CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

1

Probability Theory

1.1 Introduction

In this chapter we introduce probability theory using a measure-theoretic approach. There are two main subjects that are closely related to economics. First, the concept of σ-algebra is closely related to the notion of information set used widely in economics. We shall formalize it. Second, the concepts of conditional probability and conditional expectation are defined in terms of the underlying σ-algebra. These are background materials for understanding Wiener processes and stochastic dynamic programming.

We keep proofs to the bare minimum. In their place, we emphasize the intuition so that the reader can gain some insights into the subject matter. In fact, we shall go over many commonly employed theorems on conditional expectation with intuitive explanations.

1.2 Stochastic Processes

1.2.1 Information Sets and σ-Algebras

Let Ω be a point set. In probability theory, it is the set of elementary events. The power set of Ω, denoted by 2^Ω, is the set of all subsets of Ω. For example, if the experiment is tossing a coin twice, then the set Ω is $\{HH, HT, TH, TT\}$. It is easy to write down all $2^4 = 16$ elements in the power set. Specifically,

$$\begin{aligned}2^\Omega = \{&\varnothing, \{HH\}, \{HT\}, \{TH\}, \{TT\}, \{HH, HT\}, \{HH, TH\},\\ &\{HH, TT\}, \{HT, TH\}, \{HT, TT\}, \{TH, TT\}, \{HH, HT, TH\},\\ &\{HH, HT, TT\}, \{HH, TH, TT\}, \{HT, TH, TT\}, \Omega\}.\end{aligned}$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

In general, the cardinality of the power set is $2^{|\Omega|}$, where $|\Omega|$ is the cardinality of the set Ω. Power sets are very large. To convince yourself, let the experiment be rolling a die twice, a rather simple experiment. In this simple experiment, $|\Omega| = 36$ and the cardinality of the power set is $2^{36} = 6.87 \times 10^{10}$. It would be impractical to write down all elements in this power set. What we are interested in is subsets of the power set with certain structure.

Definition 1.1 *A class $\mathcal{F}$ of subsets of Ω, i.e., $\mathcal{F} \subset 2^{\Omega}$, is an algebra (or a field) if:*

(i) $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$, where A^c is the complement of A in Ω.
(ii) $A, B \in \mathcal{F}$ imply that $A \cup B \in \mathcal{F}$.
(iii) $\Omega \in \mathcal{F}$ (equivalently, $\varnothing \in \mathcal{F}$).

Conditions (i) and (ii) imply $A \cap B \in \mathcal{F}$, because $A \cap B = (A^c \cup B^c)^c$.

Definition 1.2 *A class $\mathcal{F}$ of subsets of Ω is a σ-algebra if it is an algebra satisfying*

(iv) if $A_i \in \mathcal{F}$, $i = 1, 2, \ldots$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

The Greek letter "σ" simply indicates that the number of sets forming the union is *countable* (including finite numbers).

Any $A \in \mathcal{F}$ is called a *measurable* set, or simply, an $\mathcal{F}$-set. We use $\mathcal{F}$ to represent the *information set*, because it captures our economic intuition. Conditions (i) through (iv) provide a mathematical structure for an information set.

Intuitively, we can treat a measurable set as an *observable set*. An object under study ($\omega \in \Omega$) is observable if we can detect that it has certain characteristics. For example, let Ω be the set of flying objects and let A be the set of flying objects that are green. Then A^c represents the set of all flying objects that are not green. Condition (i) simply says that if, in our information set, we can observe that a flying object is green (i.e., A is observable), then we should be able to observe that other flying objects are not green. That means A^c is also observable. Another example is this: if we were able to observe when the general price level is rising, then we should be able to observe when the general price level is not rising. Formally, if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

Condition (ii) says that, if we can observe the things or objects described by characteristics A and those described by characteristics B, then we should be able to observe the objects characterized by the properties of A or B. That is, $A, B \in \mathcal{F}$ imply $A \cup B \in \mathcal{F}$. For example, if we are able to observe when the price level is rising, and if we are able to observe the unemployment level is rising, then we should be able to observe the rising of price level *or* rising unemployment. The same argument applies to countably many observable sets, which is condition (iv). These mathematical structures make σ-algebras very suitable for representing information.

