Chapter One

Fundamentals of Time Series

Tingguo Zheng (zhengtg@gmail.com)
Department of Statistics and Data Science, School of Economics, Xiamen University

(For Undergraduate Program, 2025-2026, Spring Semester)

- **Outline:**
  - Introduction to Time Series
  - Fundamental Moments
  - Stationarity
  - Sample Estimates of Correlation

1. Introduction to Time Series

1.1 What Is a Time Series?

Definition 1 (**Time Series).** A time series is a sequence of data points recorded at specific, ordered time intervals (e.g., hourly, daily, or yearly):

$$\{X_t\}_{t=1}^{T} = (X_1, X_2, \ldots, X_T)$$

where t denotes time, T is the sample size, and X_t is the value of the variable at time t.

It captures how a variable changes over time to analyze trends, seasonal patterns, and to forecast future values based on historical data, commonly used in finance and economics.

In economics and finance, time series appear everywhere:

- GDP, consumption, investment
- CPI, inflation, unemployment rate
- daily interest rates, daily exchange rates, daily stock returns
- monthly sales, quarterly earnings

Unlike cross-sectional data (which captures variation across units at one time), time series data captures variation across time for the same unit.

Time series data have *temporal dependence (时间相依性)*: observations close in time are often correlated.

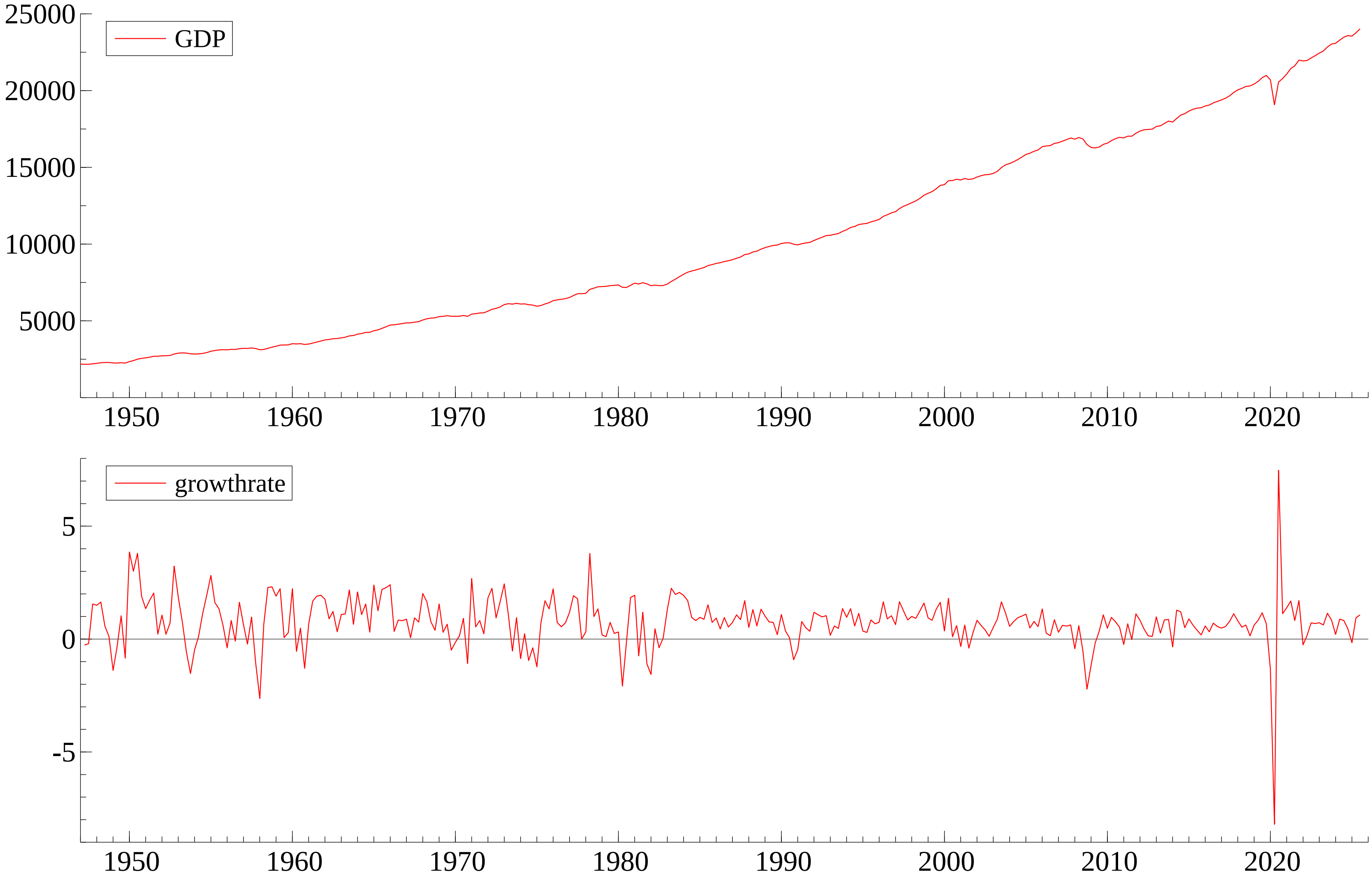


Figure 1 U.S. GDP data series sampled from 1947-2025

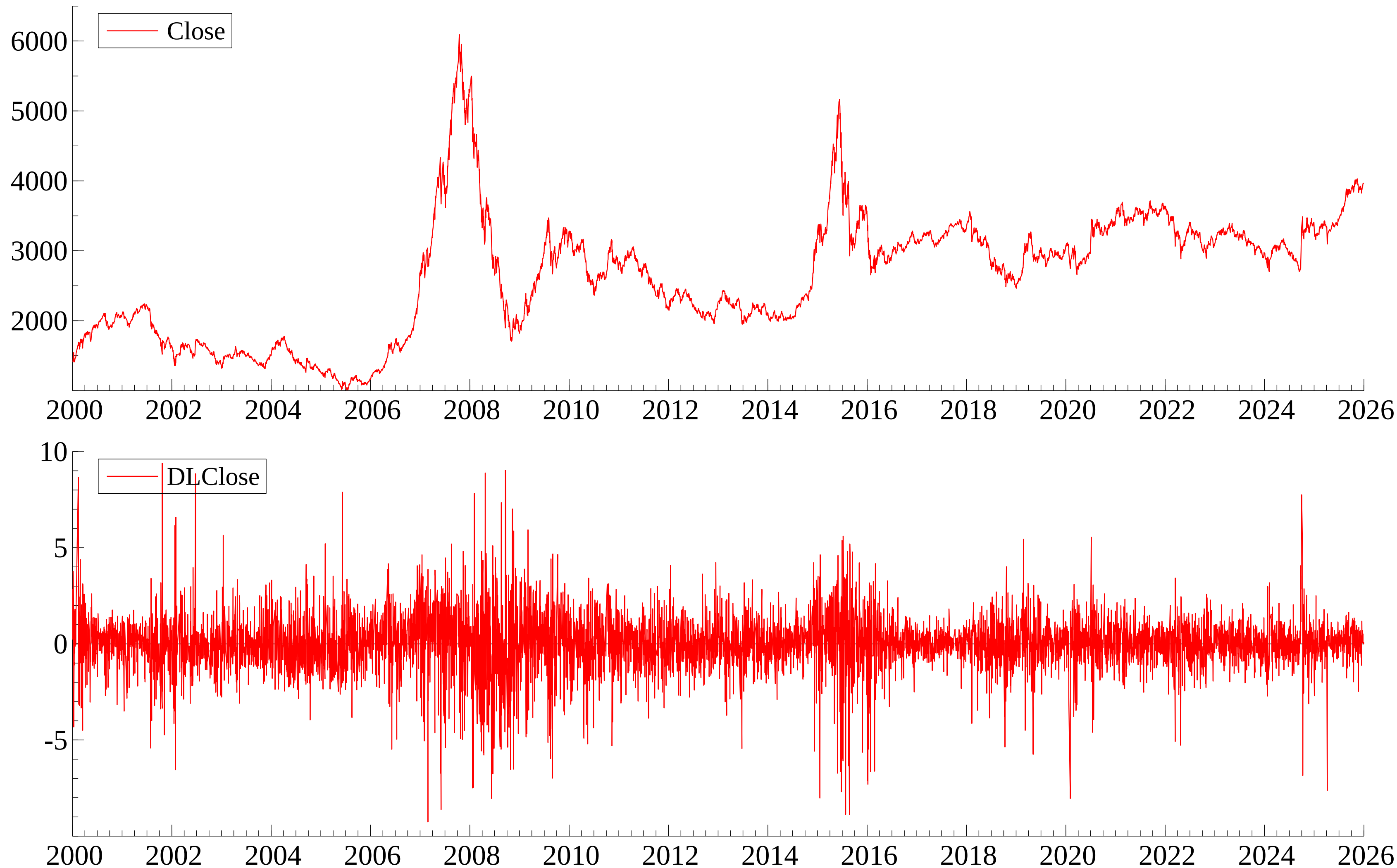


Figure 2 Shanghai Composite Index sampled from 2000/01/01-2025/12/31

1.2 Time Series vs. Stochastic Process

Definition 2 (**Stochastic Process).** A stochastic process is a *collection (集合)* of random variables $\{X_t : t \in T\}$ indexed by time T.

