743- Regression and Time Series

Mamikon S. Ginovyan

Time Series Analysis

1.Introduction:

A General Approach to Time Series Analysis

The objectives of time series analysis (TSA)

A time series refers to an **ordered sequence of observations**, where the ordering is in time.

- If $y_1, y_2, \ldots, y_T$ are observations on a RV Y_t taken at T equispaced successive points in time ($1, 2, \ldots, T$), then these observations, in general, are **not independent**.

- The **main objectives** of time series analysis are:

The objectives of time series analysis (TSA)

1. To understand the **nature of the dependence** of observations $y_1 , y_2 , \ldots, y_T$, and specification of a model which adequately describe the given data.

- **Thus,** an important part of the analysis of a time series is the selection of a **suitable probabilistic model** for the data.

2. Use the specified model to generate **forecasts (predictions)** for the future values.

The objectives of time series analysis

- To allow for the possibly unpredictable nature of future values it is natural to suppose that each observation $\boldsymbol{y_t}$ is a **realized value** of a certain RV $\boldsymbol{Y_t}$.

Definition 1.

A **time series model** for the observed data

$$y_1, y_2, \ldots, y_T = \{y_t, t = 1, 2, \ldots, T\}$$

is a specification of the joint probability distributions (or possibly only the means and covariances) of a sequence of RV's

$$\{Y_t, t = 1, 2, \ldots, T\},$$

of which $\{y_t, t = 1, 2, \ldots, T\}$ are postulated to be realizations.

The objectives of time series analysis

Remark 1.

We will use the term " time series" for both the data $\{\boldsymbol{y_t}\}$ and the process $\{\boldsymbol{Y_t}\}$.

Remark 2.

A complete probabilistic time series model for the sequence of RV's $\boldsymbol{Y_1}, \boldsymbol{Y_2}, \ldots,$ would specify **all of the joint distributions** of the random vectors

$$(Y_1, \ Y_2, \ldots, Y_T)', \quad T = 1, 2, \ldots,$$

or equivalently all of the probabilities

$$P(Y_1 \le y_1, \ldots, Y_T \le y_T), -\infty < y_1, \ldots, y_T < \infty, T = 1, 2, \ldots$$

The objectives of time series analysis

Such a specification is rarely used in time series analysis since

a)a complete specification is **difficult** to obtain;

b)even if it is possible to obtain this specification, in general, it will contain **too many parameters** to be estimated from the available data.

Instead we specify only the **first-and second-order moments of the joint distributions (second-order properties)**:

•the **expected values** $\mathbf{E}(Y_t)$, t = 1, 2, … ; and

•the **expected products** $\mathbf{E}[Y_s Y_t]$, $s, t = 1, 2, \ldots$.

The objectives of time series analysis

Remark 3.

In the special case where all the joint distributions are **multivariate normal** (in this case the process $\{Y_t\}$ is called **normal** or **Gaussian**), the **second-order properties** of $\{Y_t\}$ **completely determine the joint distributions**, and hence give a complete probabilistic characterization of the process $\{Y_t\}$.

The objectives of time series analysis

Remark 4.

In general, we will lose a certain amount of information by looking at time series **"through second-order spectacles"**;

however, we will see later that

the minimum mean squared error linear prediction problem requires only the second-order properties

of the underlying process $\boldsymbol{Y_t}$.

Time Series models: Stationary Model

Definition 1.

A time series model $\{y_t, t = 1, 2, \ldots\}$ is called **stationary time series model** if

1. Has a constant mean: Exhibits **mean reversion** in that it fluctuates around a **constant long run mean**.
2. Has a finite variance that is time invariant.
3. Has a covariance between values of y_t that depends only on the difference apart in time, that is,

$$E(y_t) = m_t = m$$

$$Var(y_t) \equiv E\left[(y_t - m)^2\right] = r(0) < \infty$$

$$Cov(y_t, y_s) \equiv E\left[(y_t - m)(y_s - m)\right] = r(t - s).$$

Time Series models: Non-Stationary Model

Definition 2.

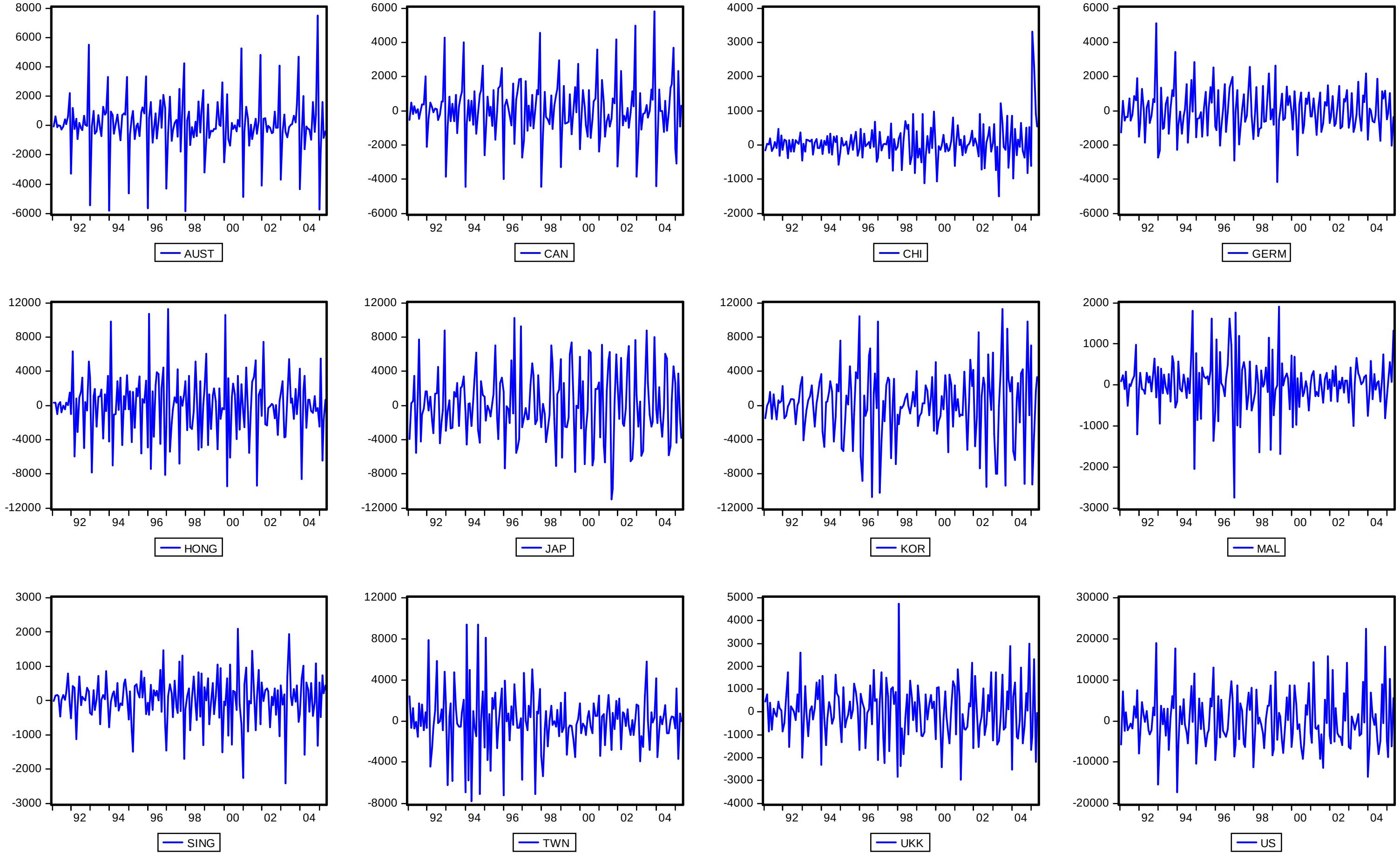
A time series model $\{y_t, t = 1, 2, \dots\}$ is called **non-stationary time series model** if at least one of the stationarity conditions is violated:

1. There is no long-run mean to which the series returns, and/or
2. The variance is time dependent and goes to infinity as time approaches to infinity