It is clear that the power set 2^{Ω} is itself a σ-algebra. But there are lots of σ-algebras that are smaller than the power set. For example, in the experiment of tossing a coin twice, $\mathcal{F}_1 = \{\Omega, \varnothing, \{HH\}, \{HT, TH, TT\}\}$ and $\mathcal{F}_2 = \{\Omega, \varnothing, \{HH, TT\}, \{HT, TH\}\}$ are both algebras. The information content of $\mathcal{F}_1$ is this: we can tell whether tossing a coin twice ends up with both heads or otherwise. The information content of $\mathcal{F}_2$ is this: we can tell whether both tosses have the same outcome or not. The reader should try to find other algebras in this setup. An obvious one is to "combine" $\mathcal{F}_1$ and $\mathcal{F}_2$. See the exercise below. We will return to these two examples in Example 1.12.

Exercise 1.2.1

(1) Verify that $\mathcal{F}_1$ and $\mathcal{F}_2$ are algebras.

(2) Show that $\mathcal{F}_1 \cup \mathcal{F}_2$, while containing $\mathcal{F}_1$ and $\mathcal{F}_2$, is not an algebra.

(3) Find the smallest algebra $\mathcal{G}$ that contains $\mathcal{F}_1$ and $\mathcal{F}_2$ in the sense that for any algebra $\mathcal{H}$ which contains $\mathcal{F}_1$ and $\mathcal{F}_2$, then $\mathcal{G} \subset \mathcal{H}$.

Definition 1.3 *A set function $P : \mathcal{F} \to \mathbb{R}$ is a probability measure if P satisfies*

(i) $0 \leq P(A) \leq 1$ for all $A \in \mathcal{F}$;

(ii) $P(\varnothing) = 0$ and $P(\Omega) = 1$;

(iii) if $A_i \in \mathcal{F}$ and the A_i's are mutually disjoint, then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

Property (iii) is called *countable additivity*. The triplet $(\Omega, \mathcal{F}, P)$ is used to denote a probability space.

Example 1.4 *(Borel Sets and Lebesgue Measure) When $\Omega = \mathbb{R}$ (the whole real line) or $\Omega = [0, 1]$ (the unit interval), and the σ-algebra*

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

is the one generated by the open sets in $\mathbb{R}$ *(or in* $[0, 1]$*), we call this* σ*-field the Borel field. It is usually denoted by* $\mathcal{B}$*. An element in the Borel field is a Borel set.*

Examples of Borel sets are open sets, closed sets, semi-open, semi-closed sets, F_σ *sets (countable unions of closed sets), and* G_δ *sets (countable intersections of open sets). When* $\Omega = [0, 1]$*,* $\mathcal{B}$ *is the* σ*-algebra, and* $P(A)$ *is the "length" (measure) of* $A \in \mathcal{F}$*, we can verify that* P *is a probability measure on* $\mathcal{B}$*. Such a measure is called the Lebesgue measure on* $[0, 1]$.

However, not all subsets of $\mathbb{R}$ *are Borel sets, i.e., not all subsets of* $\mathbb{R}$ *are observable. For example, the Vitali set is not a Borel set. See, for example, Reed and Simon (1972, p. 33). For curious souls, the Vitali set* V *is constructed as follows. Call two numbers* $x, y \in [0, 1)$ *equivalent if* $x - y$ *is rational. Let* V *be the set consists of exactly one number from each equivalent class. Then* V *is not Lebesgue measurable.*

A single point and, therefore, any set composed of countably many points are of Lebesgue measure zero. The question then is this: Are sets with uncountably many points necessarily of positive Lebesgue measure? The answer is negative, and the best-known example is the Cantor set.

