STATS 217: Introduction to Stochastic Processes I

Lecture 9

Multi-step transition probabilities

- The transition probability p_{ij} tells us the probability of going from i to j in *one* step, i.e.
$$p_{ij} = \mathbb{P}[X_1 = j \mid X_0 = i].$$

Multi-step transition probabilities

- The transition probability p_{ij} tells us the probability of going from i to j in *one* step, i.e.

$$p_{ij} = \mathbb{P}[X_1 = j \mid X_0 = i].$$

- What about the probability of going from i to j in *two* steps i.e. what is

$$p_{ij}^2 := \mathbb{P}[X_2 = j \mid X_0 = i]?$$

$(p_{ij})^2$
$= p_{ij} \cdot p_{ij}$

Multi-step transition probabilities

- The transition probability p_{ij} tells us the probability of going from i to j in *one* step, i.e.

$$p_{ij} = \mathbb{P}[X_1 = j \mid X_0 = i].$$

- What about the probability of going from i to j in *two* steps i.e. what is

$$p_{ij}^2 := \mathbb{P}[X_2 = j \mid X_0 = i]?$$

- Well, to go from i to j in two steps, we must go from i to some state $k \in S$ in one step and then from k to j in one step.

k could be j

Multi-step transition probabilities

Using the law of total probability, we have

$$\mathbb{P}[X_2 = j \mid X_0 = i] = \sum_{k \in S} \mathbb{P}[X_1 = k \wedge X_2 = j \mid X_0 = i]$$

Multi-step transition probabilities

Using the law of total probability, we have

$$\mathbb{P}[X_2 = j \mid X_0 = i] = \sum_{k \in S} \mathbb{P}[X_1 = k \wedge X_2 = j \mid X_0 = i]$$
$$= \sum_{k \in S} \underbrace{\mathbb{P}[X_1 = k \mid X_0 = i]}_{P_{ik}} \cdot \underbrace{\mathbb{P}[X_2 = j \mid X_0 = i \wedge X_1 = k]}_{}$$

" (Markov prop.)

$\mathbb{P}[X_2 = j \mid X_1 = k]$

"

P_{kj}

transition matrix.

$$P^2_{ij} = \sum_{k \in S} P_{ik} P_{kj} = (\mathbb{P}^2)_{ij}$$

so: $P^2_{ij} = (\mathbb{P}^2)_{ij}$

Multi-step transition probabilities

Using the law of total probability, we have

$$\begin{aligned}
\mathbb{P}[X_2 = j \mid X_0 = i] &= \sum_{k \in S} \mathbb{P}[X_1 = k \wedge X_2 = j \mid X_0 = i] \\
&= \sum_{k \in S} \mathbb{P}[X_1 = k \mid X_0 = i] \cdot \mathbb{P}[X_2 = j \mid X_0 = i \wedge X_1 = k] \\
&= \sum_{k \in S} \mathbb{P}[X_1 = k \mid X_0 = i] \cdot \mathbb{P}[X_2 = j \mid X_1 = k]
\end{aligned}$$

Multi-step transition probabilities

Using the law of total probability, we have

$$\begin{aligned}
\mathbb{P}[X_2 = j \mid X_0 = i] &= \sum_{k \in S} \mathbb{P}[X_1 = k \wedge X_2 = j \mid X_0 = i] \\
&= \sum_{k \in S} \mathbb{P}[X_1 = k \mid X_0 = i] \cdot \mathbb{P}[X_2 = j \mid X_0 = i \wedge X_1 = k] \\
&= \sum_{k \in S} \mathbb{P}[X_1 = k \mid X_0 = i] \cdot \mathbb{P}[X_2 = j \mid X_1 = k] \\
&= \sum_{k \in S} p_{ik} p_{kj} \\
&= (P^2)_{ij}.
\end{aligned}$$

Multi-step transition probabilities

- There is nothing special about two steps here and you should check that the same argument gives

$$p_{ij}^n := \mathbb{P}[X_n = j \mid X_0 = i] = (P^n)_{ij} \quad \forall n \geq 1.$$

$$p_{ij}^n = \sum_{k_1 \ldots k_\ell}' p_{ik_1} p_{k_1 k_2} \cdots p_{k_\ell j}$$

Multi-step transition probabilities

- There is nothing special about two steps here and you should check that the same argument gives

$$p_{ij}^n := \mathbb{P}[X_n = j \mid X_0 = i] = (P^n)_{ij} \quad \forall n \geq 1.$$

- Since for any non-negative integers ℓ, m,

$$P^{\ell+m} = P^\ell P^m, \implies (P^{\ell+m})_{ij}$$

$$= (P^\ell P^m)_{ij}$$

we obtain the **Chapman-Kolmogorov equations**

$$p_{ij}^{\ell+m} = \sum_{k \in S} p_{ik}^\ell p_{kj}^m.$$

note: if $P = U D U^{-1}$ where D is diag.

