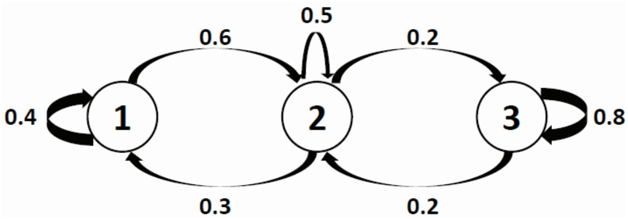
ECE 3100 **HOMEWORK ASSIGNMENT 9 - LAST HOMEWORK** **Fall 2014**

19 Nov. 2014 – Due MONDAY, 01 Dec. 2014 in Dropbox
20% grade reduction per late day submission
There are 14 problems to this homework

1. Read the book, Chapters 7 (not section 7.5), Chapter 5 (not section 5.5), Chapter 8, Chapter 9 (only section 9.1).

2. *Example 7.11.* Read and re-do Example 7.11 - this is a good example. Plot the probabilities a_i as a function of ρ (and provide the plots in your solutions) for $m = 100$ dollars for each of the following cases: $i = 1, 3, 10, 60$. Comment on what you observe. If $p = 1/3$, what is the probability of winning a_{100}? Does this make sense? Discuss.

3. *Problem 7.13.* Consider the Markov chain below. We refer to a transition that results in a state with higher (lower) index as birth (death). Calculate the following quantities, assuming that when we start observing the chain, it is already in steady-state.

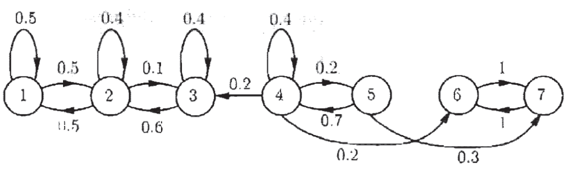


(a) For each state i, the probability that the current state is i.
(b) The probability that the first transition we observe is a birth.
(c) The probability that the first change of state we observe is a birth.
(d) The conditional probability that the process was in state 2 before the first transition that we observe, given that this transition was a birth.
(e) The conditional probability that the process was in state 2 before the first change of state that we observe, given that this change of state was a birth.
(f) The conditional probability that the first observed transition is a birth given that it resulted in a change of state.
(g) The conditional probability that the first observed transition leads to state 2 given that it resulted in a change of state.

4. *Problem 7.28.* There are m classes offered by a particular department, and each year, the students rank each class from 1 to m, in order of difficulty, with rank m being the highest. Unfortunately, the ranking is completely arbitrary. In fact, any given class is equally likely to receive any given rank on a given year (two classes may not receive the same rank). A certain professor chooses to remember only the highest ranking his class has ever gotten.

(a) Find the transition probabilities of the Markov chain that models the ranking that the professor remembers.
(b) Find the recurrent and the transient states.
(c) Find the expected number of years for the professor to achieve the highest ranking given that in the first year he achieved the ith ranking.

5. *Problem 7.30.* Consider the Markov chain below:



(a) Identify the transient and recurrent states. Also, determine the recurrent classes and indicate which ones, if any, are periodic.

(b) Does there exist steady-state probabilities given that the process starts in state 1? If so, what are they?
(c) Do there exist steady-state probabilities given that the process starts in state 6? If so, what are they?
(d) Assume that the process starts in state 2 but we begin observing it after it reaches steady-state.
   (1) Find the probability that the state increases by one during the first transition we observe.
   (2) Find the conditional probability that the process was in state 2 when we started observing it, given that the state increased by one during the first transition that we observed.
   (3) Find the probability that the state increased by one during the first change of state that we observed.
(e) Assume that the process starts in state 4.
   (1) For each recurrent class, determine the probability that we eventually reach that class.
   (2) What is the expected number of transitions up to and including the transition at which we reach a recurrent state for the first time?

6. *Problem 5.1.* A statistician wants to estimate the mean height h (in meters) of a population, based on n independent samples $X_l, ..., X_n$ chosen uniformly from the entire population. He uses the sample mean $M_n = (X_l + ... + X_n)/n$ as the estimate of h, and a rough guess of 1.0 meters for the standard deviation of the samples X_t.

(a) How large should n be so that the standard deviation of M_n is at most 1 centimeter?
(b) How large should n be so that Chebyshev's inequality guarantees that the estimate is within 5 centimeters from h, with probability at least 0.99?
(c) The statistician realizes that all persons in the population have heights between 1.4 and 2.0 meters, and revises the standard deviation figure that he uses based on the bound of Example 5.3. How should the values of n obtained in parts (a) and (b) be revised?

