LECTURES 2 - 3 : Stochastic Processes, Autocorrelation function. Stationarity.

Important points of Lecture 1:

A time series $\{X_t\}$ is a series of observations taken sequentially over time: x_t is an observation recorded at a specific time t.

Characteristics of times series data: observations are *dependent*, become available at *equally spaced time points* and are *time-ordered*. This is a *discrete time series.*

The purposes of time series analysis are to **model** and **to predict or forecast** future values of a series based on the history of that series.

2.2 Some descriptive techniques. (Based on [BD] §1.3 and §1.4)

. . . .

Take a step backwards: how do we describe a r.v. or a random vector?
• for a r.v. X:
d.f. $F_X(x) := P(X \leq x)$, mean $\mu = EX$ and variance $\sigma^2 = Var(X)$.

• for a r.vector (X_1, X_2):
joint d.f. $F_{X_1,X_2}(x_1, x_2) := P(X_1 \leq x_1, X_2 \leq x_2)$,
marginal d.f.$F_{X_1}(x_1) := P(X_1 \leq x_1) \equiv F_{X_1,X_2}(x_1, \infty)$
mean vector $(\mu_1, \mu_2) = (EX_1, EX_2)$, variances $\sigma_1^2 = Var(X_1)$, $\sigma_2^2 = Var(X_2)$, and
covariance $Cov(X_1, X_2) = E(X_1 - \mu_1)(X_2 - \mu_2) \equiv E(X_1X_2) - \mu_1\mu_2$.
Often we use correlation = normalized covariance:

$Cor(X_1, X_2) = Cov(X_1, X_2)/\{\sigma_1\sigma_2\}$

. . . .

To describe a process $X_1, X_2, \ldots$ we define

(i) **Def. Distribution function**: (fi-di) d.f.

$$F_{t_1 \ldots t_n}(x_1, \ldots, x_n) = P(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n),$$

i.e. this is the joint d.f. for the vector $(X_{t_1}, \ldots, X_{t_n})$.

(ii) **First- and Second-order moments.**

• **Mean**: $\mu_X(t) = EX_t$

• **Variance:** $\sigma_X^2(t) = E(X_t - \mu_X(t))^2 \equiv EX_t^2 - \mu_X(t)^2$

- **Autocovariance function:**

$$\gamma_X(t,s) = Cov(X_t, X_s) = E[(X_t - \mu_X(t))(X_s - \mu_X(s))] \equiv E(X_t X_s) - \mu_X(t)\mu_X(s)$$

(Note: this is an infinite matrix).

- **Autocorrelation function:**

$$\rho_X(t,s) = Cor(X_t, X_s) = \frac{Cov(X_t, X_s)}{\sqrt{Var(X_t)Var(X_s)}} = \frac{\gamma_X(t,s)}{\sigma_X(t)\sigma_X(s)}$$

Properties of the process which are determined by the first- and second- order moments are called second-order properties.
Example: stationarity. – will be defined soon!

2.3 **Meaning of the Autocorrelation function**:

In linear regression problem, PSTAT 5A : $Cor(X, Y)$ measures linear dependence.

Recall: correlation is zero if there is no linear dependence and it is close to ± 1 for variables with high linear dependence.

In TS: Autocorrelation function is used to assess numerically the dependence between two adjacent values.
Let say that for a moment we observe only two r. v.'s: X_t and X_{t+1}.

$\rho(t, t+1)$ is the correlation of X_t, X_{t+1}. The prefix *auto* is to convey the notion of self-correlation (both variables come from the same TS), correlation of the series with itself. When the TS is smooth, the autocorr. f'n is large even when t and s are quite apart, whereas very choppy series tend to have autocorr. f'ns that are nearly zero for large separations. *So, the autocorr. f'n tends to reflect the essential smoothness of the TS.*

Let us consider two "extreme" examples:

Example: Gaussian White Noise: Consider the series $\{Z_t\}$, where Z_t are Gaussian, m.z. r.v.'s and

$$\gamma_Z(s,t) = E(Z_t Z_s) = 1, \text{ if s = t }, \text{ and } 0, \text{ if } s \neq t$$

so that the autocov. f'n is 0 at all time separations. The series looks very choppy:

Example. Apply smoothing operation $X_t = \frac{1}{3}(Z_{t-1} + Z_t + Z_{t+1})$. Autocov. f'n is:

Autocor. f'n is a normalized autocov. f'n.

---

3. Stationarity

3.1 **Stationarity and Strict Stationarity** (Based on §1.4 and 2.1 of [BD])

Intuitively stationarity means that the graphs over two equal-length time intervals of a realization of the TS should exibit similar statistical characteristics.

On a graph: • no trend • no seasonality • no change of variability • no apparent sharp changes of behavior

Two approaches to stationarity:

(i) **Def.** A stochastic process is said to be **strictly stationary** if the joint probability density associated with the n r.v.'s $X_{t_1}, \ldots, X_{t_n}$ for any set of times $t_1, \ldots, t_n$, is the same as that associated with the n r.v.'s $X_{t_1+k}, \ldots, X_{t_n+k}$ for any integer k.

(ii) **Def.** A process is said to be (weakly, second-order) **stationary** if
(a) $E|X_t|^2 < \infty$
(b) $EX_t = \mu$ for all $t \in T$
(c) $\gamma_X(t,s) = \gamma_X(t+r, s+r) = \gamma_X(t-s)$ for all $t, s, r \in T$.

Note: The function $\gamma_X(k)$ is often refered as the value of the autocov. f'n at *lag* k.

(iii) The ACF (the autocorr. f'n) is

$$\rho_X(k) = \frac{\gamma_X(k)}{\gamma_X(0)}$$

(iv) Note: $\gamma_X(k) = \gamma_X(-k),\ \rho_X(k) = \rho_X(-k),\ \gamma_X(0) = \sigma_X^2$.

(v) Example: Apply smoothing operation to WN series:
$X_t = \frac{1}{3}(Z_{t-1} + Z_t + Z_{t+1})$. Autocov. f'n is:

(vi) Example: $X_t = Z_1 + \ldots + Z_t$, where Z_t are i.i.d. r.v.s with m.z. and variance σ^2.

Non-stationary.

In general: • strict stationarity + finite second moment imply stationarity

In general: • stationarity does not imply strict stationarity

Fact: For Gaussian TS *stationarity* = *strict stationarity*.
(because Gaussian distribution is determined by its mean and covariance)