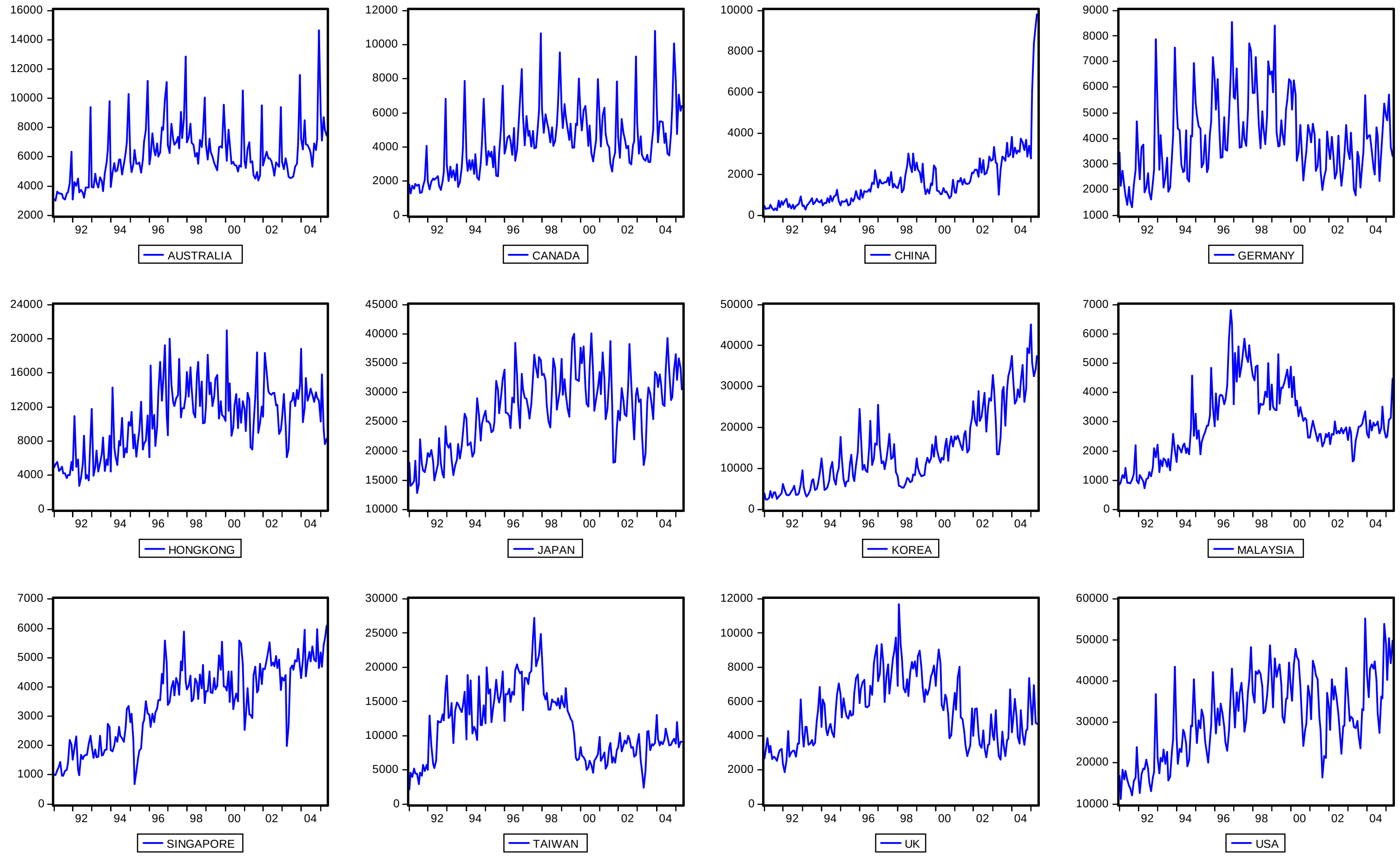
Remark.

We will consider only **stationary time series models**, the **Standard techniques are largely invalid where data is non-stationary.**

These are Examples of Stationary Time Series



These are Examples of Non-Stationary Time Series



Examples: **Some zero-mean models**

Example 1. (IID Noise).

The simplest stationary time series model is one in which:

a) the observations are **iid RV's** with zero mean, and

b) there is no trend or seasonal component.

Such defined sequence of RV's $\{\boldsymbol{Y_t = U_t}, \boldsymbol{t = 1, 2, \ldots}\}$ is called **IID Noise**.

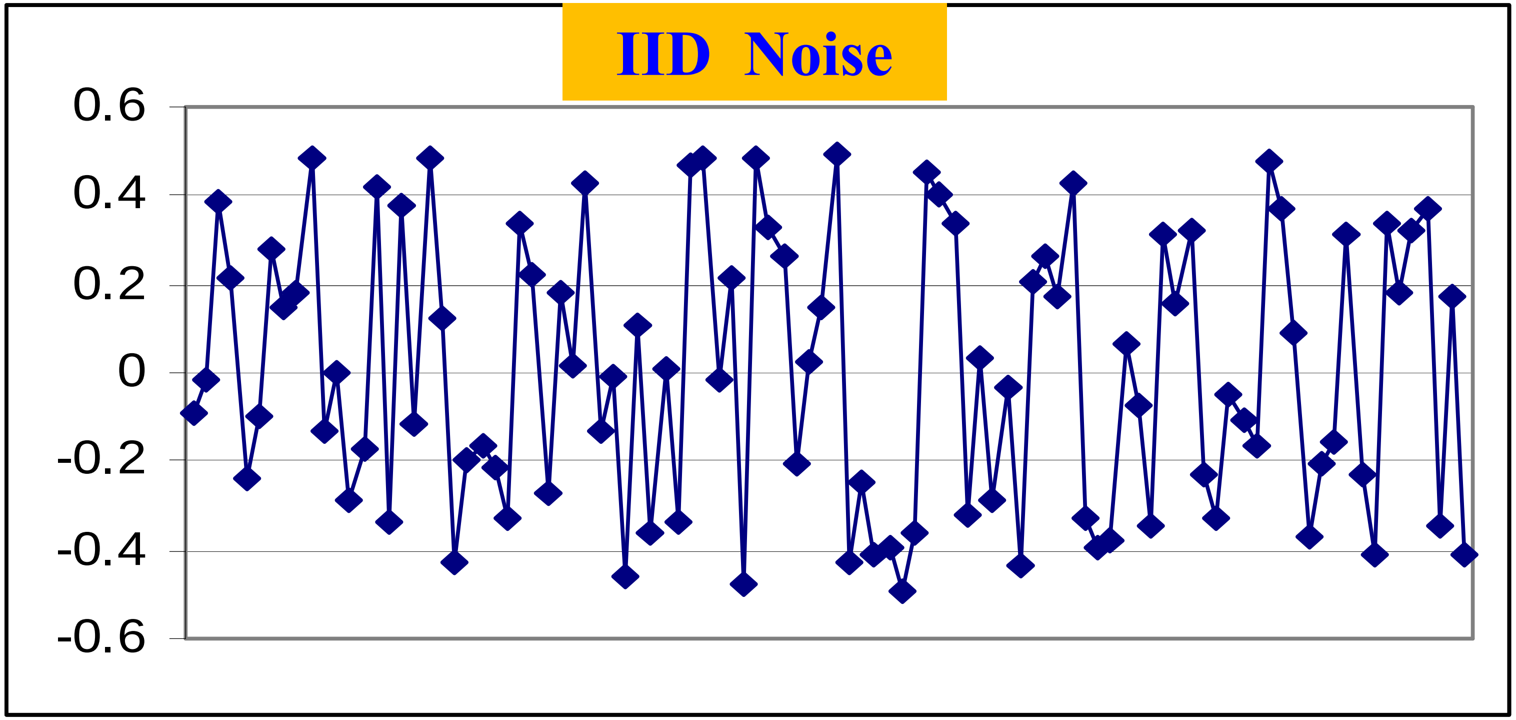
In this case, for any $\boldsymbol{T}$ and any real numbers $\boldsymbol{u_1, \ldots, u_T}$,

$$P(U_1 \le u_1, \ldots, U_T \le u_T) = P(U_1 \le u_1) \ldots P(U_T \le u_T) = \prod_{t=1}^{n} F(u_t),$$

where $\boldsymbol{F(.)}$ is the ***cdf*** of each ***iid*** RV's $\boldsymbol{U_1, U_2, \ldots}$.

Examples: Some zero-mean models

IID (WHITE) Noise: $Y_t = U_t = \varepsilon_t \sim IID(0, \sigma^2)$.



Examples: Some zero-mean models

Remark 1.

In IID- model there is **no dependence** between observations.

In particular, for all $h \geq 1$ and all $y_1, \ldots, y_T$,

$$P(Y_{T+h} \leq y \mid Y_1 = y_1, \ldots, Y_T = y_T) = P(Y_{T+h} \leq y),$$

showing that knowledge of $Y_1, \ldots, Y_T$ is of **no value for predicting the future** observation Y_{T+h}.

(Markovian Property).

Given the values $Y_1, \ldots, Y_T$, the function $f(.)$ minimizing **MSE**:

$$MSE = E[Y_{T+h} - f(Y_1, \ldots, Y_T)]^2 \equiv 0$$

is in fact **identically zero**.

Examples: Some zero-mean models

Remark 2.

Although the above considerations shows that IID Noise is a rather **uninteresting process for forecasters,** it plays an important role as a **building block** for more complicated (and now interesting) time series models.

Remark 3.

