Time Series and Related Quantitative Modeling

Agenda

#	Topic	Page
1	Introduction to Financial Time Series	3
2	Stationarity	5
3	Seasonality	8
4	Autocorrelation	10
5	Linear Time Series: AR and MA	12
6	Linear Time Series: ARMA	16
7	EWMA	18
8	Non-Stationary Time Series: ARIMA and ARIMAX	19
9	Multivariate Time Series	22
10	ARCH and GARCH	24
11	Extreme Value Theory	28
12	Error Correction Model	30
13	Current trends	32

Overview

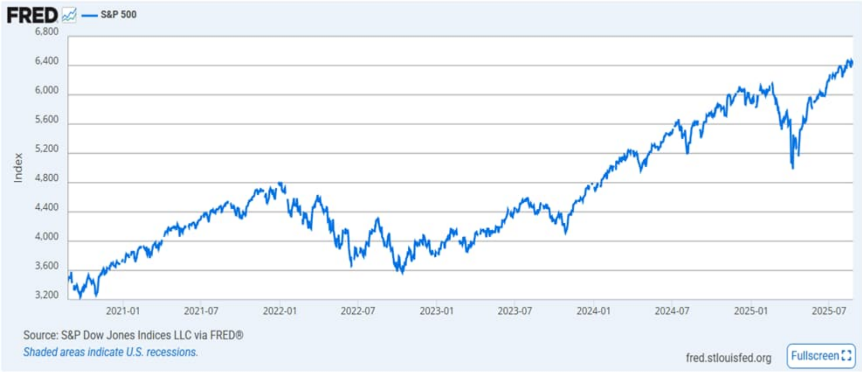
A time series in Financial markets is **a sequence of observations** of a financial variable(s) recorded at **successive timepoints**, e.g., daily, weekly, monthly. The financial variables could be the following

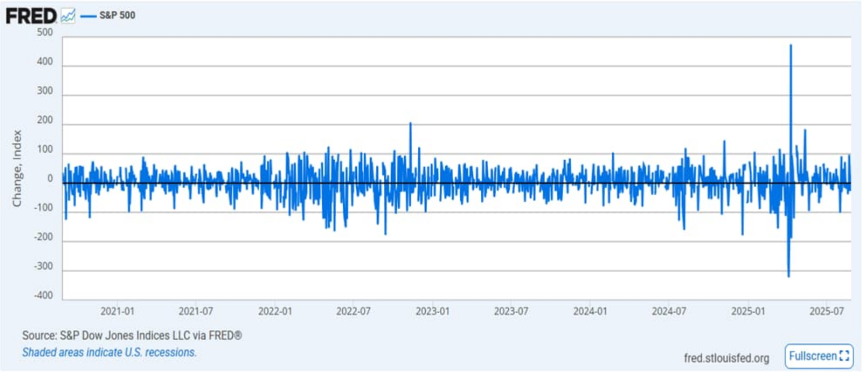
- ► Equity prices
- ► Interest rates
- ► Credit spreads
- ► Volatilities (e.g., VIX), …

► Purpose of time series analysis

- ► Modeling and forecasting (e.g., CCAR)
- ► Trend and/or pattern identification (e.g., seasonality)
- ► Scenario analysis, risk measurement (e.g., VaR), and more

► Examples (S&P 500 index vs change*)

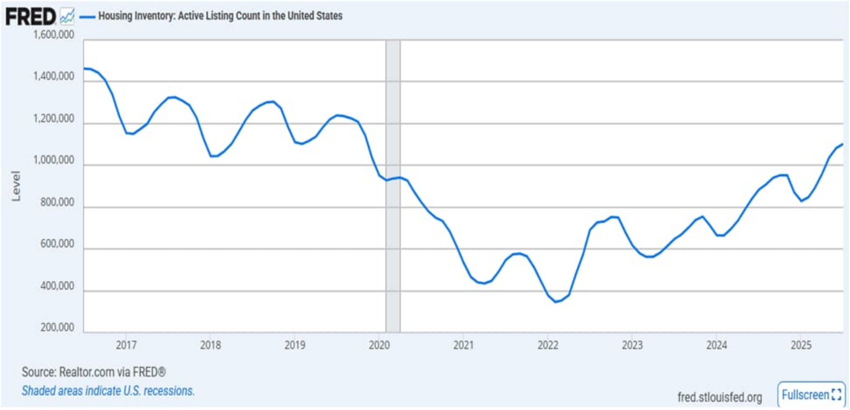


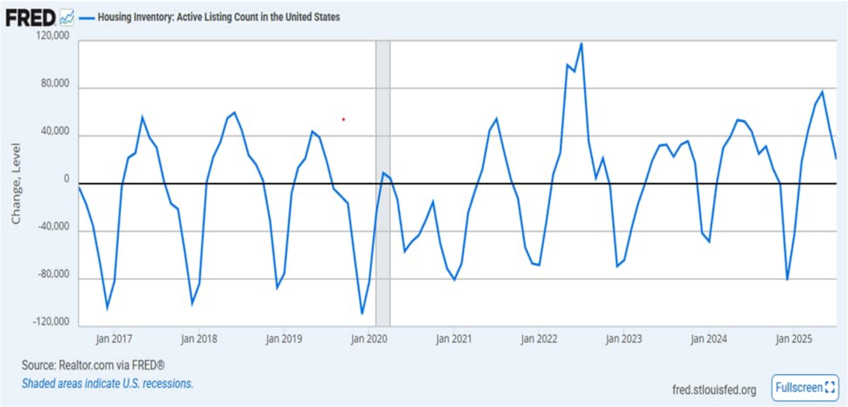


*: S&P 500 (SP500) | FRED | St. Louis Fed

Overview (cont')

- **Key components** of a time series
  - **Trend:** long-term direction or movement
    - Example: upward trend in the S&P 500
  - **Seasonality:** regular, predictable pattern that repeats over a fixed period
    - Example: peak in summer and low in winter for housing inventory*





  - **Cyclical component:** longer term, non-fixed freq fluctuations
    - Example: economic cycle
  - **Noise:** white noise – independently and identically distributed (iid) normal random numbers
  - **Other events driven irregularities**

*: Housing Inventory: Active Listing Count in the United States (ACTLISCOUUS) | FRED | St. Louis Fed

Stationarity

- **What is stationarity**
  - *Strict*: all joint distributions are time-invariant
  - *Weak*: constant mean, variance and auto-covariance, i.e., only the first two moments are time-invariant

$$E[X_t] \text{ and } Var[X_t], \text{are constant; } Cov(X_t, X_{t+h}) \text{ depends on lag } h \text{ only}$$

  - Note: Strict stationarity → Weak stationarity
- **Why it matters**

Upward trend leads to invalid relationship!

  - Unreliable statistical inference
  - Spurious regression
  - Stability of risk measure (e.g., VaR assumes each risk factor time series is stationary)

Stationarity (cont')

- **Non-stationary patterns**
  - Trend (deterministic, stochastic)
  - Seasonality
  - Regime switch
  - …

- **Diagnostic testing**
  - Augmented Dickey-Fuller (ADF)

$$\Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \sum_{i=1}^{k} \phi_i \Delta X_{t-i} + \epsilon_t$$

$H_0 : \gamma = 0$ (unit root, X_t is non-stationary)

  - KPSS

$X_t = \alpha + \beta t + Y_t + \epsilon_t$; where $Y_t = Y_{t-1} + u_t$ is random walk component and u_t is white noise

$H_0 : Var(u_t) = 0$ (X_t is stationary)

  - **Note:** KPSS is typically a complement to the ADF test. If ADF – Reject H_0 and KPSS – fail to reject H_0, there is a strong evidence on stationarity of time series

- **Co-integration:** what if there is a stable relationship between 2 non-stationary time series

Stationarity (cont')

