Discrete-Time Stationary Stochastic Processes
Lecture Notes

Marco Lippi
Dipartimento di Scienze Economiche
Università di Roma "La Sapienza"

April 23, 2006

Introduction

These Lecture Notes assume that the reader is already acquainted with the basic notions on stochastic processes and stationarity. The present text does not pretend to be self-contained. Rather, I will try to provide a lively presentation, just like in a lecture course. More precisely, the Notes contain:

1. Motivation of the definitions by detailed discussion of examples and counterexamples.
2. A guide to the proof of the main results. Again, examples and counterexamples will be very often used as substitutes of rigorous proofs, the latter being left to the readers to work out themselves or to look up in reference texts.
3. Links between results obtained by different tools or approaches.
4. A guide to the mathematical results that are necessary for a rigorous understanding of the content of the lectures.
5. References to books and articles.

The subject index of the Lecture Notes is, for the moment,

a. The spectral representation of wide sense stationary processes.

b. Linear filtering.

c. Linear prediction and the Wold representation.

d. Obtaining the Wold representation from the spectral density.

d. Fundamental and non-fundamental representations.

Chapter 1

Definitions

1.1 Stochastic Processes. Kolmogorov's Theorem

These Lecture Notes mainly deal with discrete-time stochastic processes, and only occasionally discuss continuous time. A *discrete-time stochastic process* is a family of stochastic variables parameterized on the set of integer numbers $\mathbb{Z}$:

$$x = \{x_t,\ t \in \mathbb{Z}\},$$

and a set of probability measures

$$\mu_{t_1 t_2 \ldots t_n}(H) = P\left[(x_{t_1},\ x_{t_2}, \ldots,\ x_{t_n}) \in H\right],$$

for any finite set of integers

$$t_1 < t_2 < \cdots < t_n,$$

where H is a Borel subset of $\mathbb{R}^n$.

It is important to keep in mind that the "family $\{x_t,\ t \in \mathbb{Z}\}$" has to be understood as the *function* associating the stochastic variable x_t with the integer t. Therefore the processes $x = \{x_t,\ t \in \mathbb{Z}\}$, $y = \{x_{-t},\ t \in \mathbb{Z}\}$, $z = \{x_{t-4},\ t \in \mathbb{Z}\}$ are different. Although they share the same *range*, i.e. the the same *set* of stochastic variables, the functions associating a stochastic variable with each integer t are different. For brevity, when no confusion can arise, we use expressions like "the stochastic process $\{x_t\}$", or even "the stochastic process x_t", instead of the correct expression "the stochastic process $\{x_t,\ t \in \mathbb{Z}\}$", or "the stochastic processes x".

Obviously the measures μ must fulfill a consistency condition, that is, given $t_1 < t_2 < \cdots < t_n$ and $n-1$ measurable subsets of $\mathbb{R}$, $H_1,\ H_2,\ \ldots,\ H_{n-1}$,

$$\begin{aligned}\mu_{t_1 \ldots t_n}(H_1 \times \cdots \times H_s \times \mathbb{R} \times H_{s+1} \times \cdots \times H_{n-1}) \\ = \mu_{t_1 \ldots t_{s-1} t_{s+1} \ldots t_n}(H_1 \times \cdots \times H_{n-1}).\end{aligned} \tag{1.1}$$

The following is the most familiar example of a stochastic process.

Example 1.1 At time t a coin is tossed. Outcomes at different times are assumed to be independent, while the outcomes at time t are equiprobable. The stochastic variable x_t is defined as being 1 if the outcome at time t is "Head", 0 if "Tail" ($\mathcal{H}$ and $\mathcal{T}$ henceforth). Independence implies that

$$\mu_{t_1 t_2 \ldots t_n}(H_1 \times \cdots \times H_n) = \mu(H_1) \cdots \mu(H_n). \tag{1.2}$$

Example 1.2 Define $y_t = x_{t-1} + x_t + x_{t+1}$, where x_t is defined in Example 1.1. The sample space of y_t is $\Omega_{t-1} \times \Omega_t \times \Omega_{t+1}$. Show that the distributions μ corresponding to $t_1, \ldots, t_n$ and $t_1 + k, \ldots, t_n + k$ are equal for any integer k.

