Appendix 1

Markov tasks

Task 1

AA Rentals is a small rental car agency located in Auckland. The agency currently has two pick-up/drop-off offices, one located in the central city and the other located at the airport. Customers may rent a car from either location, and return a car to either location. Over time, the manager notices that about 40% of cars rented from the central city are returned to the central city, while about 80% of cars rented from the airport are returned to the airport.

(a) Draw a transition diagram for this situation.

(b) Write down the transition matrix for this situation.

(c) Calculate the equilibrium distribution for this situation. Interpret what this means for AA rentals.

(d) If a car is picked-up in the central city, how many times would you expect it to be hired out before it arrives at the airport? Give an intuitive answer. Now sketch a distribution of the number of times the car is hired out before it arrives at the airport. *What range do you expect?*

(e) If a car is picked up in the central city, how many times would you expect it to be hired out before it is returned back to the central city? Give an intuitive answer. Now sketch a distribution of the number of times the car is hired out before it returns to the airport. *What range do you expect?*

(f) Use the software tool to model this situation *and compare with predictions*.

Task 2

AA Rentals expands their business by introducing two further pick-up/drop-off offices, one in Christchurch and one in Wellington. Over time, records suggest that rental cars picked up from Auckland central city are dropped off at locations Auckland central city, Auckland airport, Christchurch and Wellington with probabilities 0.2, 0.6, 0.1 and 0.1 respectively. Corresponding probabilities for rental cars are 0.1, 0.7, 0 and 0.2 for cars picked up from Auckland airport, 0, 0.3, 0.6 and 0.1 for cars picked up from Christchurch, and 0, 0.4, 0.1 and 0.5 for rental cars picked up from Wellington.

(a) Draw the transition diagram for this situation.

(b) Write down the transition matrix for this situation.

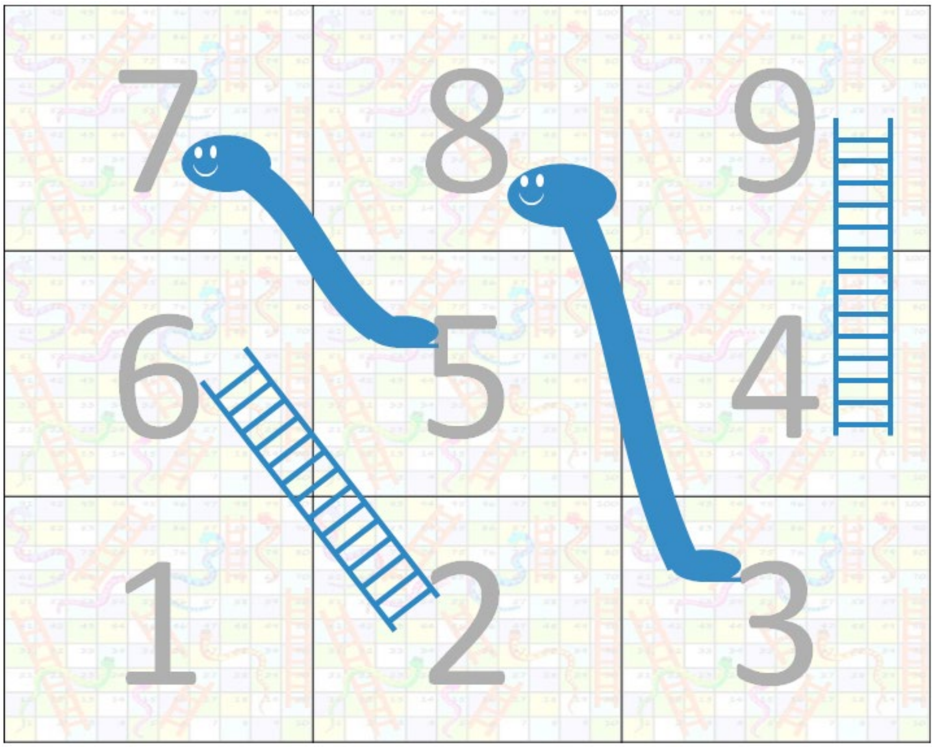
(c) Describe the equilibrium distribution; intuitively rank the probabilities from highest to lowest. Interpret what the equilibrium distribution means for AA rentals.

