UC Berkeley
Department of Electrical Engineering and Computer Science

EE 126: PROBABLITY AND RANDOM PROCESSES

Problem Set 10

Fall 2007

Issued: Monday, November 19, 2007 **Due:** Friday, November 30, 2007

Reading: Bertsekas & Tsitsiklis, Chapter 5, §6.1, 6.2

Problem 10.1

Transmitters A and B independently send messages to a single receiver in a Poisson manner with average message arrival rates of λ_A and λ_B, respectively. All messages are so brief that we may safely assume that they occupy only single points in time. The number of words in every message, regardless of its transmitting source, may be considered to be an independent experimental value of random variable W with PMF $p_W(1) = 1/3$, $p_W(2) = 1/2$, $p_W(3) = 1/6$ and $p_W(w) = 0$ otherwise.

(a) What is the probability that, during an interval of duration t, a total of exactly nine messages will be received?

(b) Determine the expected number of N be the total number of words received during an interval of duration t.

(c) Determine the PDF for X, the time from $t = 0$ until the receiver has received exactly eight three-word messages from transmitter A.

(d) Independent of what happens to all other words, a transmitter damages any particular word it sends with probability 10^{-3}. What is the probability that any particular damaged word is part of a three-word message?

(e) Probability that exactly eight of the next twelve messages received will be from A?

Solution:

(a) Let R be the total number of messages received during an interval of duration t. R is a Poisson RV with average arrival rate $\lambda_A + \lambda_B$. Therefore,

$$\begin{aligned}&\mathbb{P}(\text{exactly nine messages in time interval } t)\\ &= p_R(9)\\ &= \frac{((\lambda_A + \lambda_B)t)^9 e^{-(\lambda_A+\lambda_B)t}}{9!}\end{aligned}$$

(b) Let R be defined as in part a. Then,

$$N = W_1 + \cdots + W_R$$

We see that N is a random sum of random variables. Therefore,

$$\begin{aligned}\mathbb{E}[N] &= \mathbb{E}[W]\mathbb{E}[R]\\ &= \left(1 \cdot \frac{2}{6} + 2 \cdot \frac{3}{6} + 3 \cdot \frac{1}{6}\right)(\lambda_A + \lambda_B)t\\ &= \frac{11}{6}(\lambda_A + \lambda_B)t\end{aligned}$$

(c) Three-word messages arrive from transmitter A in a Poisson manner with average arrival rate $\lambda_A p_W(3)$. Therefore, X is $8th$ order Erlang and,

$$f_X(x) = \frac{\left(\frac{1}{6}\lambda_A\right)^8 x^7 e^{-\frac{1}{6}\lambda_A x}}{7!}$$

(d) Damage to a particular word is independent of what happens to all other words. Thus, picking a damaged word at random and determining the length of the message it came from should be the same as picking any random word (damaged or not damaged) and determining the length of the message it came from. The latter problem is more recognizable as a random incidence problem. Hence we can write the PMF of a random variable K, which describes the length of a message associated with a randomly selected word:

$$p_K(k) = \begin{cases} 2/11 & , \quad k = 1 \\ 6/11 & , \quad k = 2 \\ 3/11 & , \quad k = 3 \\ 0 & , \quad \text{otherwise} \end{cases}$$

where we obtained the PMF from the formula:

$$p_K(k) = \frac{k p_W(k)}{\mathbb{E}[W]}$$

The probability in question is 3/11.

Note that we can also obtain this result using Bayes' rule.

$$\begin{aligned} & \mathbb{P}(\text{in a message with 3 words} \mid \text{the word is damaged}) \\ &= \frac{\mathbb{P}(\text{the word is damaged} \mid \text{in a message with 3 words}) \cdot \mathbb{P}(\text{in a message with 3 words})}{\mathbb{P}(\text{the word is damaged})} \\ &= \frac{\mathbb{P}(\text{the word is damaged}) \cdot \mathbb{P}(\text{in a message with 3 words})}{\mathbb{P}(\text{the word is damaged})} \\ &= 3/11, \end{aligned}$$

where the second equality comes from the independence of damaged words and the number of words per message.

(e) Every message received either came from transmitter A or transmitter B. So, each message is a Bernoulli trial. We will say a success has occurred if a message that we receive comes from transmitter A. The probability of success in one of the Bernoulli trials is $\lambda_A/(\lambda_A + \lambda_B)$. The number of successes in a series of independent Bernoulli trials is a Binomial RV. Therefore, the probability that exactly eight of the next twelve messages received will be from transmitter A is,

$$\binom{12}{8} \left(\frac{\lambda_A}{\lambda_A + \lambda_B}\right)^8 \left(\frac{\lambda_B}{\lambda_A + \lambda_B}\right)^4$$

Problem 10.2

A bridge running East-West is so narrow that when any car is on the bridge, no cars moving in the opposite direction are allowed to use the bridge. Cars traveling from East to West arrive according to a Poisson process of rate λ, and cars traveling West to East arrive according to an independent Poisson process of rate μ. Suppose that it takes one minute to cross the bridge. Starting with an empty system, i.e. no cars on the bridge, find the distribution of:

(a) N, where N is the number of the first East-West car that experiences a conflict.

