Stat 6329 Syllabus

Stat 6329.501
Applied Probability and Stochastic Processess
Course Information

Course number/section	Stat 6329-501
Instructor	Dr. Larry P. Ammann
Email	ammann@utdallas.edu
Office	ECS 3.904
Office hours	M 2-3 pm
Phone	(972) 883-2164
Required text	Introduction to Probability Models, 9^{th} ed.
Author	Sheldon M. Ross
Suggested course materials	None

Tentative Schedule

Date	Topics	Chapters
8/24	Introduction; review of random variables	1.1-1.6; 2.1-8
9/2	Conditional probability and expectation	3.1-7
9/7	Labor Day, no classes	
9/9	Continuation of conditional probability and expectation	3.1-7
9/28	**EXAM 1**	
9/30	Markov chains	4.1-6
10/26	**EXAM 2**	
10/28	Poisson and related processes	5.1-4
11/11	Introduction to renewal processes	7.1-3
	Introduction to queuing models	8.1-3
11/26-27	Thanksgiving holiday	
12/7	**Exam 3**	

Grading Policy

Final course grade will be derived as follows:
Homework: 25%
Exam 1: 25%
Exam 2: 25%
Exam 3: 25%

Student Learning Objectives

1. Understand the fundamentals of probability theory and random variables
2. Understand the basic probability tools for stochastic processes: conditioinal probability and conditional expectation
3. Understand the basic properties and application of special classes of stochastic processes including Markov chains and Poisson and related processes

Note: the complete syllabus is available here:
`http://www.utdallas.edu/~ammann/stat6329_syllabus.pdf`

Homework Assignment 1

1. Text, p. 86: 5
2. Text, p. 87: 11
3. Text, p. 87: 13
4. Text, p. 88: 25
5. Text, p. 88: 26
6. Text, p. 90: 37
7. Text, p. 96: 77
8. Let $X_1, \cdots, X_n$ be independent r.v.'s each with the same exponential distribution with rate λ. Let $Y = \min(X_1, \cdots, X_n)$. Show that Y has an exponential distribution with rate $n\lambda$.

Solutions for Homework Assignment 1

1. (P. 86, 5) Let X_1, X_2 denote the outcomes of the two rolls.

   (a) Note that $P(max(X_1, X_2) = 1) = P(X_1 = 1, X_2 = 1) = 1/36$. For $k = 2, \cdots, 6$,

$$\begin{aligned} P(max(X_1, X_2) = k) &= P(X_1 = k, X_2 < k) + P(X_1 = k, X_2 = k) + P(X_1 < k, X_2 = k) \\ &= (1/6)((k-1)/6) + 1/36 + (1/6)((k-1)/6) \\ &= (2k-1)/36. \end{aligned}$$

   (b) Note that $P(min(X_1, X_2) = 6) = P(X_1 = 6, X_2 = 6) = 1/36$. For $k = 2, \cdots, 6$,

$$\begin{aligned} P(min(X_1, X_2) = k) &= P(X_1 = k, X_2 > k) + P(X_1 = k, X_2 = k) + P(X_1 > k, X_2 = k) \\ &= (1/6)((6-k)/6) + 1/36 + (1/6)((6-k)/6) \\ &= (13-2k)/36. \end{aligned}$$

   (c) Let $S = X_1 + X_2$. Then $P(S = 2) = P(S = 12) = 1/36$, $P(S = 3) = P(S = 11) = 2/36$, $P(S = 4) = P(S = 10) = 3/36$, $P(S = 5) = P(S = 9) = 4/36$, $P(S = 6) = P(S = 8) = 5/36$, $P(S = 7) = 6/36$.

   (d) Let $D = X_1 - X_2$. Then

$$\begin{aligned} P(D = 0) &= \sum_{k=1}^{6} P(X_1 = k, X_2 = k) = 6/36 \\ P(D = -1) &= \sum_{k=1}^{5} P(X_1 = k, X_2 = k+1) = 5/36 = P(D = 1) \\ P(D = -2) &= \sum_{k=1}^{4} P(X_1 = k, X_2 = k+2) = 4/36 = P(D = 2) \\ P(D = -3) &= \sum_{k=1}^{3} P(X_1 = k, X_2 = k+3) = 3/36 = P(D = 3) \\ P(D = -4) &= \sum_{k=1}^{2} P(X_1 = k, X_2 = k+4) = 2/36 = P(D = 4) \\ P(D = -5) &= P(X_1 = 1, X_2 = 6) = 1/36 = P(D = 5) \end{aligned}$$