- **Transformations** to make a time series as stationary
  - Remove trend

$$\text{Difference: } \Delta X_t = X_t - X_{t-1}$$

$$\text{Change: } Y_t = \frac{X_t - X_{t-1}}{X_{t-1}}$$

  - Remove seasonality

$$Y_t = X_t - X_{t-m}$$

  - Handle non-constant variance

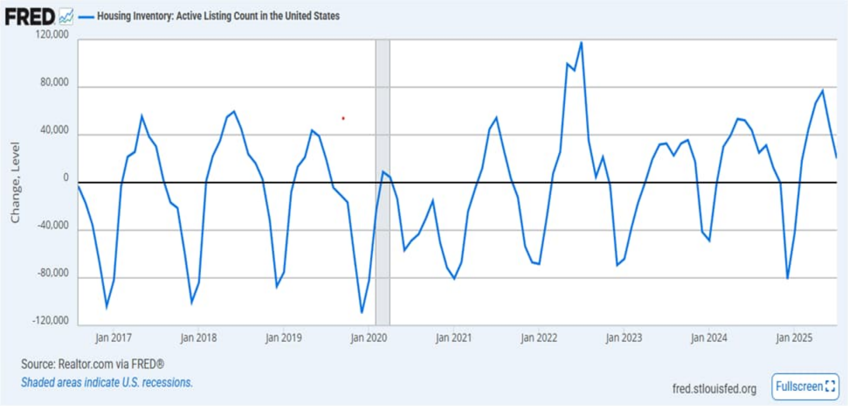
$$Y_t = \log(X_t)$$

Note: always retest transformed time series with ADF/KPSS tests to confirm stationarity has been achieved

- **Application** in Value-at-Risk (VaR)
  - Equity price: relative return
  - Interest rate: absolute change

Seasonality

- **What is seasonality**
  - Time series experiences a **predictable** and **periodic** pattern
    - Predictable: the pattern (e.g., mean) repeats in a consist manner
    - Periodic: the time between the repetitions of the pattern, e.g.*,



  - Note: periodic != cycles (fluctuations w/o a fixed known period)
  - Examples:
    - House prices
    - Gas prices
    - Retail

- **Why it matters** – refer to “Stationarity”

*: Housing Inventory: Active Listing Count in the United States (ACTLISCOUUS) | FRED | St. Louis Fed

Seasonality (cont')

- **Diagnostic testing**
  - Visual detection
    - Regular, repeating peaks and troughs at a fixed interval
    - Time series plot, boxplot by season
  - Statistical tests
    - Canova-Hansen - H_0: seasonality is stable
    - Kruskal-Wallis - H_0: all seasonal categories have the same distribution
  - X-13ARIMA-SEATS*: statistical methods for seasonality adjustment that are implemented in the US Census Bureau's software package

- **Transformations** to deseasonalize data
  - Seasonal differencing

$$Y_t = X_t - X_{t-m}$$

  - Regression on seasonal dummies: regression on dummy variables for the seasonal period, $\widehat{X_t}$,

$$Y_t = X_t - \widehat{X_t}; \widehat{X_t} = \alpha + \beta \cdot t + \sum_i^k \gamma_i \cdot DummyVariable_i$$

  - X-13ARIMA-SEATS

- **Application**
  - House Price Index

*:X-13ARIMA-SEATS Seasonal Adjustment Program

Autocorrelation

- **What is autocorrelation**
  - Correlation between a time series and its **lagged versions**
  - Autocorrelation types:
    - Serial autocorrelation: most common one, linear dependence between today's value and its past value

$$X_t = \phi X_{t-1} + \epsilon_t$$

$$\phi = \frac{Cov(X_t, X_{t-1})}{\sigma^2}$$

Note: assume **variance stationarity**

    - Partial autocorrelation

$$X_t = \sum_{i=1}^{k} \phi_i X_{t-i} + \epsilon_t$$

    - Others
  - Examples: too many
    - Equity prices
    - Commodity prices
    - Interest rates

- **Why it matters**
  - Residuals need to be **independent** identically distributed (iid)!

Autocorrelation (cont')

- **Diagnostic testing**
  - AutoCorrelation Function (ACF) and Partial ACF (PACF)

ACF and PACF for an AR(1) Time Series data set.

Note: ACF measures dependence at each lag, while PACF measures **additional dependence** at Lag k after **controlling for lags** $\mathbf{1, \ldots, k-1}$

  - Statistical tests
    - Ljung-Box - H_0: $\rho_1 = \cdots = \rho_k = 0$
    - Durbin-Watson - H_0: no $AR(1)$ in residuals
- **"Process" of model development (e.g., loss forecasting models)**
  - Visualization (e.g., ACF/PACF)
  - Transformation (e.g., differencing/log, deseasonalization) to ensure stationarity
  - Modeling framework (e.g., add exogenous variables)
  - Model selection (e.g., AIC/BIC/…)
  - Testing e.g., residual diagnostic, out-of-sample testing, monitoring)

Linear Time Series: AR and MA

- **Autoregressive (AR)**
  - An AR model of order $p \geq 0$, $AR(p)$, is defined as follows, "*memory of past values*"

$$x_t = \sum_{k=1}^{p} \phi_k x_{t-k} + \epsilon_t$$

$$(1 - \phi(L))x_t = \epsilon_t$$

$$\phi(L) = \sum_{k=1}^{p} \phi_k L^k \ and \ L^k x_t = x_{t-k} \ (shift\ operator)$$

  - Meaning: *present depends on past levels, or markets have short-term memory*
  - Examples:
    - Stock price (Geometric Brownian Motion): stationary?

$$dS_t = \mu S_t dt + \sigma S_t dW_t \rightarrow \ln(S_t) = \left(\mu - \frac{1}{2}\sigma^2\right)\Delta t + \ln(S_{t-1}) + \sigma \cdot \epsilon_t$$

    - Interest rate (Hull-White model): stationary?

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t \rightarrow r_t = \kappa \cdot \theta \cdot \Delta t + (1 - \kappa \cdot \Delta t) \cdot r_{t-1} + \sigma \cdot \epsilon_t$$

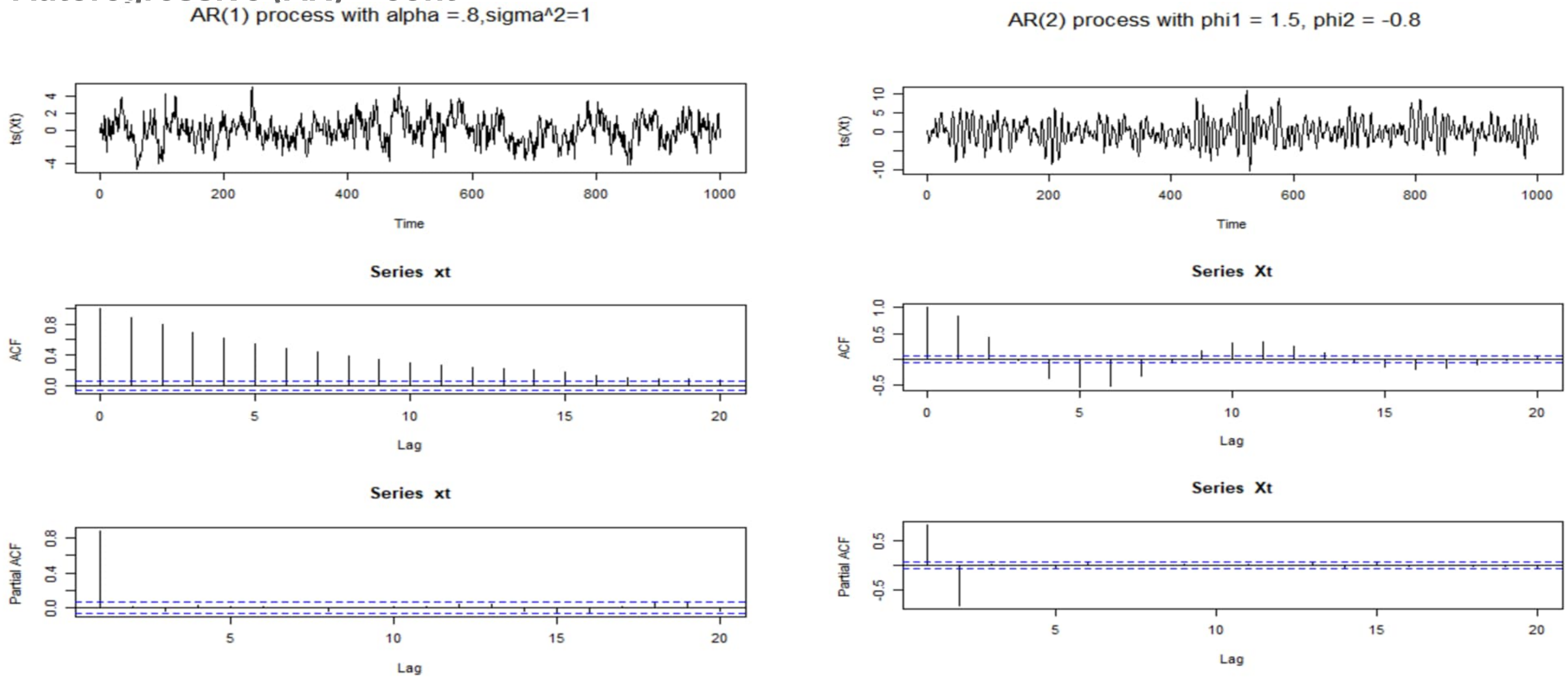
    - Others: house prices, ….
  - Stationarity: roots of $\phi(z) = 0$ outside the unit circle
    - $AR(1)$: $x_t = \phi \cdot x_{t-1} + \epsilon_t$; $|\phi| < 1$