Core Relationship:

- A **time series** is a sequence of observations ordered in time.
- A **stochastic process** is a mathematical model consisting of an indexed collection of random variables.

A time series is typically treated as a single observed realization (sample path) of an underlying stochastic process.

(1) Gaussian White Noise

Gaussian white noise is a random process characterized by a normal (Gaussian) distribution with a zero mean value in the time domain, where the noise samples are independent and identically distributed with a constant variance.

Gaussian white noise is mathematically described by its probability density function (PDF):

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\},$$

where (m) is the mean and (σ^2) is the variance of the noise. Typically, the mean is zero, and the variance is denoted by (σ^2).

We sometimes denote this process as $w_t \sim iidN(0, \sigma_w^2)$.

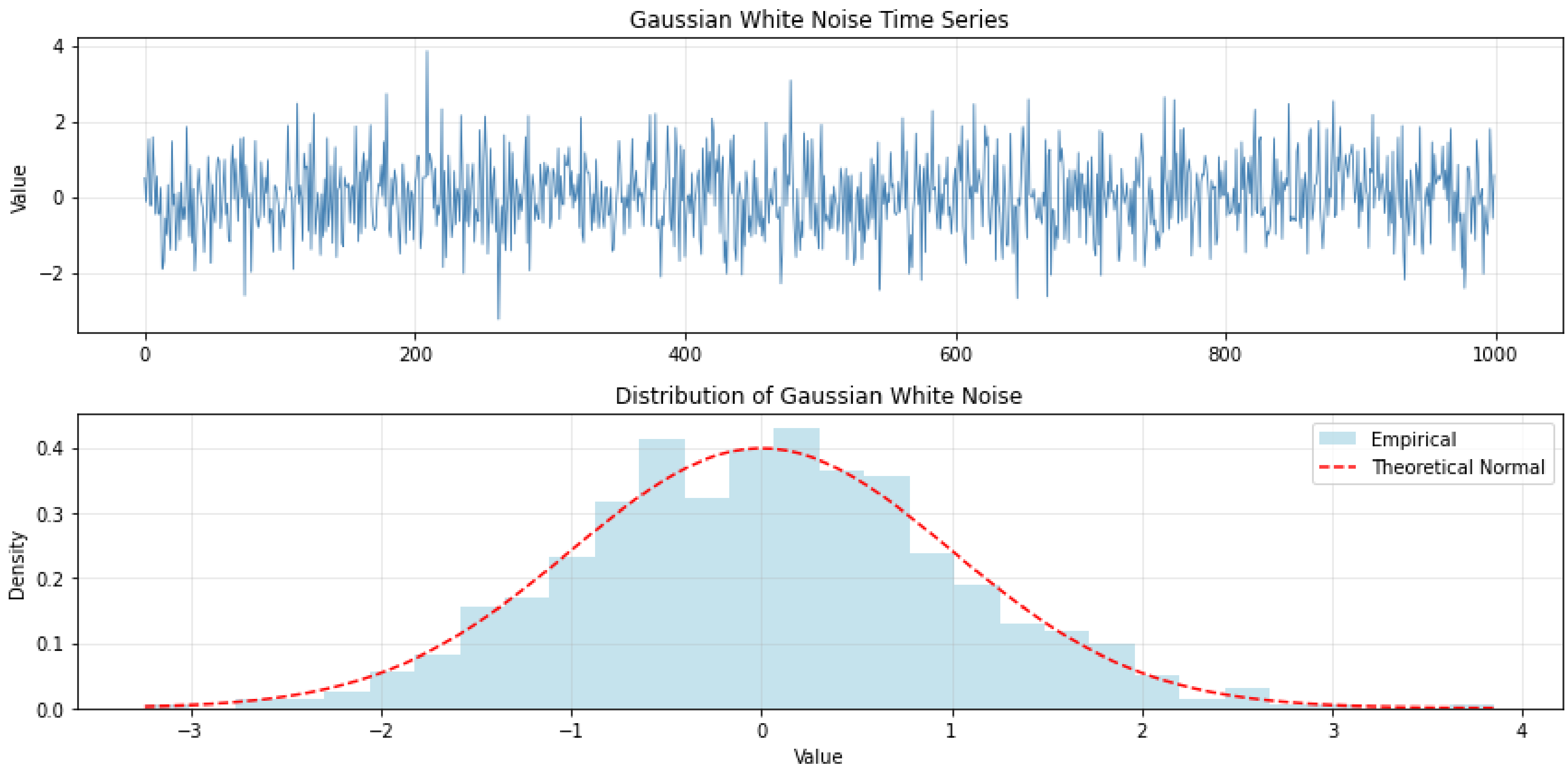


Figure 3 A time series of Gaussian White Noise: $T = 1000$

- 均值: 0.0193 (理论值: 0)
- 标准差: 0.9792 (理论值: 1)
- 是否正态分布(Shapiro-Wilk检验): p值= 0.6265

(2) Three Point Moving Average

We might replace the white noise series w_t by a moving average that smoothes the series. For example, consider replacing w_t in the previous example by an average of its current value and its immediate neighbors in the past and future. That is, let

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$$

which leads to the series shown in Figure 4.

Inspecting the series shows a smoother version of the first series, reflecting the fact that the slower *oscillations (振荡)* are more apparent and some of the faster oscillations are taken out.

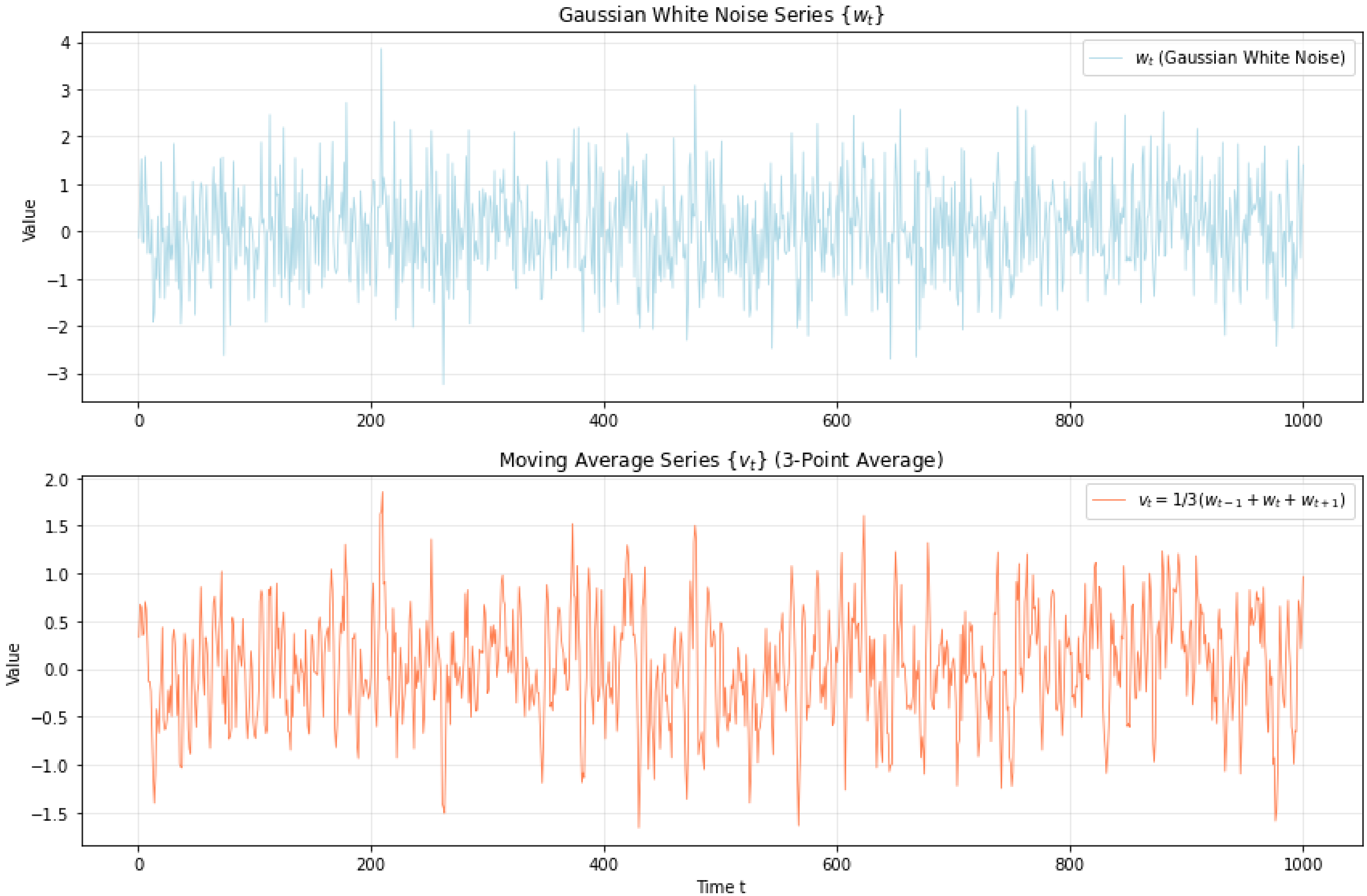


Figure 4 Gaussian white noise series (top) and three-point moving average of the Gaussian white noise series (bottom).