An example of IID Noise is a **binary process**, the sequence of ***iid*** RV's { $\boldsymbol{\varepsilon_t}$, $\boldsymbol{t = 1, 2, \ldots}$ } with

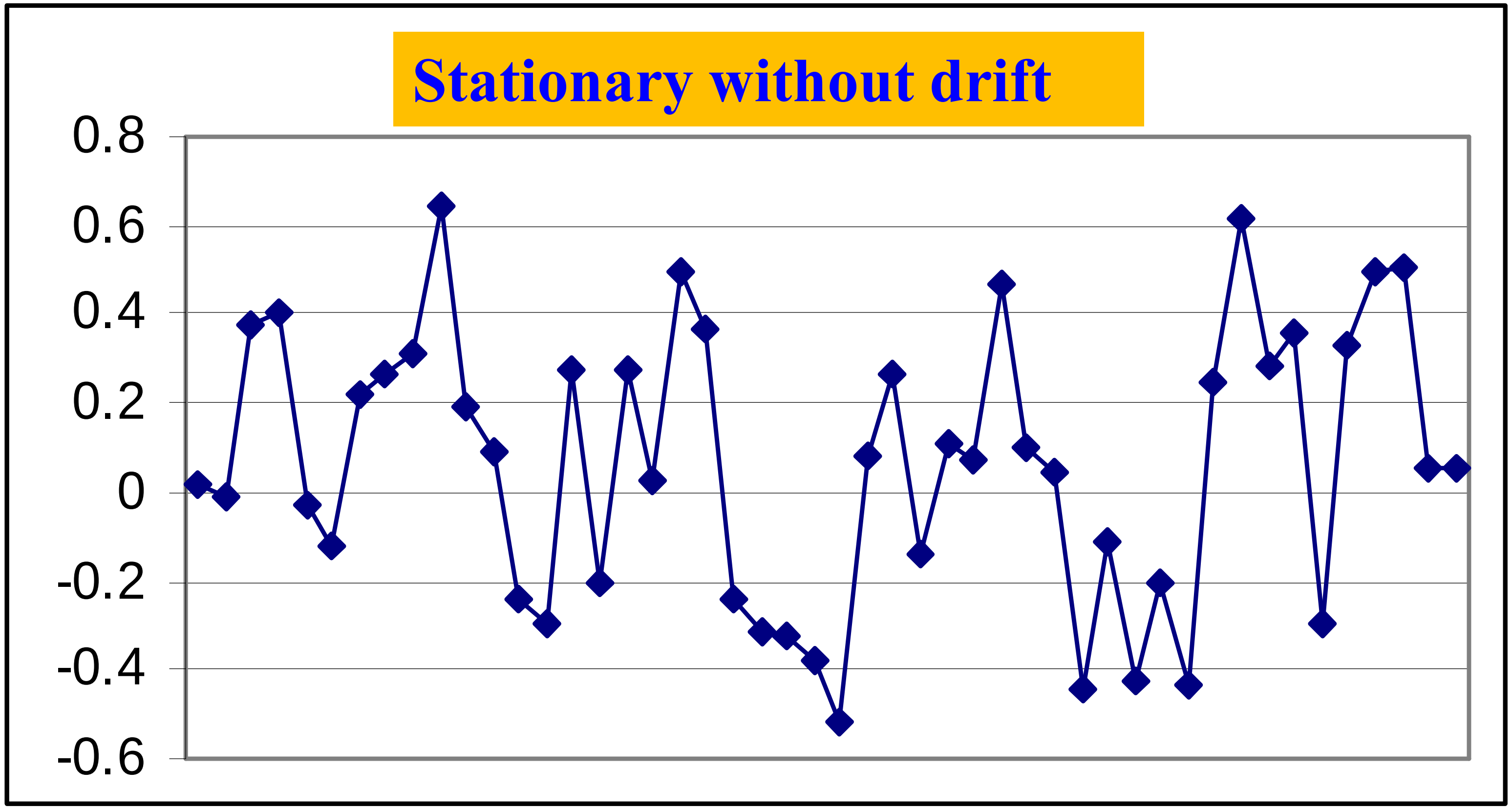
$$P(e_t = 1) = p, \quad P(e_t = -1) = 1 - p, \quad \text{where } \boldsymbol{p} = 1/2.$$

The time series obtained by tossing a fair coin repeatedly and scoring **+1** for each Head and **-1** for each Tail is modeled as a realization of this process.

Examples: Some zero-mean models

Example 2. (Stationary without drift):

$$Y_t = (0.5)Y_{t-1} + U_t, \quad U_t \sim IID(0, \sigma^2).$$



Examples: Some zero-mean models

Example 3 (Random Walk without drift).

Let $\{U_t\,, t = 1, 2, \ldots\}$ be the IID Noise.

The **Random Walk** process $\{Y_t\,, t = 1, 2, \ldots\}$ (starting at zero) is obtained by cumulatively summing (or "integrating") the iid RV's ε_t, that is, Y_t is defined by

$$Y_0 = 0 \quad \text{and}$$

$$Y_t = U_1 + U_2 + \ldots + U_t\,, \qquad \text{for} \quad t = 1, 2, \ldots \tag{1}$$

Observe that (1) is equivalent to the following

$$Y_t = Y_{t-1} + U_t\,, \quad t = 1, 2, \ldots, \quad Y_0 = 0\,, \tag{2}$$

Examples: Some zero-mean models

Remark 1.

If U_t is a **binary process**, then $\{Y_t\,, t = 1, 2, \ldots\}$ is called **simple symmetric random walk.**

Remark 2.

The **Random Walk** Y_t is a **non-stationary** process.

Indeed,

Mean $= \mathrm{E}(Y_t) = \mathrm{E}(U_1 + U_2 + \ldots + U_t) = 0.$

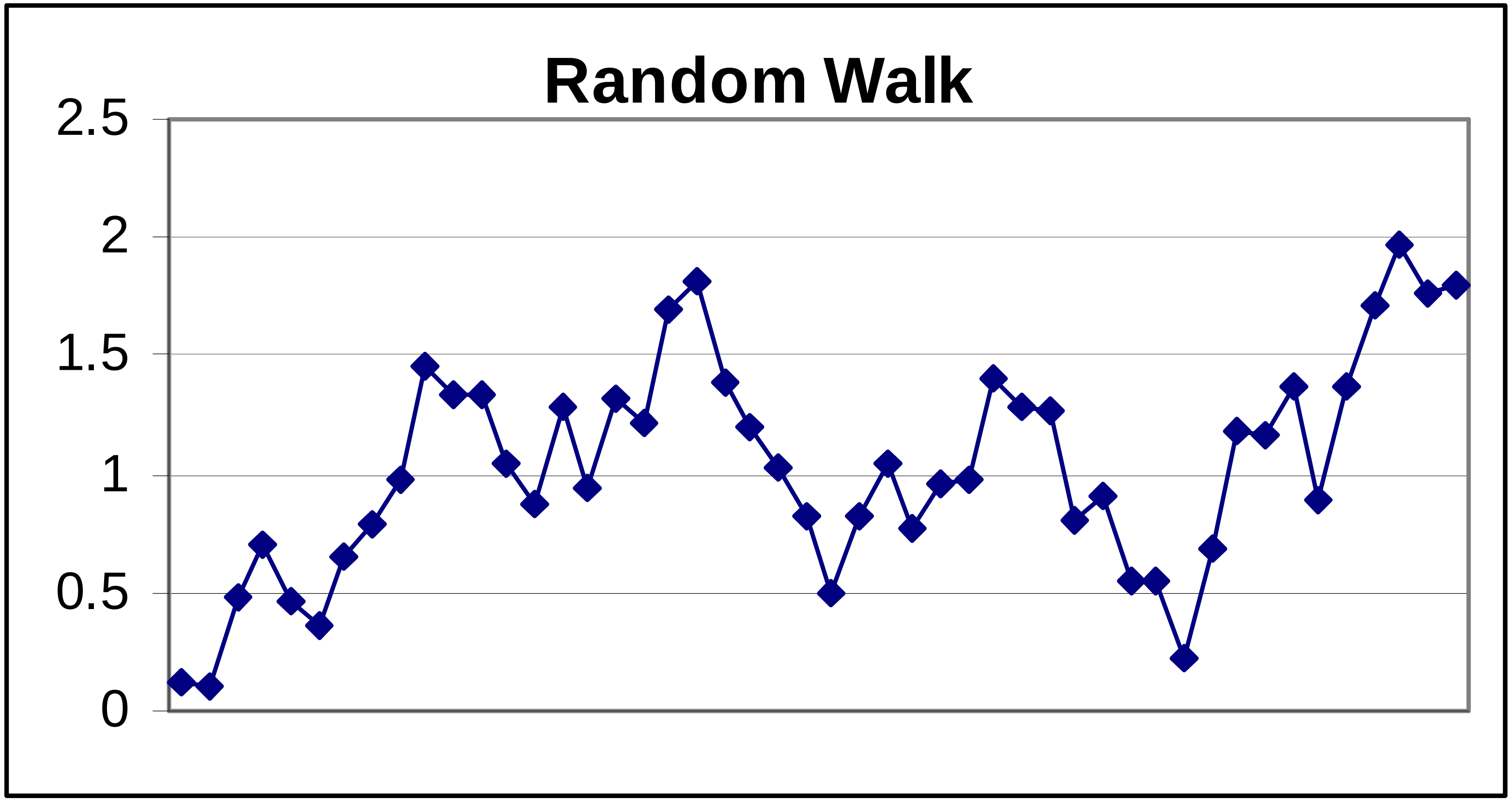
Variance $= \mathrm{Var}(Y_t) = Var(U_1) + \ldots + Var(U_t)$

$$= \sigma^2 + \ldots + \sigma^2 = t\,\sigma^2.$$

Thus, variance is not constant through time, and so, Y_t is **non-stationary**.