$$Var(x_t) = \frac{Var(\epsilon_t^2)}{1 - \phi^2}$$

$$x_t = (1 + \phi L + \phi^2 L^2 + \cdots)\epsilon_t = \frac{1}{(1 - \phi L)}\epsilon_t$$

Linear Time Series: AR and MA – cont’

- **Autoregressive (AR) – cont’**



- Intuition
  - for $AR(p)$, PACF measure correlation between x_t and x_{t-k} conditional on $x_{t-1}, \ldots, x_{t-p}$, as such, the direct impact from x_{t-k} to x_t is fully covered by the first p lags
- Estimation methods
  - Maximum-Likelihood Estimation
  - Yule-Walker (moments matching)
- Diagnostics: residual, stationarity (unit-root), stability

Linear Time Series: AR and MA – cont'

- **Moving Average (MA)**
  - An MA model of order $q \geq 0$, $MA(q)$, is defined as follows, "*memory of past shocks*"

$$x_t = \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t$$

$$x_t = (1 + \theta(L))\epsilon_t$$

$$\theta(L) = \sum_{k=1}^{q} \theta_k L^k \ \ and\ L^k \epsilon_t = \epsilon_{t-k}$$

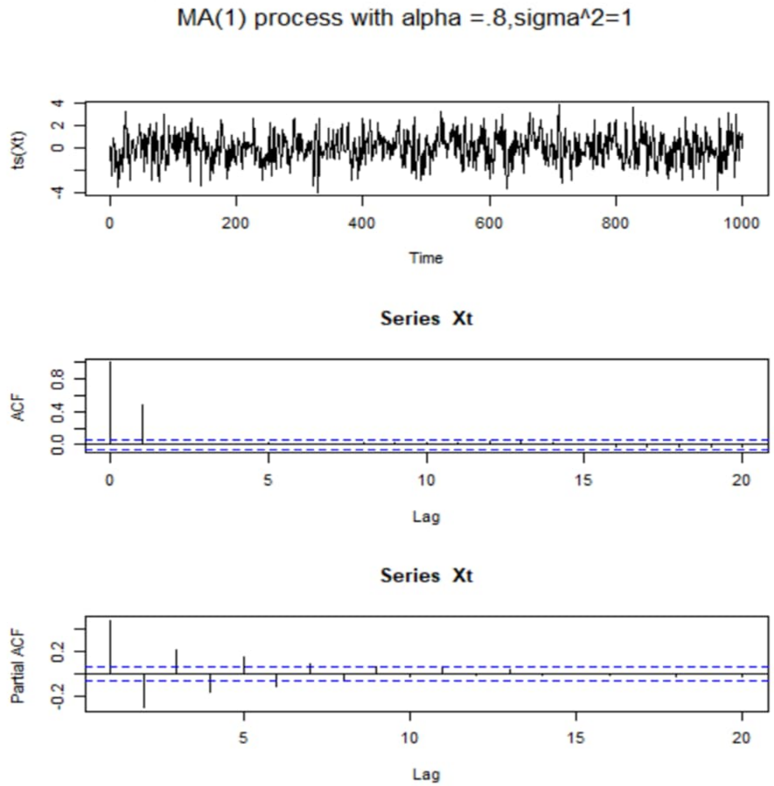
  - Meaning: *present depends on past shocks*
  - Examples:
    - Volatility indices
    - **Not used for forecasting**
  - Stationarity: always stationary, $E(x_t) = 0$, $Var(x_t) = \left(1 + \sum_{k=1}^{q} \theta_k^2\right)$
    - $MA(1)$: $x_t = \epsilon_t + \theta\ \epsilon_{t-1}$

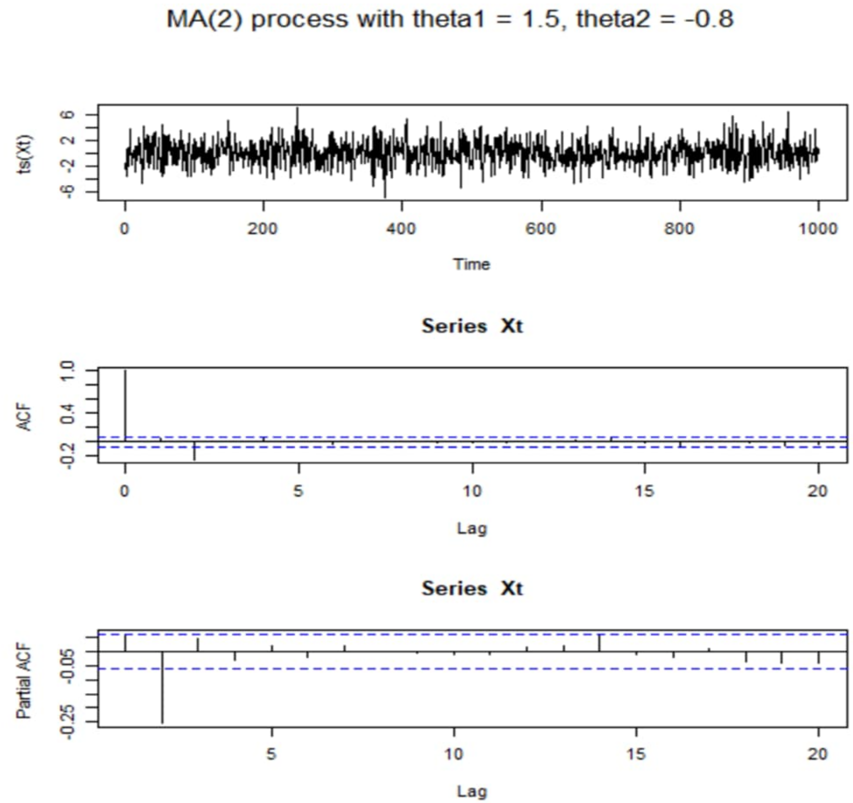
$$Var(x_t) = Var(\epsilon_t^2) \cdot (1 + \theta^2)$$

    - $AR(1)$: $x_t = \phi \cdot x_{t-1} + \epsilon_t = (1 + \phi L + \phi^2 L^2 + \cdots)\epsilon_t \rightarrow MA(\infty)$
    - Invertible: the roots of $\theta(z)$ lie outside the unit circle

Linear Time Series: AR and MA – cont’

- **Moving Average (MA) – cont’**
  - Identify lag(s)





  - Intuition
    - for $MA(q)$, ACF measure correlation between x_t and x_{t-q}, which share the same ϵ_{t-q}
  - Estimation methods
    - Maximum-Likelihood Estimation

