h = 1, 2, \ldots\}$ within the lines if the noise is indeed iid.

# A test for IID noise using the sample (ACF)

- Which of the two depicts an IID noise?



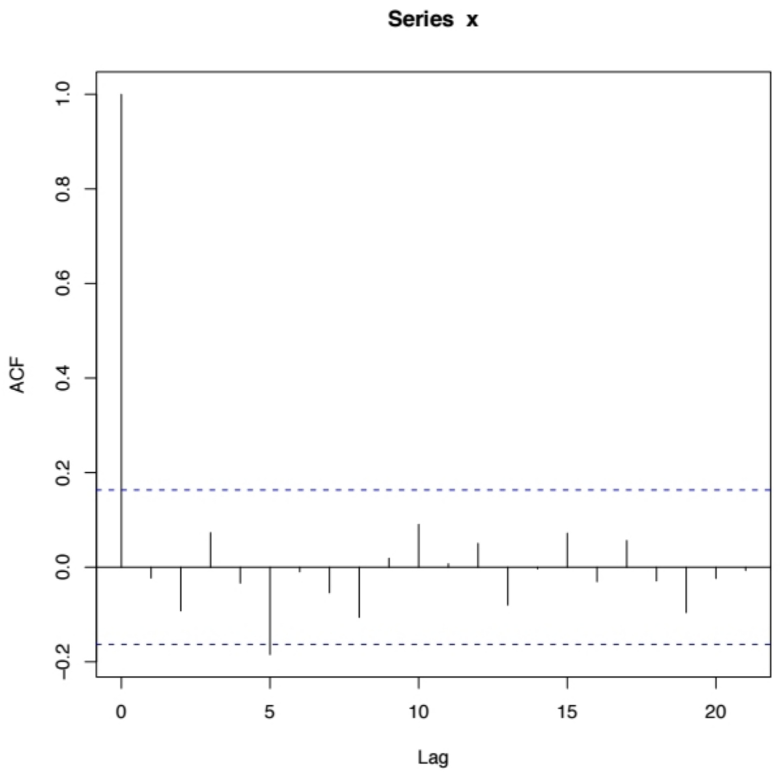




Figure: Which is more likely i.i.d?

# Asset returns

Let $P_t$ be the price of an asset at time t, and assume that there is no dividend.

- One period simple return:
  - Gross return: $1 + R_t = P_t / P_{t-1}$
  - Simple return: $R_t = \frac{P_t}{P_{t-1}} - 1$
- Multiperiod simple return:
  - k-period gross return:

$$1 + R_t(k) = \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \cdots \times \frac{P_{t-k+1}}{P_{t-k}} = (1 + R_t) \ldots (1 + R_{t-k+1})$$

  - k-period simple return: $R_t(k) = \frac{P_t}{P_{t-k}} - 1$

# Asset returns

- Continuous compound or log-return:

$$r_t = \ln(1 + R_t) = ln\frac{P_t}{P_{t-1}} \approx R_t$$

- Multiperiod log return

$$r_t(k) = \ln\frac{P_t}{P_{t-k}} = \ln\big(1 + R_t(k)\big)$$
$$= \ln(1 + R_t)\,(1 + R_{t-1}) \ldots (1 + R_{t-k+1}) = r_t + r_{t+1} + \cdots + r_{t-k+1}$$

# Example

Suppose the daily closing prices of a stock are

| Day | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Closing price | 37.84 | 38.49 | 37.12 | 37.60 | 36.30 |

The simple return from day 1 to day 2:
$R_2 = \frac{38.49-37.84}{37.84} \approx 0.017.$

The simple return from day 1 to day 5:
$R_5(4) = \frac{36.30-37.84}{37.84} \approx -0.041.$

The log return from day 1 to day 2:
$r_2 = \log 38.49 - \log 37.84 \approx 0.017.$

The log return from day 1 to day 5:
$r_5(4) = \log 36.30 - \log 37.84 \approx -0.042.$

# Common descriptive statistics of returns

1. Mean $\mu$– average return

2. Variance $\sigma^2$ – risk

3. Skewness $S(X) = E(X - \mu)^3/\sigma^3$ For symmetric distribution, $S(X) = 0$.
   More negative returns? More positive returns?

4. Kurtosis $K(X) = E(X - \mu)^4/\sigma^4$ and excess kurtosis $K(X) - 3$.

5. Quantile measures: value at risk.

# Estimates

1. Sample mean $\hat{\mu} = \frac{1}{T}\sum_{i=1}^{T} x_t$
2. Sample variance

$$\hat{\sigma}^2 = \frac{1}{T-1}\sum_{t-1}^{T}(x_t - \hat{\mu})^2$$

3. Sample skewness

$$\hat{S}(X) = \frac{1}{(T-1)\hat{\sigma}^3}\sum_{t-1}^{T}(x_t - \hat{\mu})^3$$

4. Sample kurtosis

$$\hat{K}(X) = \frac{1}{(T-1)\hat{\sigma}^4}\sum_{t-1}^{T}(x_t - \hat{\mu})^4$$

# Empirical properties of returns

The empirical distribution of asset return tends to skew to the left, has heavier tails and a higher peak than normal distribution.

```
> basicStats(y)
#Compute descriptive statistics of apple returns.
#Load fBasics library first using library(fBasics).
                      y
nobs        2515.000000
...
Mean           0.001382
Median         0.000713
...
Variance       0.001030
Stdev          0.032092
Skewness      -1.607603
Kurtosis      29.067185
```

# Distributions

- If one assumes the log return X = log(Y ) of an asset is normally distributed $N(\mu, \sigma^2)$, then Y has a log normal distribution with

$$E(Y) = \exp(\mu + \sigma^2/2), \quad V(Y) = \exp(2\mu + \sigma^2)[\exp(\sigma^2) - 1]$$

- If one wants to consider distributions with flatter tails, we can use t or skewed t distribution.
- Other popular distributions include scaled mixture and stable distribution.
- Multivariate extensions.

# Likelihood

- Given two consecutive returns $r_1$ and $r_2$, one has $f(r_1, r_2) = f(r_2|r_1)f(r_1)$. In general,

$$f(r_T, r_{T-1}, \cdots, r_2, r_1) = f(r_T|r_{T-1}, \cdots, r_1)f(r_{T-1}, \cdots, r_1)$$
$$= \ \ldots = \left[\prod_{t=2}^{T} f\ (r_t\ |r_{t-1}, \cdot\cdot\cdot, r_1)\right] f\ (r_1).$$

One can make distribution assumptions on the conditionals. For normal distribution, the log likelihood can be written in a very simple additive form.

$$\mathrm{L} = \sum_{t=2}^{T} (-\log(\sigma_t) - (r_t - \mu_t)^2/(2\sigma_t^2)) \ + \ C$$

- We will describe important models regarding μt, e.g. ARIMA models, and regarding σ_t, e.g. ARCH and GARCH.

# Recommended books

1. https://otexts.com/fpp2/
2. **A Little Book of R For Time Series, Avril Coghlan**