Probability Part II - Markov Processes

By Terry Wang

Advanced 2 - Jan 2025

This packet is a continuation from last time, so we will take for granted the definitions given before.

1 Conditional Probabilities

Definition 1 *Let $(\Omega, \mathbb{P})$ be a probability space, for some events $A, B \subseteq \Omega$, the conditional probability of A given B is defined as follows:*

$$\mathbb{P}[A \mid B] := \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

If $\mathbb{P}[B] > 0$, otherwise, it is defined to be 0. [1]

Problem 1 *Calculate the following conditional probabilities:*

- *Let Ω be the set of possible outcomes of a 6 sided dice roll, let A be the event that you roll an even number, let B the event that you roll a number greater than 3. Calculate $\mathbb{P}[A \mid B]$.*
- *Let Ω be the set of possible outcomes of a 6 sided dice roll, let A be the event that you roll a 5, let B the event that you roll a number greater than 3. Calculate $\mathbb{P}[A \mid B]$.*
- *Let Ω be the set of possible outcomes of 4 independent coin flips, let A be the event you get at least 1 head, let B be the event you get at least 2 heads. Calculate $\mathbb{P}[A \mid B]$.*
- *Let Ω be the set of possible outcomes of 4 independent coin flips, let A be the event the first coin is a head, let B be the event that the second coin is a head. Calculate $\mathbb{P}[A \mid B]$.*

[1] This is a highly restricted definition of conditional probability. A full construction permits conditioning on events with 0 probability with meaningful and applicable results.

Problem 2 *Take a moment to convince yourself that this definition is intuitive by answering the following:*

1. *Suppose A, B are disjoint: for instance, let $\Omega = \{H, T\}$ be the possible results of a coin flip, let A be the event that you get a heads and B the event that you get a tails. What is $\mathbb{P}[A \mid B]$?*
2. *Suppose $B = \Omega$, what is $\mathbb{P}[A \mid B]$?*
3. *More generally, if $A \subseteq B$ what is $\mathbb{P}[A \mid B]$?*

Problem 3 *In terms of $\mathbb{P}[A]$ and $\mathbb{P}[B]$ calculate $\mathbb{P}[A \mid B]$ if A, B are independent.*

Problem 4 *Let $(\Omega, \mathbb{P})$ be a probability space, let $B \subseteq \Omega$ be so that $\mathbb{P}[B] > 0$. Show that the function defined on subsets of Ω as $\mu(A) := [A \mid B]$ is itself a probability function/measure on Ω.*

Problem 5 *Let $(\Omega, \mathbb{P})$ be a probability space, let $B_1, \ldots, B_N \subseteq \Omega$ be events that are pairwise disjoint but:*

$$\bigcup_{i=1}^{N} B_i = \Omega$$

Prove that for any $A \subseteq \Omega$:

$$\mathbb{P}[A] = \sum_{i=1}^{N} \mathbb{P}[A \mid B_i]\mathbb{P}[B_i]$$

Problem 6 *(Bayes' formula) Let* $(\Omega, \mathbb{P})$ *be a probability space, let* $A, B \subseteq \Omega$ *with nonzero probability. Prove that:*

$$\mathbb{P}[A \mid B] = \frac{\mathbb{P}[B \mid A]\,\mathbb{P}[A]}{\mathbb{P}[B]}$$

Problem 7 *(COVID testing) Suppose that at some stage during the pandemic that* 2% *of the population is infected. You are a doctor and order an antibody test for your patient and the test comes back positive. You know that the test has a* 95% *chance of correctly returning positive for an infected individual, but also a* 10% *chance of incorrectly producing a false positive for a healthy individual. What are the chances that the patient in front of you is in fact sick?*

Problem 8 *Continuing from the last example, how many tests do you have to order before you actually can be more than* 90% *sure that the patient in front of you is sick?*

2 Conditional Expectations

Definition 2 *Let* $(\Omega, \mathbb{P})$ *be a probability space and let* X *be a random variable with* $A \subseteq \Omega$ *an event with* $\mathbb{P}[A] > 0$*. The conditional expectation of* X *with respect to* A *is:*

$$\mathbb{E}[X \mid A] := \frac{1}{\mathbb{P}[A]} \sum_{\omega \in A} X(\omega)\mathbb{P}[\{\omega\}]$$

