3) Enough to check that \( \mathbf{V} \cdot \mathbf{P} = \mathbf{V} \)

for \( \mathbf{V}_k = \mathbb{E} \left[ T_{j,j,k} \mid X_0 = j \right] \)

because \( \mathbf{V}_k \leq \mathbf{V} \)

\( \mathbf{V} \) will thereby be shown to be
"invariant measure".

\( \mathbf{V}_k \geq 0 \)

\( \mathbf{V} \cdot \mathbf{P} = \mathbf{V} \)

(but no normalization)

\[ \mathbf{V}_k = \mathbb{E} \left[ \sum \limits_{n=0}^{\infty} I \left( X_n = k \right) \mid X_0 = j \right] \]

\( \text{indicator function} = 1 \text{ if arg true} \)
\( = 0 \text{ if arg false} \)

\[ = \sum \limits_{n=0}^{\infty} \text{Prob} \left( X_n = k, T_j (\cdot) > n \mid X_0 = j \right) \]

(used: \( \mathbb{E} I \left( X_n = k \right) = 1 \cdot \text{Prob} \left( X_n = k \right) + 0 \cdot \text{Prob} \left( X_n \neq k \right) = \text{Prob} \left( X_n = k \right) \))

\[ = \mathbb{E} \left[ \sum \limits_{n=0}^{\infty} I \left( X_n = k, T_j (\cdot) > n \right) \mid X_0 = j \right] \]
Side note: Simplest version of the indicator function trick:

Consider \( N \) coin tosses, \( p = \text{prob}(H) \)
\( 1 - p = \text{prob}(T) \)

\[
\bar{X} = \text{# heads} \\
\prob(\bar{X} = j) = \binom{N}{j} p^j (1 - p)^{N - j}
\]

\[
|\mathbb{E} \bar{X} = \sum_{j=0}^N j \prob(\bar{X} = j) = a \text{ pain}
\]

Another way to calculate:

\[
\bar{X} = \sum_{n=1}^N \mathbf{1} \{ \gamma_n = 13 \}
\]

where \( \gamma_n = 1 \) if coin toss \( n \) is \( H \)
\( = 0 \) if coin toss \( n \) is \( T \)

\[
|\mathbb{E} \bar{X} = \sum_{n=1}^N \mathbb{E} \mathbf{1} \{ \gamma_n > 13 \} = \sum_{n=1}^N \mathbb{E} \mathbf{1} \{ \gamma_n = 13 \} \\
= \sum_{n=1}^N \prob(\gamma_n = 13) \\
= \sum_{n=1}^N \prob(\gamma_n = 1) \\
= \sum_{n=1}^N p = pN
\]
Consider
\[ \sum_{k \in S} \sum_{j \in S} \sum_{n=0}^\infty \Pr(X_n = k, I_j(1) \geq n \mid X_0 = j) \times \Pr(X_{n+1} = l \mid X_n = k) \]

Want to use \( \Pr(A \cap B) \Pr(B \cap C) = \Pr(A \cap B \cap C) \)

\[ \Pr(A \mid B \text{ and } C) \Pr(B \mid C) = \Pr(A \mid B \text{ and } C) \]

Try to get the expression in this form

\[ \begin{align*}
A &= \{ X_{n+1} = l \} \\
B &= \{ X_n = k \text{ and } I_j(1) > n \} \\
C &= \{ X_0 = j \}
\end{align*} \]

This will work out if I can show that

\[ \Pr(X_{n+1} = l \mid X_n = k) = \Pr(l \mid \text{Markov}) \]

\[ = \Pr(X_{n+1} = l \mid X_n = k \text{ and } I_j(1) > n \text{ and } X_0 = j) \]

This is actually true by the Markov property when \( k \neq j \).