Linear Time Series: ARMA

- **Recap: AR and MA**

	$AR(p)$	$MA(q)$
Meaning	Present depends on recent **past levels**	Present depends on the **past shocks**
Shock memory	Infinite	Finite, die out after q periods
Stationarity	Conditional, unit root	Always
ACF/PACF	ACF tails off gradually; PACF cuts at lag p	ACF cuts at lag q, PACF tails off gradually
Variance	Depends on ϕs and $Var(\epsilon)$	$Var(\epsilon) \cdot \left(1 + \cdots + \theta_q^2\right)$
Estimation	Yule-Walker/MLE	MLE
Residuals	White noise	White noise
Forecasting	Yes	No

- An ARMA model, $ARMA(p, q)$, is defined as follows,

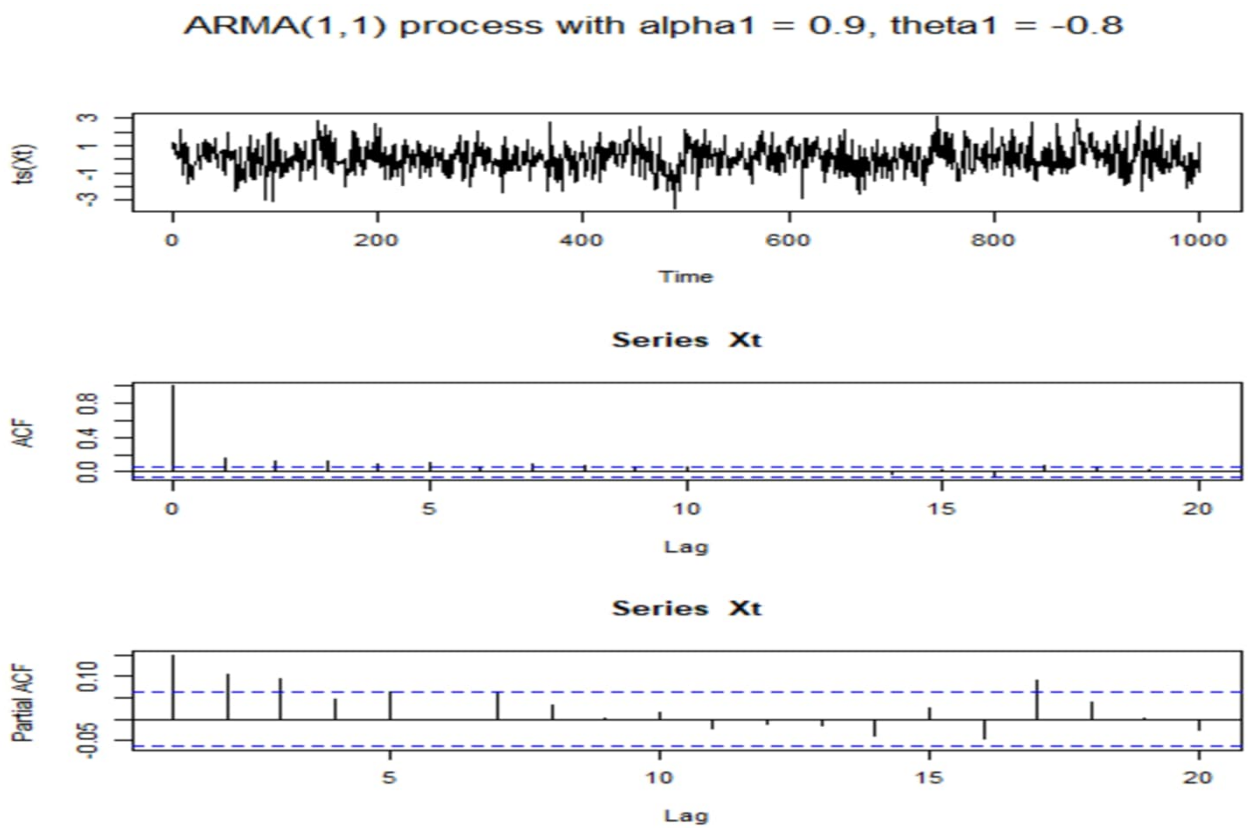
$$x_t = \sum_{k=1}^{p} \phi_k x_{t-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t$$

$$(1 - \phi(L))x_t = (1 + \theta(L))\epsilon_t$$

  - Persistent levels and impact of past shocks

Linear Time Series: ARMA – cont'

- **ARMA**
  - Identify lag(s): both ACF/PACF tail off



  - Estimation methods
    - Find (p, q) thru ACF/PACF
    - Fit model w/ MLE
    - Residual diagnostics

EWMA

- **EWMA** (Exponentially Weighted Moving Average*): a simple approach to model vol using historical vol (vs equal weights vs GARCH)
  - Variance: for a time series $\{r_t\}$, typically returns, with a decay factor $\lambda \in (0,1)$

$$\sigma_t^2 = \lambda \cdot \sigma_{t-1}^2 + (1-\lambda) \cdot r_{t-1}^2$$

$$r_t = \sigma_t \cdot \epsilon_t$$

  - Level: for a time series $\{x_t\}$ with a smoothing factor $\lambda \in (0,1)$

$$m_t = \lambda \cdot m_{t-1} + (1-\lambda) \cdot x_{t-1}$$

- Intuition: *recent matters more!*
  - Comparison w/ simple moving average (equal weights)

$$\sigma_t^2 = \frac{1}{n}\sum_{i=1}^{n} r_{t-i}^2$$

  - $\lambda \in (0,1)$: one parameter control how fast we forgot the past

$$\sigma_t^2 = (1-\lambda)\sum_{i=1}^{\infty} \lambda^i r_{t-i}^2$$

Half Life (HL): $\lambda^{HL} = \frac{1}{2}$

- Value-at-Risk (VaR) modeling
  - Monte Carlo
  - Historical Simulation

λ = 0.94 proposed by RiskMetrics**

*: EWMA Weighted Linear Ridge Regression | SOA

**:RiskMetrics Technical Document - Fourth Edition 1996, December

ARIMA and ARIMAX

- **Recap**
  - Stationarity
  - ARMA: $ARMA(p, q)$

$$x_t = \sum_{k=1}^{p} \phi_k x_{t-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t$$

$$(1 - \phi(L))x_t = (1 + \theta(L))\epsilon_t$$

- **ARIMA**
  - $ARIMA(p, 1, q)$, is defined as follows, difference of y_t is an $ARMA(p, q)$

$$x_t = y_t - y_{t-1}$$

$$x_t = \sum_{k=1}^{p} \phi_k x_{t-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t$$