If $\mathbb{P}[A] > 0$ *and* 0 *otherwise.*

Problem 9 *Calculate the following conditional expectations:*

- *Let Ω be the set of possible outcomes of a 6 sided dice roll, let A be the event that you roll an even number, let X be the random variable that is the result of the dice roll. Calculate $\mathbb{E}[X \mid A]$.*
- *Let Ω be the set of possible outcomes of a 6 sided dice roll, let A be the event that you roll a 5, let X be the random variable that is the result of the dice roll. Calculate $\mathbb{E}[X \mid A]$.*
- *Let Ω be the set of possible outcomes of 4 independent coin flips, let A be the event you get at least 1 head, let X be the number of heads you get total. Calculate $\mathbb{E}[X \mid A]$.*
- *Let Ω be the set of possible outcomes of 10 independent coin flips, let A be the event the first coin is a head, let X be the number of heads you get total after (not including) the second flip. Calculate $\mathbb{E}[X \mid A]$.*

Problem 10 *Let $(\Omega, \mathbb{P})$ be a probability space and $A, B \subseteq \Omega$ are events with $\mathbb{P}[B] > 0$. Let $\mathbb{1}_A$ be the random variable such that $\mathbb{1}_A(\omega) = 1$ if $\omega \in A \subseteq \Omega$ and 0 otherwise. Prove that:*

$$\mathbb{P}[A \mid B] = \mathbb{E}[\mathbb{1}_A \mid B]$$

Definition 3 *Let $(\Omega, \mathbb{P})$ be a probability space, a partition of this space are disjoint events $\mathcal{F} = \{B_1, \ldots, B_N\}$ so that:*

$$\bigcup_{i=1}^{N} B_i = \Omega$$

The conditional expectation with respect to a partition $\mathcal{F}$ of a random variable X is defined to be the random variable so that for $\omega \in \Omega$:

$$\mathbb{E}[X \mid \mathcal{F}](\omega) = \mathbb{E}[X \mid B_i]$$

Where B_i is the unique $B_i \in \mathcal{F}$ so that $\omega \in B_i$.

Problem 11 *Let $(\Omega, \mathbb{P})$ be a probability space and X be a random variable. Prove that:*

$$\{A \subseteq \Omega : A = X^{-1}(r) \text{ for some } r \in \mathbb{R}\}$$

Form a partition in the above sense.

Definition 4 *The previous problem justifies the following definition. Let $(\Omega, \mathbb{P})$ be a probability space and X, Y be random variables. Let:*

$$\sigma(Y) := \{Y^{-1}(r) : r \in \mathbb{R} \text{ such that } r = Y(\omega) \text{ for some } \omega \in \Omega\}$$

Be the partition generated by Y. Then the conditional expectation of X with respect to Y is:

$$\mathbb{E}[X \,|\, Y] := \mathbb{E}[X \,|\, \sigma(Y)]$$

Problem 12 *Let $(\Omega, \mathbb{P})$ be a probability space and X be a random variable and $\mathcal{F}$ a partition. Prove that for any $A \in \mathcal{F}$:*

$$\mathbb{E}\big[X \mathbb{1}_A\big] = \mathbb{E}\big[\mathbb{E}[X \,|\, \mathcal{F}] \mathbb{1}_A\big]$$

Definition 5 *Let $(\Omega, \mathbb{P})$ be a probability space and $\mathcal{F}$ and $\mathcal{G}$ be both partitions of Ω. We say that $\mathcal{F}$ is refined by $\mathcal{G}$ if for any $A \in \mathcal{F}$ there exists a finite number of sets $B_1, \ldots, B_n \in \mathcal{G}$ so that:*

$$\bigcup_{k=1}^{N} B_i = A$$

Problem 13 *(Tower Property) Let $(\Omega, \mathbb{P})$ be a probability space and X be a random variable and $\mathcal{F}, \mathcal{G}$ partitions where $\mathcal{G}$ refines $\mathcal{F}$. Then:*

$$\mathbb{E}[\mathbb{E}[X \,|\, \mathcal{F}] \,|\, \mathcal{G}] = \mathbb{E}[X \,|\, \mathcal{F}]$$