(Use \( \Pr(X_{n+1} = l \mid X_k = i \) for \( l \in I_0) \)

\[ = \Pr(X_{n+1} = l \mid X_{k^*} = i_{k^*}) \text{ with } k^* = \max_{l \in I_0} k \leq n \]

...
If \( k > j \) then \( X_n = k \) and \( T_j(1) > n \) can\( \not\) both be true, so \( \text{Prob}(X_n = k, T_j(1) > n \mid X_0 = j) = 0 \) for \( j = k \).

So I can make the other factor whatever I can make the change \( \bigcirc \).

So \( \sum \sum \limits_{k \in S} P_{ke} = \sum \sum \limits_{n=0} \text{Prob}(X_n = k, T_j(1) > n \mid X_0 = j) \times \text{Prob}(X_n+1 = l \mid X_n = k, T_j(1) > n, X_0 = j) \)

\[ = \sum \sum \limits_{n=0} \text{Prob}(X_n+1 = l, X_n = k, T_j(1) > n \mid X_0 = j) \]

\[ = \sum \sum \limits_{n=0} \text{Prob}(X_n+1 = l, X_n \in S, T_j(1) > n \mid X_0 = j) \]

\[ = \sum \sum \limits_{n=0} \text{Prob}(X_n+1 = l, T_j(1) > n \mid X_0 = j) \]

Recall \( \mathcal{V}_e = \sum \sum \limits_{n=0} \text{Prob}(X_n = l, T_j(1) > n \mid X_0 = j) \)

Let \( n' = n+1 \)

\[ \sum \sum \limits_{k \in S} P_{ke} = \sum \sum \limits_{n' = 1} \text{Prob}(X_{n'} = l, T_j(1) > n' - 1 \mid X_0 = j) \]
A) Consider first the case \( l \neq j \)

\[ T_j(1) > n - 1 \iff T_j(1) \geq n \]

\[ \text{Prob}(X_n = l, T_j(1) > n | X_0 = j) = \text{Prob}(X_n = l, T_j(1) \geq n | X_0 = j) \]

Because they differ by the event \( \{ X_n = l, T_j(1) = n \} \) which is contradictory.

Moreover, the missing \( n > 0 \) term

\[ \text{Prob}(X_0 = l, T_j(1) > 0 - 1 | X_0 = j) \neq 0 \]

There are \( V_e = \sum_{k \in S} v_k P_{ke} \) for \( l \neq j \)

because the summations are equivalent.

B) Consider case \( l = j \):

\[ V_j = \mathbb{E} \left[ T_j(1) | X_0 = j \right] = 1 \]

The expression \( \bigcirc \)

\[ = \sum_{n=0}^{\infty} \text{Prob}(X_{n+1} = j, T_j(1) > n | X_0 = j) \]

\[ = \sum_{n=0}^{\infty} \text{Prob}(T_j(1) = n + 1 | X_0 = j) \]

\[ = \text{Prob}( \exists l \leq T_j(1) < \infty | X_0 = j) = 1 \]

So \( V_e = \sum_{k \in S} v_k p_{ke} \) for \( l = j \)
This completes the argument that \( \vec{r} \cdot \vec{p} = \vec{r} \) and therefore shows that my candidate \( \vec{\pi} \) is a stationary distribution.

Existence 

Uniqueness: Resnick, Sec. 2.1d
- show that for any other stationary distribution \( \vec{\mu} \),
  that \( \vec{\pi} \leq \vec{\mu} \) \( \Rightarrow \), \( \vec{\pi} = \vec{\mu} \)

References for Finite State Markov Chains,

Ch. 1: Lawler: concise
Ch. 2+3: Karlin + Taylor: elaborate
Ch. 2: Resnick, Adventures in Stochastic Processes
What about reducible MC's (finite state) - proof doesn't work - but can modify it to prove existence for all finite state MC's but uniqueness may fail of stationary distribution

Running Markov chains backwards

If specify \( \phi_j = \text{Prob}(X_0 = j) \) and
\[
P_{ij} = \frac{\text{Prob}(X_{n+1} = j | X_n = i)}{\text{Prob}(X_{n+1} = i)}
\]

how calculate \( \phi_j \) for \( n < 0 \)?