$\Rightarrow P^n = U D^n U^{-1}$

Stopping times

- Let $(X_n)_{n \geq 0}$ be a stochastic process on a discrete state space S.
- We say that a random variable T is a **stopping time** if whether or not we stop at time k i.e., the event $\{T = k\}$, can be determined by the values of the process up to and including time k i.e., by $X_0, \ldots, X_k$.

Stopping times

- Let $(X_n)_{n \geq 0}$ be a stochastic process on a discrete state space S.
- We say that a random variable T is a **stopping time** if whether or not we stop at time k i.e., the event $\{T = k\}$, can be determined by the values of the process up to and including time k i.e., by $X_0, \ldots, X_k$.
- **Example**: let $(X_n)_{n \geq 0}$ be a symmetric simple random walk starting at 0. Then, τ_1, the first time to hit 1 is a stopping time.

Stopping times

- Let $(X_n)_{n\geq 0}$ be a stochastic process on a discrete state space S.
- We say that a random variable T is a **stopping time** if whether or not we stop at time k i.e., the event $\{T = k\}$, can be determined by the values of the process up to and including time k i.e., by $X_0, \ldots, X_k$.
- **Example**: let $(X_n)_{n\geq 0}$ be a symmetric simple random walk starting at 0. Then, τ_1, the first time to hit 1 is a stopping time.
- Indeed, for all $k \geq 0$,

$$\{\tau_1 = k\} = \{X_0 \neq 1, \ldots, X_{k-1} \neq 1, X_k = 1\}.$$

Stopping times

- Let $(X_n)_{n\geq 0}$ be a stochastic process on a discrete state space S.
- We say that a random variable T is a **stopping time** if whether or not we stop at time k i.e., the event $\{T = k\}$, can be determined by the values of the process up to and including time k i.e., by $X_0, \ldots, X_k$.
- **Example**: let $(X_n)_{n\geq 0}$ be a symmetric simple random walk starting at 0. Then, τ_1, the first time to hit 1 is a stopping time.
- Indeed, for all $k \geq 0$,

$$\{T = k\} = \{X_0 \neq 1, \ldots, X_{k-1} \neq 1, X_k = 1\}.$$

- **Non-example**: let $(X_n)_{n\geq 0}$ be a symmetric simple random walk starting at 0. Then, $\tau' = \tau_1 - 1$ is not a stopping time.

Strong Markov property

Markov prop:
$\mathbb{P}[X_{k+1} = j \mid X_k = i, X_{k-1} = x, \ldots, X_0 = x] = \mathbb{P}[X_{k+1} = j \mid X_k = i]$.

Let $(X_n)_{n \geq 0}$ be a DTMC on S and let T be a stopping time. Then, for all $k \geq 0$, all $n \geq 0$, and for all $i, j \in S$,

$$\mathbb{P}[X_{T+k} = j \mid X_T = i, T = n] = \mathbb{P}[X_k = j \mid X_0 = i].$$

- This is called the **Strong Markov Property**.

e.g. if $T = \tau_1$

$T = n$, it must be that $X_T = 1$

$\mathbb{P}[X_{T+k} = j \mid T = n] = \mathbb{P}[X_k = j \mid X_0 = 1]$

Why do we need T to be a stopping time?

e.g. $T = \tau_1' = \tau_1 - 1$

Strong Markov property

Let $(X_n)_{n\geq 0}$ be a DTMC on S and let T be a stopping time. Then, for all $k \geq 0$, all $n \geq 0$, and for all $i, j \in S$,

$$\mathbb{P}[X_{T+k} = j \mid X_T = i, T = n] = \mathbb{P}[X_k = j \mid X_0 = i].$$

- This is called the **Strong Markov Property**. Why is this true?