(b) T_k, where T_k is the arrival time of car k.

Solution:

(a) We are given that the Poisson arrival processes are independent of each other. Then since interarrival times for the East-West cars are also independent of each other, each East-West car has a certain probability of experiencing a conflict, independent of the other East-West cars. Thus we know that N must have a geometric distribution. The probability of conflict is the probability that there will be a West-East car on the bridge the moment the East-West car arrives, or that a West-East car will arrive while the East-West car is on the bridge. This then, is the probability that a West-East car arrives in an interval of 2 minutes–the minute before the East-West car arrives at the bridge, and the minute while the East-West car is on the bridge. This is simple to determine. It is the probability that the interarrival time is greater than 2 minutes. Denoting the interarrival time by I, we have:

$$\begin{aligned}\mathbb{P}(I > 2) &= 1 - \mathbb{P}(I \leq 2) \\ &= 1 - \int_0^2 \mu e^{-t\mu} dt \\ &= e^{-2\mu}.\end{aligned}$$

Thus we have that N is geometrically distributed with parameter $p = 1 - e^{-2\mu}$.

$$p_N(n) = \begin{cases} (e^{-2\mu})^{n-1}(1 - e^{-2\mu}) & n \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(b) We are looking for the time T_k of the k^{th} arrival a car at the bridge. This will be an Erlang of order k and parameter $\lambda + \mu$. The distribution will be:

$$f_{T_k}(t) = \frac{(\lambda + \mu)^k t^{k-1} e^{-(\lambda+\mu)t}}{(k-1)!}$$

Problem 10.3

Two players play the following game: they take take turns rolling a fair six-sided die. If a die comes up with n, the player is given n turns at the Gauss machine. Each turn at the

Gaussian machine generates a reward $X_i \sim N(m, \sigma^2)$, independently of all other turns, and the the score of a player after the round is the total of the outputs he has received from the Gauss machine. It takes each player 1 minute to roll the die and each turn at the Gauss machine takes 1 minute. There are no other delays in the game. We consider that a play starts with the rolling of the die and ends with the next roll of the die.

(a) What is the PDF for player 1's score just before he rolls the die for the $(n+1)^{th}$ time?

(b) An observer enters the room at a random time T uniformly distributed between the beginning and the end of the game, independently of the game. What is the distribution of the score accumulated during the play she observes?

(c) The players play the following variant: each round consists of rolling the die once and getting one turn at the Gauss machine, with the score per round equal to the product of the die number and the Gauss machine output. Repeat parts (a) and (b) for this new game.

(d) Let σ_k be the standard deviation of the difference between the scores of the two players after each has played k turns on the old game. The players decide to stop the game after the difference between their scores reaches σ_k. Which of the old game and the new game is more likely to stop with the fewer number of plays? Explain.

Solution:

(a) Let X be the scores of player 1 in the first play (i.e., right before rolling the die for the 2^{nd} time), and let D be the outcome of the first die roll. Then the PDF of X is given by

$$f_X(x) = \sum_{i=1}^{6} \mathbb{P}(D = i) f_{X|D}(x|i).$$

If $D = i$, then X is a sum of i independent Gaussian random variables G_l with mean m and variance σ^2, so X is Gaussian with mean im and variance $i\sigma^2$. Therefore the above equation simplifies to

$$f_X(x) = \frac{1}{6} \sum_{i=1}^{6} \frac{1}{\sqrt{2\pi i \sigma^2}} e^{-\frac{(x-im)^2}{2i\sigma^2}}.$$

If T_1 is the score of player 1 right before he rolls the die for the $(n+1)^{th}$ time, then

$$T_1 = X_1 + \ldots + X_n$$

where each of the X_i are independent and distributed identically to X.

Therefore the PDF of T_1 is the convolution of n PDFs identical to f_X. The resulting PDF can be expressed as a sum of Gaussians. To see this, note that T_1 can be expressed as

$$T_1 = G_1 + \ldots + G_N,$$

where now N is a random variable (representing the sum of the outcomes of the first n die rolls).