(3) Autoregressions

Suppose we consider the white noise series w_t as input and calculate the output using the second-order equation

$$x_t = x_{t-1} - 0.90x_{t-2} + w_t \tag{1}$$

successively for $t = 1, 2, \ldots, 1000$.

Equation (1) represents a regression or prediction of the current value x_t of a time series as a function of the past two values of the series, and, hence, the term autoregression is suggested for this model.

The resulting output series is shown in Figure 6, and we note the periodic behavior of the series.

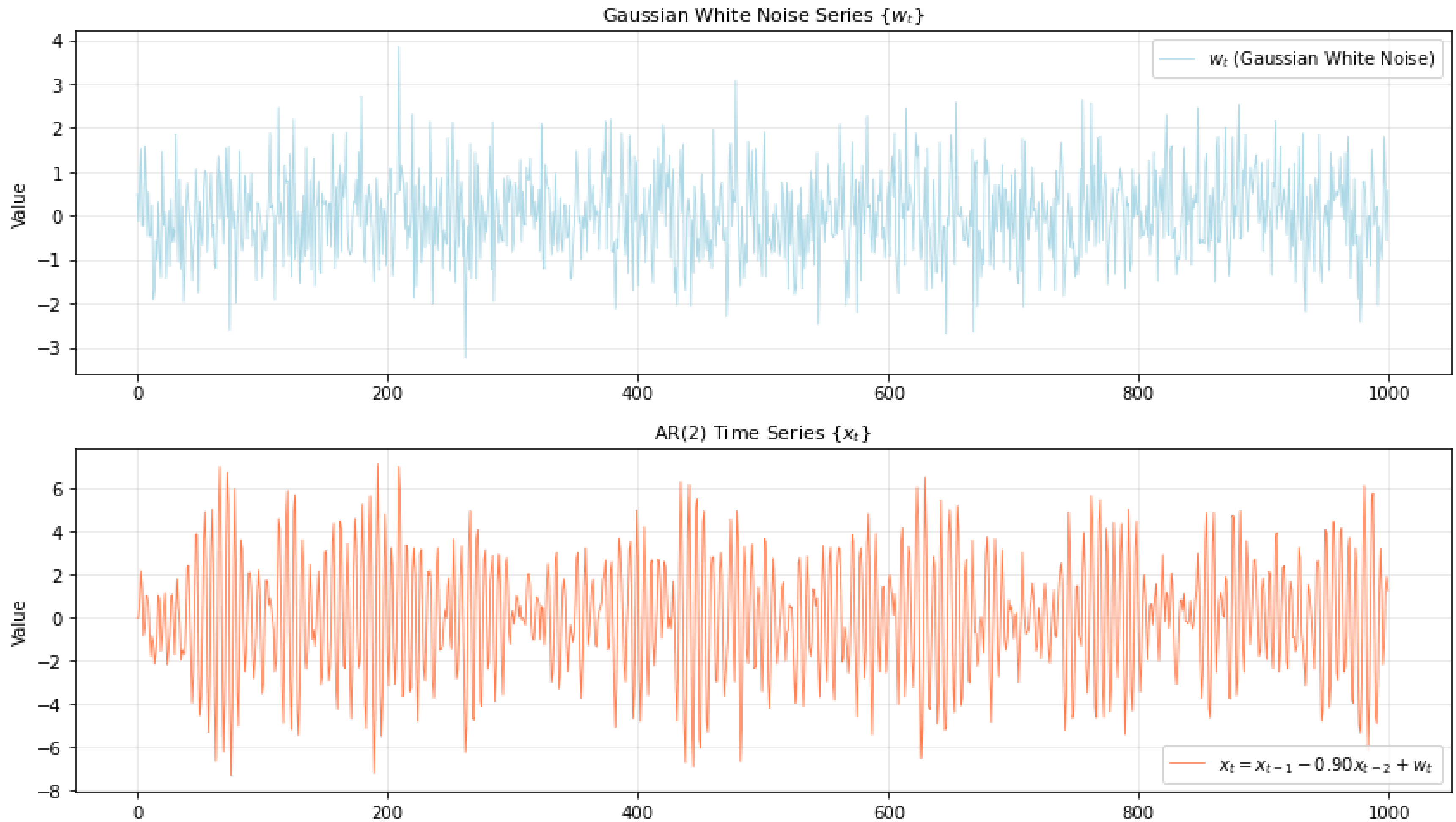


Figure 5 Autoregressive series generated from model (1).

(4) Random Walk

A model for analyzing trend is the random walk with drift given by

$$x_t = \delta + x_{t-1} + w_t \tag{2}$$

for $t = 1, 2, \ldots, 1000$, with initial condition $x_0 = 0$, and where w_t is white noise. The constant δ is called the drift, and when $\delta = 0$, (2) is called simply a random walk.

Note that we may rewrite (2) as a cumulative sum of white noise variates. That is,

$$x_t = \delta t + \sum_{j=1}^{t} w_j$$

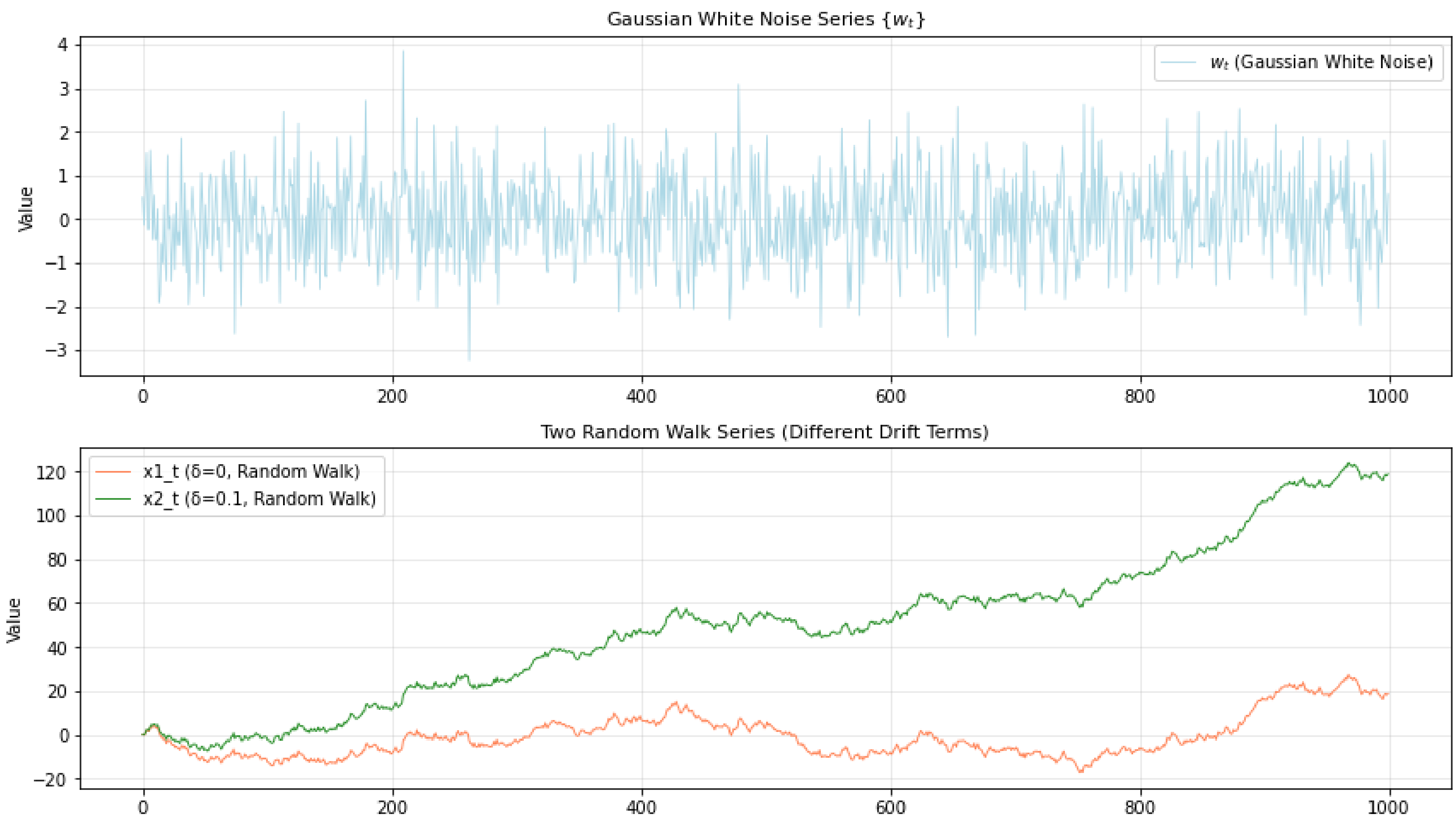


Figure 6 Random walk, $\sigma_w = 1$, with drift $\delta = 0$ (red line), without drift, and $\delta = 0.1$ (green line).