Given a stochastic process x we are interested in other stochastic processes that are defined as functions of the variables of x, linear functions in particular. Given x, the process

$$y_t = \sum_{k=-m}^{m} a_k x_{t-k}, \tag{1.3}$$

with coefficients independent of t, is well known as a *finite moving average* of x. Example 1.2 is a finite moving average, one that has a smoothing effect (this will be discussed later on, with linear filters). We are also interested in processes that are implicitly defined as solutions of stochastic difference equations . The simplest example is the following:

$$y_t = \alpha y_{t-1} + x_t, \tag{1.4}$$

where x_t is a stochastic process. If $|\alpha| < 1$, a natural candidate for a solution is

$$y_t^* = x_t + \alpha x_{t-1} + \alpha^2 x_{t-2} + \cdots. \tag{1.5}$$

However, this *infinite moving average* does not make sense unless we (1) define its sample space and (2) specify the kind of convergence in (1.5) (in probability, in mean square, almost surely).

Let us discuss this point a little further. Assume that in (1.4) x_t is the process defined in Example 1.1. The variable x_t has sample space

$$\Omega_t = \{\mathcal{H}_t, \mathcal{T}_t\},$$

i.e. the set whose elements are "$\mathcal{H}$ at time t", and "$\mathcal{T}$ at time t", or, if you prefer that different coins C_t be used at different times, "$\mathcal{H}$ with coin C_t' and "$\mathcal{T}$ with coin C_t", the probability P_t being $1/2$ for both sample points. The infinite sum in

(1.5), if existing, would be the limit, for $k \to \infty$, of the sum $S_{kt} = x_t + \alpha x_{t-1} + \cdots + \alpha^k x_{t-k}$, whose sample space is

$$\Omega_t \times \Omega_{t-1} \times \cdots \times \Omega_{t-k},$$

with the product probability (1.2).

We may easily check that the variance of the difference between S_{kt} and S_{ht} becomes negligible when h and k tend to infinity, this being the Cauchy condition for convergence in mean square. For, independence of the variables x_t implies that, assuming $k > h$,

$$\mathrm{var}(S_{kt} - S_{ht}) = \mathrm{var}(x_t)\alpha^{h+1}\frac{1-\alpha^{2(k-h)}}{1-\alpha^2},$$

which converges to zero as $h \to \infty$ (note that $\mathrm{var}(x_t)$ is independent of t). However, the mean-square Cauchy condition is necessary for convergence in mean square, not sufficient in general. Moreover, the limit, if existing, should be a stochastic variable defined on the infinite product

$$\Omega_t \times \Omega_{t-1} \times \cdots \times \Omega_{t-k} \times \cdots,$$

and the latter should therefore be endowed with a σ-field and a probability measure. The situation is fairly uncomfortable, the solution of a simple problem like equation (1.4) requiring the introduction of a new, infinite-dimensional probability space.

This difficulty can be overcome by an elegant re-definition of the process x in such a way that the stochastic variables are defined all on the same probability space. This construction is known as the Kolmogorov's Existence Theorem.

We do not give a proof of the theorem. The following is an illustration based on the tossing-coin example. Consider the infinite product space

$$\Omega^{\mathbb{Z}} = \cdots \times \Omega_{t-1} \times \Omega_t \times \Omega_{t+1} \times \cdots,$$

whose elements are the bilateral sequences $\{r_t,\ r_t \in \Omega_t\}$. Now, for given $\tau \in \mathbb{Z}$ and $A_\tau \subseteq \Omega_\tau$, consider

$$K(\tau, A_\tau) = \{r \in \Omega^{\mathbb{Z}},\ r_\tau \in A_\tau\},$$

i.e. the set of all bilateral sequences of $\mathcal{H}$ and $\mathcal{T}$ such that at $t = \tau$ the outcome belongs to A_τ. For example, if $\tau = 1$ and $A_1 = \{\mathcal{H}_1\}$, then $K(1, A_1)$ is the set of all sequences that have an $\mathcal{H}$ at $t = 1$. Then consider the σ-field $\mathcal{G}$ generated by all the sets K. Firstly, observe that all finite-dimensional events have their "copies" in $\mathcal{G}$. For example, $\mathcal{H}$ at times $t_1, \ldots, t_m$, corresponds to the set of all