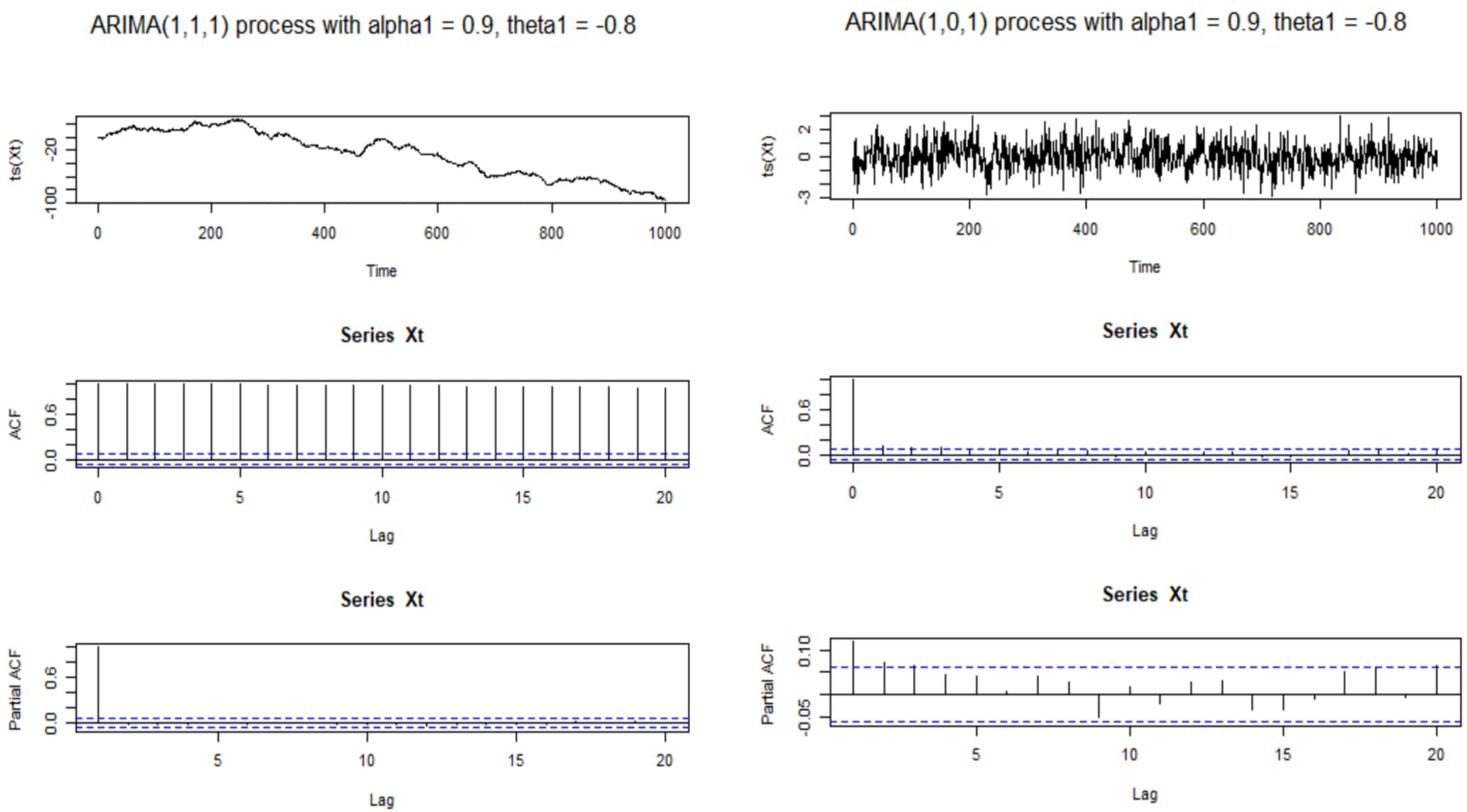
$$\big(1 - \phi(L)\big)x_t = \big(1 + \theta(L)\big)\epsilon_t; \text{ or}$$

$$\big(1 - \phi(L)\big)(1 - L)y_t = (1 + \theta(L))\epsilon_t$$

    - $ARIMA(p, d, q)$: $\big(1 - \phi(L)\big)(1 - L)^d y_t = (1 + \theta(L))\epsilon_t$
  - **I:** “integrated” (I); d: the # of nonseasonal differences
  - The most general class of forecasting models for time series that can be stationarized* by differencing

ARIMA and ARIMAX (cont')

- **ARIMA**



- Estimation methods
  - Identify order of differencing (d), the lowest differencing level that makes a time series stationary
  - Follow ARMA model estimation methods

ARIMA and ARIMAX (cont')

- **ARIMAX**: ARIMA model with **exogenous variable(s)**
  - $ARIMAX(p, 1, q)$, is defined as follows

$$x_t = y_t - y_{t-1}$$

$$x_t = \sum_{k=1}^{p} \phi_k x_{t-k} + \sum_{k=1}^{q} \theta_k \epsilon_{t-k} + \epsilon_t + \sum_{k=1}^{N} \beta_k z_{k,t}$$

    - $ARIMA(p, d, q)$: $\big(1 - \phi(L)\big)(1 - L)^d y_t = \big(1 + \theta(L)\big)\epsilon_t + \vec{\beta} \cdot \vec{z_t}$
  - Key assumptions
    - Stationarity: x_t is stationary
    - Linearity: linear relationship between x_t and its lags, past residuals and exogenous variables
    - No multicollinearity: exogenous variables, $\vec{z_t}$, are not linear combination of one another
    - No-autocorrelation/normality for residuals
- Model development (see p11 for a simple version)

Visualization	Data preparation (e.g., missing data, outlier)
Transformation	Exploratory analysis (e.g., univariate) thru transformation and preliminary variable selection
Modeling framework	Modeling framework
Model selection	Model selection (statistical testing and business intuition)
Testing	Testing (e.g., diagnostic testing, in-sample, out-of-sample, ongoing monitoring)

Multivariate Time Series

- Why multivariate time series
  - Limitation of univariate modeling
    - cannot capture cross-asset effects (e.g., equity vs rates vs credit)
    - Static vs Dynamic dependencies
  - Business needs: CCAR/DFAST*, IFRS9/CECL**, IFRS17/LDTI***, …
- Key multivariate modeling
  - Vector AutoRegression (VAR), Structural VAR (SVAR), VAR w/ Exogenous variables (VARX)
  - Vector Error Correction Model (VECM), …
- VAR: a natural extension of univariate AR model, also one of the most popular multivariate time series models
  - Christopher A. Sims – Wikipedia; Nobel laureates
  - $VAR(p)$, is defined as follows

$$\begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{n,t} \end{pmatrix} = \sum_{k=1}^{p} \begin{pmatrix} a_{k,1,1}, \cdots, a_{k,1,n} \\ a_{k,2,1}, \cdots, a_{k,2,n} \\ \vdots \\ a_{k,n,1}, \cdots, a_{k,n,n} \end{pmatrix} \cdot \begin{pmatrix} x_{1,t-k} \\ x_{2,t-k} \\ \vdots \\ x_{n,t-k} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \vdots \\ \epsilon_{n,t} \end{pmatrix}, or$$

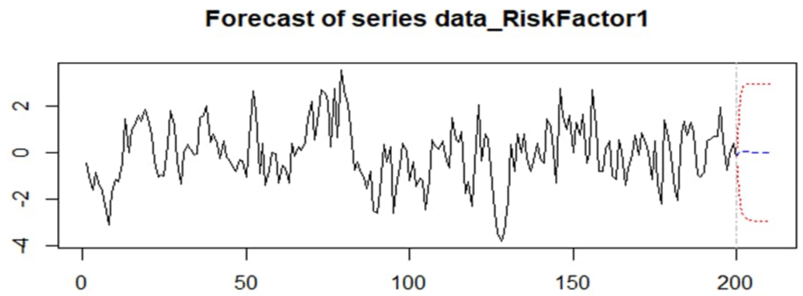
$$\overrightarrow{x_t} = \sum_{k=1}^{p} A_k \cdot \overrightarrow{x_{t-k}} + \overrightarrow{\epsilon_t}$$

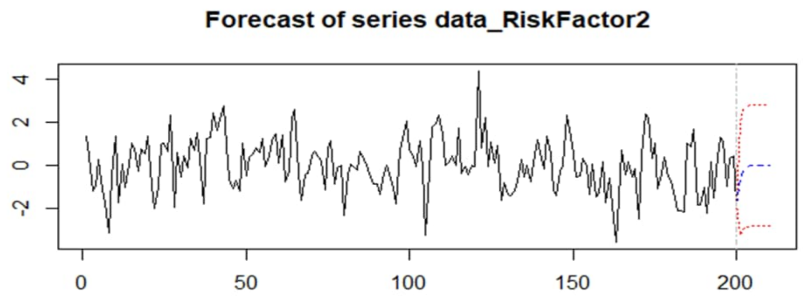
  - $\overrightarrow{x_t}$ is stationary, 1st and 2nd moments are independent of t; $\overrightarrow{\epsilon_t} \sim N(0, \Sigma)$
  - Interestingly, VARMA is rarely considered (?)