2. (P. 87, 11) The draws are independent since the balls are replaced after each draw. So number white balls in 4 draws has Binomial(4,.5) distribution.

$$P(X = 2) = \binom{4}{2}(.5)^4 = 3/8.$$

3. (P. 87, 13) Under the assumption of no ESP, probability of correct prediction of toss is 0.5, so expected number of correct predictions would be 5. Getting 7 correct is higher

than expected and so is some evidence that there may be some ESP. Getting more than 7 correct would be even stronger evidence, so the relevant event to judge the strength of evidence for ESP is the event: getting at least 7 correct. Under the assumption of no ESP, the number correct has Binomial(10,.5) distribution.

$$\begin{aligned} P(X \geq 7) &= \sum_{k=7}^{10} \binom{10}{k} (.5)^{10} \\ &= \frac{179}{1024} = 0.172. \end{aligned}$$

4. (P. 88, 25) Let N denote the number of games played. Then the event that 7 games are played is the event that each team wins 3 games after 6 games are played. This probability is the Binomial probability of 3 successes out of 6 trials, and so is given by

$$\binom{6}{3} p^3(1-p)^3 = 20[p(1-p)]^3.$$

The function $h(p) = p(1-p)$, $0 < p < 1$, is maximized at the solution to $1 - 2p = 0$ (solution to derivative of h equals 0), which has solution $p = 1/2$. This is the maximum since 2^{nd} derivative of h is -2.

5. (P. 88, 26) For $i = 2$,

$$E(N) = 2P(N = 2) + 3P(N = 3) = 2(p^2 + (1-p)^2) + 3(2p(1-p)) = 2 + 2p(1-p).$$

For $i = 3$,

$$\begin{aligned} E(N) &= 3P(N = 3) + 4P(N = 4) + 5P(N = 5) \\ &= 3(p^3 + (1-p)^3) + 4(3p^3(1-p) + 3p(1-p)^3) + 5(6p^2(1-p)^2) \\ &= 3 + 3p(1-p) + 6p^2(1-p)^2. \end{aligned}$$

Since both are non-decreasing functions of $p(1-p)$, then they are maximized at $p = 1/2$.

6. (P. 89, 32) Use the Poisson approximation since n is large and p is small. This approximation is Poisson(np=0.5).

   (a) $P(X \geq 1) = 1 - P(X = 0) \approx 1 - e^{-0.5} = 0.393.$

   (b) $P(X = 1) \approx (0.5)e^{-0.5} = 0.303.$

   (c) $P(X \geq 2) = 1 - P(X \leq 1) \approx 1 - [P(X = 0) + P(X = 1)] = 0.090.$

7. (P. 90, 37) For $0 \le x \le 1$,

$$\begin{aligned} P(\max(X_1, \cdots, X_n) \le x) &= P(X_1 \le x, \cdots, X_n \le x) \\ &= \prod_{k=1}^{n} P(X_k \le x) \\ &= x^n. \end{aligned}$$

So, $F_M(x) = x^n$, $0 \le x \le 1$. Density function is the derivative,

$$f_M(x) = nx^{n-1}, \quad 0 \le x \le 1.$$

8. (P96, 77) Use moment generating functions. Let $U = X + Y$, $V = X - Y$. Then joint MGF of U, V is

$$\begin{aligned} M(\theta_1, \theta_2) &= E\exp(\theta_1 U + \theta_2 V) \\ &= E\exp(\theta_1(X+Y) + \theta_2(X-Y)) \\ &= E\exp((\theta_1+\theta_2)X + (\theta_1-\theta_2)Y) \\ &= E\exp((\theta_1+\theta_2)X)E((\theta_1-\theta_2)Y) \\ &= \exp((\theta_1+\theta_2)\mu + (\theta_1+\theta_2)^2(\sigma^2/2) + \exp((\theta_1-\theta_2)\mu + (\theta_1-\theta_2)^2(\sigma^2/2) \\ &= \exp(\theta_1(2\mu) + \theta_1^2(2\sigma^2/2) + \theta_2^2(2\sigma^2/2)) \\ &= M_1(\theta_1)M_2(\theta_2), \end{aligned}$$

where M_1 is MGF of $N(2\mu, 2\sigma^2)$ and M_2 is MGF of $N(0, 2\sigma^2)$. Hence, U, V are independent r.v.'s.