Problem 14 *Give an example where:*

$$\mathbb{E}\big[\mathbb{E}[X \,|\, \mathcal{F}] \,|\, \mathcal{G}\big] = \mathbb{E}\big[\mathbb{E}[X \,|\, \mathcal{G}] \,|\, \mathcal{F}\big]$$

Problem 15 *(Independence) Suppose X, Y are independent random variables. Show that:*

$$\mathbb{E}[X \,|\, Y] = \mathbb{E}[X]$$

Problem 16 *(Conditional Jensen's inequality) Let $(\Omega, \mathbb{P})$ be a probability space with X a random variable and $\mathcal{F}$ a partition, show that:*

$$\mathbb{E}[X^2 \,|\, \mathcal{F}] \geq \mathbb{E}[X \,|\, \mathcal{F}]^2$$

Problem 17 *Let $X_1, \ldots, X_n$ be independent and identically distributed random variables (iid). Let $S_n = \sum_{k=1}^{n} X_k$ be another random variable. Prove that:*

$$\mathbb{E}[X_1 \,|\, S_n] = \frac{1}{n} S_n$$

3 Markov Processes

Definition 6 *A Markov process is a sequence of random variables $X_1, \dots$ such that:*

$$\mathbb{E}[X_{k+1} \mid X_1, \dots, X_k] = \mathbb{E}[X_{k+1} \mid X_k]$$

Here $\mathbb{E}[X_{k+1} \mid X_1, \dots, X_k]$ means conditioning on the partition generated by sets of the form $X_1^{-1}(s_1) \times \cdots \times X_k^{-1}(s_k)$.

Problem 18 *Which of the following are Markov processes?*

1. *The random variables $X_1, X_2, \dots$ where each X_k is an independent roll of a 6 sided dice.*
2. *The random variables $X_1, X_2, \dots$ where each $X_k = \sum_{n=1}^{k} A_n$ and each A_n is an independent 6 sided dice roll.*
3. *Consider the following setup, there are two boxes and one ball. At each step, the ball has a $\frac{1}{2}$ chance of being moved to the other box and a $\frac{1}{2}$ chance of staying. Let X_k be the random variable which is 1 if the ball is in box 1 at step k and 0 otherwise. The ball starts with a $\frac{1}{4}$ chance of being in box 1*
4. *The random variables $X_1, X_2, \dots$ where each $X_1 = 1$ and $X_2 = 2$ always and $X_{k+1} = X_{k-1}$ with probability $\frac{1}{2}$ and $X_{k+1} = X_k$ otherwise.*

Problem 19 *Suppose for each of the above sequences of random variables, the process is stopped at X_N for some large N (ie. the sequence of random variables is finite). Write down (a possible) probability space for each. Are these options unique?*

Definition 7 *We will mainly be working with **time homogeneous** Markov processes. Namely, there exits a set E, called a state space, so that the Markov process can be constructed on $(\Omega = E^N, \mathbb{P})$, with $X_k(e_1, \dots, e_n) = {e_k}$[2], and for any $m, k \in \mathbb{N}$:*

$$\mathbb{P}[X_{k+m} = e \mid X_k = e'] \text{ is independent of } k$$

With this, we define the transition probabilities[3]*:*

$$p_m(x, y) := \mathbb{P}[X_{m+1} = y \mid X_1 = x]$$

[2]Technically, we have only discussed real valued random variables, but it is better to think of e_k here not as a number but as just a label

[3]In our current framework, this definition requires an additional condition to make sense. Namely: we need the initial distribution to have a nonzero chance of being at any given $x \in E$.

Problem 20 *For each of the previous processes which you determined to be a Markov process, write down the transition probabilities $p_1(x, y)$ for all x, y.*

Problem 21 *Prove that:*

$$p_n(x, y) = \sum_{z \in E} p_{n-1}(x, z) p_1(z, y)$$

Now if we enumerate $E = \{x_1, \ldots, x_n\}$ and set:

$$P_n := \begin{bmatrix} p_n(x_1, x_1) & \ldots & p_n(x_n, x_1) \\ \vdots & \ddots & \vdots \\ p_n(x_1, x_n) & \ldots & p_n(x_n, x_n) \end{bmatrix}$$

show that there exists a matrix Λ so that $P_{n+1} = \Lambda P_n$, for any n.