$r \in \Omega^{\mathbb{Z}}$ such that $r_{t_j} = \mathcal{H}_{t_j}$ for $j = 1, \ldots, m$. However, there are also infinite dimensional events in $\mathcal{G}$. For example, the countable union of the sets "$\mathcal{H}$ at time t", for all $t \in \mathbb{Z}$, gives the set "$\mathcal{H}$ at least once", whose complement, also in $\mathcal{G}$, is the singleton "$\mathcal{T}$ for all t". The probability of the sets K is defined as

$$P(K(t, A_t)) = P_t(A_t).$$

This law is then extended to $\mathcal{G}$. It is quite obvious that such probability is consistent with the probability (1.2) on the finite dimensional sets $\Omega_{t_1} \times \cdots \times \Omega_{t_n}$. It is also consistent with fairly elementary limit results; for example, the probability of the set "$\mathcal{T}$ for all t" is zero (for that matter, the probability of any elementary event in $\Omega^{\mathbb{Z}}$ is zero). In conclusion, we may say that $(\Omega^{\mathbb{Z}}, \mathcal{G}, P)$ is an extension of the set of all probability spaces $\Omega_{t_1} \times \cdots \times \Omega_{t_n}$, each endowed with its product probability.

Lastly, we define copies of the stochastic variables $x_t : \Omega_t \to \mathbb{R}$. Precisely, let $Z_t : \Omega^{\mathbb{Z}} \to \mathbb{R}$ be defined by $Z_t(r) = x_t(r_t)$. Given the sequence r, Z_t takes on the value 1 if r_t is $\mathcal{H}$, 0 otherwise. Thus Z_t does exactly what x_t does. Z_t is called the t-th coordinate variable. The distribution of Z_t is the same as that of x_t, 1 with probability $1/2$ and zero with probability $1/2$. Moreover, the probability measure of $(Z_{t_1}, \ldots, Z_{t_m})$ is the same as that of $(x_{t_1}, \ldots, x_{t_m})$. The big difference, the result of the construction above, is that the variables Z_t are all defined on *the same probability space.* Equation (1.4) can be rewritten as

$$y_t = \alpha y_{t-1} + Z_t,$$

and the tentative solution as

$$Y_t^* = Z_t + \alpha Z_{t-1} + \alpha^2 Z_{t-2} + \cdots . \tag{1.6}$$

Still we need a proof of convergence for the partial sums $\sum_{j=0}^{k} \alpha^j Z_{t-j}$, but they are now all defined on the same probability space. A consequence of the results reported in Section 1.3.9 is that the Cauchy condition is also sufficient in this case, see Observation 1.10 in particular. The sequence S_{kt} also converges surely, see Exercise 1.1 below.

Exercise 1.1 Prove that $\sum_{j=0}^{k} \alpha^j Z_{t-j}$ converges *surely* (not only *almost surely*), i.e. for any $r \in \Omega^{\mathbb{Z}}$, so that (1.6) defines the stochastic variable Y_t^*. Prove that $\sum_{j=0}^{k} \alpha^j Z_{t-j}$ also converges to Y_t^* in mean square.

Exercise 1.2 Discuss the distribution of the stochastic variable Y_t^*. Show firstly that it is independent of t. The Matlab program

```
function y=autor(ALPHA,OBS,REPL)
a=floor(2*rand(OBS,REPL));
% 2*rand generates a uniform
% stochastic variable in (0 2).
% Taking floor we obtain 0 if the uniform
% lies  in (0 1), 1 if it lies in (1 2).
% The matrix a  has OBS rows and REPL columns
y=(ALPHA.^(0:(OBS-1)))*a;
% ALPHA.^(0:(OBS-1))  has 1 row and OBS columns,
% so that y has 1 row and REPL columns.
```

computes, for a given ALPHA, a number REPL of replications of Y_t^* by generating sequences

$$Z_t,\ Z_{t-1},\ \ldots,\ Z_{t-\mathrm{OBS}+1}$$

(the series Y_t^* is truncated at $\mathrm{ALPHA}^{\mathrm{OBS}-1} Z_{t-\mathrm{OBS}+1}$). The output y contains the replications. Draw the histogram by hist(y,100), and comment on the result.