Strong Markov property

Let $(X_n)_{n \geq 0}$ be a DTMC on S and let T be a stopping time. Then, for all $k \geq 0$, all $n \geq 0$, and for all $i, j \in S$,

$$\mathbb{P}[X_{T+k} = j \mid X_T = i, T = n] = \mathbb{P}[X_k = j \mid X_0 = i].$$

- This is called the **Strong Markov Property**. Why is this true?
- Let V_n be the set of all vectors $x = (x_0, \ldots, x_n) \in S^{n+1}$ such that

$$X_0 = x_0, \ldots, X_n = x_n \implies T = n \text{ and } X_T = i.$$

Strong Markov property

Let $(X_n)_{n\geq 0}$ be a DTMC on S and let T be a stopping time. Then, for all $k \geq 0$, all $n \geq 0$, and for all $i, j \in S$,

$$\mathbb{P}[X_{T+k} = j \mid X_T = i, T = n] = \mathbb{P}[X_k = j \mid X_0 = i].$$

- This is called the **Strong Markov Property**. Why is this true?
- Let V_n be the set of all vectors $x = (x_0, \ldots, x_n) \in S^{n+1}$ such that

$$X_0 = x_0, \ldots, X_n = x_n \implies T = n \text{ and } X_T = i.$$

- Since T is a stopping time,

e.g. before shows that this isn't true for non- s.times.

$$\mathbb{P}[X_T = i, T = n] = \mathbb{P}[(X_0, \ldots, X_n) \in V_n].$$

Strong Markov property

$x = (x_0, \ldots, x_n)$

By the law of total probability, $\mathbb{P}(X_{T+k} = j \wedge (X_0 \ldots X_n) \in V_n)$

$$\mathbb{P}[X_{T+k} = j, X_T = i, T = n] = \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j, X_n = x_n, \ldots, X_0 = x_0]$$

$$= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j \mid X_n = x_n, \ldots, X_0 = x_0] \times \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0].$$

by markov: $\mathbb{P}[X_{n+k} = j \mid X_n = x_n, \ldots, X_0 = x_0]$

$$= \mathbb{P}[X_{n+k} = j \mid X_n = x_n]$$

Strong Markov property

By the law of total probability,

$$\mathbb{P}[X_{T+k} = j, X_T = i, T = n] = \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j, X_n = x_n, \ldots, X_0 = x_0]$$

markov prop. →

$$= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j \mid X_n = x_n] \cdot \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0]$$

time homogeneity

$$= \mathbb{P}[X_k = j \mid X_0 = x_n]$$

$$x_n = i$$

$$= \mathbb{P}[X_k = j \mid X_0 = i]$$

Strong Markov property

By the law of total probability,

$$\begin{aligned}\mathbb{P}[X_{T+k} = j, X_T = i, T = n] &= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j, X_n = x_n, \ldots, X_0 = x_0] \\ &= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j \mid X_n = x_n] \cdot \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0] \\ &= \mathbb{P}[X_k = j \mid X_0 = i] \sum_{x \in V_n} \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0]\end{aligned}$$

Strong Markov property

By the law of total probability,

$$
\begin{aligned}
\mathbb{P}[X_{T+k} = j, X_T = i, T = n] &= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j, X_n = x_n, \ldots, X_0 = x_0] \\
&= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j \mid X_n = x_n] \cdot \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0] \\
&= \mathbb{P}[X_k = j \mid X_0 = i] \sum_{x \in V_n} \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0] \\
&= \mathbb{P}[X_k = j \mid X_0 = i] \cdot \mathbb{P}[(X_0, \ldots, X_n) \in V_n]
\end{aligned}
$$

Strong Markov property

By the law of total probability,

$$\begin{aligned}
\mathbb{P}[X_{T+k} = j, X_T = i, T = n] &= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j, X_n = x_n, \ldots, X_0 = x_0] \\
&= \sum_{x \in V_n} \mathbb{P}[X_{n+k} = j \mid X_n = x_n] \cdot \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0] \\
&= \mathbb{P}[X_k = j \mid X_0 = i] \sum_{x \in V_n} \mathbb{P}[X_n = x_n, \ldots, X_0 = x_0] \\
&= \mathbb{P}[X_k = j \mid X_0 = i] \cdot \mathbb{P}[(X_0, \ldots, X_n) \in V_n] \\
&= \mathbb{P}[X_k = j \mid X_0 = i] \cdot \mathbb{P}[X_T = i, T = n].
\end{aligned}$$

Hitting times

- Let $(X_n)_{n\geq 0}$ be a Markov chain on S with $X_0 \sim \mu_0$.
- This just means that the initial state X_0 is a random variable with

$$\mathbb{P}[X_0 = j] = \mu_0(j) \quad \forall j \in S.$$

Hitting times

- Let $(X_n)_{n\geq 0}$ be a Markov chain on S with $X_0 \sim \mu_0$.
- This just means that the initial state X_0 is a random variable with

$$\mathbb{P}[X_0 = j] = \mu_0(j) \quad \forall j \in S.$$

- For a subset $A \subset S$, we define the first A-hitting time by

$$\tau_{A,\mu_0} := \min\{n \geq 1 : X_n \in A\}.$$

e.g. $S = \mathbb{Z}$, $A = \{1\}$

- If $\mu_0(s) = 1$ for some $s \in S$ (i.e., $\mu_0 = \delta_s$) then we lighten the notation a bit and write this as $\tau_{A,s}$.