*: Federal Reserve Board Publication
**: Current Expected Credit Losses (CECL) | FDIC.gov
***: LDTIwhitePaper.3.23.pdf

Multivariate Time Series (cont')

- VAR
  - Identify dependencies of $(x_{1,t}, x_{2,t}, \cdots, x_{n,t})$
    - Granger Causality Test: null hypothesis - $x_{2,t}$ does not "Granger-cause"* $x_{1,t}$
  - Estimation method
    - Find max p; ensure $\#\ of\ observations\ >> n \cdot p_{max}$
    - Fit a series of VAR models for each candidate w/ MLE (use package)
    - Residual diagnostics





*: Granger Causality: A Review and Recent Advances - PMC

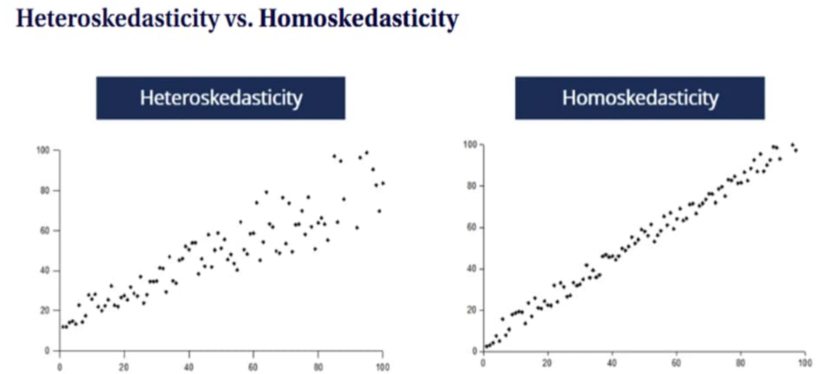
ARCH and GARCH

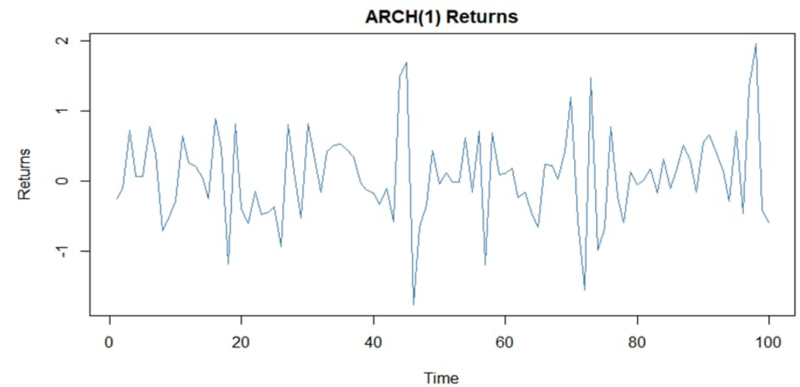
- Why AutoRegressive Conditional Heteroscedasticity (ARCH)
  - Heteroscedasticity*: volatility clustering, big moves are followed by big moves; fat tail
  - Note: AR model for conditional mean, ARCH model for conditional variance
  - [Robert F. Engle - Wikipedia]
- An ARCH model, $ARCH(q)$, is defined as follows,

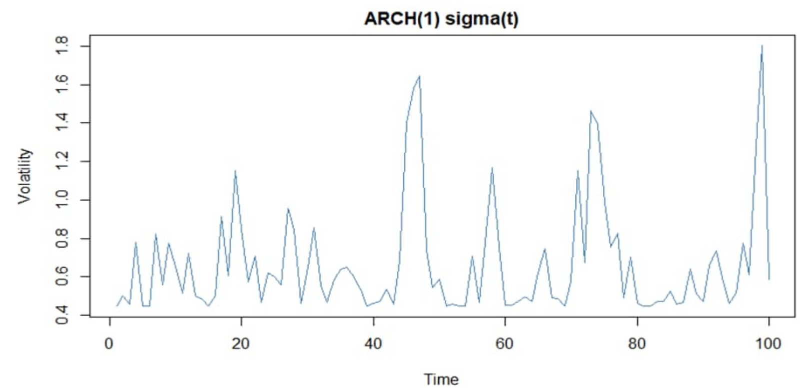
$$x_t = \sigma_t \cdot \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{k=1}^{q} \alpha_k \cdot x_{t-k}^2, \epsilon_t \sim N(0,1)$$

  - $E(x_t | x_{t-1}, x_{t-2}, \ldots) = 0$; $Cov(x_t, x_{t-1}) = 0$
  - $Var(x_t | x_{t-1}, x_{t-2}, \ldots) = \sigma_t^2 = \alpha_0 + \sum_{k=1}^{q} \alpha_k \cdot x_{t-k}^2$
  - $x_t^2 \sim AR(q)$ (?)
  - Unconditional variance $Var(x_t)$ =?
- Estimation methods
  - Find p thru ACF/PACF of x_t^2
  - Fit model w/ MLE
  - Residual diagnostics



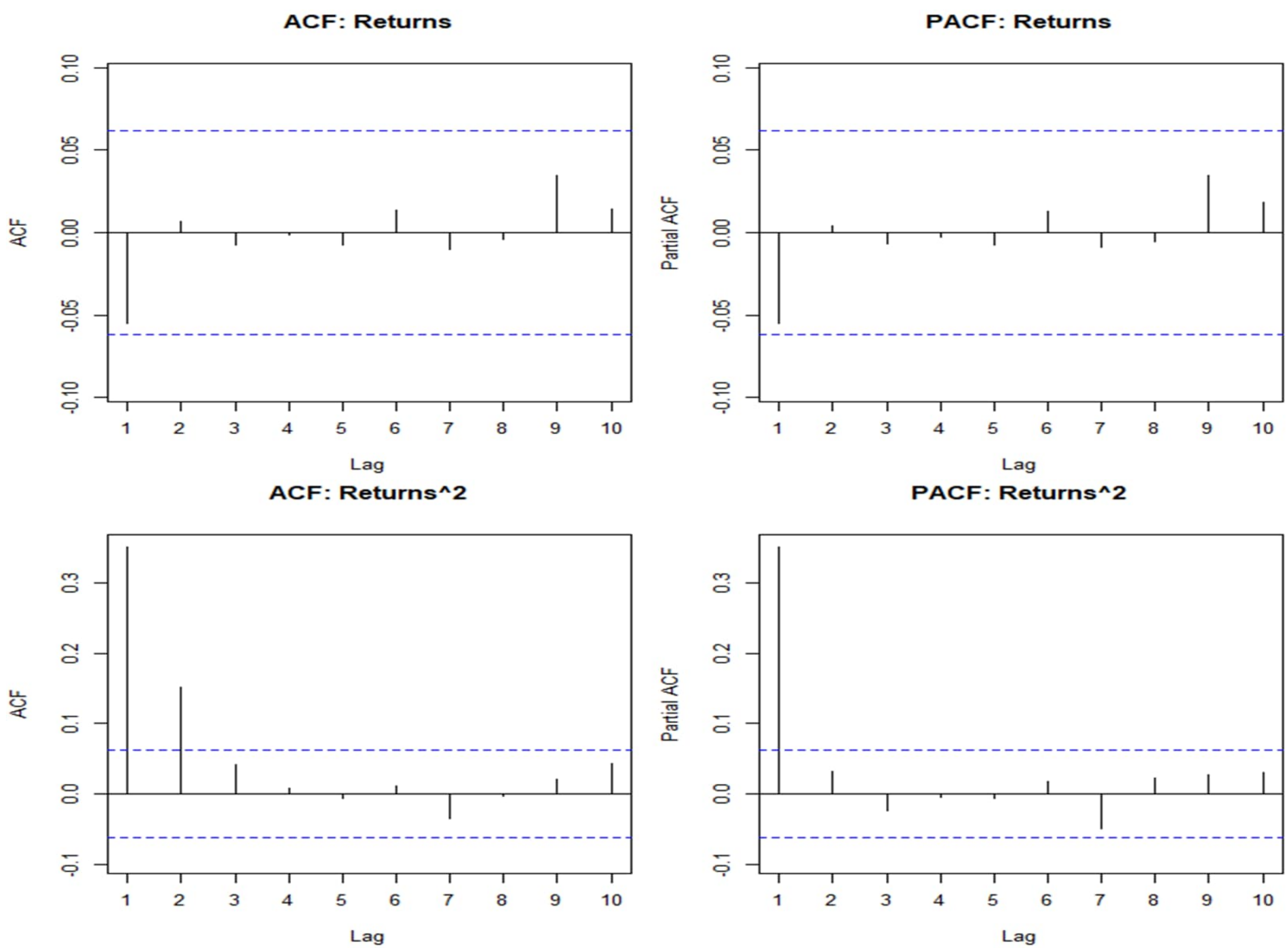




*: Heteroskedasticity - Overview, Causes and Real-World Example

ARCH and GARCH (con't)

