Homework 4 due on Thursday, December 15 at 5 PM (hard deadline).

How are formulas for large-time behavior of discrete-time Markov chains modified for continuous-time Markov chains?

Classification of Markov chains (transience vs. recurrence)

- Communication classes are determined the same way as for discrete-time Markov chains by looking at topology.
- Transience or recurrence is determined by whether the embedded discrete-time Markov chain is transient or recurrent.

Positive recurrence vs. null recurrence

- Again, an irreducible positive recurrent CTMC is precisely an irreducible CTMC with a stationary distribution. (And if there is not a stationary distribution, then the CTMC is not positive recurrent.)
- Stationary distribution has to be computed differently than for the embedded discrete time Markov chain because the amount of time spent in states is relevant.

Stationary distribution for a CTMC:

\[ \pi_j = P[X(t) = i] \]

Should be a time-independent solution to the forward Kolmogorov equation (which describes how the
probability distribution for a state evolves).

\[ \frac{\partial \mathbf{P}}{\partial t} = \mathbf{0} \cdot \mathbf{A} \]

\[ \mathbf{P}_t \mathbf{i} = \mathbf{p}(X(t) = \mathbf{i}) \]

\[ \mathbf{0} = \mathbf{\pi} \cdot \mathbf{A} \]

As for discrete-time Markov chains, one can try to find detailed balance solutions:

\[ \mathbf{\pi}_i A_{ij} = \mathbf{\pi}_j A_{ji} \quad \text{for all } i, j \in S \]

This is overdetermined, so solution is not guaranteed to exist, but it's usually easy to check if a detailed balance solution can be found, and if so, it does give a good stationary distribution (assuming the other two conditions about nonnegativity and normalization).

Once this classification is done, then we can compute the long-time properties by the following extensions to discrete-time theory:

1. For positive recurrent classes, one can determine the long-time behavior by using the stationary
distribution and the following continuous-time version of the law of large number for continuous-time Markov chains:

\[ \lim_{T \to \infty} \frac{1}{T} \int_0^T f(X(t)) \, dt = \sum_{j \in S} \pi_j \cdot f(j) \]

2. When starting in a transient state, one can do an absorption probability calculation to determine which (if any) of the recurrent classes the Markov chain will move into (with what probabilities). This is done by applying the absorption probability calculations to the embedded discrete-time Markov chain.

3. If one wants to compute the expected reward/cost accumulated while moving through transient states, until one hits a recurrent class, then one can use the following modification to the discrete-time formula:

\[
W_i = \mathbb{E} \left( \int_{0}^{\tau} f(X(t)) \, dt \mid X(0) = i \right)
\]

\[
\tau = \inf \left\{ t \geq 0; t \notin T \right\}
\]

Note that here \( f \) is the rate at which the cost/reward is accumulated.
Proper derivation using first-step analysis can be found in Lawler Sec. 3.3, Karlin and Taylor Ch. 4.

4. And as for discrete-time Markov chains, not much more can be said quantitatively about null recurrent classes beyond the general properties already described for them in the discrete-time setting.

Similar formulas apply when one extends to continuous state space (stochastic differential equations) but now the infinitesimal generator $A$ is given by a second order differential operator.

Renewal processes

Reading: Karlin and Taylor Ch. 5
Resnick Ch. 3

Renewal processes generalize the Poisson process by relaxing the Markovian property of the associated counting process. In particular, the incidents generated by a renewal process only lose memory when a new incident occurs. But while waiting between incidents, the process does have memory of how long it has been waiting.

This relaxation of the Markov property (of the associated counting process) means that the time intervals between successive incidents in a renewal process are
independent but not necessarily exponentially distributed. (Usually the times between successive incidents are identically distributed except maybe the time until the first event).

Applications:

- Equipment replacement
- Neuronal signals: particularly if one is looking at outgoing signals.