Examples: Some zero-mean models

Non-Stationary without drift: $Y_t = Y_{t-1} + U_t, \quad U_t \sim IID(0, \sigma^2).$



Examples: Some zero-mean models

Suppose we wanted to make a forecast for random walk process $Y_t = U_1 + U_2 + \ldots + U_t$.
The **one-step forecast** is given by

$$\hat{Y}_{T+1} = E[Y_{T+1} | Y_{T,} \ldots, Y_1] = Y_T + E[U_{T+1}] = Y_T. \qquad (3)$$

The **two-step forecast** (or forecast two periods ahead) is

$$\hat{Y}_{T+2} = E[Y_{T+2} | Y_{T,} \ldots, Y_1] = E[Y_{T+1} + U_{T+2}]$$
$$= E[Y_T + U_{T+1} + U_{T+2}] = Y_T.$$

Similarly, the k -step forecast is also Y_T, that is,

$$\hat{Y}_{T+k} = Y_T, \quad k = 1, 2, \ldots. \qquad (4)$$

Examples: Some zero-mean models

Observe that although

the forecast $\hat{Y}_{T+k} = Y_T$ for any $\boldsymbol{k}$ = 1, 2, …

the **variance of the forecast error will grow** as $\boldsymbol{k}$ **becomes larger**.

For **one-step** forecast the error $\boldsymbol{e_1}$ is given by

$$e_1 = Y_{T+1} - \hat{Y}_{T+1} = Y_T + U_{T+1} - Y_T = U_{T+1}.$$

Hence

$$Var(e_1) = Var(U_{T+1}) = s_U^2.$$

Examples: Some zero-mean models

For **two-step** forecast the error $\boldsymbol{e_2}$ we have

$$e_2 = Y_{T+2} - \hat{Y}_{T+2} = Y_T + U_{T+1} + U_{T+2} - Y_T = U_{T+1} + U_{T+2}.$$

Hence

$$Var(e_2) = Var(U_{T+1} + U_{T+2}) = Var(U_{T+1}) + Var(U_{T+2}) = 2s_U^2.$$

Similarly, for ***k*-step** forecast the error $\boldsymbol{e_k}$ we have

$$VAR(e_k) = ks_U^2.$$

Thus, the standard error of forecast increase with $\sqrt{k}$.

Therefore, we can obtain **confidence intervals** for our forecast, and these intervals will become wider as the forecast horizon increases.

Models with Trend and Seasonality

1.Models with Trend

In some situations, there is a **clear trend** in the time series data. In such cases, a **zero-mean model** for the data is clearly **inappropriate**.

A model of the form

$$X_t = m_t + Y_t \qquad (1)$$

will fit the data appropriately,

where

$\boldsymbol{m_t}$ is a **slowly varying function known as a trend component**,

and $\boldsymbol{Y_t}$ is a **zero-mean model**:

$$\mathbf{E}[\boldsymbol{Y_t}] = \mathbf{0}.$$

Models with Trend

Remark 1.

- The long-term tendency (**Trend**) is usually one of three: **growth, decline, or constant.**
- Reasons for trends include:
- Population growth -- greater demand for products and services, and greater supply of products and services.
- Technology -- impacts on efficiency, supply, and demand.
- Innovation -- impacts efficiency as well as supply and demand.

Models with Trend

Remark 2.

A useful technique for estimating $\boldsymbol{m_t}$ is the **least squares method** (some other methods will be considered later).

In the least squares procedure we attempt to fit a parametric family of functions, for example,

$$m_t = a_0 + a_1 t + a_2 t^2$$

to the data $\boldsymbol{x_1, \ldots, x_T}$ by choosing the parameters $(\boldsymbol{a_0, a_1, a_2})$ to minimize the sum

$$\sum_{t=1}^{T} (x_t - m_x)^2 .$$

This is least squares regression method of curve fitting.

Models with Trend

A simple example of a time series model with trend is the **random walk with drift**, given by

$$Y_t = Y_{t-1} + d + e_t. \qquad (2)$$

Observe that the process will trend to move upward if **d > 0**, and downward if **d < 0**.

For model (2) , the **one-step forecast** is

$$\hat{Y}_{T+1} = E[Y_{T+1} | Y_T, \mathbf{L}, Y_1] = Y_T + d,$$

and the ***k*-step forecast** is

$$\hat{Y}_{T+1} = Y_T + kd.$$

The **standard error** of forecast will be the same as before:

$$VAR(e_k) = ks_U^2.$$

Models with Trend

For **one-step forecast**

$$e_1 = Y_{T+1} - \hat{Y}_{T+1} = Y_T + d + e_{T+1} - Y_T - d = e_{T+1}.$$

Thus, for the model random walk with drift, the standard error of forecast again will increase with $\sqrt{k}$, and we can obtain CI's for forecasts.

Models with Trend: Stationary vs. Non-stationary

Assume that a time series model is given by equation:

$$Y_t = \alpha + \rho Y_{t-1} + u_t , \qquad u_t \sim IID(0, \sigma^2).$$

Then

$\rho < 1$ stationary process

- "process forgets past"

$\rho = 1$ non-stationary process

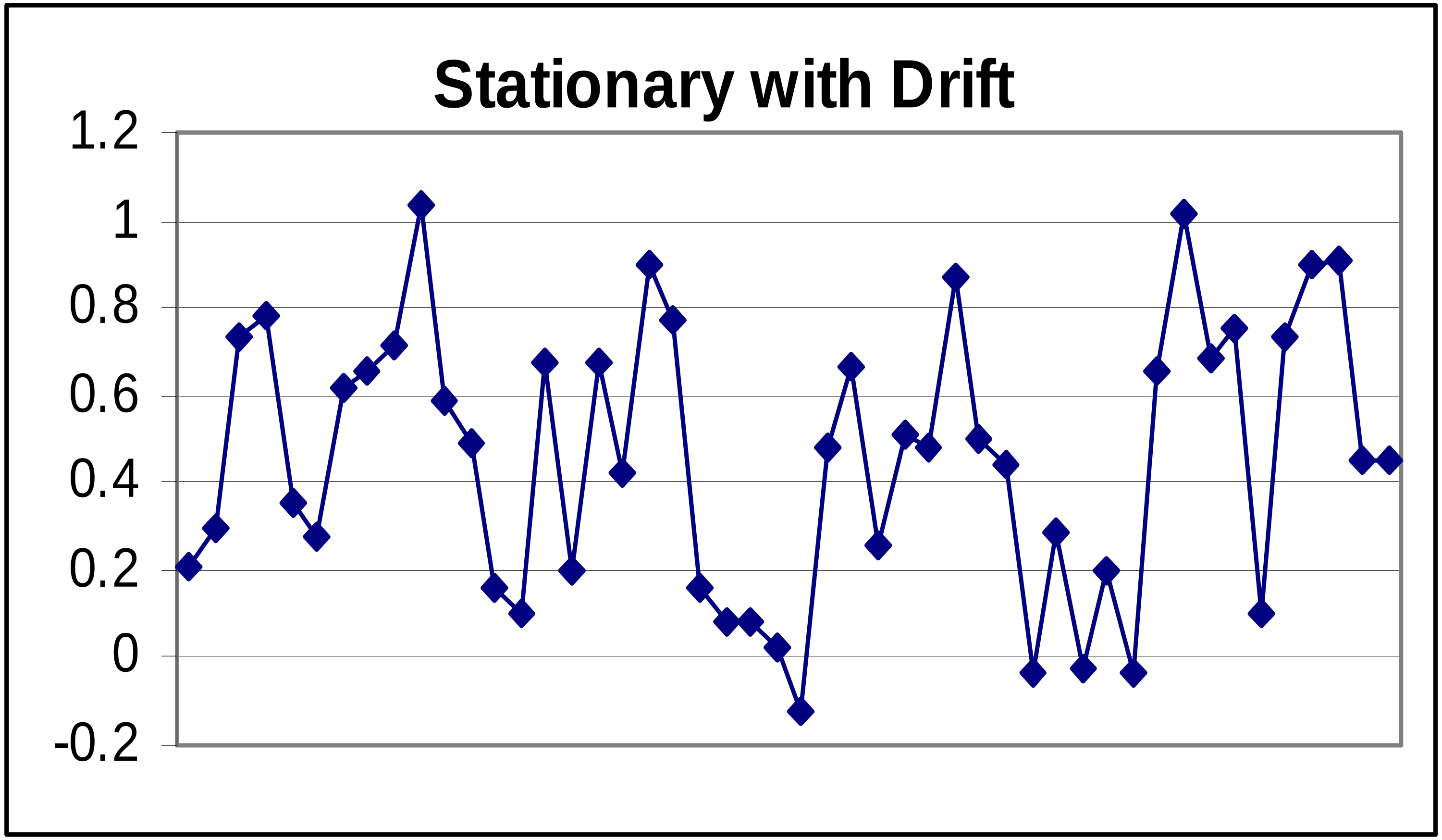
- "process does not forget past"