9. (Extra 1)

$$\begin{aligned} P(Y > y) &= P(\min(X_1, \cdots, X_n) > y) \\ &= P(X_1 > y, \cdots, X_n > y) \\ &= \prod_{k=1}^{n} P(X_k > y) \\ &= e^{-n\lambda}. \end{aligned}$$

So, $F_Y(y) = 1 - e^{-n\lambda}$, which is d.f. of exponential distribution with rate $n\lambda$. Density is the dervative,

$$f_Y(y) = ny^{n-1}, \; y \ge 0.$$

Homework Assignment 2

1. Text, p. 165: 3, 4
2. Text, p. 166: 11, 15
3. Text, p. 167: 17
4. Text, p. 169: 24. Use first step analysis.

Solutions for Homework Assignment 2

1. (p.165, 3).

$$\begin{aligned} E(X|Y=1) &= \frac{1\cdot\frac{1}{9}+2\cdot\frac{1}{3}+3\cdot\frac{1}{9}}{\frac{1}{9}+\frac{1}{3}+\frac{1}{9}} \\ &= \frac{\frac{10}{9}}{\frac{5}{9}} \\ &= 2. \end{aligned}$$

$$\begin{aligned} E(X|Y=2) &= \frac{1\cdot\frac{1}{9}+2\cdot 0+3\cdot\frac{1}{18}}{\frac{1}{9}+0+\frac{1}{18}} \\ &= \frac{\frac{5}{18}}{\frac{3}{18}} \\ &= \frac{5}{3}. \end{aligned}$$

$$\begin{aligned} E(X|Y=3) &= \frac{1\cdot 0+2\cdot\frac{1}{6}+3\cdot\frac{1}{9}}{0+\frac{1}{6}+\frac{1}{9}} \\ &= \frac{\frac{12}{18}}{\frac{5}{18}} \\ &= 2.4. \end{aligned}$$

2. (p.165, 4). Not independent because, for example, $P(X=1,Y=3)=0$, but

$$P(X=1)=2/9,\ P(Y=3)=5/18,$$

and so

$$P(X=1)P(Y=3)=5/81\neq P(X=1,Y=3).$$

3. (p.166, 11).

$$\begin{aligned} f_Y(y) &= \int_{-y}^{y} f(x,y)dx \\ &= \frac{1}{8}e^{-y}\int_{-y}^{y}(y^2-x^2)dx \\ &= \frac{1}{8}e^{-y}[y^2x-\frac{1}{3}x^3]|_{-y}^{y} \\ &= \frac{1}{8}e^{-y}\frac{4}{3}y^3 \\ &= \frac{1}{6}y^3e^{-y}. \end{aligned}$$

This is $Gamma(4,1)$ density. So,

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{\frac{1}{8}e^{-y}(y^2 - x^2)}{\frac{1}{6}y^3e^{-y}} \\ &= \frac{3}{4}y^{-3}(y^2 - x^2), \end{aligned}$$

for $|x| \leq y,\ 0 < y < \infty$. Since this conditional density is symmetric about $x = 0$ for each $0 < y < \infty$, then $E(X|Y = y) = 0$.

4. (p.166, 15).

$$f_Y(y) = \int_0^y f(x, y)dx = e^{-y},$$

$0 < y < \infty$. This is $exponential(1)$ density, so

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} = y^{-1},$$

$0 < x < y$, $0 < y < \infty$. This is $Uniform(0, y)$, and so

$$E(X^2|Y = y) = y^{-1}\int_0^y x^2 dx = \frac{1}{3}y^2.$$

5. (p.167, 17). The joint density of X, Y is given by

$$f_{X,Y}(i, y) = \frac{y^i}{i!}e^{-y}Cy^{s-1}e^{-\alpha y} = \frac{Cy^i}{i!}y^{i+s-1}e^{-(1+\alpha)y}.$$

Hence,

$$\begin{aligned} P(X = i) &= \int f_{X,Y}(i, y)dy \\ &= \frac{Cy^i}{i!}\frac{\Gamma(i + s)}{(1 + \alpha)^{s+i}}, \end{aligned}$$

and so,

$$\begin{aligned} f_{Y|X}(y|i) &= \frac{f_{X,Y}(i, y)}{P(X = i)} \\ &= Dy^{i+s-1}e^{-(1+\alpha)y}, \end{aligned}$$

where D is a constant that does not depend on y and is the value that makes this conditional density integrate to 1. Therefore, this conditional density is $Gamma(i + s, 1 + \alpha)$.