Of course there exist stochastic processes, whose variables are defined directly on the same probability space, so that the construction sketched above is not necessary.

Example 1.3 Let ϕ be a given real number, and a and b stochastic variables on the probability space $(\Omega, \mathcal{F}, P)$, and let

$$x_t = a\cos\phi t + b\sin\phi t.$$

What is stochastic in x_t does not depend on t. In this case values for a and b are drawn, so to speak, before the beginning of time. When time "starts going" no further stochastic events occur. As particular cases we have $x_t = a$, the *constant process*, for $\phi = 0$, and $x_t = (-1)^t a$, for $\phi = \pi$.

Example 1.4 Same as in Example 1.3, but now ϕ also is a stochastic variable. Same considerations apply.

The ideas illustrated by means of the tossing-coin example possess a rigorous and general formulation in the Kolmogorov's Existence Theorem. Given a stochastic process $x = \{x_t,\ t \in \mathbb{Z}\}$, with probability measures $\mu_{t_1 t_2 \cdots t_n}$, fulfilling the consistency condition (1.1), an "equivalent" process can be constructed on the set $\mathbb{R}^{\mathbb{Z}}$. Precisely, define $\Omega = \mathbb{R}^{\mathbb{Z}}$ and $\mathcal{F}$ as the σ-field generated by the subsets

$$K(t_1, t_2,\ \ldots\ , t_n, A_1, A_2,\ \ldots,\ A_n) = \{r \in \mathbb{R}^{\mathbb{Z}},\ r_{t_s} \in A_s,\ s = 1, 2,\ \ldots,\ n\},$$

where t is any integer and A_s a Borel subset of $\mathbb{R}$; then define the probability measure on Ω extending $\mu_{t_1 t_2 \cdots t_n}$ on the whole $\mathcal{F}$. Lastly, define y_t as the t-th coordinate variable Z_t, which is defined by $Z_t(r) = r_t$, for all $r \in \mathbb{R}^{\mathbb{Z}}$. (Here the infinite dimensional space is constructed on $\mathbb{R}^{\mathbb{Z}}$, while in the tossing-coin example we have used the product of the spaces Ω_t, for $t \in \mathbb{Z}$; of course the results are equivalent.)

Thus, with no loss of generality, the definition of a stochastic process can be restated as follows:

Definition 1.1 A discrete time stochastic process is a family of stochastic variables $\{x_t,\ t \in \mathbb{Z}\}$, where x_t belongs to the set of all stochastic variables defined on a probability space $(\Omega, \mathcal{F}, P)$.

Remark that the construction of the equivalent process $\{Z_t,\ t \in \mathbb{Z}\}$ on $\mathbb{R}^{\mathbb{Z}}$ can be done irrespectively of the original definition of $\{x_t,\ t \in \mathbb{Z}\}$, thus even when all the x's are directly defined on the same probability space. The sample points of $\mathbb{R}^{\mathbb{Z}}$ are called *sample sequences* (in the continuous-time case the term *sample function* is used), or *realizations* of the process x. If the set Ω, whereon the variables x_t are defined, coincides with $\mathbb{R}^{\mathbb{Z}}$, then the realizations of the process x and the ω's are the same objects. Note also that realizations belong to a probability space, so that probability statements apply to sets of realizations. For example, the realizations of the processes in Examples 1.3 and 1.4 are bounded with probability 1, the realizations of the process in Example 1.1 have an infinite number of 1's with probability 1.

Summary. We start with the definition of a stochastic process as a collection of stochastic variables x_t, each on its own probability space, together with the joint probability measures of vectors $(x_{t_1} \ \cdots \ x_{t_n})$. We show that even in the simple case of a linear difference equation of order one the solution requires the definition of a stochastic variable on the infinite-dimensional product of the spaces Ω_t. The Kolmogorov's Theorem proves that there exists a process, equivalent to the process x, whose variables are all defined on the same probability space. Thus we can use Definition 1.1. On the Kolmogorov's Theorem see [4, pp. 506-517], containing a general and detailed proof, and [7, pp. 9-12].