$\alpha = 0$ without drift

$\alpha \neq 0$ with drift

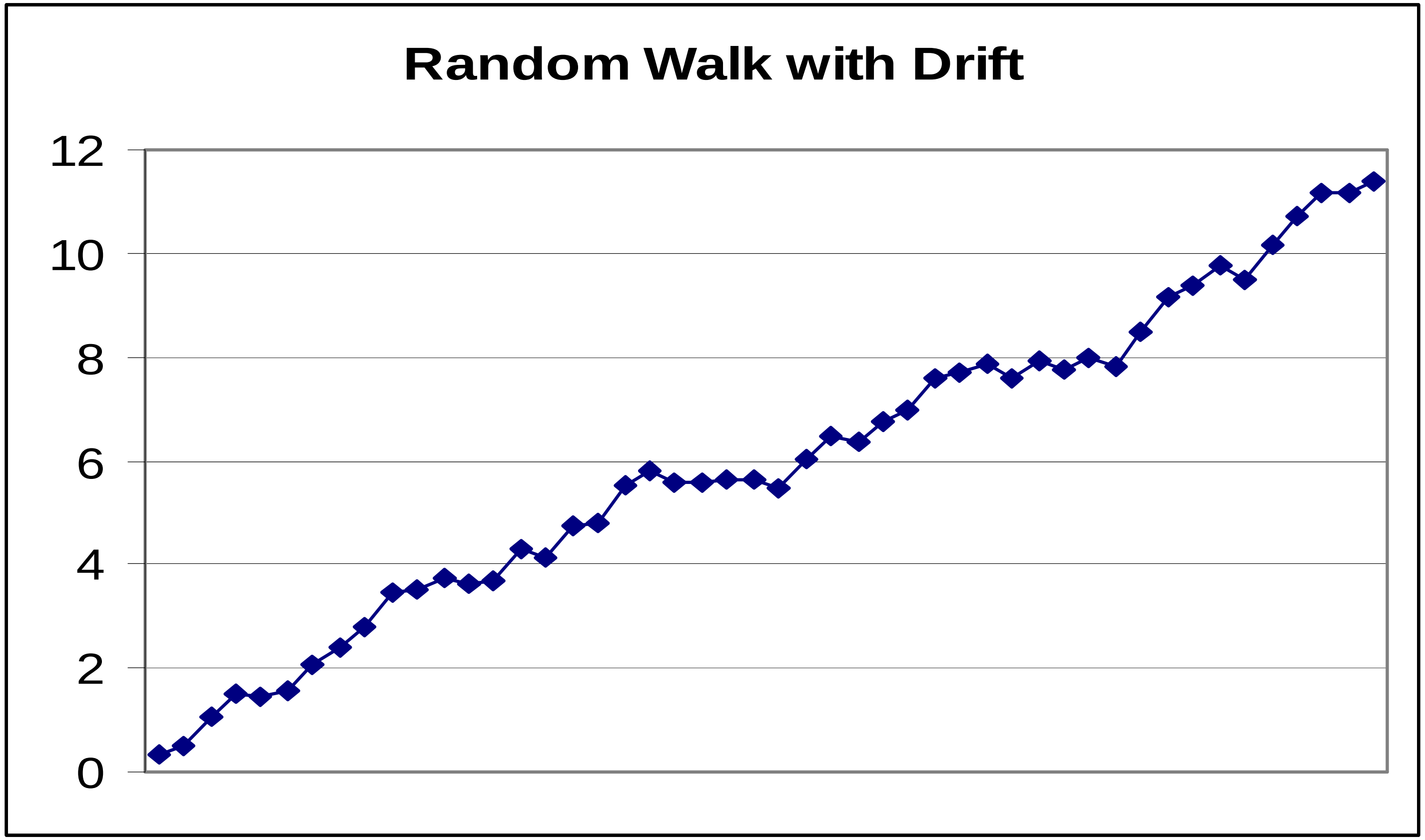
Example: Stationary time series with drift

$$Y_t = \alpha + (0.5)Y_{t-1} + U_t \qquad U_t \sim IID(0, \sigma^2)$$



Example: Non-stationary time series with drift

$$Y_t = \alpha + Y_{t-1} + U_t\ , \qquad U_t \sim IID(0, \sigma^2\ )$$



Example: Types of Models

$$Y_t = m + f_1 Y_{t-1} + b\,t + e_t$$

	Parameter Set	Description
1	$m \neq 0, \lvert f_1 \rvert < 1, b \neq 0$	Deterministic Trend With Stationary AR(1) components
2	$m \neq 0, f_1 = 1, b \neq 0$	Random Walk with Drift and Deterministic Trend
3	$m \neq 0, f_1 = 1, b = 0$	Random Walk with Drift
4	$m \neq 0, f_1 = 0, b \neq 0$	Deterministic Trend
5	$m = 0, f_1 = 1, b = 0$	Pure Random Walk

Models with seasonality (Harmonic Regression)

Many time series models are influenced by **seasonally varying factors** such as the weather, the effect of which can be modeled by a **periodic component** with fixed known period.

In order to represent such a **seasonal effect, allowing for noise but assuming no trend**, we can use the so-called **Harmonic Regression** model given by

$$X_t = S_t + Y_t,$$

where $\boldsymbol{S_t}$ is a periodic function of $\boldsymbol{t}$ with period $\boldsymbol{l}$.

Models with seasonality

Remark 1.

A convenient choice for $\boldsymbol{S_t}$ is a sum of **harmonics** **(or sine - waves)** given by

$$S_t = a_0 + \sum_{j=1}^{k} (a_j \cos(l_j t) + b_j \sin(l_j t)),$$

where $\boldsymbol{a_0, \ldots, a_k}$ and $\boldsymbol{b_0, \ldots, b_k}$ are **unknown** parameters to be estimated from the data, and $\boldsymbol{\lambda_0, \ldots, \lambda_k}$ are fixed frequencies, each being some integer multiple of $2\,\boldsymbol{\pi / l}$.

Models with seasonality

Remark 2.

- Upward and downward movements which repeat **at the same time each year.**
- Reasons for seasonal influences include:

  Weather -- both outdoor and indoor activities can impact demand because of the number of people involved

  -- supplies of products and services may depend on the weather
- Events, Holidays -- often impact supply and demand.

Models with Cyclical Component

- Similar to seasonal variations except that there is likely not a relationship to the time of the year.

Examples of cyclical influences include:

- Inflation/deflation -- energy costs, wages and salaries, and government spending
- Stock market prices -- bull markets, bear markets
- Consequences of unique events -- severe weather, law suits.

It is important for economic time series models.

Models with Irregular Component

- Unexplained variations which we usually treat as **randomness.** This is the equivalent of the **error term** in the analysis of variance model and the regression model.

- These are short-term effects, usually. We treat them as independent from one time period to the next.
- The length of the duration of these effects would then be shorter than one time period, that is, one month for monthly data, one year for annual data.