6. (p.169, 24). Let N_{ij} denote the number of trials required to obtain at least i heads and at least j tails in a sequence of i.i.d. Bernoulli trials. Let $X_1, X_2, \cdots$ denote the Bernoulli trials. Note that N_{10} has a geometric distribution with success probability p and N_{01} has a geometric distribution with success probability $1-p$.

(a)

$$\begin{aligned}
E(N_{11}) &= E(N_{11}|X_1 = H)P(X_1 = H) + E(N_{11}|X_1 = T)P(X_1 = T) \\
&= E(1 + N_{01})p + E(1 + N_{10})(1-p) \\
&= p(1 + 1/(1-p)) + (1-p)(1 + 1/p) \\
&= \frac{1}{p} + \frac{p}{1-p} \\
&= \frac{1 - p + p^2}{p(1-p)}.
\end{aligned}$$

(b) Let M_{11} denote the number of heads during N_{11} trials. Then

$$\begin{aligned}
E(M_{11}) &= E(M_{11}|X_1 = H)P(X_1 = H) + E(M_{11}|X_1 = T)P(X_1 = T) \\
&= E(1 + N_{01} - 1)p + 1(1-p) \\
&= (1-p) + \frac{p}{1-p} \\
&= \frac{1 - p + p^2}{1-p}
\end{aligned}$$

(c) Let L_{11} denote the number of tails during N_{11} trials. Then $L_{11} = N_{11} - M_{11}$, and so,

$$\begin{aligned}
E(L_{11}) &= E(N_{11}) - E(M_{11}) \\
&= \frac{1 - p + p^2}{p}.
\end{aligned}$$

(d) First note that $E(N_{20}) = 2/p$.

$$\begin{aligned}
E(N_{21}) &= E(N_{21}|X_1 = H)P(X_1 = H) + E(N_{21}|X_1 = T)P(X_1 = T) \\
&= E(1 + N_{11})p + E(1 + N_{20})(1-p) \\
&= p + \frac{1 - p + p^2}{1-p} + (1-p) + \frac{2(1-p)}{p} \\
&= \frac{p^2}{1-p} + \frac{2}{p}.
\end{aligned}$$

Homework Assignment 3

1. Text, p. 264: 6.

2. Text, p. 265: 14.

3. Text, p. 266: 16.

4. Text, p. 266: 20.

5. Text, p. 266: 21.

6. Text, p. 266: 22. **Hint, continued**. Use *modulo 13* arithmetic for the states of Y_n.

7. A computer program consists of a sequence of addresses that must be fetched from one of three locations, local memory (RAM), cache memory, or virtual memory (swap). A simple memory model can be expressed as follows: if the current address is in RAM, then the next address will be in RAM, cache, or swap with probabilities $0.8, 0.15, 0.05$, respectively; if the current address is in cache, then the next address will be in RAM, cache, or swap with probabilities $0.05, 0.9, 0.05$, respectively; if the current address is in swap, then the next address will be in RAM, cache, or swap with probabilities $0.2, 0.3, 0.5$, respectively. Let X_n denote the location of the n^{th} address and assume that $X_n,\ n \geq 1$ is a Markov chain. Find the proportion of time spent in each memory location over a long period of time.

8. Consider a component that begins operation with 0 damage. Suppose that at the end of a period of operation, it has accumulated damage N_1, where

$$P(N_1 = k) = pq^k,\ k \geq 0,$$

$0 < p < 1$, and $q = 1 - p$. If $N_1 \geq m$, where $m > 0$ is some fixed integer, then the component is replaced with an identical spare so that the damage at the beginning of the next period of operation for the component in use would be 0. Otherwise, the component begins the next period of operation with damage N_1. Damage is cumulative. That is, damage that occurs during a period of operation is added to the damage the component had at the beginning of the period, with the understanding that if the cumulative damage is m or higher at the end of a period, the component is replaced. Let N_k denote the damage that occurs during period k and assume that $\{N_k,\ k \geq 1\}$ are i.i.d. r.v.'s having the same geometric distribution given for N_1. Let $X_n,\ n \geq 1$ denote the damage of the component in use at the beginning of period n, and note that

$$\begin{aligned} X_{k+1} &= X_k + N_k,\ \text{if } X_k + N_k \leq m - 1, \\ &\quad 0,\ \text{otherwise.} \end{aligned}$$

(a) What is the transition probability matrix of this Markov chain?

(b) Find its stationary distribution.

(c) What is the mean time between visits to state 0?

Note that the states are $0, 1, \cdots, m-1$.

