02/09/04 Stochastic Processes: Fundamentals
basically a random function $X(t)$

$X$ takes values in state space $S$
and $t \in T$ ("time domain")

Brownian motion: $T = \mathbb{R}$
$S = \mathbb{R}^3$

Population models: $T = \mathbb{R}_{\geq 0}$ (one observation per year)
$S = \mathbb{R}_{\geq 0}$

A few ways to view a stochastic process

i) $X(t)$ as a random object

ii) $X(t, u)$ is a function of explicit time and of the

   elementary event $u$ determining

   the future of interest

iii) "Intuitive": discretize the time domain $T$

   somehow and then view $X(t)$ as

   a collection of random vars that

   are related to each other.

iv) Any measurement (functional) of $X(t)$

    is a random variable,

\[
X(1, \frac{1}{b-a}\int_a^b X(t) dt, \frac{dx}{dt}\bigg|_{t=1})
\]

What's the first time that $X(t) \leq 60$?
How rigorously formulate a stochastic process?

- challenge is mostly in that one is dealing with a \( \infty \) number of r.v.s which can be arbitrarily close to each other if \( T \) continuous.

i) Describe or construct the stochastic process so that it only involves countably many random variables.
- Markov chains, Poisson process

ii) Define stochastic process as a limit of
stochastic processes w/ finite/countable # r.v.s.
- Diffusion processes, SDE's

iii) Describe finite-dimensional distributions (all observations involving finite subsets of \( T \)) in a self-consistent manner, and declare the process to be separable (nice).

- stationary random fields

\[ x(t) \]

\[ \begin{align*}
  y(t) &= 0 \quad \text{for} \; x \neq Z \\
  &= 1 \quad \text{for} \; x = Z
\end{align*} \]

where \( Z \) is a continuously distributed r.v.

\((X, Y)\) have same finite dim dists,
The procedure using finite-dim dists + separability is called "Kolmogorov extension theorem"

- Øksendal
- Billingsley Sec. 36 + 38
- Karatzas & Shreve: Brownian Motion and Stochastic Calculus
  - super technical
Finite State Markov Chains

- finite state space \( S \)
- time is discrete \( T = \mathbb{Z} \) (to begin)

Usually write Markov chain in discrete time as

\[ X_n = \text{value of the Markov chain} \]
\[ \text{at time } n \quad (n \in T = \mathbb{Z}) \]

Markov property: Future values of the Markov chain may depend on present value, but past information has no extra predictive value.

Examples that might be modeled by finite state, discrete time Markov chains:

- cellular automaton w/ finite # states
- conformations of a protein molecule
- random walk among a finite # sites
- queues/inventories w/ finite capacity

- genetic
- genetic regulatory networks
- simulated annealing/ Metropolis algorithm
- base-pair sequence or non-coding region of DNA
Mathematical characterization of a Markov chain \( \{X_n\}_{n=0}^{\infty} \) with finite states, discrete time

\( X_n \in S \) (finite) for each \( n \in \mathbb{N} \)

Markov:

1) \[
\Pr(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \Pr(X_{n+1} = j \mid X_n = i_n)
\]

2) \( X_{n+1} = f_n(X_n, Z_n) \) with \( f_n \) deterministic and \( \{Z_n\}_{n=0}^{\infty} \) are independent rvs

\( (\text{Time homogenous: } X_{n+1} = f(X_n, Z_n) \) universal stochastic update rule

and \( Z_n \) are i.i.d.

3) Let \( T_0 \leq T \) with \( l_* = \max \{ l \mid l \in T \} \)

Let \( T_1 \leq T \) with \( m_* = \min \{ m \mid m \in T \} \)

If \( l_* < m_* \) then

\[
\Pr(\exists X_k = j_k \text{ for each } k \in T_1 \mid X_k = i_k \text{ for each } k \in T_0) = \Pr(\exists X_k = j_k \text{ for each } k \in T_1 \mid X_{l_*} = i_{l_*})
\]

(This works in reverse time too)
4) Let $T_0$ be "past times" and $T_1$ be "future times" as before. Suppose $k_x < n < m_x$.

\[
\Pr(\{X_k = j_k \text{ for each } k \in T_1 \text{ and } X_e = i_e \text{ for each } e \in T_0 \mid X_n = i_n\})
\]

\[
= \Pr(\{X_k = j_k \text{ for each } k \in T_1 \mid X_n = i_n\})
\times \Pr(\{X_e = i_e \text{ for each } e \in T_0 \mid X_n = i_n\})
\]

Conditioned on the present, future is independent of the past.
Time-homogenous Markov chains

- finite state discrete time

All probabilistic information connecting the \( \{X_n\}_{n=-\infty}^{\infty} \) at different values of \( n \) is encoded in:

i) Probability Transition Matrix:

\[
P_{ij} = \text{Prob}(X_{n+1} = j | X_n = i)
\]

(no dependence on \( n \))

ii) Probability distribution at a given point (usually \( n = 0 \))

\[
\phi_j = \text{Prob}(X_0 = j)
\]

---

(\( n \))

0 1 2 3

(can also go backwards... later)
Probability transition matrix \( P \) can be any matrix w/ the properties:

- a) all \( P_{ij} \geq 0 \)
- b) \( \sum_{j \in S} P_{ij} = 1 \)

Initial distribution \( \text{\( \phi \)} \) can be any vector w/:

- a) \( \phi_j \geq 0 \)
- b) \( \sum_{j \in S} \phi_j = 1 \)

Specifying \( P \) and \( \text{\( \phi \)} \) completely & uniquely defines the Markov chain.