Due date is extended until Tuesday, Feb. 13.

Office hours for Wed. Feb. 7 moved to Thursday, Feb. 8 at 3 PM.

More on defining stochastic processes through finite-dimensional distributions. The idea is to think of the stochastic process as a collection of random variables parameterized by the time index set T. So the mathematical issue is how do we extend the notion of talking about a finite set of (dependent) random variables to an infinite set of dependent random variables particularly if these random variables are indexed by a continuous parameter. The way this is done mathematically is to only make explicit reference to finite-dimensional distributions, and then appeal to an abstract extension procedure (Kolmogorov extension theorem) that guarantees that you are implicitly defining a stochastic process on the whole time index set in a well-posed way. Some dangers in this procedure...there is in general many stochastic processes consistent with a given prescription for the finite-dimensional distributions.

Here is a trivial example that shows the problem in making this leap from finite-dimensional distributions to a stochastic process with respect to a continuous time parameter:

Define $X(t) \equiv 0$, for $t \in [0,1]$

Define $Y(t) = \left\{ \begin{array}{ll} 1 & \text{when } t = \frac{1}{2} \\ 0 & \text{otherwise} \end{array} \right.$
Both $X(t)$ and $Y(t)$ have the same finite-dimensional distributions, namely, that when observed at any finite number of pre-specified points, the observations will all be 0 with probability 1.

To avoid this kind of pathology, one usually prescribes the finite-dimensional distributions for the stochastic process, and then declares that the stochastic process to be defined is separable. Separability means that the stochastic process is completely determined by its behavior on some countable dense subset of the time index set. (See Oksendal, Billingsley Probability theory, Karatzas and Shreve). Imposing separability means that if there is a nice (continuous, etc.) version of the stochastic process with the given finite-dimensional distributions, then that is what will be defined. How do I know when the separable version of the stochastic process is "nice" in a certain way? We'll get to that in a bit.

But for now: Kolmogorov extension theorem says that if you provide self-consistent rules for all finite-dimensional distributions, then there exists (usually many) versions of a stochastic process with these properties. Imposing separability then focuses on the one version which is most nice.
focuses on the one version which is most nice.

A more abstract formulation of conditional expectation and probability:
(Grigoriu Sec. 2.17)

Given probability space: \(( \Omega, \mathcal{F}, \mathbb{P} )\)

Let \( \mathcal{G} \subseteq \mathcal{F} \) be a sub \( \sigma \)-algebra corresponding to coarse-graining of information. The full \( \sigma \)-algebra corresponds to all events which are decided after all uncertainty is resolved, and for which a probability can be defined. Sub \( \sigma \)-algebras contain those events that are decided upon knowledge of only some of the uncertainty. (See Minier and Peiran.)

\[ \mathcal{F}(X_1) \] is defined to be the sub \( \sigma \)-algebra which is generated by the random variable \( X_1 \),

\[ \mathcal{F}(X_1) = \{ X_1^{-1}(B) : B \in \mathcal{B} \} \]

where \( \mathcal{B} \) is the \( \sigma \)-algebra of measurable sets in state space \( \mathcal{S} \).
Conditional expectation w.r.t.
sub $\sigma$-algebra

Given prob. triplet $(\mathcal{F}, \mathcal{F}, \mu)$
and $X$ is a r.v. w.r.t. this triplet.
Let $\mathcal{A} \subseteq \mathcal{F}$ be a sub $\sigma$-algebra

$\mathbb{E}(X|\mathcal{A})$, the conditional expectation of
the random variable, conditioned on the sub
sigma-algebra $\mathcal{A}$ is defined to be a random
variable on a (coarse-grained) probability triplet
$(\mathcal{A}, \mathcal{F}, \mu)$ such that the defining relation
holds:

For every $S \in \mathcal{A}$, makes $\mathbb{E}(X|\mathcal{A})$

$\int_S \mathbb{E}(X|\mathcal{A}) \, d\mu = \int_S X \, d\mu$

Intuitive idea is that
$\mathbb{E}(X|\mathcal{A})$ is an intermediate
between the random variable $X$
which refers to a detailed description
of the uncertainty, and $\mathbb{E}X$
which relates to complete ignorance
about the uncertainty.
The reason this is done is to handle cases where the partial information contained in the sub sigma-algebra corresponds to an infinite number of random variables. So, if one wants to talk about conditional probability, one does it in terms of conditional expectation!

In this framework we define $\mathbb{P}(A|\mathcal{A})$ but avoids really explicit reference to values $\mathbb{E}(X|A)$ which is $\mathbb{E}(x|A)$, which can be written $\mathbb{E}(X|\mathcal{A})$. Then $\mathbb{E}(X|A)$ gives the best guess for value of $X$ given information at the coarse-grained level corresponding to $A$. Important special case: $\mathbb{E}(X|\mathcal{A})$.
Markov processes: simple, classical notation first

A stochastic process $\mathcal{X}(t)$ has the **Markov property** when the following relation holds:

$$
\Pr \left( \mathcal{X}(t) \in B \bigg| \mathcal{X}(t_1) = x_1, \mathcal{X}(t_2) = x_2, \ldots, \mathcal{X}(t_n) = x_n \right)
= \Pr \left( \mathcal{X}(t) \in B \bigg| \mathcal{X}(t_n) = x_n \right)
$$

whenever $t_1 < t_2 < \cdots < t_n < t$

and $B \in \mathcal{B}$.

Intuitively, if I have information about the past and present values of the stochastic process and I want to calculate statistics about the (unknown) future, then I will obtain the same answer if I use arbitrary amount of information about the stochastic process in the past and present as I will if I just refer to the current state of the system. Given the present state, knowing the past states has no influence on my statistical predictions about the future.