VISUALISING MARKOV PROCESSES

Stephanie Budgett, Azam Asanjarani, and Heti Afimeimounga
The University of Auckland
azam.asanjarani@auckland.ac.nz

Researchers and educators have long been aware of the misconceptions prevalent in people's probabilistic reasoning processes. Calls to reform the teaching of probability from a traditional and predominantly mathematical approach to include an emphasis on modelling using technology have been heeded by many. The purpose of this paper is to present our experiences of including an activity based on an interactive visualisation tool in the Markov processes module of a first-year probability course. Initial feedback suggests that the tool may support students' understanding of the equilibrium distribution and points to certain aspects of the tool that may be beneficial. A targeted survey, to be administered in Semester 1, 2022, aims to provide more insight.

INTRODUCTION AND BACKGROUND LITERATURE

At no time has it been more apparent that people's innate probabilistic reasoning can lead to misleading conclusions. David Spiegelhalter, one of the UK's most respected risk communicators, has been called upon many times to unpack COVID-19 statistics and clarify their interpretation (e.g., addressing the confusion of the inverse fallacy, Spiegelhalter & Masters, 2021). Our intuitive reactions to probabilistic situations lead to muddled thinking. Kahneman (2011) uses the terms System 1 and System 2 to describe two types of thinking commonly used by humans. Briefly, System 1 is based on gut instinct and intuition, whereas System 2 is based on logical reasoning. Misconceptions such as the base rate fallacy and confusion of the inverse are deeply ingrained and, even if we have an awareness of them, our fast, instinctive and emotional System 1 brain often overrules our slower, more deliberate and effortful System 2 brain, leading us to erroneous judgments.

Making the wrong decision, or coming to an ill-formed conclusion, in the face of uncertainty can have far-reaching consequences (e.g., Gigerenzer et al., 1998; Nance & Morris, 2005). Although public debate, involving Spiegelhalter, Royal Statistical Society Statistical Ambassador Masters, and many other excellent communicators, may raise self-awareness about our own weaknesses when dealing with probabilistic situations, it is arguable that we should be laying better foundations in educational settings.

As has been widely reported, the traditional approach to teaching probability, relying on mathematical formulae and symbols, has led to many students struggling to grasp fundamental concepts (e.g., Greer & Mukhopadhyay, 2005). Furthermore, such an approach can obscure the underpinning random phenomena described by mathematical representations (Moore, 1997). In the last 20 years, in response to recommendations from probability and statistics educators (e.g., Batanero et al., 2016; Pfannkuch & Ziedins, 2014), a modelling approach to teaching probability, incorporating both static and dynamic visualisations, is being introduced in many classrooms (e.g., Fitch & Budgett, 2018). Research suggests that the process of understanding a statistical concept, such as the equilibrium distribution of a Markov chain, involves the creation of a mental image of that concept (Tall & Vinner, 1981). The ability to represent the concept through dynamic interactive technology has the potential to help students deepen their understanding of a concept (Burrill, 2018). For example, such technology allows for the exploration of 'what-if' scenarios whereby students can predict what might happen to the equilibrium distribution if the initial distribution of the Markov chain is changed.

Markov processes are typically introduced at the tertiary level. The combination of counter-intuitive probability concepts and advanced mathematical techniques mean that many students struggle with the topic (Wang, 2001b). When considering the training of researchers in the use of statistics, noting the difficulties experienced by novices, it was recommended that technology-based simulation and visualisation could be used to facilitate access to the important concepts underpinning Markov processes (Wang, 2001a). In line with prior research in other probability topics (Pfannkuch et al., 2015), a technology-based approach has the potential to allow students to link abstract concepts and representations such as transition matrices, transition probabilities, state diagrams, and equilibrium distributions.

In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the Gap: Empowering & Educating Today's Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11 2022), Rosario, Argentina.* International Association for Statistical Education. iase-web.org

OUR TEACHING CONTEXT

Our introductory probability course has an algebra/calculus co-requisite and attracts over 300 students each year, distributed across two 12-week semesters. The course serves as a prerequisite for our intermediate (second year) probability course, which is a requirement for all students wishing to pursue postgraduate studies in statistics. The course format includes three lectures and one tutorial per week, each lasting 50 minutes. In addition to a set of probability problems to work through, each tutorial session involves a discussion question. The discussion question provides a team-working opportunity for students to consider some more complex or chaotic probability problems with each other and to deepen their understanding about probability and randomness.