9. A prisoner wakes up in a structure that contains 4 rooms.
Room 1 has 3 doors. 1 door leads to Room 2, 1 door leads to Room 3, 1 door leads to a maximum security prison from which he cannot leave.
Room 2 has 4 doors. 1 door leads to Room 1, 1 door leads to Room 3, 1 door leads to Room 4, and 1 door leads to the maximum security prison.
Room 3 has 3 doors. 1 door leads to Room 1, 1 door leads to Room 2, 1 door leads back to Room 3.
Room 4 has 5 doors. 1 door leads to Room 1, 1 door leads to Room 2, 1 door leads to Room 3, 1 door leads to the maximum security prison, and 1 door leads to freedom.
Suppose he is unable to determine which door he used previously and that he is equally likely to use each of the doors in a room.

   (a) Briefly explain (mathematical proof not required) why the prisoner's location can be represented by a Markov chain.

   (b) Obtain the transition probability matrix of this Markov chain and classify its states (recurrent or transient).

   (c) Find the mean number of visits to each room given that he wakes up in Room i, $1 \leq i \leq 4$.

   (d) Find the mean time he spends in this structure given that he wakes up in Room i, $1 \leq i \leq 4$.

   (e) Find the probability he gets free given that he wakes up in Room i, $1 \leq i \leq 4$.

Bernoulli and Poisson Processes

Bernoulli Processes

Let $\{X_n\}$, $n \geq 1$, denote a sequence of independent Bernoulli random variables with the same success probability p. This sequence can be used to represent a two-state system in which a state occurs randomly and independently in discrete time. For example, the occurrence or non-occurrence of a particular type of defect in a wafer could be modelled in this way. Departures from this model might indicate the presence of a systematic error in the production process.

There are two other ways to represent the Bernoulli Process. Define $S_0 = 0$ and for $n \geq 1$, let

$$S_n = \sum_{k=1}^{n} X_k.$$

The collection, $\{S_n,\ n \geq 0\}$, defines a discrete time, discrete state space stochastic process that is the counting process version of the Bernoulli process since S_n represents the number of successes up to time n. Note that these random variables are not independent. The other way to represent this process is to record the numbers of failures between successes. Let D_1 denote the number of failures before the first success and for $n > 1$, let D_n denote the number of failures after the $(n-1)^{th}$ success but before the n^{th} success. This sequence of random variables is referred to as the inter-arrival time representation of the Bernoulli process. Note that if one of these three sequences of random variables is known, then the other two can be derived from the known sequence. Note that if T_N is the trial on which the n^{th} success occurs, then,

$$T_n = n + \sum_{k=1}^{n} D_k.$$

Properties of S_n and D_n

1. The distribution of S_n is $Binomial(n, p)$. This follows directly from the definition of S_n.

2. If $m > n \geq 0$ then $S_m - S_n$ and S_n are independent random variables. Note that

$$S_n = \sum_{k=1}^{n} X_k, \quad S_m - S_n = \sum_{k=n+1}^{m} X_k.$$

   Since S_n and $S_m - S_n$ are functions of non-overlapping sets of X_k's, then S_n and $S_m - S_n$ are independent. This argument can be extended to prove that if $r \geq 1$ and $0 \leq n_0 < n_1 < \cdots < n_r < \infty$, then $S_{n_0}, S_{n_1} - S_{n_0}, \cdots, S_{n_r} - S_{n_{r-1}}$ are independent random variables. If this property is satisfied for every $r \geq 1$, then we say that the stocastic process has **independent increments**.

3. Let $m > n$. Then the distribution of $S_m - S_n$ is $Binomial(m-n, p)$. This follows from the fact that $S_m - S_n$ is the sum of $m-n$ independent Bernoulli random variables all with the same success probability p. In particular note that this distribution does not depend directly on the time points m or n, but instead only depends on the length of the time interval, $m-n$. Any stochastic process that possesses this property for all $m > n$ is said to have **stationary increments**.

4. Let $m > n$. Then $E(S_m|S_n) = S_n + (m-n)p$.
   Proof:

$$\begin{aligned} E(S_m|S_n = k) &= E(S_m - S_n + S_n|S_n = k) \\ &= E(S_m - S_n|S_n = k) + E(S_n|S_n = k) \\ &= E(S_m - S_n) + k = (m-n)p + k. \end{aligned}$$

   This follows since $S_m - S_n$ and S_n are independent. Now substitute S_n for k to obtain the desired result.