![Graph of interspike intervals]

- Following in distance as a spatial renewal process
- Generic events associated to continuous-time Markov chains with the strong Markov property:
  - Times of successive visits to a certain state

More generally, if continuous-time Markov chain models are modified to have nonexponential distributions for times between state changes, this is called a *semi-Markov process*, but is not necessarily a renewal process because the state change may not refresh the system.

- Queueing theory with nonexponential times for requests quitting the queue
- Branching processes with age considerations, so the branching events are not separated by exponentially distributed times
- Epidemiological models where the times between infection and recovery are not exponentially distributed.

The theory of semi-Markov processes is relatively
difficult, and I haven't really seen a unified treatment. But for specific models, it's possible to develop a good deal of analysis (and renewal processes are just the simplest example.)

Let's begin by viewing a renewal process as a point process on the positive real line.

The interincident (interarrival) times:

\[ T_j = T_j - T_{j-1}, \quad j \geq 2 \]

\[ T_1 = T_1 - 0 = T_1 \]

For a standard (pure) renewal process, all the interincident times are assumed to be independent and identically distributed. (A commonly studied variation, called a delayed renewal process, allows \( T_1 \) to have a different probability distribution, but we won't go into this.)

So pure renewal processes have their statistical dynamics
completely described by the probability distribution for the interincident times, and this is usually expressed in terms of a cumulative distribution function (CDF):

\[ F(t) = P(T_j \leq t) \]

For continuously distributed interincident times, one can express the CDF in terms of the PDF of the interincident times by:

\[ F(t) = \int_0^t p_T(t') \, dt' \]

\[ p_T(t) = \frac{d}{dt} F(t) \]

The reason why CDFs are used in renewal process theory (rather than PDFs) is that the interincident time often is not purely continuous; it could be discrete or even a hybrid discrete-continuous random variable. CDFs are a unified way to describe discrete, continuous, hybrid random variables.

For a Poisson point process, this would be a renewal process with interincident PDF:

\[ p_T(t) = \begin{cases} 
\lambda e^{-\lambda t} & \text{for } t \geq 0 \\
0 & \text{for } t < 0
\end{cases} \]
What random variables associated to a renewal process are of fundamental interest in applications?

- The associated counting process:

\[ N(t) = \sum_{j=1}^{\infty} I\{t > t_j\} \]

Counts how many incidents have happened before time \( t \).

- Total life \( \beta_t \)
- Current life \( \delta_t \)
- Residual life \( \varrho_t \)

To warm up, let's compute these quantities for Poisson point process. This will lead us to the Poisson paradox.
We already have computed the properties of the Poisson counting process \( N(t) \); it's given by a Poisson distribution with mean \( \lambda t \).

From the Markov property and associated memorylessness of how long one has spent since the last Poisson event, the residual life of a Poisson point process must be exponentially distributed with mean \( 1/\lambda \).

\[
F_{\delta_t}(t') = P(\delta_t \leq t') = \begin{cases} 
1 - e^{-\lambda t'} & \text{for } t' \geq 0 \\
0 & \text{for } t' < 0
\end{cases}
\]

For computing the current life, this involves looking at the past of time \( t \). Recall that the Markov property can be applied in either time direction. But there's a catch, that the current life can't be bigger than \( t \).

If one had extended the renewal process back to negative times, then the current life would have the same probability distribution as the residual life, but if we are only looking at the renewal process on the positive time axis (as is the convention), then one has to apply a cutoff at time \( t \).

\[
F_{\delta_t}(t') = P(\delta_t \leq t') = \begin{cases} 
1 - e^{-\lambda t'} & \text{for } 0 \leq t' \leq t \\
1 & \text{for } t' > t \\
0 & \text{for } t' < 0
\end{cases}
\]
The total life $\beta_t$ can be computed as:

$$\beta_t = \delta_t + \delta_{t'}$$

By the Markov property, one of these summands is a random variable of the past, and one is a random variable of the future, and so they must be independent (for Poisson point process!)