(d) If a car is picked up in Auckland city centre, how many times would you expect it to be hired out before it arrived at Auckland airport (Christchurch, Wellington) – or was returned to Auckland city centre? Intuitively rank and estimate the expected times of these scenarios from highest to lowest. Sketch the distributions. What range do you expect?

(e) Use the software tool to model this situation *and compare with predictions*.

Task 3

A cut-down version of the popular board game Snakes and Ladders has been devised. Consider the following board.



The rules are as follows:

Each player is required to roll a '6' before being eligible to move on to the Start box. Once a player is in the Start box, their next roll of the die determines how many squares they travel. The first person to reach the ninth square is deemed the winner. A player is not required to reach the ninth square **exactly**, i.e. a player on square 6 who throws a 3 or higher will reach square 9.

(a) Play the game several times, each time recording the number of throws of the die required for a win.

Game	Number of rolls of die
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	

Here is the transition matrix for the Snakes and Ladders game.

(b) Why are there only seven states for this game? Identify the states and label them on the board.

(c) Explain the first line of the transition matrix. Explain the fifth line of the transition matrix.

(d) How many rolls of the die do you expect to make before you win the game? Give an intuitive answer. Now sketch a distribution of the number of times the die is rolled before you win the game. What range do you expect?

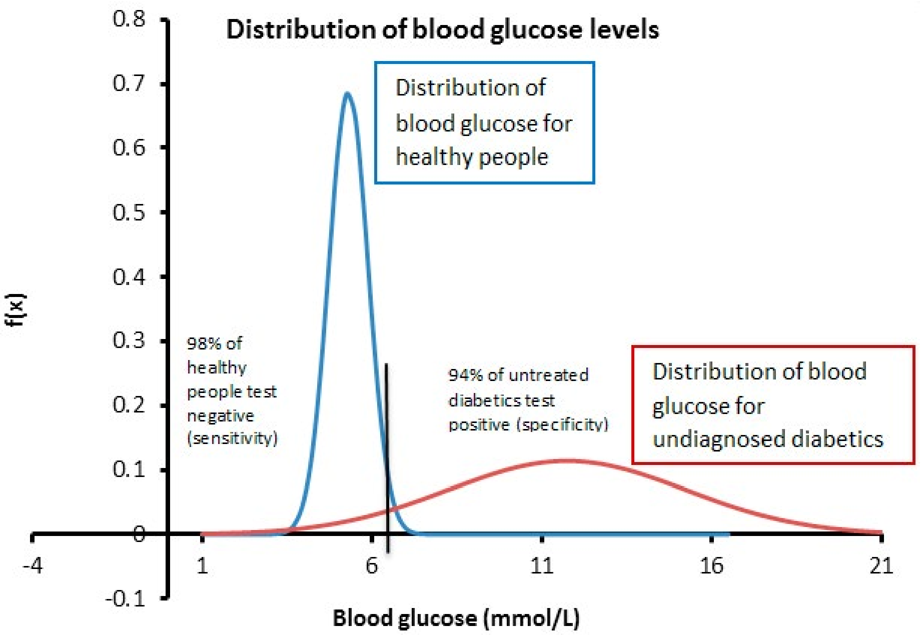
(e) Use the software tool to model this situation and compare with predictions.

Appendix 2

Pachinkogram task

A new housing development has been built in your neighbourhood. In order to service the needs of this new community, a new health clinic has opened. As part of the health clinic's enrolment procedure, new patients are required to undergo health check-ups which include, among other things, a series of blood tests. One such test is designed to measure the amount of glucose in an individual's blood. This measurement is recorded after the individual fasts (abstains from eating) for a prescribed period of time. Fasting blood glucose levels in excess of 6.5mmol/L are deemed to be indicative of diabetes. This threshold of 6.5mmol/L works most of the time with about 94% of people who have diabetes being correctly classified as diabetics and about 98% of those not having diabetes being correctly classified as non-diabetics.