Summary: Components of a Time Series

- ***Long Term Trend***
  - A time series may be stationary or exhibit trend over time.
  - Long term trend is typically modeled as a linear, quadratic or exponential function.
- ***Seasonal Variation***
  - When a repetitive pattern is observed over some time horizon, the series is said to have seasonal behavior.
  - Seasonal effects are usually associated with calendar or climatic changes.
  - Seasonal variation is frequently tied to yearly cycles.
- ***Cyclical Variation***
  - An upturn or downturn not tied to seasonal variation.
  - Usually results from changes in economic conditions.
- ***Random effects***

General Time-Series Model

- The four components of time series come together to form a time series model.
- There are two popular time series models:
  - **Additive Model:**

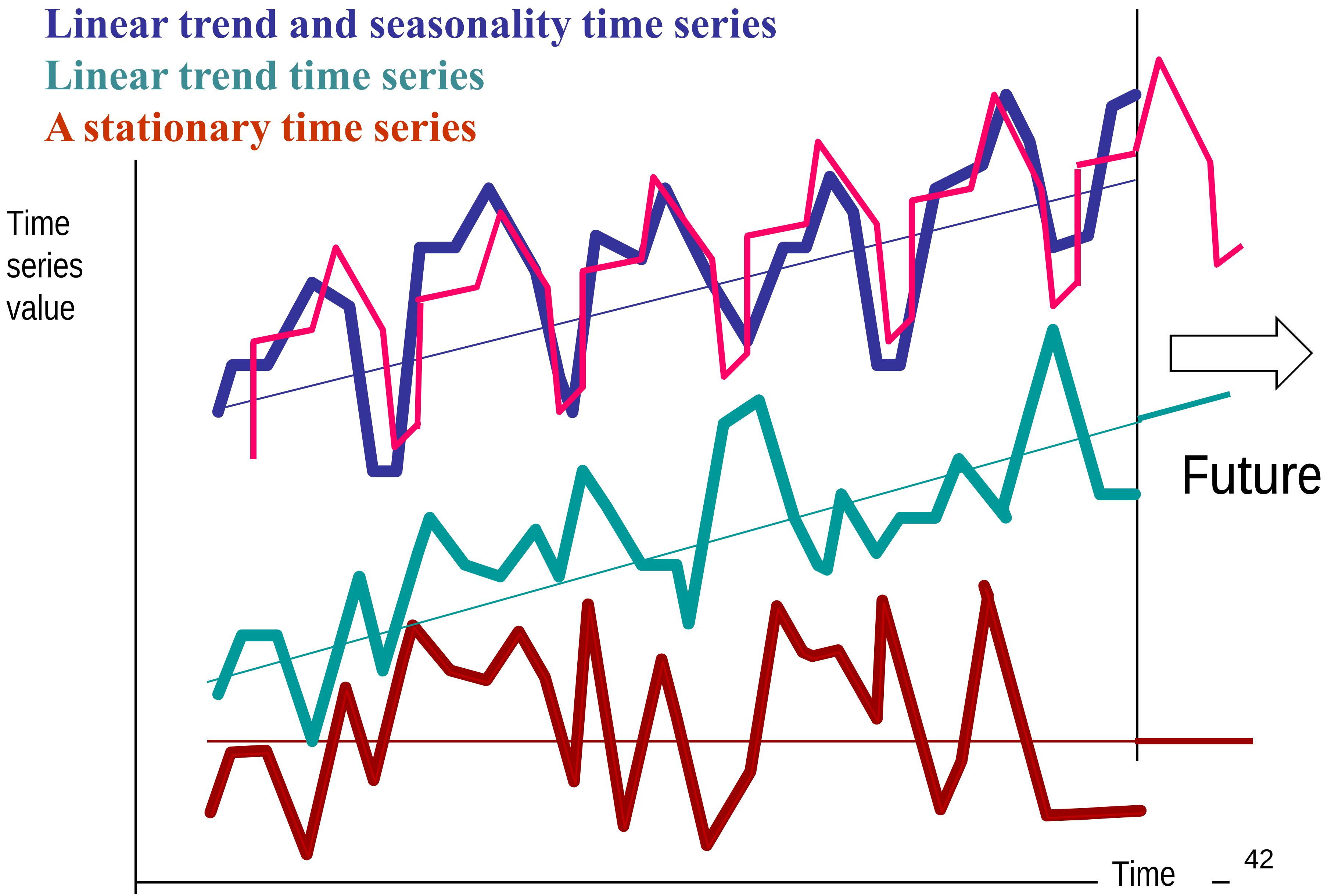
$$Y_t = T_t + S_t + C_t + I_t$$

  - **Multiplicative Model:**

$$Y_t = (T_t)(S_t)(C_t)(I_t),$$

where $\mathbf{T_t}$ is the trend, $\mathbf{S_t}$ is the seasonal, $\mathbf{C_t}$ is the cyclical and $\mathbf{I_t}$ is the irregular components.

Components of a Time Series



A General Approach to Time Series Modeling

The following recommendation can be useful in the Time Series Modeling and Analysis.

1.Plot the series (data) and examine the main futures of the graph, checking in particular whether there is:

- **a) a trend component**
- **b) a seasonal component**
- **c) any apparent sharp changes in behavior,**
- **d) any outlying observations.**

2.Remove the **trend** and **seasonal** components to get **stationary residuals.** To achieve this goal it may be necessary to transform the underlying data.

A General Approach to Time Series Modeling

3. **Chose a model** to fit the residuals, making use of various sample statistics including the sample autocorrelation function to be defined below.

4. **Forecasting** will be achieved by forecasting the residuals and then inverting the transformations to arrive at forecasts of the original series $\{Y_t\}$.

5. **An extremely** useful alternative approach is to express the series in terms of its **Fourier components,** which are sinusoidal waves of different frequencies.

2. Stochastic Processes

Stochastic Processes

1. Definition and Remarks.

Definition 1.

Let $(\Omega, \Im, \mathrm{P})$ be a given probability space and $\boldsymbol{T}$ be a given index set.

A **stochastic (or random) process (SP)** indexed by $\boldsymbol{T}$ is a family (or collection) of random variables

$$\{X(t) = X(t,w), \quad t \in T\}$$

defined on $(\Omega, \Im, \mathrm{P})$ and taking values in a set $\boldsymbol{S}$ which is called the **state space** of the process.

Thus, a SP is a function of two variables:

$$X(t,w) : \Omega \times T \to S.$$

Definition and Remarks

- **Remark 1.**

If $t \in T$ represents time (in seconds, minutes, hours, days, months, years, etc.), then the process $\boldsymbol{X(t\,)}$ is called a **"time-series"**.

- **Remark 2.**

If $w \in \Omega$ is fixed, say $\boldsymbol{\omega = \omega_0}$, then

$$X(w_0, t) = X_0(t) = g(t)$$

is a **deterministic (non-random)** function, which is called the **realization** or **sample function** or **path** of the process

$$\boldsymbol{X(t) = X(t, \omega)}.$$

In this case we have **deterministic** time series.

Definition and Remarks

- **Remark 3.**

If $t \in T$ is fixed, say $\boldsymbol{t = t_0}$, then

$$X(w, t_0) = X_0(w) = x(w)$$

is a **random variable,** and in this case we have **no** **"time-series"**.

- **Remark 4 (Example).**

A simple way to construct a SP is:
take any RV $x(w)$, $w \in \Omega$,
and any "usual" function $g(t)$, $t \in T$ and multiply to get

$$X(w, t) = x(w)g(t).$$

Specification of Stochastic Processes

- In the theory of stochastic process the **Basic Question** are:

1. **How to specify the stochastic process *X(t)*?**
2. **How to describe the distribution of *X(t)*?**

The main elements, distinguishing stochastic processes are in the nature of

- **the index set *T***
- **the state space *S***, and
- **the dependence relations among the random variables**

$$X(t) = X(t, w), t \in T.$$

Specification of Stochastic Processes

(A) The Index set :

(A -1). If $\boldsymbol{T} = \boldsymbol{Z} = \{0, \pm 1, \pm 2, \ldots\}$ or $\boldsymbol{T} = \boldsymbol{N} = \{1, 2, \ldots\}$, then $\boldsymbol{X}(\mathbf{t})$ is called **discrete-time** (-parameter) process, or **time-series**.

(A -2). If $\mathbf{T} = \mathbf{R} = (-\infty, \infty)$ or $T = (a, b) \subset R$, then $\boldsymbol{X}(\mathbf{t})$ is called **continuous-time** (-parameter) process, or **continuous-time time-series**.

(A -3). If $\mathrm{T} = \boldsymbol{Z}^{m}$ or $\mathrm{T} = \boldsymbol{R}^{m}$, then $\boldsymbol{X}(\mathbf{t}) = \boldsymbol{X}(\boldsymbol{t}_1, \ldots, \boldsymbol{t}_m)$ is called **discrete-time** or **continuous-time** (respectively)

$\boldsymbol{m}$ - dimensional **Random Field**.

Specification of Stochastic Processes

(B) The State Space *S*:
(the space of possible values of ***X*(t)**.)

(B -1). If $\boldsymbol{S} = \boldsymbol{R} = (-\infty, \infty)$, then ***X*(t)** is called **real-valued** process.

(B -2). If $\boldsymbol{S} = \boldsymbol{C}$ (the complex numbers), then ***X*(t)** is called **complex-valued** process.

(B -3). If $\boldsymbol{S} = \boldsymbol{Z} = \{0, \pm 1, \pm 2, \ldots\}$, then ***X*(t)** is called **integer-valued** or **discrete-state** process.