- ACF/PACF
  - Decay ACF and single spike in PACF for $ARCH(1)$ - x_t^2



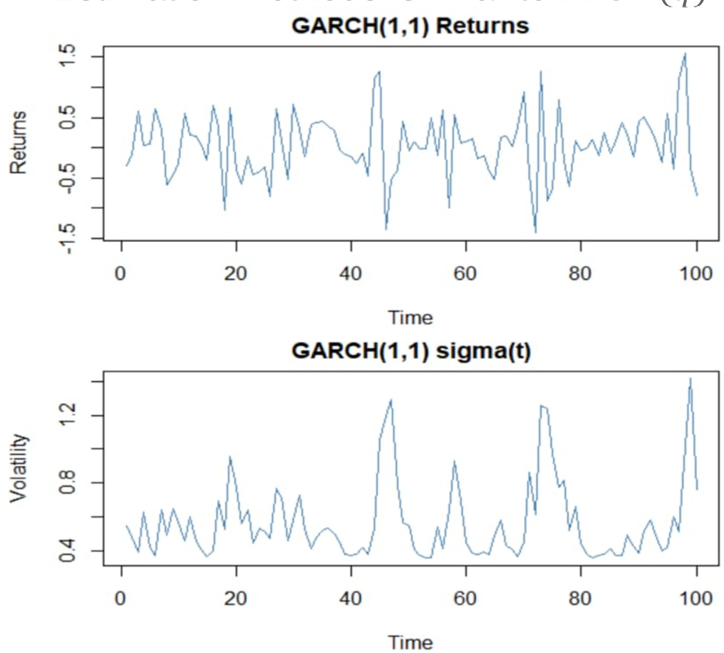
ARCH and GARCH (con't)

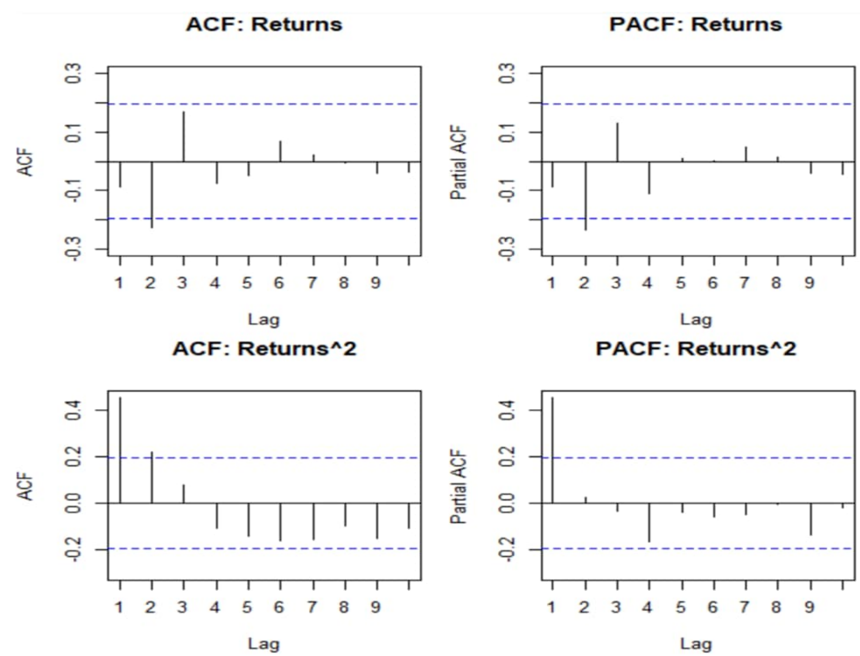
- A Generalized ARCH (GARCH) model, $GARCH(p,q)$, is defined as follows,

$$x_t = \sigma_t \cdot \epsilon_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot x_{t-i}^2 + \sum_{j=1}^{p} \beta_j \cdot \sigma_{t-j}^2 \quad \epsilon_t \sim N(0,1)$$

  - $E(x_t | x_{t-1}, x_{t-2}, \dots) = 0$; $Cov(x_t, x_{t-1}) = 0$
  - $Var(x_t | x_{t-1}, x_{t-2}, \dots) = \sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot x_{t-i}^2 + \sum_{j=1}^{p} \beta_j \cdot \sigma_{t-j}^2$
  - $x_t^2 \sim ARMA(q,p)$ (?)
  - Unconditional variance $Var(x_t)$ =?
- Estimation methods: similar to $ARCH(q)$





ARCH and GARCH (con't)

- Variation of GARCH model*
  - Exponential GARCH (EGARCH)

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{q} g(x_{t-i}) + \sum_{j=1}^{p} \beta_j \cdot \log\left(\sigma_{t-j}^2\right); \;\; \epsilon_t \sim N(0,1)$$

e.g., $g(x_{t-i}) = \alpha_i \cdot \left(\left|\frac{x_{t-i}}{\sigma_{t-i}}\right| - E\left|\frac{x_{t-i}}{\sigma_{t-i}}\right|\right) + \gamma_i \cdot \frac{x_{t-i}}{\sigma_{t-i}}$

  - Integrated GARCH (IGARCH)

$$\sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j = 1$$

  - Glosten-Jagannathan-Runkle GARCH (GJR-GARCH)

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \cdot x_{t-i}^2 + \sum_{i=1}^{q} \gamma_i \cdot x_{t-i}^2 \cdot I_{t-i} + \sum_{j=1}^{p} \beta_j \cdot \sigma_{t-j}^2$$

$$I_{t-i} = 1 \; if \; \epsilon_{t-i} < 0; 0 \; if \epsilon_{t-i} \geq 0$$

  - Others

*: Amazon.com: Analysis of Financial Time Series: 9780470414354: Tsay, Ruey S.: Books

Extreme Value Theory

- Motivation
  - *how to model events that have **never or rarely** happened but could be **catastrophic***
  - Tail risk: risk of high impact events that lie in the tails of a probability distribution
- **Extreme Value Theory**: describe tail behavior of a distribution and **estimate likelihood/outcome** of extreme events
- Fisher–Tippett–Gnedenko (FTG) theorem*: Let $X_1, \ldots, X_n$ be iid random variables, $M_n =$ max$\{X_1, \ldots, X_n\}$. Assume $\exists(a_n, b_n)$, such that

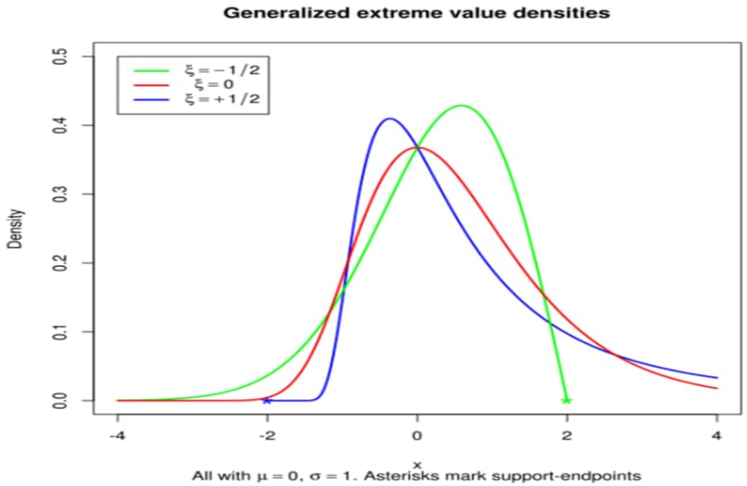
$$\lim_n P\left(\frac{M_n - b_n}{a_n} \leq x\right) = G(x)$$