(5) Signal in Noise

Consider the model

$$x_t = 2\cos(2\pi t/50 + 0.6\pi) + w_t \tag{3}$$

for $t = 1, 2, \ldots, 300$, where the first term is regarded as the signal, shown in the upper panel of Figure 7.

We note that a sinusoidal waveform (正弦波形) can be written as

$$A\cos(2\pi\omega t + \phi)$$

where A is the *amplitude (振幅)*, ω is the *frequency of oscillation (振荡频率，$1/\omega$为周期)*, and ϕ is a *phase shift (相位偏移)*.

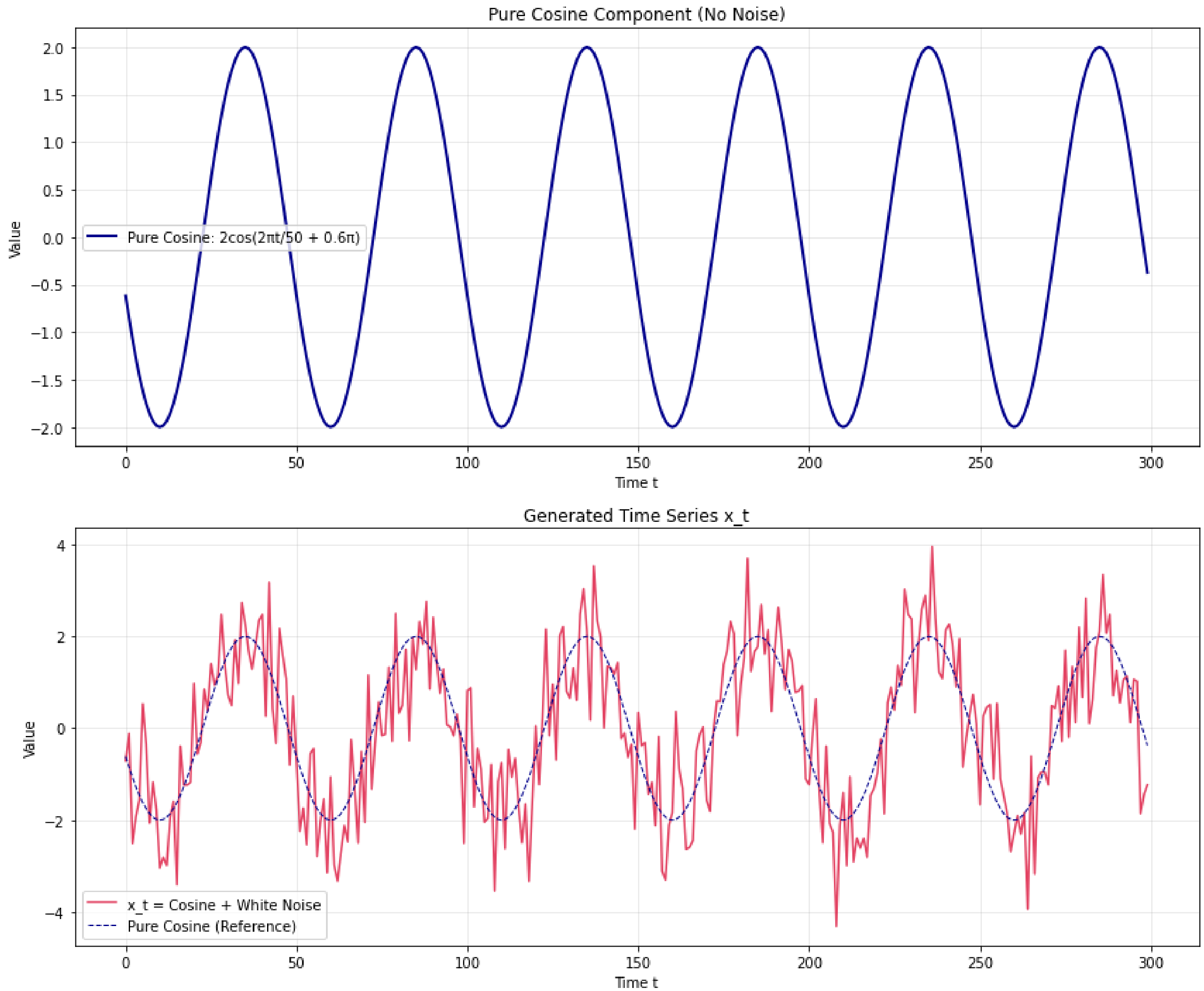


Figure 7 Pure cosine wave compared with the cosine wave contaminated with additive Gaussian white noise.

1.3 Goals of Time Series Analysis

- Description

  Describe patterns: trend, seasonality, cycles, volatility.

- Explanation

  Understand how past values affect current values.

- Prediction / Forecasting

  Predict future values using past information.

- Control / Policy Analysis

  Estimate dynamic effects of shocks or policies.

2. Fundamental Moments

2.1 Mean Function: The “Central Tendency” Over Time

Definition 3 (Mean Function). The mean function of a time series $\{X_t\}$ is the expected value of the random variable at each time point t:

$$\mu_t = \mathbb{E}[X_t] = \int_{-\infty}^{\infty} x \cdot f_t(x) dx$$

where $f_t(x)$ is the marginal probability density function (pdf) of X_t at time t.

- If $\mu_t = \mu$ (constant), the process has constant mean.
- In economics: trending series have μ_t changing over time.

Example 1 (Mean Function of a Random Walk with Drift)**.** Consider the random walk with drift model given in (2),

$$x_t = \delta t + \sum_{j=1}^{t} w_j, \quad t = 1, 2, \ldots$$

Because $E(w_t) = 0$ for all t, and δ is a constant, we have

$$\mu_{xt} = E(x_t) = \delta t + \sum_{j=1}^{t} E(w_j) = \delta t$$

Example of a Trending Mean Function:

U.S. Real GDP (1947 – 2023) has μ_t increasing over time because of population growth, technological progress, and capital accumulation.

A simple linear trend model for μ_t might be:

$$\mu_t = \alpha + \beta t$$

where $\beta > 0$ (positive growth rate).

2.2 Variance Function: Volatility Over Time

Definition 4 (Variance Function). The variance function measures the dispersion (散度) of X_t around its mean μ_t:

$$\sigma_t^2 = \text{Var}(X_t) = \mathbb{E}\left[(X_t - \mu_t)^2\right] = \mathbb{E}[X_t^2] - (\mathbb{E}[X_t])^2$$

Key Properties:

- Non-negativity: $\sigma_t^2 \geq 0$ (variance cannot be negative).
- Units: Squared units of X_t (e.g., if X_t is GDP in billions of dollars, σ_t^2 is (billions of dollars)^2).
- Standard deviation (more interpretable): $\sigma_t = \sqrt{\sigma_t^2}$ (same units as X_t).

Economic Volatility: Time-Varying Variance in Practice

Most economic and financial time series have heteroskedastic variance (time-varying volatility):

Example 2 (Stock Returns (Financial Volatility))**.** S&P 500 daily returns have σ_t^2 that spikes during crises (e.g., 2008 Financial Crisis, 2020 COVID Crash) and is low during calm periods (e.g., 2017 “low-volatility” year).

Why? Investors react to uncertainty: during crises, fear leads to large price swings (high variance); during stability, trades are more predictable (low variance).

Example 3 (Macroeconomic Volatility (Great Moderation)). U.S. GDP growth variance fell sharply after the 1980s (the ”Great Moderation”) due to better monetary policy, reduced oil shocks, and more flexible labor markets—then rose again during the 2008 crisis and 2020 pandemic.