1.2.2 The Cantor Set

Since the Cantor set contains many important properties that are essential to understanding the nature of a Wiener process, we shall elaborate on this celebrated set. The construction of the Cantor set proceeds as follows. Evenly divide the unit interval $[0, 1]$ into three subintervals. Remove the middle *open* interval, $(1/3, 2/3)$, from $[0, 1]$. The remaining two closed intervals are $[0, 1/3]$ and $[2/3, 1]$. Then remove the two middle open intervals, $(1/9, 2/9)$ and $(7/9, 8/9)$, from $[0, 1/3]$ and $[2/3, 1]$ respectively. Continue to remove the four middle open intervals from the remaining four closed intervals, $[0, 1/9]$, $[2/9, 1/3]$, $[2/3, 7/9]$, and $[8/9, 1]$, and so on indefinitely. The set of points that are not removed is called the *Cantor set*, $\mathfrak{C}$.

Any point in the Cantor set can be represented by

$$\sum_{n=1}^{\infty} \frac{i_n}{3^n}, \qquad \text{where } i_n = 0 \text{ or } 2.$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

For example,

$$\frac{7}{9} = \frac{2}{3} + \frac{0}{9} + \frac{2}{27} + \frac{2}{81} + \frac{2}{243} + \cdots,$$

i.e., $7/9 = (2, 0, 2, 2, 2, \ldots)$. Similarly, $8/9 = (2, 2, 0, 0, \ldots)$, $0 = (0, 0, 0, \ldots)$, $8/27 = (0, 2, 2, 0, 0, 0, \ldots)$, and $1 = (2, 2, 2, \ldots)$. Therefore, the cardinality of the Cantor set is that of the continuum. Since the Lebesgue measure of the intervals removed through this process is

$$\frac{1}{3} + \frac{1}{9} \cdot 2 + \frac{1}{27} \cdot 4 + \cdots = \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = 1,$$

the Cantor set must be of Lebesgue measure zero.

The main properties that are of interest to us are three. First, here is a set with uncountably many elements that has a zero Lebesgue measure. Second, every point in the Cantor set can be approached by a sequence of subintervals that were removed. In other words, every point in the Cantor set is a limit point. Such a set is called a *perfect set*. Third, for any interval $I \subset [0, 1]$, it must contain some subinterval that was eventually removed, i.e., we can find a subinterval $J \subset I$ such that J and the Cantor set $\mathfrak{C}$ are disjoint: $J \cap \mathfrak{C} = \varnothing$. That is, $\mathfrak{C}$ is *nowhere dense* in [0, 1]. These three properties are the basic features of the zero set of a Wiener process, as we shall see later.

1.2.3 Borel–Cantelli Lemmas

Definition 1.5 *The limit superior and the limit inferior of a sequence of sets* $\{A_n\}$ *are*

$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k,$$

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k.$$

Simply put, $x \in \limsup_{n\to\infty} A_n$ means x belongs to infinitely many A_k. In contrast, $x \in \liminf_{n\to\infty} A_n$ means x belongs to virtually all A_k, in the sense that there exists N such that $x \in A_k$ for $k \geq N$. Since $\mathcal{F}$ is a

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

σ-algebra, $\limsup_{n\to\infty} A_n \in \mathcal{F}$ and $\liminf_{n\to\infty} A_n \in \mathcal{F}$ if $A_n \in \mathcal{F}$. By definition, $\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n$.

Exercise 1.2.2 *Let*

$$A_n = \begin{cases} \left[0, 1 + \frac{1}{n}\right] & \text{if } n \text{ is even,} \\ \left[0, 2 + \frac{1}{n}\right] & \text{if } n \text{ is odd.} \end{cases}$$

Show that $\liminf\limits_{n\to\infty} A_n = [0, 1]$ *and* $\limsup\limits_{n\to\infty} A_n = [0, 2]$.