7. *Problem 5.4.* In order to estimate f, the true fraction of smokers in a large population, Alvin selects n people at random. His estimator M_n is obtained by diviging S_n, the number of smokers in his sample, by n., i.e., $M_n = S_n/n$. Alvin chooses the sample size n to be the smallest possible number for which the Chebyshev inequality yields a guarantee that

$$P(|M_n - f| \geq \epsilon) \leq \delta,$$

Where ϵ and δ are some pre-specified tolerances. Determine how the value of n recommended by the Chebyshev inequality changes in the following cases:

(a) The value of ϵ is reduced to half of its original value.
(b) The probability δ is reduced to half of its original value.

8. *Problem 5.8.* Before starting to play the roulette in a casino, you want to look for biases that you can exploit (i.e., increase your odds of winning). You therefore watch 100 rounds that result in a number between 1 and 36 and count the number of rounds for which the result is odd. If the count exceeds 55, you decide that the roulette is not fair. Assuming that the roulette is fair, find an approximation for the probability that you will make the wrong decision. *Explain clearly your steps/calculations.*

9. *Problem 5.9.* During each day, the probability that your computer's OS crashes at least once is 5%, independent of every other day. You are interested in the probability of at least 45 crash-free days out of the next 50 days.

(a) Find the probability of interest by using the normal approximation to the binomial. *Explain clearly your steps/calculations.*
(b) Repeat part (a), this time using the Poisson approximation of the binomial. *Explain clearly your steps/calculations.*

10. *Problem 5.10.* A factory produces X_n gadgets on day n, where the X_n are independent and identically distributed random variables, with mean 5 and variance 9.

(a) Find an approximation to the probability that the total number of gadgets produced in 100 days is less than 440. *Explain clearly your steps/calculations.*
(b) Find (approximately) the largest value of n such that

$$P(X_1 + ...X_n \geq 200 + 5n) \leq 0.05.$$

Explain clearly your steps/calculations.

(c) Let N be the first day on which the total number of gadgets produced exceeds 1000. Calculate an approximation to the probability that $N \geq 220$. *Explain clearly your steps/calculations.*

11. *Problem 5.11.* Let $X_1, Y_1, X_2, Y_2, ...$ be independent random variables, uniformly distributed in the unit interval $[0, 1]$, and let

$$W = \frac{(X_1 + ... + X_{16}) - (Y_1 + ... + Y_{16})}{16}.$$

Find a numerical approximation to the quantity:

$$P(|W - E[W]| < 0.001).$$

Explain clearly your steps/calculations.

12. *Problem 8.2.* Nefeli, a student in probability class, takes a multiple-choice test with 10 questions and 3 choices per question. For each question, there are two equally likely possibilities, independent of other questions: either Nefeli knows the answer, in which case she answers the question correctly, or she doesn't know the answer and she guesses the answer with probability of success 1/3.

(a) Given that Nefeli answered correctly the first question, what is the probability that she knew the answer to that question?

(b) Given that Nefeli answered correctly 6 out of the 10 questions, what is the posterior PMF of the number of questions of which she knew the answer?

Explain clearly your steps/calculations.

13. *Problem 8.10.* A police radar always overestimates the speed of incoming cars by an amount that is uniformly distributed between 0 and 5 miles/hour. Assume that car speeds are uniformly distributed between 55 and 75 miles/hour. What is the Least Mean Squares estimate of a car's speed based on the radar's measurement?

14. In investment analysis, given two stocks chosen at random from the NYSE, analysts can be rated (i.e., get a bad or good rating) based on whether or not he or she can recommend correctly which of the two stocks will have better yield (i.e., greater returns) in the coming year. Let $B = P(\text{the recommendation is correct})$, and assume that B is a uniform random variable with PDF:

$$f_B(b) = 1, \; 0 \; \leq b \leq 1.$$

You are now given three different pairs of stocks (all are selected at random from the NYSE). You are asked to give your recommendation on which stock from each pair (one stock per pair) will have the better performance in the coming year. One year passes, and you find that G of your predictions were accurate, and $3 - G$ were not accurate (i.e., the stock under-performed), with of course $0 \leq G \leq 3$. Assume for the following that the outcomes of all of your recommendations are conditionally independent (i.e., given that $B = b$).

(a) If the value that B takes is given to you, you are asked to find the most probable value for G. Is this a hypothesis testing or estimation problem?

(b) Now, if the value that G takes is given to you, you are now asked to calculate the conditional mean of B. Is this a hypothesis testing or estimation problem?

(c) Calculate the Least Mean Squares (LMS) estimate for B given that $G = 2$.

(d) Assume now that the distribution of B is given by the following PDF instead:

$$f_B(b) = \begin{cases} 6b(1-b), & 0 \leq b \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

Given that $G = 2$, will the LMS estimate of B be more or less than the answer that you found in part (c)? You can answer this question by reasoning (without any calculations) - do so and explain your reasoning. Then, calculate the exact value of the LMS of B given that G=2. Discuss your answer.