Prior to the pandemic, the lectures and tutorials took place in-person, with recordings of lectures provided via the course learning management system, Canvas (https://www.instructure.com/canvas). Due to the pandemic, in-person lectures and tutorials were the exception, rather than the rule. Furthermore, over 150 students who were unable to travel to New Zealand due to border restrictions were enrolled remotely. We cater for students with a wide range of mathematical backgrounds, from the diffident and anxious to the proficient and confident. Discrete-time, discrete-state Markov processes, referred to as Markov chains throughout this article, is the final topic in the course and includes equilibrium distributions, limiting distributions, hitting/reaching probabilities, and expected hitting times.

Course instructors are connected with statistics and probability education research and consequently have been proactive in experimenting with a modelling approach incorporating visual representations where appropriate. The students are also encouraged to make connections between the different representations used in the course, for example, words, mathematical symbols, and visual imagery, in order to promote versatile thinking, which is the hallmark of an expert mathematical and statistical thinker (Thomas, 2004). Our experiences of using the eikosogram (a visual representation of the two-way table, based on the unit square) and the pachinkogram (a dynamic, interactive probability tree) in an earlier part of the course have been discussed previously (Fitch & Budgett, 2018). We now report on the Markov Chains tool that is based on a dynamic visualisation created by Victor Powell on the Setosa blog (see: https://setosa.io/blog/2014/07/26/markov-chains/index.html).

THE MARKOV CHAINS TOOL

The Markov Chains tool (https://www.stat.auckland.ac.nz/~wild/MarkovChains/) is a dynamic visual representation of a Markov chain (Pfannkuch & Budgett, 2016). Probabilities are entered into a transition matrix, which results in the creation of a state diagram (Figure 1).

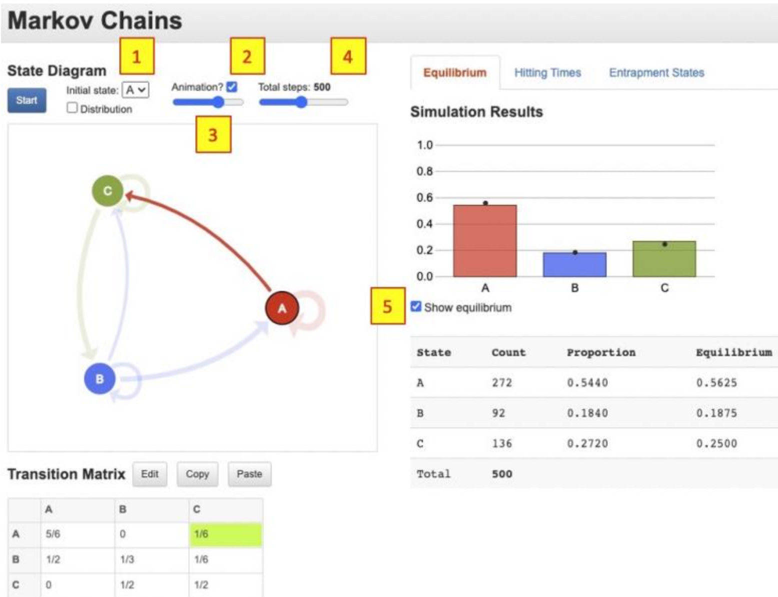


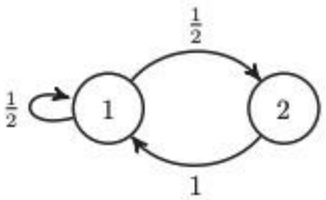
Figure 1. Screenshot of Markov Chains tool showing visualisation of state diagram, transition matrix, and resulting equilibrium distribution

The arrows connecting the states in the state diagram are proportional to the probabilities associated with moving from one state to another. By selecting the tabs to the right, the equilibrium distribution, hitting times, or entrapment states can be explored. The equilibrium tab allows the user to specify (1) the initial state (or a distribution across all possible initial states), (2) whether to animate the process or not, (3) the speed of the simulation, (4) the number of steps to simulate, and (5) the theoretical equilibrium distribution. The highlighted cell in the transition matrix represents the final step in this particular simulation, a step from state A to state C (see Figure 1).

MARKOV CHAINS AND EQUILIBRIUM DISTRIBUTION

To motivate the topic, the Markov Chains module developed for the class begins with a discussion of some real-life stochastic processes such as those in biology, the stock market, risk, and insurance. Following introduction of the Markov property, transition diagrams, transition matrices, sample path behaviours, and n-step transition probabilities are presented. The concepts of equilibrium and limiting distributions are then introduced, including how to evaluate the equilibrium distribution mathematically via a set of full (or global) and detailed balance equations. Additionally, the conditions under which the limiting distribution exists and is the same as the equilibrium distribution are provided and discussed. Students are then guided through the calculation of equilibrium distributions for several Markov chains and provided with worked examples, two of which are shown in Figure 2.

Example 1
Consider the Markov chain with the following transition diagram:



Find all possible equilibrium distributions.