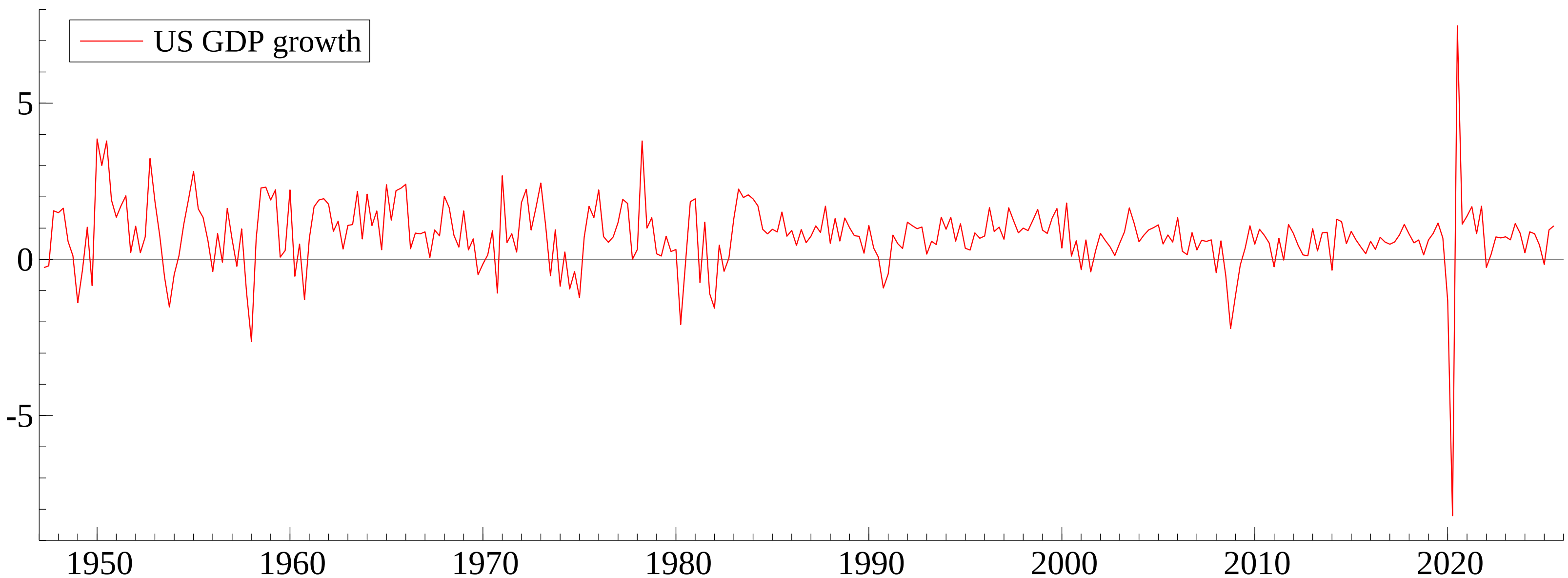


Figure 8 US GDP growth sampled from 1947-2025.

Consequence of Ignoring Time-Varying Variance

If you assume $\sigma_t^2 = \sigma^2$ (homoskedasticity) for a volatile series (e.g., stock returns), your model will underestimate risk during crises and overestimate it during calm periods.

This is critical for:

- Risk management (e.g., Value-at-Risk calculations for banks).
- Forecasting (e.g., inflation volatility affects central bank policy decisions).

2.3 Autocovariance Function (ACVF)

Definition 5. The autocovariance function (ACVF) measures the linear dependence between X_t (value at time t) and X_{t+k} (value at time $t+k$), where k is the lag (time gap between observations):

$$\gamma(t, t+k) = \text{Cov}(X_t, X_{t+k}) = \mathbb{E}\left[(X_t - \mu_t)(X_{t+k} - \mu_{t+k})\right]$$

- Positive if X_t above μ_t tends to be followed by X_{t+k} above μ_{t+k} (positive dependence).
- Negative if X_t above μ_t tends to be followed by X_{t+k} below μ_{t+k} (negative dependence).
- Zero if no linear relationship (no autocorrelation).

Stationary vs. Non-Stationary ACVF

For a weakly stationary process, the ACVF depends only on lag k, not on the specific time t —we simplify notation to $\gamma(k)$.

Example 4 (White Noise (Zero Autocovariance))**.** A white noise process has $\gamma(k) = 0$ for all $k \neq 0$: No temporal dependence—each observation is independent of the previous ones.

For non-stationary processes, the ACVF depends on both t and k:
Example 5. Trending GDP: $\gamma(t, 1)$ (covariance between GDP at t and $t + 1$) increases over time because GDP levels are higher in later years, so their deviations from the trending mean are larger.

Key Properties of ACVF (for Stationary Processes)

- $\gamma(0) = \text{Var}(X_t)$: Lag-0 autocovariance is the variance of the series (dependence between X_t and itself).

- Symmetry:

$$\gamma(k) = \gamma(-k)$$

  Dependence between X_t and X_{t+k} is the same as between X_{t+k} and X_t (time reversibility).

- Cauchy-Schwarz Inequality:

$$|\gamma(k)| \leq \gamma(0)$$

  The autocovariance at any lag k cannot exceed the variance.

2.4 Autocorrelation Function (ACF)

Definition 6. The autocorrelation function (ACF) normalizes the ACVF by the variance to eliminate units and provide a interpretable measure of dependence (between -1 and 1):

$$\rho_k := \rho(k) = \frac{\gamma(k)}{\gamma(0)}$$

- ACVF has units (e.g., (billions of dollars)^2 for GDP), making it hard to compare dependence across different series.

- ACF is unitless: A $\rho(1) = 0.8$ means the same strength of lag-1 dependence for GDP, inflation, or stock returns.

Key Properties of ACF (for Stationary Processes)

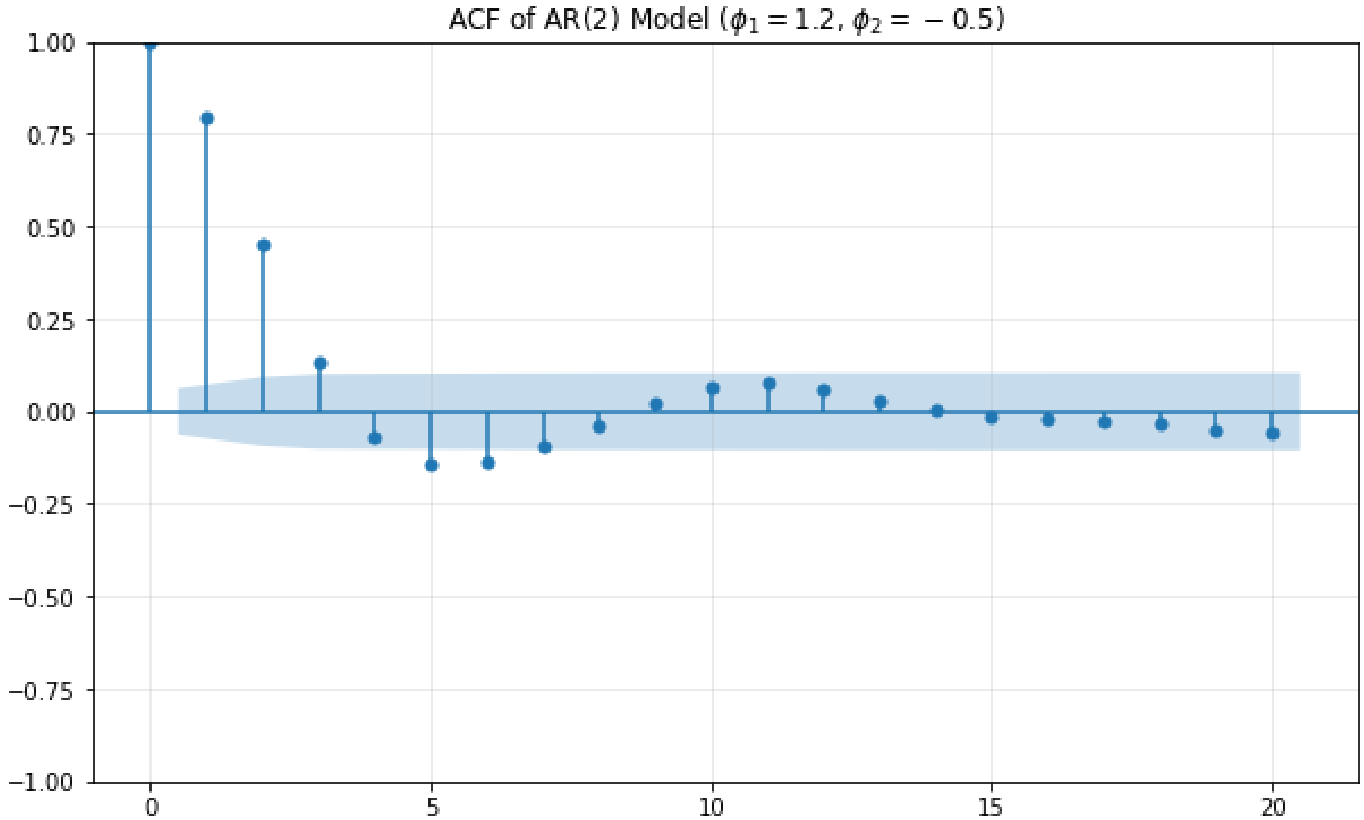
- $\rho(0) = 1$: Perfect correlation between X_t and itself.
- Bounded range: $-1 \leq \rho(k) \leq 1$:
- $\rho(k) = 1$: Perfect positive dependence (e.g., a constant series: $X_t = c$ for all t).
- $\rho(k) = -1$: Perfect negative dependence (e.g., $X_{t+1} = -X_t$).
- $\rho(k) = 0$: No linear dependence.
- Symmetry: $\rho(k) = \rho(-k)$ (same as ACVF).