The prevalence of diabetes in the NZ population is about 7%[1] (i.e. approximately 7% of the NZ population are estimated to have diabetes).



**Graph above taken from* Pfannkuch, Seber, & Wild, 2002

1. As part of enrolment in this health clinic, an individual has a fasting blood test. He/she is told that his/her blood glucose level is higher than 6.5mmol/L. What are the chances he/she has diabetes?
2. As part of enrolment in this health clinic, a male aged between 65 and 74 has a fasting blood test. He is told that his blood glucose level is higher than 6.5mmol/L. What are the chances that he has diabetes?
3. As part of his enrolment in this health clinic, a person of Pacific ethnicity and aged over 75 has a fasting blood test. He/she is told that his/her blood glucose level is higher than 6.5mmol/L. What are the chances that he/she has diabetes?
4. Suppose the threshold for diagnosing a person with diabetes is changed to 6mmol/L. Someone from the general population (no further information available) is told that his/her blood glucose level is higher than 6mmol/L. What are the chances that he/she has diabetes?
5. The Ministry of Health has a fixed amount of financial resources for a new diabetes treatment initiative and needs to decide how to allocate funds in order to optimize the treatment of New Zealanders. Part of this optimization process requires information on the expected number of people with diabetes enrolling in a particular health clinic.

A health clinic for approximately 10,000 patients is situated in Mangere-Otahuhu. Suppose that everyone enrolled in the health clinic has a blood test for diabetes.

(i) Estimate the number of people enrolled in this health clinic that test positive for diabetes.

(ii) Estimate the number of people enrolled in this health clinic that are diabetic but test negative.

6. Repeat question 5 where a health clinic for approximately 10,000 patients is situated in Devonport-Takapuna.

Appendix 3

Poisson processes tasks

1. Large amounts of data are routinely collected from large sporting events such as the FIFA World Cup. For example, data on the times at which goals were scored in all games in the 2014 World Cup are available from www.fifa.com/worldcup/archive/brazil2014/matches/index.html. We have access to inter-goal waiting time data from the 1990, 1994, 1998 and 2002 World Cups held in Italy, the USA, France and Korea/Japan respectively. (The data on inter-goal waiting times for the four World Cups held between 1990 and 2002 was kindly provided by Singfat Chu-Chun-Lin, National University of Singapore.) Note that any goals scored in extra time, that is, after the 90-minute match duration, have been excluded.
   (a) Using this data and the software tool, explore the distribution of the waiting times between goals. Comment on what you see.
   (b) By manipulating the Multiply count interval box, predict the expected number of goals scored per match.
   (c) Observe the distribution of the number of goals scored per match. Is the Poisson distribution a reasonable model for this situation? Explain your answer, including discussion of the underlying assumptions of the Poisson model.
2. Students in an introductory statistics course at a US university were set the task of collecting data on bus schedules in Chicago. Each student was allocated a 2-hour block during the rush hour (3pm to 7pm, Monday to Friday) and sat at the same designated stop noting the time of every arrival of the #55 bus along with other various pieces of information.
   (a) Using this data and the software tool, explore the distribution of the waiting time between buses. Comment on what you see.
   (b) By manipulating the Multiply count interval box, predict the expected number of buses arriving in a 1-hour period.
   (c) Observe the distribution of the number of buses arriving per hour. Is the Poisson distribution a reasonable model for this situation? Explain your answer, including discussion of the underlying assumptions of the Poisson model.
3. In Questions 1 and 2, you used actual data to discover whether or not the Poisson model provided a reasonable approximation to two realities; the first being the expected number of goals in a World Cup match, the second being the expected number of buses to arrive at a particular bus stop in a 1-hour period. You should now have an estimate of the expected number of goals scored in a World Cup match (your answer to 1(b)).

   The software tool is able to simulate waiting times with different generating distributions. Simulate waiting times between goals in World Cup matches for each of the five generating distributions (Constant, Uniform, Exponential, Triangular, and Symmetric Triangular) using your answer to 1(b) as the input for Rate λ per match. Which, if any, provides the best approximation to reality? What conclusions can you make about the waiting time between successive goals in World Cup matches?