Exercise 1.2.3 *Let*

$$A_n = \begin{cases} \left[0, \frac{1}{n}\right] & \text{if } n \text{ is even,} \\ \left[-\frac{1}{n}, 0\right] & \text{if } n \text{ is odd.} \end{cases}$$

Show that $\liminf\limits_{n\to\infty} A_n = \limsup\limits_{n\to\infty} A_n = \{0\}$.

Exercise 1.2.4 *Let A and B be subsets of Ω. Define*

$$A_n = \begin{cases} A & \text{if } n \text{ is even,} \\ B & \text{if } n \text{ is odd.} \end{cases}$$

Show that $\liminf\limits_{n\to\infty} A_n = A \cap B$, $\limsup\limits_{n\to\infty} A_n = A \cup B$.

Theorem 1.6 *Let $A_n \in \mathcal{F}$. Then we have the following inequalities:*

$$P\left(\liminf_{n\to\infty} A_n\right) \leq \liminf_{n\to\infty} P(A_n) \leq \limsup_{n\to\infty} P(A_n) \leq P\left(\limsup_{n\to\infty} A_n\right).$$

Proof. See Billingsley (1995, Theorem 4.1). ■

Corollary 1.7 *If* $\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A$, *then* $\lim_{n\to\infty} P(A_n) = P(A)$, *i.e., if* $A_n \to A$, *then* $P(A_n) \to P(A)$.

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

Theorem 1.8

(i) (The First Borel–Cantelli Lemma):

$$\sum_{n=1}^{\infty} P(A_n) < \infty \quad \Rightarrow \quad P\left(\limsup_{n\to\infty} A_n\right) = 0.$$

(ii) (The Second Borel–Cantelli Lemma) Suppose A_n are independent events. Then

$$\sum_{n=1}^{\infty} P(A_n) = \infty \quad \Rightarrow \quad P\left(\limsup_{n\to\infty} A_n\right) = 1.$$

Proof. See Billingsley (1995, Theorems 4.3 and 4.4). ■

Example 1.9 *(Value Loss Assumption and Asymptotic Stability of Optimal Growth) While the Borel–Cantelli lemmas may appear abstract, they have interesting economic applications. Recall that a standard model in first-year graduate macroeconomics is the discrete-time optimal growth problem: Let c_t, k_t, β, and $f(k)$ be, respectively, per capita consumption, capital–labor ratio, subjective discount factor, and production function. Then the problem is to find an optimal consumption program $\{c_t\}$ that solves*

$$\max_{\{c_t\}} \sum_{t=1}^{\infty} \beta^{t-1} u(c_t), \qquad \text{s.t. } k_t = f(k_{t-1}) - c_t, \quad t = 1, 2, \ldots, \text{ given } k_0 > 0.$$

The value loss assumption is employed to ensure the asymptotic stability of optimal growth. By asymptotic stability we mean that the difference between two optimal growth programs under two different initial stocks will converge to zero. To prove such a strong result (independent of the initial stock), a value loss assumption is employed. Specifically, we assign a minimum value loss $\delta > 0$ to each time period in which the two optimal programs are parted by a distance at least $\varepsilon > 0$. The value loss assumption makes it impossible to have infinitely many such periods (otherwise, the program is not optimal), so that the asymptotic stability is possible.