ACF Decay Patterns

The shape of the ACF plot (vs. lag k) is critical for identifying time series models later (AR/MA/ARMA). Key patterns:

- Slow Decay (e.g., $\rho(k) = 0.9^k$):
  - Interpretation: Past values have a long-lasting effect on current values.
- Abrupt Cutoff:
  - Example: MA(1) process (e.g., $X_t = \epsilon_t + 0.5\epsilon_{t-1}$, where ϵ_t is white noise).
  - ACF is non-zero only at lag 1 $(\rho(1) = 0.5/(1+0.25) = 0.4)$ and zero for $k \geq 2$.

- Oscillating Decay:
  - AR(2) DGP: $X_t = 1.2X_{t-1} - 0.5X_{t-2} + \epsilon_t$.
  - ACF alternates positive/negative as it decays.



2.5 Partial Autocorrelation Function (PACF)

Definition 7. The partial autocorrelation function (PACF) at lag k, denoted ϕ_{kk}, measures the direct linear dependence between X_t and X_{t-k} after removing the linear effects of all intermediate lags $(X_{t-1}, X_{t-2}, ..., X_{t-k+1})$.

Mathematical Derivation (Linear Regression Approach):

(S1) Regress X_t on $X_{t-1}, X_{t-2}, \ldots, X_{t-k+1}$:

$$X_t = \hat{\beta}_1 X_{t-1} + \hat{\beta}_2 X_{t-2} + \cdots + \hat{\beta}_{k-1} X_{t-k+1} + \hat{e}_t$$

where $\hat{e}_t$ is the residual (variation in X_t not explained by lags 1 to $k-1$).

(S2) Regress X_{t-k} on $X_{t-1}, X_{t-2}, \ldots, X_{t-k+1}$:

$$X_{t-k} = \hat{\delta}_1 X_{t-1} + \hat{\delta}_2 X_{t-2} + \cdots + \hat{\delta}_{k-1} X_{t-k+1} + \hat{u}_t$$

where $\hat{u}_t$ is the residual (variation in X_{t-k} not explained by lags 1 to $k-1$).

(S3) The PACF ϕ_{kk} is the correlation between $\hat{e}_t$ and $\hat{u}_t$:

$$\phi_{kk} = \mathrm{Corr}(\hat{e}_t, \hat{u}_t)$$

Simple Case: Lag-1 PACF (ϕ_{11})

- No intermediate lags, so $\phi_{11} = \rho(1)$ (same as lag-1 ACF).

Simple Case: Lag-2 PACF (ϕ_{22})

- Removes the effect of X_{t-1} (the intermediate lag) from both X_t and X_{t-2}, then measures their correlation.

- Formula (Yule-Walker Equation):

$$\phi_{22} = \frac{\rho(2) - \rho(1)^2}{1 - \rho^2(1)}$$

3. Stationarity

3.1 Why Stationarity Is the “Backbone” of Time Series Analysis

Stationarity is the most critical concept in classical time series analysis (AR/MA/ARMA/ARIMA).

Without stationarity, the fundamental properties of the series (mean, variance, autocorrelation) change over time—making it impossible to reliably estimate models, forecast future values, or draw valid inferences.

Stationarity in Economic Context

Most raw economic time series are non-stationary:

- GDP (levels) grows over time (trending mean).
- Stock prices tend to rise long-term (trending mean) and have volatile periods (changing variance).

However, many transformed economic series are stationary:

- GDP growth rates (percent change in GDP).
- Stock returns (percent change in stock prices).

3.2 Types of Stationarity: Strict vs. Weak

Stationarity has two key definitions —strict (strong) and weak (covariance) —with strict stationarity being a stronger condition.

Definition 8 (**Strict (Strong) Stationarity).** A stochastic process $\{X_t\}$ is strictly stationary if for any integer $m \geq 1$, any time indices $t_1 < \cdots < t_m$, and any time shift h (integer), the joint probability distribution of $(X_{t_1}, X_{t_2}, ..., X_{t_m})$ is identical to the joint distribution of $(X_{t_1+h}, X_{t_2+h}, ..., X_{t_m+h})$:

$$F_{X_{t_1},X_{t_2},...,X_{t_m}}(x_1, x_2, ..., x_m) = F_{X_{t_1+h},X_{t_2+h},...,X_{t_m+h}}(x_1, x_2, ..., x_m)$$

for all $x_1, x_2, \ldots, x_m \in \mathbb{R}$, where F denotes the joint cumulative distribution function (CDF).

Key Implications

- All marginal distributions are identical (e.g., X_t has the same CDF as X_{t+h} for any t, h).

- All joint distributions depend only on the time gaps between observations, not their absolute positions in time.

- A white noise process with i.i.d. normal random variables ($X_t \sim N(0, \sigma^2)$) is strictly stationary: every observation is independent and has the same normal distribution, so joint distributions are invariant to time shifts.

Definition 9 (**Weak (Covariance) Stationarity).** A stochastic process $\{X_t\}$ is weakly stationary (or covariance stationary) if it satisfies three conditions:

- Constant mean: $\mathbb{E}[X_t] = \mu$ for all t (mean does not depend on time).
- Constant variance: Var$(X_t) = \gamma(0)$ for all t (variance is finite and time-invariant).
- Time-invariant autocovariance: Cov$(X_t, X_{t+k}) = \gamma(k)$ for all t, k (autocovariance depends only on lag k, not on the absolute time t).

Stationarity of a Moving Average

The three-point moving average process $v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$ with $w_t \sim WN(0, \sigma_w^2)$ ($\sigma_w = 1$) is stationary because we may write the autocovariance function and the ACF as

$$\gamma(k) = \begin{cases} 3/9, & \text{if } k = 0; \\ 2/9, & \text{if } k = \pm 1; \\ 1/9, & \text{if } k = \pm 2; \\ 0, & \text{if } |k| \geq 3. \end{cases}$$

$$\rho(k) = \begin{cases} 1, & \text{if } k = 0; \\ 2/3, & \text{if } k = \pm 1; \\ 1/3, & \text{if } k = \pm 2; \\ 0, & \text{if } |k| \geq 3. \end{cases}$$