Specification of Stochastic Processes

(C) <u>Specification of Stochastic Processes.</u>
<u>Dependence relations among the random variables</u>,

$$X(t) = X(t,w), t \in T.$$

<u>Dependence</u> relations among the random variables $X(t) = X(t,w), t \in T$ are specified by the

<u>family of finite-dimensional distribution functions</u>

of the process ***X(t)***.

Specification of Stochastic Processes

Remark.

Observe that

(a) A RV $\boldsymbol{X(\omega)}$ is completely specified by its ***cdf***

$$F(x) = P(X(w) \le x);$$

(b) A random vector $(\boldsymbol{X_1}, \ldots, \boldsymbol{X_n})$ is completely specified by its ***cdf***

$$F_n(x_1, \ldots, x_n) = P(X_1(w) \le x_1, \ldots, X_n(w) \le x_n).$$

If $\boldsymbol{T}$ is **infinite** it is not easy to specify the process

$$X(t) = X(t, w), t \in T.$$

Specification of Stochastic Processes

We shall regard the SP $X(t), t \in T$, as being specified, if

1.for each $t_1 \in T$ there is defined the distribution function

$F_{t_1}(x)$ of RV $X(t_1)$:

$$F_{t_1}(x) = P\{X(t_1) \le x\}.$$

2.for each $t_1, t_2 \in T$ we are given the ***cdf*** of 2-dimensional

random vector $(X(t_1), X(t_2))$.

$$F_{t_1, t_2}(x_1, x_2) = P(X(t_1) \le x_1, X(t_2) \le x_2)$$

and so on….

Specification of Stochastic Processes

3. for **any** finite number of elements $t_1, \ldots t_n \in T \quad (n \in N)$, we are given the ***cdf*** of ***n*** -dimensional random vector **(*X*(t$_1$)** ,…, ***X*(t$_n$))** :

$$F_{t_1, \ldots, t_n}(x_1, \ldots x_n) = P(X(t_1) \le x_1, \ldots, X(t_n) \le x_n).$$

Thus for an **infinite** set ***T*** the SP $X(t), t \in T$ is completely specified by the **whole family of finite *n* -dimensional *cdf*`s**

$$\left\{F_{t_1, \ldots, t_n}(x_1, \ldots x_n), \quad n = 1, 2, \ldots\right\}.$$

Specification of Stochastic Processes

Definition 1.

The **finite-dimensional distribution functions** of the process $X(t), t \in T$, are the functions

$$\left\{ F_{\bar{t}}(\bar{x}) = F_{t_1 \ldots t_n}(x_1, \ldots x_n), \quad \bar{t} = (t_1, \ldots, t_n)' \in \Im, \quad n = 1, 2, \ldots \right\}$$

where

$$F_{\bar{t}}(\bar{x}) = F_{t_1 \ldots t_n}(x_1, \ldots x_n) = P\left\{ X(t_1) \leq x_1, \ldots, X(t_n) \leq x_n \right\}. \qquad (1)$$

Thus, the **finite-dimensional distribution functions** of the stochastic process ***X*(t)** are the distributions of the finite-dimensional vectors

$$(X(t_1), \ldots, X(t_n)), \quad t_1, \ldots, t_n \in T,$$

for all possible choices of times $t_1, \ldots, t_n \in T$ and every $n \geq 1$.

Specification of Stochastic Processes

The Consistency Condition

Notation:

For a vector $\overline{h} = (h_1, \ldots, h_n)$ we denote

$$\overline{h}(k) = (h_1, \ldots, h_{k-1}, h_{k+1}, \ldots, h_n),$$

that is, $\overline{h}(k)$ is a $(\boldsymbol{n-1})$ -component vector obtained by $\overline{h}$ deleting the $\boldsymbol{k}$ -th component of $\overline{h}$.

It follows from the Def. 1 that the finite-dimensional distribution functions $F_{\overline{t}}(\overline{x})$ of the process $X(t), t \in T$ satisfy the following **Consistency Condition:**

$$\lim_{x_k \to \infty} F_{\overline{t}}(\overline{x}) = F_{\overline{t}(k)}\left(\overline{x}(k)\right). \qquad (2)$$

Specification of Stochastic Processes

The Consistency Condition

Remark.

The Consistency condition states that:

Each function $F_{\bar{t}}(\bar{x})$ should have marginal distributions which coincide with the specified lower dimensional distribution functions.

Specification of Stochastic Processes

Thus, we have seen that the family of finite-dimensional distribution functions of a process $X(t), t \in T$, necessarily satisfies the **consistency condition**.

Now we ask the following **question:**
Let a family of finite-dimensional distribution functions $\{F_{\bar{t}}(\bar{x})\}$ with the parameter set $\boldsymbol{T}$ is apriori given,
that is, for any $\boldsymbol{n}$, and $t_1, \ldots, t_n \in T$,

$$F_{t_1, \ldots, t_n}(x_1, \ldots x_n) = F(x_1, \ldots x_n)$$

is an $\boldsymbol{n}$ -dimensional distribution function,
under what conditions can this family be the family of finite-dimensional distribution functions of a stochastic process?

Specification of Stochastic Processes

Or more precisely,

Given $\{F_{\bar{t}}(\bar{x})\}$ with parameter set **T**, under what conditions do there exist

- a probability $(\Omega, \Im, \mathrm{P})$ space and
- a stochastic process $X(t) = X(t,w), t \in T,$ defined on $(\Omega, \Im, \mathrm{P})$, such that

$$P(X(t_1) \le x_1, \ldots, X(t_n) \le x_n) = F_{t_1,\ldots,t_n}(x_1,\ldots x_n) \quad (= F_{\bar{t}}(\bar{x}))$$

Kolmogorov's famous theorem is the answer to this question, and it is remarkable that all such conditions are covered by a single condition: the consistency condition.

Specification of Stochastic Processes

Theorem (Kolmogorov's Theorem on existence of a stochastic Process).

Let a family of distributions $\{F_{\bar{t}}(\bar{x}), t \in \Im\}$ with the parameter set ***T*** be apriori given.

Then a **necessary and sufficient condition** for the existence of

a probability space $(\Omega, \Im, \mathrm{P})$ and

a stochastic process $X(t) = X(t, w)$, $t \in T$

defined on $(\Omega, \Im, \mathrm{P})$, such that

$$P\left(X(t_1) \le x_1, X(t_2) \le x_2, ..., X(t_n) \le x_n\right) = F_{\bar{t}}(\bar{x})$$

is that given family $\{F_{\bar{t}}(\bar{x})\}$ satisfies the **consistency condition:**

$$\lim_{x_k \to \infty} F_{\bar{t}}(\bar{x}) = F_{\bar{t}(k)}(\bar{x}(k)).$$

3. Some Classes of Stochastic Processes

Some Classes of Stochastic Processes

1. The Class of Strictly Stationary SP's.

Definition 1.

Let $X(t) = X(t, w), t \in T$ be a stochastic process defined on a probability space $(\Omega, \Im, \mathrm{P})$ with finite-dimensional distribution functions:

$$F_{t_1, \ldots, t_n}(x_1, \ldots x_n) = P(X(t_1) \le x_1, \ldots, X(t_n) \le x_n).$$

Then ***X*(t)** is called

- **1st order Stationary** SP if

$$F_{t_1 + t}(x_1) = F_{t_1}(x_1) \quad \text{for all } t_1 \in T \text{ and } t > 0.$$

- **2nd order Stationary** SP if

$$F_{t_1 + t, t_2 + t}(x_1, x_2) = F_{t_1, t_2}(x_1, x_2) \text{ for all } t_1, t_2 \in T \text{ and } t > 0.$$

The Class of Strictly Stationary Processes

- **n** **- order stationary** SP if

$$F_{t_1+t \ldots t_n+t}(x_1,\ldots x_n) = F_{t_1 \ldots t_n}(x_1,\ldots x_n) \text{ for all } t_1,\ldots,t_n \in T \text{ and } t > 0.$$

Definition 2.

A stochastic process $X(t) = X(t,w), t \in T$ is called **Strictly (or Strongly) Stationary** SP if it is **n - order stationary** process for any ***n***.

Thus, a SP ***X*(t)** is **strictly stationary** if the finite-dimensional distribution functions of ***X*(t)** are

invariant with respect to time-shift.

The Class of Second Order Processes

2. The Class of Second Order SP's.

Definition 1.

A stochastic process $X(t) = X(t,w), t \in T$, defined on a probability space $(\Omega, \Im, \mathrm{P})$ is called **Second Order** SP if for all $t \in T$

$$E[X(t)]^2 = \int_\Omega |X(t,w)|^2 \, dP(w) < \infty.$$

We now define the **mean** and **covariance** functions of a stochastic process ***X(t)***.

The Class of Second Order Processes

Definition 2.

The **mean** function ***m(t)*** of the process ***X(t)*** is defined to be

$$m(t) = E[X(t)],\ t \in T.$$

Definition 3.