To extend this theorem to stochastic cases, the first Borel–Cantelli lemma comes in handy. Given $\varepsilon > 0$, let A_n be the set of points in Ω such that, at time n, the difference between the realizations of these two

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

optimal programs is at least $\varepsilon > 0$. *We shall assign for each* $\omega \in A_n$ *a minimal value loss* $\delta > 0$. *Then the* expected *value loss at time n is at least* $\delta P(A_n)$, *and the total expected value loss is at least* $\delta \sum_{n=1}^{\infty} P(A_n)$. *This being an optimal program, the expected value loss cannot be infinity. That is to say, the premise of the first Borel–Cantelli lemma is valid:* $\sum_{n=1}^{\infty} P(A_n) < \infty$. *It follows that* $P(\limsup_{n\to\infty} A_n) = 0$. *In words, the probability of those* $\omega \in \Omega$ *that belong to infinitely many* A_n *is zero. Thus, the time path of the difference between two optimal programs converges to zero with probability one. The interested reader is referred to Chang (1982) for more details.*

1.2.4 Distribution Functions and Stochastic Processes

Definition 1.10 *Let* $(\Omega, \mathcal{F}, P)$ *and* $(\Omega', \mathcal{F}', P')$ *be two probability spaces. Then* $T : (\Omega, \mathcal{F}, P) \to (\Omega', \mathcal{F}', P')$ *is measurable if* $T^{-1}(B) \in \mathcal{F}$ *for all* $B \in \mathcal{F}'$. *In particular, if* $\Omega' = \mathbb{R}$, $\mathcal{F}' = \mathcal{B}$ *(Borel field), and* P' *is Lebesgue measure, then* T *is a random variable.*

The term $T^{-1}(B)$ represents the *preimage* of B, or the *pullback* of set B. Recall that continuity of a mapping between two topological spaces is defined as follows: the pullback of any open set in the image space must be an open set in the domain. The concept of a measurable function is defined similarly, i.e., the pullback of a measurable set in the image space is a measurable set in the domain. A random variable is simply a special case of measurable functions.

Definition 1.11 *The distribution function of a random variable* X, *denoted by* $F(X)$, *is defined by*

$$F(x) = P(\{\omega \in \Omega : X(\omega) \le x\}) = P[X \le x]$$

(following Billingsley's (1995) notation) or, for any Borel set A,

$$F(A) = P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\}).$$

Clearly, $F(\cdot)$ is defined on Borel field with image in $\mathbb{R}$. Such a function is called a *Borel function*.

Given a random variable X, there are many σ-algebras that can make X measurable. For example, $\mathcal{F}$ is one. The σ-algebra generated by X,

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

denoted by $\sigma(X)$, is the smallest σ-algebra with respect to which X is measurable; that is, $\sigma(X)$ is the intersection of all σ-algebras with respect to which X is measurable. For a finite or infinite sequence of random variables $\{X_1, X_2, \ldots\}$, $\sigma(X_1, X_2, \ldots)$ is the smallest σ-algebra with respect to which each X_i is measurable.

Example 1.12 *(Tossing a coin twice) Let*

$$X_1 = \begin{cases} 0 & \text{if heads in both tosses,} \\ 1 & \text{otherwise,} \end{cases}$$

$$X_2 = \begin{cases} 0 & \text{if same occurrence in both tosses,} \\ 1 & \text{otherwise,} \end{cases} .$$

Then $\sigma(X_1) = \mathcal{F}_1 = \{\Omega, \varnothing, \{HH\}, \{HT, TH, TT\}\}$ *and* $\sigma(X_2) = \mathcal{F}_2 = \{\Omega, \varnothing, \{HH, TT\}, \{HT, TH\}\}$. *It is easy to show that* $\sigma(X_1, X_2) = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \{\{HH, HT, TH\}, \{TT\}\}$. *For example, if* $A = \{HH\} \in \mathcal{F}_1$ *and* $B = \{HH, TT\} \in \mathcal{F}_2$, *then* $A^c \cap B = \{TT\} \in \sigma(X_1, X_2)$. *In words, if the result is "same occurrence in both tosses" but not "heads in both tosses," then it must be "tails in both tosses."*

Exercise 1.2.5 *Let the experiment be rolling a die twice. Let X be the random variable of the sum of the two rolls.*

(1) Describe the probability space $(\Omega, \mathcal{F}, P)$, i.e., spell out Ω, find the smallest algebra $\mathcal{F}$ which makes X measurable, and let P be the usual probability.