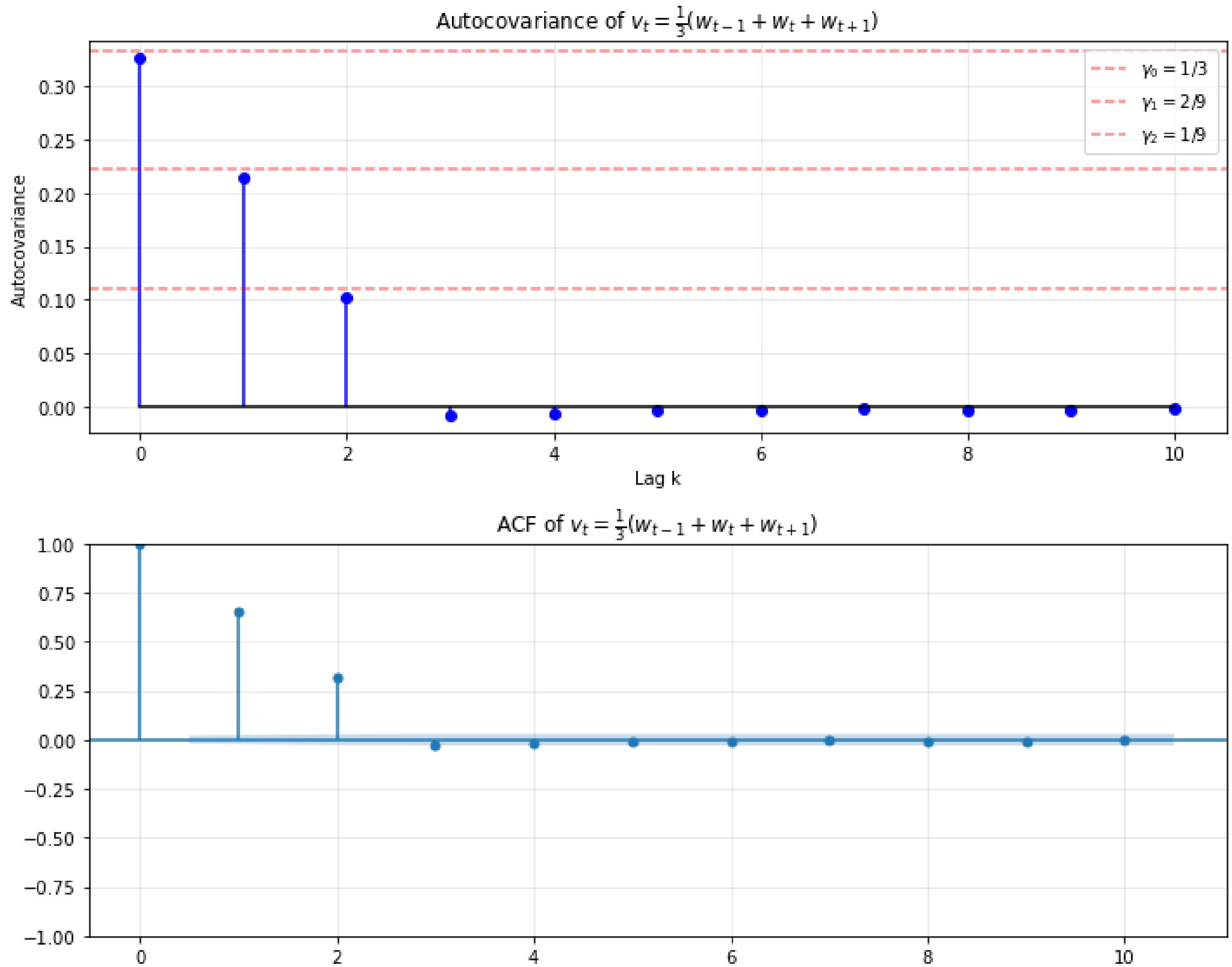


Figure 9 Autocovariance function and ACF of a three-point moving average

Critical Notes on Weak Stationarity

Weak stationarity does not require identical marginal distributions—only that first and second moments (mean, variance, autocovariance) are time-invariant.

For normal processes (all joint distributions are normal), weak stationarity implies strict stationarity (since normal distributions are fully characterized by first and second moments).

All models in this course (AR, MA, ARMA, ARIMA) require weak stationarity.

Relationship Between Strict and Weak Stationarity

(1) Strict stationarity $\Rightarrow$ **Weak stationarity** (if first and second moments exist and are finite).

Proof: If joint distributions are invariant to time shifts, then marginal distributions are also invariant—so mean and variance are constant. Autocovariance depends only on lag (since Cov(X_t, X_{t+k}) is a function of the joint distribution of X_t and X_{t+k}, which is invariant to t).

(2) Weak stationarity $\nRightarrow$ **Strict stationarity** (unless the process is normal).

4. Sample Estimates of Correlation

4.1 Sample Mean

For observed data $x_1, x_2, ..., x_T$, the sample mean is:

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

- **Unbiasedness:** $\mathbb{E}[\bar{x}] = \mu$ (population mean) for any T, assuming the process is stationary (constant mean μ).
  Proof: $\mathbb{E}[\bar{x}] = \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}[x_t] = \frac{1}{T} \sum_{t=1}^{T} \mu = \mu$.

- **Consistency:** $\bar{x} \xrightarrow{p} \mu$ as $T \to \infty$ (sample mean converges in probability to population mean), by the Law of Large Numbers (LLN) for stationary ergodic processes.

Definition 10 (Ergodicity). A stationary stochastic process $\{X_t\}$ with $E|X_t| \leq \infty$ is *ergodic* if and only if, for almost every sample path, the time average (时间平均) converges almost surely to the ensemble mean (总体均值):

$$\frac{1}{T}\sum_{t=1}^{T} X_t \xrightarrow{a.s.} E[X_t] \quad \text{as} \quad T \to \infty$$

A weakly stationary process $\{X_t\}$ with mean μ, autocovariance function $\gamma(k)$, and $\sum_{k=-\infty}^{\infty} |\gamma(k)| < \infty$ is *mean ergodic* (ergodic in the mean) if its sample mean satisfies

$$\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t \xrightarrow{p} \mu \quad \text{as} \quad T \to \infty.$$

Ergodic Process with $x_t = 2\cos(2\pi t/50 + 0.6\pi) + w_t$

Time averages at $T = 10000$ ===

Path 1: -0.0096 (converges to 0)

Path 2: 0.0043 (converges to 0)

Path 3: -0.0058 (converges to 0)

Path 4: 0.0155 (converges to 0)

Path 5: -0.0308 (converges to 0)

Non-ergodic Process with $x_t = Z$, where $Z \sim N(0,1)$

Time averages at $T = 10000$ ===

Path 1: -1.0374 (random value, not 0)

Path 2: -0.6893 (random value, not 0)

Path 3: -0.4901 (random value, not 0)

Path 4: -1.4919 (random value, not 0)

Path 5: 1.2487 (random value, not 0)

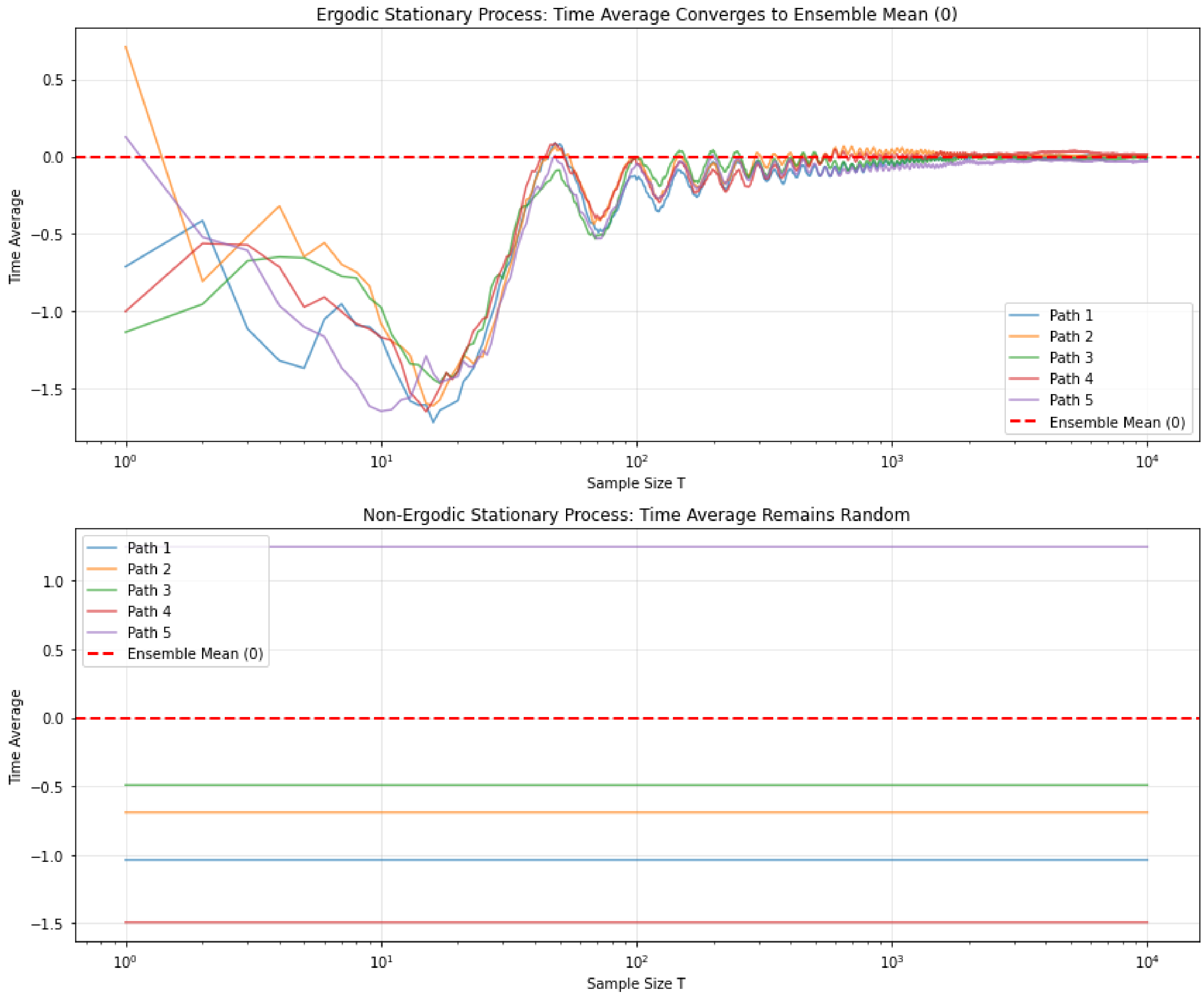


Figure 10 Time average of Ergodic Process and Non-ergodic Process

Finite Sample Bias for Non-Stationary Series

If the series is non-stationary (e.g., trending GDP levels), $\mathbb{E}[x_t] = \mu_t$ (time-varying mean), so:

$$\mathbb{E}[\bar{x}] = \frac{1}{T}\sum_{t=1}^{T}\mu_t \neq \mu_t \quad \text{for any } t$$

This is why stationarity is required for valid sample mean (and thus ACF/PACF) estimation—non-stationary series must be transformed first.