5. Let $m > n$. Then $S_M - S_n$ and S_n are independent since they involve non-overlapping time intervals, and so,

$$\begin{aligned} E(S_m S_n) &= E((S_m - S_n + S_n)S_n) \\ &= E((S_m - S_n)S_n) + E(S_n^2) \\ &= E(S_m - S_n)E(S_n) + Var(S_n) + (E(S_n))^2 \\ &= (m-n)pnp + np(1-p) + n^2p^2 \\ &= n(m-n)p^2 + np(1-p) + n^2p^2 \\ &= nmp^2 + np(1-p). \end{aligned}$$

   Note that this implies that

$$Cov(S_m, S_n) = E(S_m S_n) - E(S_m)E(S_n) = np(1-p).$$

6. The distribution of D_1 is $Geometric(p)$. This follows directly from the definition of D_1. Note that the version of the geometric distribution for this property is the one that does not count the terminating success.

7. $P(D_1 = k, D_2 = r) = pq^k pq^r$, $k \geq 0$, $r \geq 0$.
   Proof:

$$\begin{aligned} P(D_1 = k, D_2 = r) &= P(X_1 = 0, \cdots, X_k = 0, X_{k+1} = 1, X_{k+2} = 0, \cdots, \\ &\qquad X_{k+r+1} = 0, X_{k+r+2} = 1) \\ &= q^k pq^r p. \end{aligned}$$

This property implies that D_1, D_2 are independent geometric random variables. This argument can be extended to show that $\{D_n,\ n \geq 0\}$ is a sequence of independent geometric random variables with the same success probability p.

8.

$$\begin{aligned}
P(D_1 = r | S_n = 1) &= \frac{1}{n},\ 0 \leq r < n, \\
P(T_1 = k | S_n = 1) &= P(D_1 = k-1 | S_n = 1) = \frac{1}{n},\ 1 \leq k \leq n,
\end{aligned}$$

In English this property says that given there was 1 success up to time n, the time of that success is uniformly distributed over the time points, $1, \cdots, n$.

Proof:

$$\begin{aligned}
P(D_1 = r | S_n = 1) &= \frac{P(X_1 = 0, \cdots, X_r = 0, X_{r+1} = 1, X_{r+2} = 0, \cdots, X_n = 0)}{P(S_n = 1)} \\
&= \frac{q^r p q^{n-r-1}}{n p^1 q^{n-1}} \\
&= \frac{1}{n}.
\end{aligned}$$

This also shows that $E(D_1 | S_n = 1) = (n+1)/2$.

9. **Renewal Property**. Suppose that a success has occurred at time r. We can define a new stochastic process by effectively resetting the number of successes to 0 and resetting the clock to 0. The resulting stochastic process would be $S_n^* = S_{n+r} - S_r$. Since

$$S_{n+r} - S_r = \sum_{k=1}^{n} X_{r+k},$$

then this new process is defined in the same way as the original process, and so $\{S_n^*\}$ also is a Bernoulli process with the same properties as the original process. Therefore, if we are interested in observing an ongoing Bernoulli process and we synchronize the start of our observations at the time of a success, then the process we observe will be equivalent to the original Bernoulli process.

10. **Renewal Property, continued**. Now suppose that we begin our observation of the Bernoulli process at some arbitrary time r. We know from the previous result that starting from the time of the next success, we will be observing a Bernoulli process with the same properties as the original. Therefore, the only possible difference between the properties of the process we observe and the properties of the original process would

be the time until the first success in the process we observe. Let D_1^* denote the number of failures until the first success after time r. Then

$$\begin{aligned} P(D_1^* = k) &= P(X_{r+1} = 0, \cdots, X_{r+k} = 0, X_{r+k+1} = 1) \\ &= pq^k. \end{aligned}$$

Hence, the distribution of D_1^* is the same geometric distribution as all the other inter-arrival times, and so the process we observe beginning at an arbitrary time r is a Bernoulli process with the same properties as the original process. Note that this also shows that if we begin observing an ongoing Bernoulli process at an arbitrary time, the inter-arrival time that contains the start of our observations is special. In particular, the time between the success prior to the start of our observations and the time of the next success has a different distribution than the other inter-arrival times since there may have been some failures between the previous success and the start of our observations.

