Non-stationary process

Example

Consider the random signal $x[k] = A\cos(\omega k + \phi)$, where ϕ is a random variable with uniform distribution in $[0, \pi]$. The mean of this signal (process) is

$$E(x[k]) = E(A\cos(\omega k + \phi)) = A\int_0^\pi \cos(\omega k + \phi) f(\phi)\, d\phi = \frac{A}{\pi}\int_0^\pi \cos(\omega k + \phi)\, d\phi$$
$$= -\frac{2A}{\pi}\sin(\omega k)$$

which is a function of time. Thus, the process is **non-stationary.**

Order stationarity in distribution

A stochastic process is said to be N^{th}-**order stationary** (in distribution) if the joint distribution of N observations is invariant to shifts in time,

$$F_{X_{k_1},\cdots,X_{k_N}}(x_1,\cdots,x_N) = F_{X_{T+k_1},\cdots,X_{T+k_N}}(x_1,\cdots,x_N) \quad \forall T, k_1, \cdots, k_N \in \mathcal{Z}^+ \tag{2}$$

where $X_{k_1},\cdots,X_{k_N}$ are the RVs associated with the observations at $k = k_1, \cdots, k_N$, respectively and $x_1, \cdots, x_N$ are any real numbers.

Special cases

1. **First-order stationarity in distribution:**

$$F_{X_{k_1}}(x_1) = F_{X_{k_1+T}}(x_1) \qquad \forall T, k_1 \in \mathcal{Z}^+ \tag{3}$$

   Every observation should fall out of the same distribution.

2. **Second-order stationarity in distribution:**

$$F_{X_{k_1},X_{k_2}}(x_1, x_2) = F_{X_{k_1+T},X_{k_2+T}}(x_1, x_2) \qquad \forall T, k_1, k_2 \in \mathcal{Z}^+ \tag{4}$$

   The distribution depends **only on the time-difference** $k_2 - k_1$.

Relaxation of strict stationarity

The requirement of strict stationarity is similar to the strict requirement of time-invariance or linearity (in deterministic processes), which are also academically convenient assumptions, but rarely satisfied in practice.

In reality, rarely will we find a process that satisfies the strict requirements of stationarity defined above.

A **weaker** requirement is that certain key statistical properties of interest such as mean, variance and a few others at least, remain invariant with time

Weak stationarity

A common relaxation, is to require invariance up to second-order moments.

Weak or wide-sense or second-order stationarity

A process is said to be weakly or wide-sense or second-order stationary if:

i. The **mean of the process is independent of time**, *i.e.*, invariant w.r.t. time.

ii. It has **finite variance**.

iii. The **auto-covariance function** of the process

$$\sigma_{xx}[k_1, k_2] = \text{cov}(X_{k_1}, X_{k_2}) = E((X_{k_1} - \mu_1)(X_{k_2} - \mu_2)) \tag{5}$$

is only a function of the "time-difference" (lag $l = k_2 - k_1$) but not the time.

On wide-sense stationarity (WSS)

Q: Under what conditions is the weak stationarity assumption justified?

Where linear models are concerned, the optimal parameter estimates are fully determined by the first- and second-order properties of the joint p.d.f.

Example

For a stationary process, suppose a linear predictor of the form

$$\hat{x}[k] = -d_1 x[k-1] - d_2 x[k-2]$$

is considered. Determine the optimal (in the m.s. prediction error sense) estimates.

Gaussian WSS process

$\boxed{\textbf{A WSS multivariate Gaussian process is also strictly stationary.}}$

Why?

- A joint Gaussian distribution is completely characterized by the first two moments. Therefore, when the first two moments remain invariant, the joint pdf also remains invariant.

Non-stationarities

Just as with non-linearities, there are different types of non-stationarities, for e.g., mean non-stationarity, variance non-stationarity, and so on. It is useful to categorize them, for working purposes, into two classes:

1. **Deterministic:** Polynomial trend, variance non-stationarity, periodic, etc.
2. **Stochastic:** Integrating type, i.e., random walk, heteroskedastic processes, etc.

We shall be particularly interested in two types of non-stationarities, namely, *trend-type* and *random walk* or *integrating* type non-stationarities.

Trend-type non-stationarity

A suitable mathematical model for such a process is,

$$x[k] = \mu_k + w[k] \tag{6}$$

where μ_k is a polynomial function of time and $w[k]$ is a stationary process.