The **covariance** function ***r(t,s)*** of the process ***X(t)*** is defined to be

$$r(t,s) = E\Big[\big(X(t) - m(t)\big)\big(X(s) - m(s)\big)\Big],\ t,s \in T.$$

Remark 1.

It follows from Cauchy-Schwarz inequality that the **mean** function ***m(t)*** and **covariance** function ***r(t,s)*** of the second order process ***X(t)*** exist ($\mathbf{E}[X(\mathbf{t})]^2 < \infty$).

Some Classes of Stochastic Processes

3. The Class of Second Order (Wide) Stationary SP's.

Definition 1.

Let $X(t), t \in T$, be a second order SP defined on a probability space $(\Omega, \Im, \mathrm{P})$ with **mean** and **covariance** functions ***m(t)*** and ***r(t,s)***, respectively.

Then ***X*(t)** is called **Second Order Stationary** if

$$(a) \quad m(t) = m = const (= 0) \text{ for all } t \in T.$$

$$(b) \quad r(t, s) = R(t - s) \text{ for all } t, s \in T.$$

The Class of Second Order Stationary Processes

•**Remark 1.**

In the class of second order process ($\mathbf{E[X(t)]^2 < \infty}$)
strict stationarity implies wide stationarity.

The converse, in general, is **not true**.

For Gaussian processes, however, the **converse is true**, that is,
in the class of Gaussian processes
"strict stationarity" and **"wide stationarity" coincide.**

$\Leftrightarrow$

4. The Class of Gaussian Processes.

Definition 1.

We say that an ***n*** -dimensional random vector **X** = (X_1 ,…, X_n) has a MND with mean vector **μ** = (μ_1 ,…, μ_n)’ and variance-covariance matrix $R = \left\| r_{kj} \right\|_{kj=\overline{1.n}}$, and write $X \sim N_n(\mu, R)$ if for all $u = (u_1, \ldots, u_n) \in R^n$ the mgf $M_X(u)$ (or CF $j_X(u)$) of X is given by

$$M_X(u) = \exp\left\{(u,m) + \frac{1}{2}(Ru,u)\right\} = \exp\left\{\sum_{k=1}^{n} u_k m_k + \frac{1}{2}\sum_{kj=1}^{n} r_{kj} u_k u_j\right\}$$

$$j_X(u) = \exp\left\{i\sum_{k=1}^{n} u_k m(t_k) - \frac{1}{2}\sum_{kj=1}^{n} r_{kj} u_k u_j\right\}.$$

The Class of Gaussian Processes

where $m_k = E\left[X_k\right]$ and $r_{kj} = E\left[X_k - m_k\right]\left[X_j - m_j\right]$.

For ***pdf*** of $\boldsymbol{X}$ we have for $x = (x_1, \ldots, x_n) \in R^n$,

$$f_X(x) = \frac{1}{(2p)^{n/2}\left|R\right|^{1/2}} \exp\left\{-\frac{1}{2}\sum_{kj=1}^{n} a_{kj}(x_k - m_k)(x_j - m_j)\right\},$$

where $\boldsymbol{\alpha_{kj}}$ are the elements of the inverse matrix $\mathbf{R^{-1}}$:

$$R^{-1} = \left\| a_{kj} \right\|_{kj=\overline{1.n}},$$

and

$$| R | = \det R.$$

The Class of Gaussian Processes

Definition 2.

Let $X(t), t \in T$ be a second order SP defined on a probability space $(\Omega, \Im, \mathrm{P})$ with **mean** and **covariance** functions ***m***(***t***) and ***r***(***t, s***), respectively.

Then we say that ***X***(***t***) is a **Gaussian Process** with **mean** function ***m(t)*** and **covariance** function ***r(t, s)***, and write

$$\mathbf{X(t) \sim N\ (m(t),\ r(t,s),}$$

if the finite-dimensional distribution functions of ***X(t)*** are **Gaussian.**

The Class of Gaussian Processes

That is, for any $n \in N$ and any $t_1, .., t_n \in T$ the $\boldsymbol{n}$ -dimensional random vector $\boldsymbol{X} = (\ \mathbf{X(t_1)}\ ,\ldots,\ \mathbf{X(t_n)}\ )$ has a MND with mean vector $\boldsymbol{\mu} = (\ \mathbf{m(t_1)}\ ,\ldots,\ \mathbf{m(t_n)})'$ and variance-covariance matrix

$$R = \left\| r_{kj} \right\|_{kj=\overline{1.n}}, r_{kj} = r(t_k, s_j) = E\left[\left(X(t_k) - m(t_k) \right) \left(X(s_j) - m(s_j) \right) \right]$$

So the Characteristic Function $\boldsymbol{\varphi_X (u)}$ of $\boldsymbol{X}\mathbf{(t)}$ is given by

$$j_X(u) = \exp\left\{ i \sum_{k=1}^{n} u_k m(t_k) - \frac{1}{2} \sum_{kj=1}^{n} r_{kj} u_k u_j \right\}$$

Remark 1.

A Gaussian process $\boldsymbol{X}\mathbf{(t)}$ is completely specified by the **mean** function $\boldsymbol{m}\mathbf{(t)}$ and **covariance** function $\boldsymbol{r}\mathbf{(t, s)}$.

The Class of Gaussian Processes

Remark 2.

W.l.o.g we can assume that **mean** function $\boldsymbol{m}(\mathbf{t}) = \mathbf{0}$.
Otherwise instead of $\boldsymbol{X}(\mathbf{t})$ we can consider a new process $\boldsymbol{Y}(\mathbf{t})$ defined by $\boldsymbol{Y}(\mathbf{t}) = \boldsymbol{X}(\mathbf{t}) - \boldsymbol{m}(\mathbf{t})$ for which

$$m_Y(t) = E[Y(t)] = E[X(t) - m(t)]$$
$$= E[X(t)] - m(t) = m(t) - m(t) = 0.$$

The Class of Gaussian Processes

Definition 3.

A real-valued function of two variables ***h(t, s)*** is called **non-negative definite** (***nnd***), if for any $n \in N$, for any $t = (t_1, .., t_n) \in T$ and any numbers $a_1, .., a_n \in R$,

$$\sum_{kj=1}^{n} h(t_k, s_j) a_k a_j \geq 0. \qquad (3)$$

Theorem 2.

The covariance function ***r(t, s)*** of any second order process ***X(t)*** (**E[X(t)]²** $< \infty$) is a ***nnd***-function, that is, ***r(t, s)*** satisfies (3).

The Class of Gaussian Processes

Proof.

Indeed, since $r(t,s) = Cov\left(X(t), X(s)\right) = E\left[\left(X(t)X(s)\right)\right]$

we have

$$\sum_{kj=1}^{n} r(t_k, s_j)a_k a_j = \sum_{kj=1}^{n} E\left[\left(X(t_k)X(s_j)\right)\right]a_k a_j = E\left[\sum_{kj=1}^{n} X(t_k)X(s_j)a_k a_j\right]$$

$$= E\left[\sum_{k=1}^{n} X(t_k)a_k \sum_{j=1}^{n} X(s_j)a_j\right] = E\left[\sum_{k=1}^{n} X(t_k)a_k\right]^2 \geq 0.$$

The next result shows that the **converse also is true**.

The Class of Gaussian Processes

Theorem 3

(Existence of a Gaussian Processes with given Covariance).

Assume that we are given a parameter set $\boldsymbol{T}$ and a real-valued ***nnd***-function $r(t,s),\ t,s\in T.$

Then there exist a probability space $(\Omega,\Im,\mathrm{P})$ and a real-valued Gaussian process $X(t)=X(t,w)$, $t\in T$, defined on $(\Omega,\Im,\mathrm{P})$, such that

$$Cov\big(X(t),X(s)\big)=E\Big[\big(X(t)X(s)\big)\Big]=r(t,s).$$

Examples of Gaussian Process

1. Brownian motion (or Wiener process).

Let T = [0, ∞). The process $X(t),\ t \in T$, with $\boldsymbol{X}(0)=0$, and covariance $\boldsymbol{R(t,s) = min(t,s)}$ is Gaussian.

Observe that $\boldsymbol{X}(\mathbf{t})$ has independent increment, i.e., for arbitrary, $\boldsymbol{t_1 < t_2 < \ldots < t_n}$,

$$\boldsymbol{X}(t_2) - \boldsymbol{X}(t_1), \ldots, \boldsymbol{X}(t_n) - \boldsymbol{X}(t_{n-1}) \quad \text{are independent.}$$

2. Conditional Wiener Process (Brownian Bridge Process).

$$\boldsymbol{T = [0, 1], \quad X(0) = 0, \quad R(s, t) = min(s, t) - st\,.}$$

3. Gauss-Markov Process.

$$T = R,\ R(s,t) = e^{-|t-s|}.$$