4.2 Sample Autocovariance Function (ACVF)

The population ACVF is $\gamma(k) = \mathbb{E}[(X_t - \mu)(X_{t+k} - \mu)]$. For sample estimation, we replace μ with $\bar{x}$, but the choice of denominator (T vs. $T - k$) leads to two estimators:

- Biased Estimator (Most Common)

$$\hat{\gamma}(k) = \frac{1}{T} \sum_{t=1}^{T-k} (x_t - \bar{x})(x_{t+k} - \bar{x})$$

- Unbiased Estimator

$$\hat{\gamma}_{\text{unbiased}}(k) = \frac{1}{T-k} \sum_{t=1}^{T-k} (x_t - \bar{x})(x_{t+k} - \bar{x})$$

Key Properties of Sample ACVF

- Symmetry: $\hat{\gamma}(k) = \hat{\gamma}(-k)$ (since $(x_t - \bar{x})(x_{t+k} - \bar{x}) = (x_{t+k} - \bar{x})(x_t - \bar{x})$).

- $\hat{\gamma}(0)$ = Sample variance (if using T denominator) or unbiased sample variance (if using $T-1$ denominator).

- Note: Some textbooks define sample variance as

$$s^2 = \frac{1}{T-1} \sum_{t=1}^{T} (x_t - \bar{x})^2$$

  (unbiased for σ^2). In this case, $\hat{\gamma}(0) = \frac{T-1}{T} s^2$ (biased but consistent).

- Boundedness: $|\hat{\gamma}(k)| \leq \hat{\gamma}(0)$ (by Cauchy-Schwarz inequality for sample moments).

4.3 Sample Autocorrelation Function (ACF)

Definition 11. The sample ACF is the normalized sample ACVF—dividing by $\hat{\gamma}(0)$ eliminates units and bounds the estimate between -1 and 1:

$$\hat{\rho}(k) = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} = \frac{\sum_{t=1}^{T-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{T}(x_t - \bar{x})^2}$$

For finite T, $\hat{\rho}(k)$ is biased—the bias arises from two sources:

- Replacement of μ with $\bar{x}$ (introduces sampling error).
- Finite sample size (the denominator $\hat{\gamma}(0)$ is itself a biased estimate of $\gamma(0)$).

Confidence Intervals for Sample ACF

To judge if a sample ACF $\hat{\rho}(k)$ is statistically significant (i.e., different from 0), we use confidence intervals (CIs). The key assumption is that under the null hypothesis of white noise ($\rho(k) = 0$ for all $k \neq 0$), the sample ACF is approximately normally distributed for large T.

Result: Under iid assumption (Fuller, 1976):

$$\hat{\rho}(k) \xrightarrow{d} N\left(0, \frac{1}{T}\right) \quad \text{as } T \to \infty$$

- Standard Error (SE): SE$(\hat{\rho}(k)) \approx \frac{1}{\sqrt{T}}$.
- 95% Confidence Interval: $\hat{\rho}(k) \pm 2 \times \frac{1}{\sqrt{T}}$ (using the 95% quantile of the standard normal distribution, $z_{0.025} = 1.96 \approx 2$).

Box-Pierce Q-statistic (Box and Pierce, 1970): Consider testing H_0 : $\rho_1 = \ldots = \rho_m = 0$. Under iid assumption

$$Q_m = T \sum_{s=1}^{m} \hat{\rho}_s^2 \xrightarrow{d} \chi_m^2. \tag{4}$$

Modified Q-statistic (Ljung and Box, 1978): (finite-sample correction) Under iid assumption

$$Q'_m = T(T+2) \sum_{s=1}^{m} \frac{\hat{\rho}_s^2}{T-s} \xrightarrow{d} \chi_m^2.$$

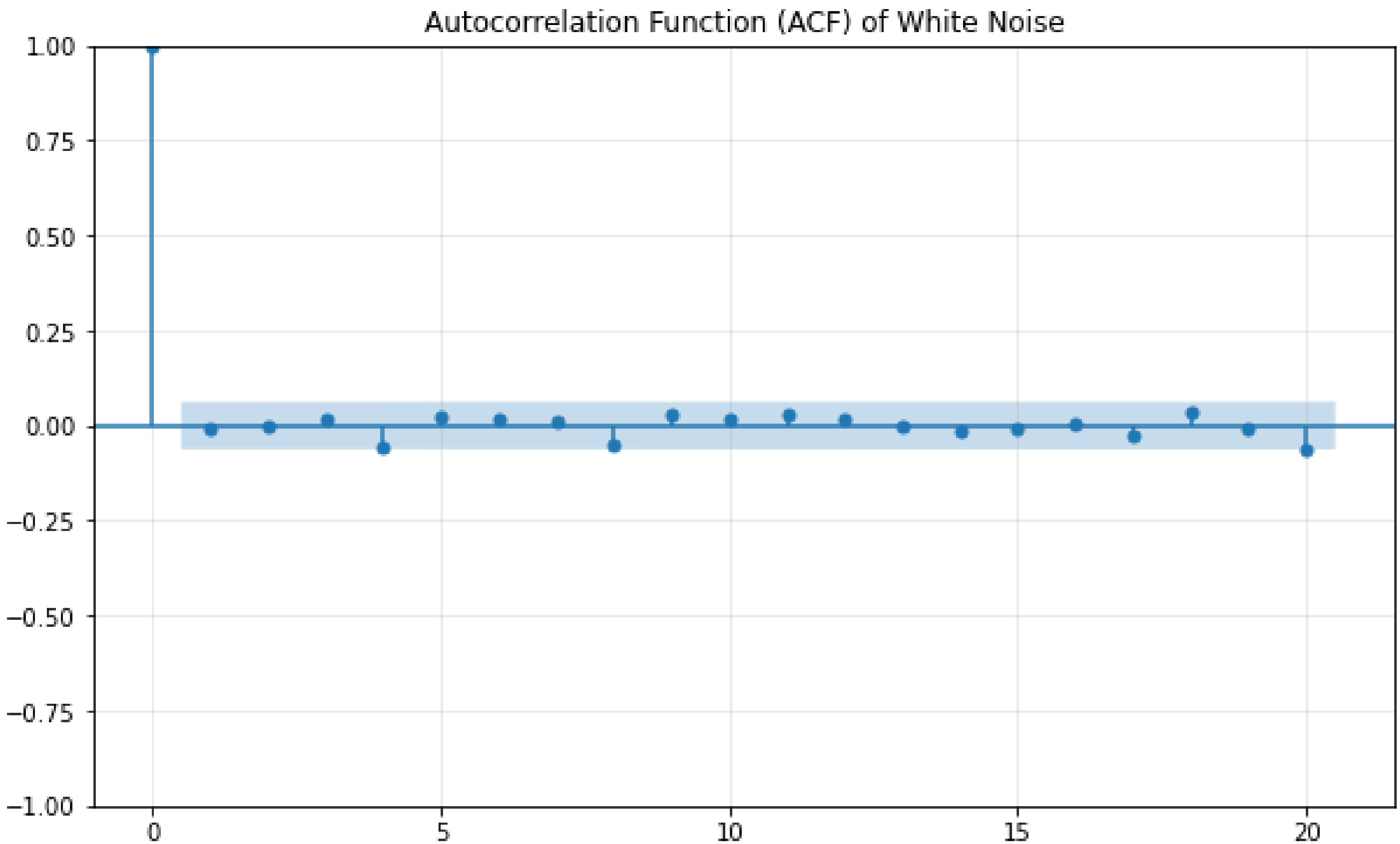


Figure 11 Autocorrelation function of a Gaussian White Noise series: $T = 1000$