Example 2
What is/are the equilibrium distribution(s) of the Markov chain with state space S = {0, 1, 2} and transition matrix:

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}?$$

Figure 2. Two worked examples requiring students to find the equilibrium distributions

The Markov Chains tool is presented to students after the introduction of the transition matrix and the transition diagram. At this point, the purpose is to demonstrate the stochastic nature of a Markov process and to illustrate how varying probabilities in the transition matrix results in varying trajectories through the process. Observing the numerical values associated with the equilibrium distribution (Figure 1, item 5) allows students to note, following a long run of the process, the proportion of time in which the process spends in each state. When students are familiar with both equilibrium and limiting distributions, and are able to evaluate them mathematically, they are again shown the Markov Chains tool. They can use it to check their calculations for different examples involving between three and six states. As part of a weekly tutorial, students are asked to manually calculate the equilibrium distribution(s) for a given Markov chain, and to show their work. They are then asked to use the Markov Chains tool to check their result and to provide feedback on their use of the tool.

STUDENT FEEDBACK

In October 2021, the weekly tutorial session involved a discussion question that required students to find all possible equilibrium distributions for a given Markov chain[1] using the full balance equations. They were then required to check their answer using the Markov Chains tool. Follow-up questions were asked in order to elicit the students' opinions as to if and how the Markov tool helped them to understand how an equilibrium distribution is calculated, and what it represents. Responses were received from 83 students (Figure 3).

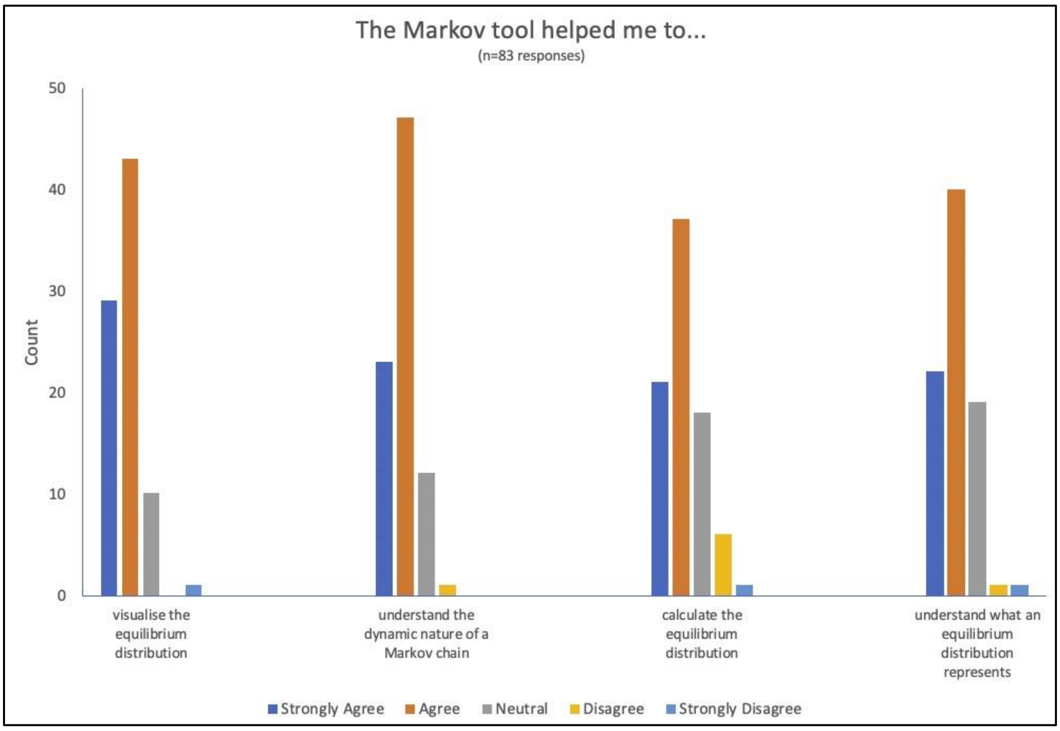


Figure 3. Student responses (n = 83) to four questions regarding the use of the Markov Chains tool

Of those who responded, the majority either agreed or strongly agreed that the tool was beneficial in terms of helping to understand the dynamic nature of a Markov chain. Furthermore, students were generally in agreement that the tool assisted them to understand, visualize, and calculate the equilibrium distribution.

In addition, the following open-ended question was asked: *For you, what was the best feature of the tool?* A word-cloud was generated from the student responses (Figure 4).