(b) We will assume that the game has been going on for some time. Because all die outcomes are equally likely, the probability that the observer enters the room during any given play is proportional to the length of that play. The length of a play (in turns) is equal to the outcome of the die plus 1. The probability that the observer enters during a play on which the die had outcome i is given by $\frac{i+1}{\sum_{j=1}^{6}(j+1)} = \frac{i+1}{27}$. Therefore, the distribution of the score accumulated during the play she observes, Y, is given by

$$f_Y(y) = \frac{1}{27}\sum_{i=1}^{6}(i+1)\frac{1}{\sqrt{2\pi i\sigma^2}}e^{-\frac{(x-im)^2}{2i\sigma^2}}.$$

(c) Let Z be the score of player 1 in the first play (i.e., right before rolling the die for the 2^{nd} time), and let D be the outcome of the first die roll. Then the PDF of Z is given by

$$f_Z(z) = \sum_{i=1}^{6}\mathbb{P}(D = i)f_{Z|D}(z|i).$$

If $D = i$, then Z is equal to i times a Gaussian random variable with mean m and variance σ^2, so Z is Gaussian with mean im and variance $i^2\sigma^2$. Therefore the above equation simplifies to

$$f_Z(z) = \frac{1}{6}\sum_{i=1}^{6}\frac{1}{\sqrt{2\pi i^2\sigma^2}}e^{-\frac{(z-im)^2}{2i^2\sigma^2}}.$$

If S_1 is the score of player 1 right before he rolls the die for the $(n+1)^{th}$ time, then

$$S_1 = Z_1 + \ldots + Z_n$$

where each of the Z_i are independent and distributed identically to Z.

Therefore the PDF of S_1 is the convolution of n PDFs identical to f_Z. The resulting PDF can be expressed as a sum of Gaussians, because S_1 can be written as

$$S_1 = \alpha G_1 + \ldots + \zeta G_n,$$

where $\alpha, \ldots, \zeta$ are IID, uniformly distributed between 1 and 6.

In this case, the oberver is equally likely to enter during a play corresponding to any die outcome, because they all have the same length. Therefore the distribution of the score accumulated during the play she observes, W, is given by $f_W(w) = f_Z(w)$.

(d) We assume here that k is fixed. For the first game σ_k^2 is the variance of the random variable $T_1 - T_2$, which is equal to $2var(T_1)$. In the second game, $\hat{\sigma}_k^2$ is the variance of $S_1 - S_2$, and hence equal to $2var(S_1)$.

Notice that X and Z (the outcomes of a single play of the games described in parts a and c, respectively) are both given by sums of Gaussians which are identical except that the Gaussians in the PDF of Z have larger variance. Therefore, $\hat{\sigma}_k^2$ is bigger, and the game described in part c is more likely to stop in a fewer number of plays (because, for k turns, the probability of the score difference beging bigger than σ_k is higher in the second game, than in the first.)

Problem 10.4

The EECS department offers n classes per year, which the students rank each class from 1 to n, in order of difficulty. Unfortunately, the ranking is completely arbitrary, so that 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 conveniently chooses to remember only the highest ranking his class has ever gotten. Show that the system described by the ranking that the professor remembers is a Markov chain, and specify its transition matrix.

Solution:

If we can show that this process has the Markov property, then it is a Markov chain since it clearly satisfies the other conditions. Thus we need only show:

$$\mathbb{P}(X_{m+1} = j | X_m = i, X_{m-1} = k, \ldots, X_1 = q) = \mathbb{P}(X_{m+1} = j | X_m = i).$$

But this is clear from the statement of the problem, since the only relevant information is the highest ranking ever attained, and this information is precisely the information contained by X_m.

For $i > j$, $p_{ij} = 0$. Since the professor will continue to remember the highest ranking, even if he gets a lower ranking in a subsequent year, we have: $p_{ii} = \frac{i}{n}$. Finally, for $j > i$, $p_{ij} = \frac{1}{n}$ since the class is equally likely to receive any given rating.

Since the professor only remembers the *highest* ranking received by his class, only the highest possible rank is recurrent (i.e. the state { Professor recalls rank n } is recurrent). All of the lower-rank states are transient, since the class will eventually receive a higher rank, and then the professor will never recall that lower ranking again.

Problem 10.5

For each of the following definitions of state X_k at time k ($k = 1, 2, \ldots$), determine whether the Markov property is satisfied and, when it is, specify the transition probabilities p_{ij}:

(a) A six-sided die is rolled repeatedly.

  (i) Let X_k denote the largest number rolled in the first k rolls.

  (ii) Let X_k denote the number of sixes in the first k rolls.

  (iii) At time k, let X_k be the number of rolls since the most recent six.

(b) Let Y_k be the state of some discrete-time Markov process at time k (i.e., it is known Y_k satisfies the Markov property) with known transition probabilities q_{ij}.