Then $G(x)$ is the CDF of a dist. belonging to Weibull/Gumbel/Frechet distribution – Generalized Extreme Value distribution (GEV):

$$f(x; \mu, \sigma, \xi) = \frac{1}{2} t(x)^{\xi+1} \cdot e^{-t(x)}$$

where $t(x) = \begin{cases} \left[1 + \xi \cdot \left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}, & \xi \neq 0 \\ e^{-\frac{x-\mu}{\sigma}}, & \xi = 0 \end{cases}$

- μ: location; σ: scale
- ξ: shape
  - $\xi < 0$, Weibull (bounded tail)
  - $\xi = 0$, Gumbel (exponential tail)
  - $\xi > 0$, Frechet (heavy tail)



*: Fisher-Tippett-Gnedenko Theorem: Generalizing Three Types of Extreme Value Distributions | Wolfram Demonstrations Project

Extreme Value Theory (cont')

- Block Maxima approach – GEV model estimation
  - Divide a time series (stationary?) into non-overlapping blocks
  - Select maximum (or minimum) from each block, series of block maxima
  - Fit GEV distribution using MLE to the series of block maxima
  - Run diagnostic testing (e.g., QQ plot, goodness-of-fit tests – Kolmogorov - Smirnov)
- Usages: operational risk*
  - EVT was considered in the following areas Market risk (VaR/ES), Credit risk (e.g., loss forecasting)
- Others
  - Pickands–Balkema–De Haan Theroem

$$F_u(y) = P(X - \mu \leq y | X > \mu) \sim G(y; k, \sigma) - Generalized\ Pareto\ Distribution\ (GPD)$$

$$g(y; \beta, \xi) = \begin{cases} \frac{1}{\beta} \cdot \left[1 + \xi \cdot \frac{y}{\beta}\right]^{-\left(\frac{1}{\xi}+1\right)}, y \geq 0\ if\ \xi > 0; 0 \leq t \leq -\frac{\beta}{\xi}\ if\ \xi < 0;\ \xi \neq 0 \\ \frac{1}{\beta} e^{-\frac{y}{\beta}}, y \geq 0\,; \qquad \xi = 0 \end{cases}$$

    - ξ: shape parameter (see FTG theorem)
  - Peaks Over Threshold (POT) Approach
    - Choose a high threshold, μ; calculate exceedances, $X_i - \mu$
    - Fit GPD distribution using MLE to the series of exceedances
    - Run diagnostic testing

*: Extreme value theory for operational risk in insurance: a case study - Journal of Operational Risk

Error Correction Model

- Motivation
  - Prevent spurious regression on levels (non-stationary)
  - Combination of short-run (SR) dynamics and long-run (LR) equilibrium
    - Cointegration, linear combination of two $I(1)$ times series is stationary ($I(0)$)
  - Examples: spot vs future prices, OIS vs UST, bond yields vs CDS spreads, …
- Error Correction Model (ECM)
  - x_t, y_t are $I(1)$ time series
  - Long-run equilibrium: $y_t^E \sim \alpha + \beta x_t^E$; co-integration
  - Short-run dynamics: $\Delta y_t \sim \gamma_1 \Delta x_t$
  - ECM:

$$\Delta y_t = \gamma_0 + \gamma_1 \Delta x_t - \lambda(y_{t-1} - \alpha - \beta x_{t-1}) + \epsilon_t$$

    - $y_{t-1} - \alpha - \beta x_{t-1}$: error-correction term
    - Derivation:

$$y_t = \alpha + \gamma_1 x_t + \gamma_2 x_{t-1} + \mu y_{t-1} + \epsilon_t$$

$$\Longrightarrow \Delta y_t = \alpha + \gamma_1 \Delta x_t + (\gamma_1 + \gamma_2) x_{t-1} + (\mu - 1) y_{t-1} + \epsilon_t$$

$$\Longrightarrow \Delta y_t = \alpha' + \gamma_1 \Delta x_t - (1-\mu)\left(y_{t-1} - \alpha - \frac{\gamma_1 + \gamma_2}{1-\mu} x_{t-1}\right) + \epsilon_t$$

  - Key assumptions
    - Each series (y_t, x_t) is $I(1)$, i.e., $\Delta y_t, \Delta x_t$ are stationary
    - Co-integration exist, i.e., $\varepsilon_t = y_t - \alpha - \beta x_t$ is stationary

Error Correction Model (cont’)

- Error Correction Model
  - Estimation methods
    - Engle-Granger: bivariate or 2 variables
      - Estimate long-run relationship: estimate $y_t = \alpha + \beta x_t + \epsilon_t$ using MLE
      - Cointegration testing: run ADF on ϵ_t (using *Engle-Granger critical values*); $H0\!: residual\ has\ a\ unit\ root, or\ no\ cointegration$
      - Estimate ECM: estimate $\Delta y_t = \gamma_0 + \gamma_1 \Delta x_t - \lambda(y_{t-1} - \alpha - \beta x_{t-1}) + \epsilon_t$ using MLE
      - Residual diagnostics
    - Johansen: multivariate
- ECM vs ARIMA vs VAR

	ECM	ARIMA	VAR
Purpose	Model **LR equilibrium** + SR dynamics for **cointegrated non-stationary series (>=2)**	Model a **single** time series after **differencing** based on lags and errors	Model multivariate stationary time series
Time series	$I(1)$, **cointegrated**	One series, $I(d)$	All series stationary
LR relationship	Yes	No	No
Model interpretation	**Medium to High**, speed of adj. to LR equilibrium and SR dynamics	**Low**, often lack direct theoretical interpretation	**Medium**, system’s interdependencies and capture Granger-causality
Primary use cases	**Econometrics analysis**, observe a stable spread	**Forecasting**, short-term forecast when a theoretical model is absent	**Structural analysis**, interaction across diff variables
Estimation	EG or Johansen (for VECM)	MLE	MLE
Limitations	Require cointegration, sensitive to structure breaks	Cannot model interdependencies, lack LR equilibrium	Lack LR equilibrium, require many parameters

Current trends

- Recap
  - Key components of time series: trend, seasonality, noise
    - Cyclical component, irregularities
    - Irregularities:
      - Non-recurring (e.g., 9/11, 2020 pandemic): $D_{t=t_0} = 1, if\ t = t_0; 0, o/w$
      - Level shift (e.g., post-pandemic inflation): $S_{t\geq t_0} = 1, if\ t \geq t_0; 0, o/w$
  - AR/MA terms
  - Stationarity: log-return/differencing
- How about regime changes, non-linearity, ….
- Regime changes
  - Non-stationarity: mean/variance change over time
    - E.g., pre/post global financial crisis (2008) – low vs high vol, flat vs volatile basis
  - Implications of ignoring regime shifts
    - Poor out-of-sample forecasts
    - Underestimation of tail risk
  - Regine change (e.g., structural break) detection
    - A known break date - Chow test
      - $H_0: no\ strural\ break\ at\ T$, i.e., same reg coef apply to both subsamples $t \leq T$ and $t > T$
    - A unknown break date – Quandt-Andrews test (or sup-Wald test)
      - $H_0: no\ strural\ break\ over\ \ the\ whole\ sample$

Current trends (cont')

- Regime changes
  - Regine change (e.g., structural break) modeling
    - Time-Varying Parameter (TVP)*

$$y_t = \boldsymbol{x_t} \cdot \boldsymbol{\beta_t} + \epsilon_t$$

    - Markov-Switching models

$$y_t = \boldsymbol{x_t} \cdot \boldsymbol{\beta_{S_t}} + \epsilon_t$$

    Note: S_t follows a Markov chain

    - Others: observable-trigger regime models

- Machine Learning (ML) approaches for time series**
  - Pros/Cons:
    - Pros: nonlinearity, feature richness, can model distributional targets (e.g., quantile), automatic regime sensitivity,
    - Cons: explainability, overfitting, computational overhead
  - ML approaches
    - Decision trees (e.g., XGBoost, LightGBM), Random forest
    - Deep learning: CNN/RNN/…

*: Time-Varying Parameter Vector Autoregressions: Specication, Estimation, and an Application lubik.pdf
**: M5 accuracy competition: Results, findings, and conclusions - ScienceDirect