11. **Markov Property**. Think of time $n+1$ as the future and time n as the present, and consider the conditional distribution of S_{n+1} given the present and the past,

$$P(S_{n+1} = k | S_1 = k_1, \cdots, S_n = k_n).$$

Since $S_{n+1} = X_{n+1} + S_n$ and since X_{n+1} and $\{S_1, \cdots, S_n\}$ are independent, then

$$\begin{aligned} P(S_{n+1} = k | S_1 = k_1, \cdots, S_n = k_n) &= P(X_{n+1} + S_n = k | S_1 = k_1, \cdots, S_n = k_n) \\ &= P(X_{n+1} + k_n = k | S_1 = k_1, \cdots, S_n = k_n) \\ &= P(X_{n+1} = k - k_n). \end{aligned}$$

Therefore,

$$P(S_{n+1} = k | S_1, \cdots, S_n) = P(S_{n+1} = k | S_n).$$

This is the Markov property for stochastic processes. In English it means that the conditional distribution of S_{n+1} given the present and the past depends only on the present. Discrete-time stochastic processes that satisfy this property are said to be **Markov processes**.

Poisson Processes

The next example of a stochastic process we will examine is the Poisson process. This stocastic process has many important applications in addition to providing the basis for extensions that make it even more widely applicable. The Poisson process is a continuous time, discrete state space process, $\{N(t),\ t \geq 0\}$, that represents the number of arrivals of some entity up to time t. It is defined by

1. $\{N(t),\ t \geq 0\}$ is a counting process; that is, $N(t)$ is integer-valued, $N(0) = 0$, and if $t > s$, then $N(t) \geq N(s)$ with probability 1.

2. $\{N(t),\ t \geq 0\}$ has stationary, independent increments. Specifically, for every $r \geq 1$ and for every collection $0 \leq t_0 < t_1 < \cdots < t_r$, $N(t_1) - N(t_0), \cdots, N(t_r) - N(t_{r-1})$ is a collection of independent random variables. For every $0 \leq s < t < \infty$, the distribution of $N(t) - N(s)$ depends only on $t - s$ and therefore is the same as the distribution of $N(t - s)$.

3. The likelihood that there is exactly one arrival during a small interval of time is proportional to the length of the time interval. That is, there exists $0 < \lambda < \infty$ such that

$$\lim_{h \searrow 0} \frac{P(N(t+h) - N(t) = 1)}{h} = \lambda.$$

The parameter λ is called the intensity of the Poisson process. The likelihood that there is more than 1 arrival during a small interval of time is vanishingly small. Specifically,

$$\lim_{h \searrow 0} \frac{P(N(t+h) - N(t) > 1)}{h} = 0.$$

It can be shown that condition (3) is equivalent to the condition that the distribution of $N(t)$ is Poisson with mean λt. This condition also shows that the Poisson process is not an appropriate model for congested arrivals such as traffic on LBJ Expressway.

$N(t)$ is the counting process representation of the Poisson process. There is also an arrival time representation. Let T_n denote the time of the n^{th} arrival and let $D_1 = T_1$, $D_n = T_n - T_{n-1},\ n > 1$ denote the inter-arrival times. The correspondence between the counting process representation and the arrival times is given by,

$$\{N(t) \geq n\} = \{T_n \leq t\}.$$

In particular,

$$P(D_1 \leq t) = P(T_1 \leq t) = P(N(t) \geq 1) = 1 - e^{-\lambda t}.$$

This shows that the time of the first arrival has an exponential distribution with mean $\mu = 1/\lambda$.

Properties of $N(t)$ and T_n.

1. Let $0 < s < t$. Then

$$\begin{aligned} P(T_1 > s, T_2 > t) &= P(N(s) < 1, N(t) < 2) \\ &= P(N(s) = 0, N(t) - N(s) \leq 1) \\ &= P(N(s) = 0) P(N(t) - N(s) \leq 1) \\ &= e^{-\lambda t}[1 + \lambda(t - s)]. \end{aligned}$$

Differentiating with respect to s, t gives the joint density of T_1, T_2,

$$f(s,t) = \lambda^2 e^{-\lambda t},\ 0 < s < t < \infty.$$

The joint moment generating function of T_1, T_2 can be obtained from this joint density and is equal to

$$M(\theta, \eta) = E[e^{\theta T_1 + \eta T_2}] = \frac{\lambda^2}{(\lambda - \theta - \eta)(\lambda - \eta)}.$$

Finally, the joint moment generating function of D_1, D_2 is then

$$\begin{aligned} E[e^{\theta_1 D_1 + \theta_2 D_2}] &= E[e^{\theta_1 T_1 + \theta_2 (T_2 - T1)}] \\ &= M(\theta_1 - \theta_2, \theta_2) \\ &= \frac{\lambda^2}{(\lambda - \theta_1)(\lambda - \theta_2)} \\ &= \left(\frac{\lambda}{\lambda - \theta_1}\right)\left(\frac{\lambda}{\lambda - \theta_2}\right). \end{aligned}$$

This shows that D_1, D_2 are independent *exponential*(λ) random variables. This argument can be extended to show that the inter-arrival times, $\{D_n,\ n \geq 1\}$ are independent *exponential*(λ) random variables.