(want the forward evolution for \( n < 0 \) to look like the forward evolution for \( n > 0 \))

To go backwards want to calculate
\[
\hat{P}^{(n)}_{ij} = \frac{\text{Prob}(X_n = j | X_{n+1} = i)}{\text{Prob}(X_{n+1} = i)}
\]
\[
= \frac{\text{Prob}(X_n = j \text{ and } X_{n+1} = i)}{\text{Prob}(X_{n+1} = i)}
\]
\[ P_{i,j} = \frac{P_{j,j} P_{j}^{(n)}}{\pi_{i}} \]

Get contamination by "undoing" forgetting of initial data.

But one can get useful backward evolution by getting rid of initial data
- initialize with stationary distribution
- because then \( p_{j}^{(n)} = \Pi j \) for all \( n \)

\[ \hat{p}_{i,j} = \frac{P_{j,j} \Pi j}{\Pi i} \]
Physical meaning of this?
Suppose I have a "microreversible system".

\[ P_{ij} = \hat{P}_{ij} P_{ij} \]

That means

\[ \sum_i P_{ij} = P_{ji} \sum_j \]

"detailed balance"

apply to systems in thermal equilibrium

Non-detailed balance
Some comments about computing stationary distribution

A) \[ \vec{\Pi} \cdot \vec{P} = \vec{\Pi} \]
\[ \sum_{j \in S} \Pi_j = 1 \]
\[ \Pi_j \geq 0 \]

B) Haken Synergetics Sec. 4.6-4.8
- easy to solve for \( \vec{\Pi} \) when detailed balance
- graphical Kirchoff procedure for calculating \( \vec{\Pi} \) w/ detailed balance

C) Resnick Sec. 2.14:
\[ \vec{\Pi} = \vec{1} \cdot (I - \vec{P} + \text{ONE})^{-1} \]

D) Resnick Sec. 2.13:
Calculate \( \vec{\Pi} \) by using stochastic update rule, manipulate w/ gen funs.
- useful for \( \alpha \) states
Brief note on statistical questions:
- How should I choose $\theta$ and $p$? (given data)
- How do I know if my model is "good"?

Ideas: Maximum likelihood method

1) Suppose we try to model the given data as a set of i.i.d. variables with distribution $\theta$.

   How choose $\theta$? (after choosing state space)

   What's the likelihood to observe data given my model?

   $L = \prod_{i=1}^{M} \theta^{q_i}$, where $q_i = \#$ occurrences of state $i$ in the data.

   Choose $\theta$ to maximize $L$.

   $\Rightarrow \theta_i = \frac{q_i}{\sum_{i \in S} q_i}$
II) Markov chain model: How choose $\theta$ and $P$?

$$L = \prod_{i \in S} P_{ij}^{a_{ij}} \theta_k$$

where $k = \text{initial state}$

$a_{ij} = \# \text{transitions } i \rightarrow j \text{ in the data}$

$$1 \ 3 \ 5 \ 4 \ 2 : \theta_1 \ P_{13} \ P_{35} \ P_{54} \ P_{42}$$

Optimize $L$ with respect to $P$ (usually ignore $\theta$)

$$P_{ij} = \frac{a_{ij}}{\sum_{j \in S} a_{ij}}$$

Was the use of a Markov chain a significant improvement over using iid rvs,

- use $\chi^2$ statistic to see if

$$L^{mc}/L^{iid} \text{ is large enough}$$

...to have statistically significant value, give extra degrees of freedom,
How calculate long-time properties of $M(t)$?

Need some classification of Markov chains.

Period of a state $i$:

$d(i) = \gcd \{ n \geq 0, (p^n)_{ii} > 0 \}$

Prob to come back at time $n$. 

$\uparrow$