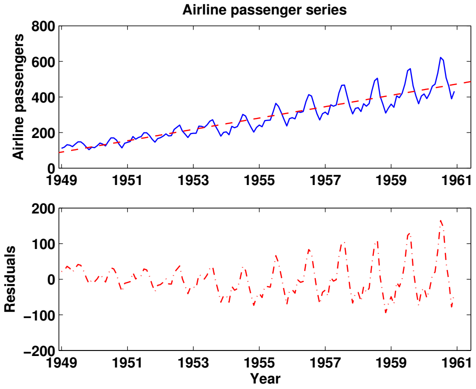
For example, a linear trend is modeled as $\mu_k = a + bk$.

Trend type non-stationarity ...contd.

- When the removal of a trend (e.g., linear ,quadratic) results in stationary residuals, the process is said to be **trend non-stationary.**
- Trends may be removed by either fitting a suitable polynomial (parametric approach) or by applying an appropriate smoothing filter (non-parametric approach).

Example: Trend-type non-stationarity

- A linear trend is fit to the series.
- Residuals show variance type nonstationarity. This is typical of a growth series.
- Two approaches:
    1. Trend fit + transformation OR
    2. Advanced models known as GARCH (generalized auto-regressive conditional heteroskedastic) models.



Monthly airline passenger series.

Integrating type non-stationarity

One of the most commonly encountered non-stationary processes is the *random walk* process (special case of Brownian motion).

The simplest random walk process is an *integrating* process,

$$x[k] = \sum_{n=0}^{k} e[k] \tag{7}$$

where $e[k]$ is the (unpredictable) *white-noise* affecting the process at the k^{th} instant.

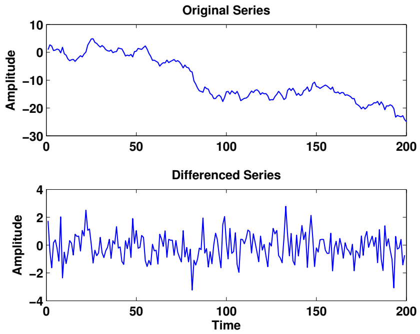
Integrating processes

- At any instant the signal is the accumulation of all shock-wave like changes from the beginning. Hence the name *integrating*.
- It is also known as a difference stationary process because

$$x[k] - x[k-1] = e[k] \tag{8}$$

Example

- Non-stationarity can be easily discerned by a visual inspection.
- The differenced series appears to be stationary.
- In general, a single degree of differencing is capable of removing a linear trend, two degrees removes quadratic trends and so on.



Top panel: $N = 200$ samples of a random walk series. Bottom plot: differenced series.

Caution

Despite its capability in handling a wide range of stationarities, the differencing approach also has potentially a few detrimental effects.

- Excessive or unnecessary differencing can lead to spurious correlations in the series.
- Amplification of noise, i.e., decrease in SNR in system identification.

Ergodicity

When the process is stationary, the second requirement on the time-series stems from the fact that in practice we work with only a single record of data.

Estimates computed from averages of time (or other domain) samples should serve as suitable representatives of the theoretical statistical properties, which are defined as averages in the outcome space (**ensemble**).

Ergodicity: Formal statement

Ergodicity

A process is said to be **ergodic** if the (time averaged) estimate **converges** to the true value (statistical average) when the number of observations $N \to \infty$.

Examples

1. A stationary i.i.d. random process is ergodic (by the strong LLN)

$$\frac{1}{N}\sum_{k=1}^{N} x[k] \xrightarrow{a.s.} E(X_k) \qquad \text{as } N \to \infty \tag{9}$$

2. A process such that $x[k] = A,\ \forall k,$ s.t.$E(A) = 0.$ Is it ergodic?

Remarks

- We can speak of ergodicity only when the process is stationary!
- *Loose interpretation:* **given sufficient time, the process would have unravelled nearly all possibilities that exist at any time instant (regardless of the starting point)**,
- Ergodicity is not necessarily a characteristic of the process, but can also be of the experiment that is carried out to obtain the time-series.
- Ergodicity is difficult to verify in practice; however, can be ensured by a careful experimentation, particularly through a proper selection and configuration of sensors and instrumentation.