" τ_{A,δ_s}

delta dist. at s

e.g. gambler's Ruin with -100 & $+200$, $A = \{-100, 200\}$.

Hitting times

- Let $(X_n)_{n\geq 0}$ be a Markov chain on S with $X_0 \sim \mu_0$.
- This just means that the initial state X_0 is a random variable with

$$\mathbb{P}[X_0 = j] = \mu_0(j) \quad \forall j \in S.$$

- For a subset $A \subset S$, we define the first A-hitting time by

$$\tau_{A,\mu_0} := \min\{n \geq 1 : X_n \in A\}.$$

- If $\mu_0(s) = 1$ for some $s \in S$ (i.e., $\mu_0 = \delta_s$) then we lighten the notation a bit and write this as $\tau_{A,s}$.
- Note that the minimum is taken over $n \geq 1$. Therefore,

$$\tau_{\{s\},s} =: T_s$$

is the first time we "return" to s, starting from s.

Number of visits

- For every $s \in S$, we let
$$f_s := \mathbb{P}[T_s < \infty].$$
In words, f_s is the probability that chain will ever return to s, provided that it starts at s.

Number of visits

- For every $s \in S$, we let
$$f_s := \mathbb{P}[T_s < \infty].$$
In words, f_s is the probability that chain will ever return to s, provided that it starts at s.
- For every $s \in S$ and initial distribution μ_0, we let
$$N_{\mu_0}(s) = \sum_{n=1}^{\infty} 1[X_n = s].$$
In words, $N_{\mu_0}(s)$ is the number of times we visit s (counting from time 1 onwards), starting with an initial state distributed according to μ_0.

Number of visits

- For every $s \in S$, we let
$$f_s := \mathbb{P}[T_s < \infty].$$
In words, f_s is the probability that chain will ever return to s, provided that it starts at s.
- For every $s \in S$ and initial distribution μ_0, we let
$$N_{\mu_0}(s) = \sum_{n=1}^{\infty} 1[X_n = s].$$
In words, $N_{\mu_0}(s)$ is the number of times we visit s (counting from time 1 onwards), starting with an initial state distributed according to μ_0.
- For lightness of notation, we set
$$N(s) := N_{\delta_s}(s).$$

Number of returns

- Just as for the symmetric simple random walk, we have for any $k \geq 1$ that

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = \underbrace{\mathbb{P}[N(s) \geq k-1 \mid X_0 = s]}_{\text{must return} \geq k-1 \text{ times}} \cdot \underbrace{f_s}_{\text{must return} \geq \text{once after } k-1^{\text{st}} \text{ return.}}.$$

Number of returns

- Just as for the symmetric simple random walk, we have for any $k \geq 1$ that

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = \mathbb{P}[N(s) \geq k-1 \mid X_0 = s] \cdot f_s.$$

$f_s = \mathbb{P}(T_s < \infty)$

- By induction, this shows that for all $k \geq 1$,

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = f_s^k.$$

Number of returns

- Just as for the symmetric simple random walk, we have for any $k \geq 1$ that

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = \mathbb{P}[N(s) \geq k - 1 \mid X_0 = s] \cdot f_s.$$

- By induction, this shows that for all $k \geq 1$,

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = f_s^k.$$

- Therefore, for all $k \geq 0$

$= \mathbb{P}(N(s) \geq k \mid X_0 = s] - \mathbb{P}[N(s) \geq k+1 \mid X_0 = s]$

$$\mathbb{P}[N(s) = k \mid X_0 = s] = f_s^k(1 - f_s),$$

and

$$\mathbb{E}[N(s) \mid X_0 = s] = \frac{f_s}{(1 - f_s)}.$$

note: $f_s = 1 \iff \mathbb{E}[N(s) \mid X_0 = s] = \infty.$

Number of returns

How do we show that

$$\mathbb{P}[N(s) \geq \tilde{k} \mid X_0 = s] = \mathbb{P}[N(s) \geq \widetilde{k-1} \mid X_0 = s] \cdot f_s?$$

* idea: define "correct" stopping time & use strong markov.