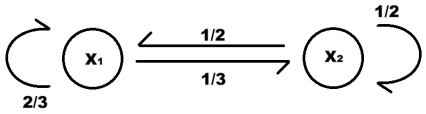
Problem 22 *(Chapman–Kolmogorov equation) Prove that:*

$$p_{n+m}(x, y) = \sum_{z \in E} p_n(x, z) p_m(z, y)$$

Definition 8 *One common Markov process is a random walk on a graph. Let $V = \{x_1, \ldots, x_N\}$ be the vertex set, then the random walk on V with transition matrix Λ is a Markov process $X_1, \ldots$ with:*

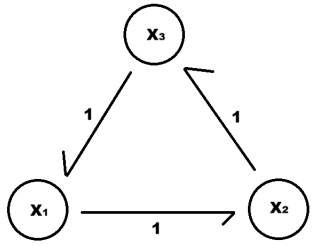
$$p_1(x_i, x_j) = \Lambda_{j,i} = \mathbb{P}[X_2 = x_j \mid X_1 = x_i]$$

These are frequently graphically represented, for instance, the process with transition matrix $\begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{bmatrix}$ may be represented as:

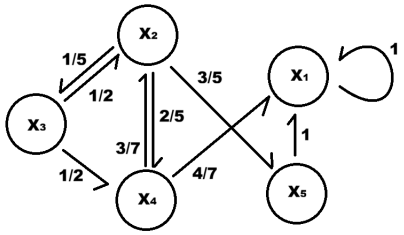


Problem 23 *Write down the associated transition matrices to the following random walks on graphs.*

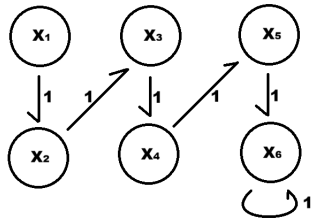
1.



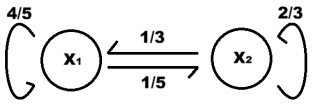
2.



3.



Problem 24 *For he following random walk on a graph, calculate the distribution after n steps in terms of n. Initially the distribution for X_1 is $\frac{1}{2}$ at x_1 and $\frac{1}{2}$ at x_2*



4 Bonus questions

Problem 25 *(Reflection Principle) Let $Y_1, \ldots, Y_n$ be independent and identically distributed random variables with centered distributions, that is $\mathbb{P}[Y_k = s] = \mathbb{P}[Y_k = -s]$ for all $1 \leq k \leq n$ and $s \in \mathbb{R}$. Let $X_0 = 0$ and X_k be the random variable defined as:*

$$X_k = \sum_{i=1}^{k} Y_i$$

Then:

$$\mathbb{P}[\max_{1 \leq k \leq n} X_k \geq a] \leq 2\mathbb{P}[X_n \geq a] - \mathbb{P}[X_n = a]$$

For any $a > 0$

Problem 26 *(Bernstein–Chernov bound) Let $p_1, \ldots, p_n$ be real numbers in $(0, 1)$ and for each $1 \leq k \leq n$ toss a (unfair) coin with probability p_k of landing on heads and $1 - p_k$ tails, let X_k be 1 if it is heads and 0 otherwise. Set:*

$$S := \sum_{k=1}^{n} X_k \text{ and } m = \mathbb{E}[S]$$

Prove that for any $\delta > 0$ we have:

$$\mathbb{P}[S_n \geq (1+\delta)m] \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \right)^m$$

$$\mathbb{P}[S_n \leq (1-\delta)m] \leq e^{-\frac{\delta^2 m}{2}}$$

Problem 27 *Suppose you are flipping a fair coin. What is the probability that you see the sequence HHT before you see the sequence TTH? Now what about the probability of seeing HHT before THT?*

Problem 28 *(FKG inequality, special case) Let X be a random variable that takes values in $0, 1, \ldots, N$ for some large integer N, suppose that $f, g : \mathbb{N} \to \mathbb{N}$ be two functions with the property that $f(a) > f(b)$ and $g(a) > g(b)$ whenever $a > b$ are natural numbers. Prove that the random variables $f(X)$ and $g(X)$ satisfy the following*[4]*:*

$$\mathbb{E}[f(X)g(X)] \geq \mathbb{E}[f(X)]\mathbb{E}[g(X)]$$

[4]Yes, this is a repeat of last weeks question but I do not believe anyone got it so I would like to pose it again. Good luck!