Figure 4. Word cloud generated from student responses (n = 83)

Most of the responses to the open-ended question were favourable, with many making explicit mention of the visualisation and animation aspects of the tool. Some student responses are given below:

- *Animation of the path taken was helpful for making sense of what was actually happening*
- *The best feature is the dynamic nature*
- *The animation was a great visualisation of what was going on*

CONCLUSION

Information gleaned from the student responses to follow-up questions suggest that the Markov Chains tool was helpful in a variety of ways. The prominence of the words, *animation* and *visualisation*, and their derivatives suggest that the dynamic visual aspects were of particular benefit for students. Anecdotally, the students enjoyed interacting with the Markov Chains tool. They thought that it was easy to use; it helped them to check their calculations; and they found it aesthetically pleasing.

This exploratory work has provided the motivation to investigate introductory probability students' use of the Markov Chains tool in more depth. It remains to be determined if the characteristics of the Markov Chains tool complement and deepen understanding of the necessary mathematical and statistical concepts underpinning the equilibrium and limiting distributions. For example, does interacting with the Markov Chains tool allow students to make sense of the underpinning full balance equations? How does the visualization aspect of the tool contribute to students' understanding of the stochastic nature of Markov Chains? To attempt to answer these questions, ethics permission will be sought in order to carry out focus group sessions with previous students of this introductory probability course in order to understand more about how the Markov Chains tool may have impacted their learning.

NOTES

[1]Examples each had one unique equilibrium distribution.

REFERENCES

Batanero, C., Chernoff, E., Engel, J., Lee, H., & Sanchez, E. (2016). *Research on teaching and learning probability*. Springer. https://doi.org/10.1007/978-3-319-31625-3_1

Burrill, G. (2018). Concept images and statistical thinking: The role of interactive dynamic technology. In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward*. *Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018) Kyoto, Japan*. ISI/IASE. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_9F3.pdf?1531364300

Fitch, A. M., & Budgett, S. (2018). Conditional probability: Friend or foe? In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018) Kyoto, Japan*. ISI/IASE. https://iase-web.org/icots/10/proceedings/pdfs/ICOTS10_6A1.pdf?1531364278

Gigerenzer, G., Hoffrage, U., & Ehert, A. (1998). AIDS counseling for low-risk clients. *AIDS Care, 10*(2), 197–211. https://doi.org/10.1080/09540129850124451

Greer, B., & Mukhopadhyay, S. (2005). Teaching and learning the mathematization of uncertainty: Historical, cultural, social and political contexts. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp.297–324). Springer. https://doi.org/10.1007/0-387-24530-8_13

Kahneman, D. (2011). *Thinking, fast and slow.* Allen Lane.

Moore, D. (1997). Probability and statistics in the core curriculum. In J. Dossey (Ed.), *Confronting the core curriculum* (pp. 93–98). Mathematical Association of America.

Nance, D. A., & Morris, S. B. (2005). Juror understanding of DNA evidence: An empirical assessment of presentation formats for trace evidence with a relatively small random-match probability. *Journal of Legal Studies, 34*(2), 395–444. https://doi.org/10.1086/428020

Pfannkuch, M., & Budgett, S. (2016). Markov processes: Exploring the use of dynamic visualizations to enhance student understanding. *Journal of Statistics Education, 24*(2), 63–73. https://doi.org/10.1080/10691898.2016.1207404

Pfannkuch, M., Budgett, S., & Arnold, P. (2015). Experiment-to-causation inference: Understanding causality in a probabilistic setting. In A. Zieffler & E. Fry (Eds.), *Reasoning about uncertainty: Learning and teaching informal inferential reasoning* (pp. 95–127). Catalyst Press.

Pfannkuch, M., & Ziedins, I. (2014). A modelling perspective on probability. In E. J. Chernoff, & B. Sriraman (Eds.), *Probabilistic thinking: Presenting plural perspectives* (pp. 101–116). Springer. https://doi.org/10.1007/978-94-007-7155-0_5

Spiegelhalter, D., & Masters, A. (2021, April 25). *Covid, false positives and conditional probabilities*... The Guardian. https://www.theguardian.com/theobserver/commentisfree/2021/apr/25/covid-false-positives-and-conditional-probabilities

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151–169. https://doi.org/10.1007/BF00305619

Thomas, M. O. (2004, July 4–11). Conceptual representations and versatile mathematical thinking [Regular lecture]. *Tenth International Congress on Mathematical Education*. Copenhagen, Denmark.

Wang, A. L. (2001a). How much can be taught about stochastic processes and to whom? In C. Batanero (Ed.), *Training researchers in the use of statistics* (pp. 73–85). International Association for Statistical Education and International Statistical Institute.

Wang, A. L. (2001b). Introducing Markov chains models to undergraduates. *Proceedings of the 53rd Session of the International Statistical Institute*. ISI. https://iase-web.org/documents/papers/isi53/596.pdf?1402524976