* natural: T is the time of $(k-1)^{st}$ return to s.

Number of returns

How do we show that

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = \mathbb{P}[N(s) \geq k-1 \mid X_0 = s] \cdot f_s?$$

- Let T be the time of the $(k-1)^{st}$ return to s. Note that T is a stopping time.

Number of returns

How do we show that

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = \mathbb{P}[N(s) \geq k-1 \mid X_0 = s] \cdot f_s?$$

- Let T be the time of the $(k-1)^{st}$ return to s. Note that T is a stopping time.
- By the Strong Markov property, we have

$$\mathbb{P}[N(s) \geq k \mid T = n, X_T = s] = \mathbb{P}[N(s) \geq 1 \mid X_0 = s] = f_s.$$

Number of returns

How do we show that

$$\mathbb{P}[N(s) \geq k \mid X_0 = s] = \mathbb{P}[N(s) \geq k-1 \mid X_0 = s] \cdot f_s?$$

- Let T be the time of the $(k-1)^{st}$ return to s. Note that T is a stopping time.
- By the Strong Markov property, we have

$$\mathbb{P}[N(s) \geq k \mid T = n, X_T = s] = \mathbb{P}[N(s) \geq 1 \mid X_0 = s] = f_s.$$

- By Bayes' rule,

$$\mathbb{P}[N(s) \geq k, T = n, X_T = s] = f_s \cdot \mathbb{P}[T = n, X_T = s].$$

Number of returns

- Summing this over n and using the law of total probability, we have

$$\mathbb{P}[N(s) \geq k] = \mathbb{P}[N(s) \geq k, T < \infty]$$

Number of returns

- Summing this over n and using the law of total probability, we have

$$\begin{aligned}\mathbb{P}[N(s) \geq k] &= \mathbb{P}[N(s) \geq k, T < \infty] \\ &= \sum_{n=0}^{\infty} \mathbb{P}[N(s) \geq k, T = n, X_T = s]\end{aligned}$$

Number of returns

- Summing this over n and using the law of total probability, we have

$$\begin{aligned}
\mathbb{P}[N(s) \geq k] &= \mathbb{P}[N(s) \geq k, T < \infty] \\
&= \sum_{n=0}^{\infty} \mathbb{P}[N(s) \geq k, T = n, X_T = s] \\
&= f_s \cdot \sum_{n=0}^{\infty} \mathbb{P}[T = n, X_T = s]
\end{aligned}$$

strong markov prop.

$\mathbb{P}(T < \infty) = \mathbb{P}(N(s) \geq k-1 \mid X_0 = s)$

Number of returns

- Summing this over n and using the law of total probability, we have

$$\begin{aligned}\mathbb{P}[N(s) \geq k] &= \mathbb{P}[N(s) \geq k, T < \infty] \\ &= \sum_{n=0}^{\infty} \mathbb{P}[N(s) \geq k, T = n, X_T = s] \\ &= f_s \cdot \sum_{n=0}^{\infty} \mathbb{P}[T = n, X_T = s] \\ &= f_s \cdot \mathbb{P}[T < \infty]\end{aligned}$$

Number of returns

- Summing this over n and using the law of total probability, we have

$$
\begin{aligned}
\mathbb{P}[N(s) \geq k] &= \mathbb{P}[N(s) \geq k, T < \infty] \\
&= \sum_{n=0}^{\infty} \mathbb{P}[N(s) \geq k, T = n, X_T = s] \\
&= f_s \cdot \sum_{n=0}^{\infty} \mathbb{P}[T = n, X_T = s] \\
&= f_s \cdot \mathbb{P}[T < \infty] \\
&= f_s \cdot \mathbb{P}[N(s) \geq k - 1].
\end{aligned}
$$

Recurrence and transience

Let $(X_n)_{n \geq 0}$ be a DTMC on S.

- $s \in S$ is a **recurrent state** if $f_s = 1$.
- $s \in S$ is a **transient state** if $f_s < 1$.

Recurrence and transience

Let $(X_n)_{n\geq 0}$ be a DTMC on S.

- $s \in S$ is a **recurrent state** if $f_s = 1$.
- $s \in S$ is a **transient state** if $f_s < 1$.

- By the formula

$$\mathbb{E}[N(s) \mid X_0 = s] = \frac{f_s}{1 - f_s},$$

  we see that

  - f_s is recurrent $\iff \mathbb{E}[N(s) \mid X_0 = s] = \infty$.
  - f_s if transient $\iff \mathbb{E}[N(s) \mid X_0 = s] < \infty$.