Lecture 2-3

1 Main Definitions

Definition 1: The mean function of a real-valued time series is the function

$$\mu_X : T \to \mathbb{R}$$

defined by:

$$\mu_X(t) \equiv E[X_t]$$

The mean function is one of the features of a time series model that allows us to understand how the marginal distributions of a time series changes over time.

Definition 2: Two random variables X, Y with a joint p.d.f. $f(x, y)$ are independent if the joint p.d.f. f can be written as:

$$f(x, y) = f_1(x) f_2(y)$$

where f_1 and f_2 are two non-negative functions that integrate to one. Two random variables are said to be *dependent* if they are not independent.

Here are some properties of independent random variables:

Claim: For any sets A, B

$$\mathbb{P}(X \in A \text{ and } Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Proof:

$$\mathbb{P}(X \in A \text{ and } Y \in B) = \int_A \left(\int_B f(x, y) dy \right) dx$$

$$\begin{aligned}
&= \int_A \left(\int_B f_1(x) f_2(y) dy \right) dx \\
&= \int_A \left(\int_B f_2(y) dy \right) f_1(x) dx \\
&= \mathbb{P}(X \in A)\mathbb{P}(Y \in B).
\end{aligned}$$

This means that the product of any joint event, can be written as the product of the marginal events.

Corollary: $\mathbb{P}(X \in A | Y \in B) = \mathbb{P}(X \in A)$

How do we know if two random variables are dependent? Well, based on the definition above this means that all we need to do is find a set A and a set B for which:

$$\mathbb{P}(X \in A | Y \in B) \neq \mathbb{P}(X \in A)$$

Computing conditional probabilities is difficult. But here is a trick we can use. Remember that for any two random variables X and Y the covariance is defined as:

$$\mathrm{Cov}(X, Y) \equiv E[(X - \mu_x)(Y - \mu_y)]$$

Claim: If (X, Y) are independent then $\mathrm{Cov}(X, Y) = 0$.

Proof: For any functions g, h:

$$\begin{aligned}
E[g(x)h(y)] &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x)h(y) f(x, y) dx \right) dy \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x)h(y) f_1(x) f_2(y) dx \right) dy \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} g(x) f_1(x) dx \right) h(y) f_2(y) dy \\
&= E[g(x)]E[h(y)]
\end{aligned}$$

Taking $g(x) = x - \mu_x$ completes the proof, as $E[x - \mu_x] = 0$.

This means that in order to show that two random variables are not independent, it suffices to find a non-zero covariance. Thus, the covariance can give us some superficial and preliminary glimpse into the dependence embedded in a time series model.

Definition 3: The covariance function of a real-valued time series is the function:

$$\gamma_X : T \times T \to \mathbb{R}$$

defined by:

$$\gamma_X(s,t) \equiv \mathrm{Cov}(X_s, X_t)$$

The autocorrelation function is simply defined as

$$\rho_X(r,s) = \gamma_X(r,s) / \sqrt{\gamma_X(r,r)}\sqrt{\gamma_X(s,s)}$$

Let us compute the covariance function for the cosine model. Note that:

$$\begin{aligned}
\gamma_X(r,s) &= \mathbb{E}[(X_r - \mu_X(r))(X_s - \mu_X(s))] \\
&= E[\varepsilon_r \varepsilon_s] \\
&= 0
\end{aligned}$$

In general, zero covariance does not imply independence (remember this?). However, if the joint distribution is normal, then it does.

Complement these notes with Section 1.4 in the book.

Definition 4: A times series model is weakly stationary if

i) The mean function does not depend on t

ii) The covariance function $\gamma_X(t, t+h)$ is independent of t for each h.

Comment: Therefore in a stationary model we can write the autocovariance function in terms of h only:

$$\gamma_X(h) \equiv \gamma_X(t+h, t) = E[(X_{t+h} - \mu)(X_t - \mu)]$$

We will refer to $\gamma_X(h)$ **as the h-th order autocovariance.**