(2) Write down the distribution function $F(x)$, and find $F(A)$ when $A = \{3, 4\}$.

(3) Let Y be the random variable that designates the larger of the two rolls. Repeat (1) and (2).

Hint. In (1), first find $X^{-1}(\{i\})$, $i = 2, 3, \ldots, 12$, and then choose $\mathcal{F} = \sigma(X) =$ the smallest algebra generated by X.

If the distribution function $F(x)$ has a derivative $f(x)$, which is called the density function of X, then $f(x)$ satisfies the equation

$$F(A) = \int_A f(x)\, dx.$$

© in this web service Cambridge University Press www.cambridge.org

CAMBRIDGE

Cambridge University Press
978-0-521-54194-7 - Stochastic Optimization in Continuous Time
Fwu-Ranq Chang
Excerpt
More information

Definition 1.13 *A stochastic process* $\{X(t) : t \in I\}$ *is a family of random variables, where* I *is the index set. If* $I = \mathbb{Z}$ *(integers), then* $\{X(t) : t \in I\}$ *is called a discrete (time) stochastic process, a time series, or simply a stochastic sequence. If* $I = [0, \infty)$ *or* $[0, 1]$, *then* $\{X(t) : t \in I\}$ *is called a continuous (time) stochastic process.*

A stochastic process can be thought of as a function defined on $I \times \Omega$, i.e., $X : I \times \Omega \to \mathbb{R}$ such that $X : (t, \omega) \mapsto X(t, \omega) = X_t(\omega)$. For a given $t \in I$, $X_t(\cdot) : \Omega \to \mathbb{R}$ is a random variable. For a given $\omega \in \Omega$, $X(\cdot, \omega) : I \to \mathbb{R}$ is a function mapping from the index set I to $\mathbb{R}$. Such a function is called a *sample function* (a *sample path*, a *realization*, or a *trajectory*). The range of the random variable X is called the *state space*, and the value $X(t, \omega)$ is called the *state* at time t for a given draw $\omega \in \Omega$. Thus, a sample function describes the states at different times for a given draw $\omega \in \Omega$. Sample functions play an important role in understanding the nature of a Wiener process.

Recall that the distribution function $F(x)$ of a random variable X is defined by $F(x) = P[X \le x]$. Similarly, the finite-dimensional distribution function of a stochastic process $\{X(t) : t \in I\}$ is given by

$$\begin{aligned} F_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) &= P\left[X_{t_1} \le x_1, X_{t_2} \le x_2, \ldots, X_{t_n} \le x_n\right] \\ &= P\left[\bigcap_{i=1}^{n} \left\{\omega \in \Omega : X_{t_i} \le x_i\right\}\right]. \end{aligned}$$

In short, given a stochastic process, we can derive the entire family of finite-dimensional distribution functions and there are uncountably many of them. Is the converse true? That is, given a family of finite-dimensional distribution functions, $\{F_{t_1, t_2, \ldots, t_n}(x_1, x_2, \ldots, x_n) : t_i \in I, \;\; n \ge 1\}$, satisfying some regularity conditions (symmetry and compatibility), can we reconstruct a stochastic process and the underlying probability space such that the distribution functions generated by the reconstructed stochastic process are the ones originally given? The answer is affirmative and is formulated as Kolmogorov's existence theorem. See Billingsley (1995, section 36).

There is another issue here. Two stochastic processes $\{X(t) : t \in I\}$ and $\{Y(t) : t \in I\}$ are *stochastically equivalent* if, for all $t \in I$, $X(t) = Y(t)$ w.p.1, where "w.p.1" stands for "with probability one." The problem is: even with Kolmogorov's theorem, we may come up with

© in this web service Cambridge University Press www.cambridge.org