  (i) For a fixed integer $r > 0$, let $X_k = Y_{r+k}$.

  (ii) Let $X_k = Y_{2k}$.

(iii) Let $X_k = (Y_k, Y_{k+1})$; that is, the state X_k is defined by the sequence of state *pairs* in a given Markov process.

Solution:

(a) (i) Since the state X_k is the largest number rolled in k rolls, the set of states S $=\{1,2,3,4,5,6\}$. The probability of the largest number rolled in the first (k+1) trials is only dependent to the what the largest number that was rolled in the first k trials. This satisfies the Markov property. The transition probabilities are given by

$$p_{ij} = \begin{cases} 0 & , \quad j < i \\ \frac{i}{6} & , \quad j = i \\ \frac{1}{6} & , \quad j > i \end{cases}$$

(ii) Since the state X_k is the number of sixes in the first k rolls, the set of states S=$\{0,1,2,.......\}$. The probability of getting a six in a given trial is 1/6. The number of sixes rolled in the first (k+1) trials is only dependent to the number of sixes rolled in the first k trials. This satisfies the Markov property. The transition probabilities are given by

$$p_{ij} = \begin{cases} \frac{1}{6} & , \quad j = i+1 \\ \frac{5}{6} & , \quad j = i \\ 0 & , \quad otherwise \end{cases}$$

(iii) Since the state X_k is the number of rolls since the most recent six, the set of states S=$\{0,1,2,.......\}$. If the roll of the die is 6 on the next trial the chain goes to state 0. If not, the state goes to the next higher state. Therefore, the probability of the next state depends on the past only through the present state. Clearly, this satisfies the Markov property. The transition probabilities are given by

$$p_{ij} = \begin{cases} \frac{1}{6} & , \quad j = 0 \\ \frac{5}{6} & , \quad j = i+1 \\ 0 & , \quad otherwise \end{cases}$$

(b) (a) For $X_k = Y_{r+k}$,

$$\mathbb{P}(X_{k+1} = j | X_k = i, ...X_0 = i_0) = \mathbb{P}(Y_{r+k+1} = j | Y_{r+k} = i, ...Y_r = i_r)$$

by the Markov property of Y

$$= \mathbb{P}(Y_{r+k+1} = j | Y_{r+k} = i)$$

$$\mathbb{P}(X_{k+1} = j | X_k = i, ...X_0 = i_0) = \mathbb{P}(X_{k+1} = j | X_k = i)$$

This satisfies the Markov Property for X. Also we can see that, X_k is a delayed process by r of Y_k. Therefore, they should have the same transition probability p_{ij}, So, we have

$$p_{ij} = q_{ij}$$

(b) For $X_k = Y_{2k}$,

$$\mathbb{P}(X_{k+1} = j | X_k = i, X_{k-1} = i_{k-1}.....X_0 = i_0) = \mathbb{P}(Y_{2k+2} = j | Y_{2k} = i, Y_{2k-2} = i_{2k-2}.....Y_0 = i_0)$$

$$by \quad the \quad Markov \quad property \quad of \quad Y$$

$$= \mathbb{P}(Y_{2k+2} = j | Y_{2k} = i)$$

$$= \mathbb{P}(X_{k+1} = j | X_k = i)$$

This satisfies the Markov Property for X. The transition probabilities p_{ij} is given by

$$\begin{aligned} p_{ij} &= \mathbb{P}(X_{k+1} = j | X_k = i) \\ &= \mathbb{P}(Y_{2k+2} = j | Y_{2k} = i) \\ &= r^y_{ij}(2) \end{aligned}$$

where $r^y_{ij}(n)$ is the n step transition probability of Y

(c)

$$\begin{aligned} &\mathbb{P}(X_{k+1} = (n,l) | X_0 = (i_0, i_1), X_1 = (i_1, i_2), ..., X_k = (i_k, n)) \\ &= \mathbb{P}(X_{k+1} = (n,l) | Y_0 = i_0, Y_1 = i_1, Y_2 = i_2........., Y_k = i_k, Y_{k+1} = n) \\ &by \quad the \quad Markov \quad property \quad of \quad Y \\ &= \mathbb{P}(X_{k+1} = (n,l) | Y_{k+1} = n) \\ &= \mathbb{P}(X_{k+1} = (n,l) | X_k = (i_k, n)) \end{aligned}$$

The transition probabilities $p_{i,j}$ is given by:
Let $i = (i_k, i_{k+1})$ and $j = (n,l)$,

$$p_{ij} = \mathbb{P}(X_{k+1} = (n,l) | X_k = (i_k, i_{k+1})) = \begin{cases} q_{nl} & , \quad i_{k+1} = n \\ 0 & , \quad i_{k+1} \neq n \end{cases}$$