2. The distribution of T_n, the time of the n^{th} arrival, is *Gamma*(n, λ). This follows from the previous property since

$$T_n = \sum_{k=1}^{n} D_k$$

and the sum of independent exponential random variables with the same parameter has a *gamma* distribution. This result gives an alternative definition of a Poisson process: a Poisson process is a counting process on $[0, \infty)$ with inter-arrival times that are independent, exponentially distributed with parameter λ.

3. Let $0 < s < t$. Then

$$P(T_1 \leq s | N(t) = 1) = \frac{s}{t}.$$

This says that conditioned on the event that there is exactly 1 arrival up to time t, the time of that arrival is uniformly distributed over the interval $[0, t]$.

$$
\begin{aligned}
P(T_1 \leq s|N(t)=1) &= \frac{P(T_1 \leq s, N(t)=1)}{P(N(t)=1)} \\
&= \frac{P(N(s)=1, N(t)=1)}{P(N(t)=1)} \\
&= \frac{P(N(s)=1, N(t)-N(s)=0)}{P(N(t)=1)} \\
&= \frac{P(N(s)=1)P(N(t)-N(s)=0)}{P(N(t)=1)} \\
&= \frac{\lambda s e^{-\lambda t}}{\lambda t e^{-\lambda t}} \\
&= \frac{s}{t}.
\end{aligned}
$$

This argument can be extended to show that conditioned on the event that $N(t) = n$, the n arrival times have a distribution that is the distribution of the ordered values from a sample of size n uniformly distributed random variables over the interval $[0, t]$.

4. For $0 < s < t$,

$$P(N(s) = k|N(t) = n) = \binom{n}{k} p^k q^{n-k},$$

where $p = s/t$ and $q = 1 - p$. That is, if $0 < s < t$, then the conditional distribution of $N(s)$ given $N(t)$ is $Binomial(n, p)$.

5. **Renewal Property**. Suppose that we begin observing a Poisson process at the time of an arrival, say T_k. Let N_k denote this new process, $N_k(t) = N(T_k + t) - k$. Since the inter-arrival times of this new process are independent, exponentially distributed with the same parameter λ, then N_k is a Poisson process with the same properties as the original process.

6. **Renewal Property, continued**. Now suppose that we begin observing the Poisson process at an arbitrary time, t_0, and define a new process by $N_0(t) = N(t_0+t)-N(t_0)$. The inter-arrival times after the first arrival of N_0 are independent, exponentially distributed, so the only possible difference between this process and the original would be the distribution of the time to the first arrival of the new process. This corresponds to $T_{01} = T_{N(t_0)+1} - t_0$, called the residual waiting time. It can be shown using the memory-less property of the exponential distribution that this time also has the same exponential distribution as the other inter-arrival times and is independent of those times, so that this new process is also a Poisson process with the same properties as the original. Note that as was the case in the Bernoulli process, this implies that the inter-arrival time that contains our starting time, t_0, is special in that it is stochastically longer than the other inter-arrival times.

7. **Markov Property**. Consider time points $0 \leq t_0 < t_1 < \cdots < t_{n+1}$ and the conditional distribution of the future given the present and the past,

$$P(N(t_{n+1}) = k | N(t_1) = k_1, \cdots, N(t_n) = k_n).$$

Since $N(t_{n+1}) = [N(t_{n+1}) - N(t_n)] + N(t_n)$, and since $N(t_{n+1}) - N(t_n)$ is independent of $\{N(t_1), \cdots, N(t_n)\}$ (independent increments), then

$$\begin{aligned} P(N(t_{n+1}) = k | N(t_1) = k_1, \cdots, N(t_n) = k_n) &= \\ P([N(t_{n+1}) - N(t_n)] + N(t_n) = k | N(t_1) = k_1, \cdots, N(t_n) = k_n) \\ = P([N(t_{n+1}) - N(t_n)] + k_n = k | N(t_1) = k_1, \cdots, N(t_n) = k_n) \\ = P(N(t_{n+1}) - N(t_n) = k - k_n). \end{aligned}$$

Hence,

$$P(N(t_{n+1}) = k | N(t_1), \cdots, N(t_n)) = P(N(t_{n+1}) - N(t_n) = k - N(t_n)) = P(N(t_{n+1}) = k | N(t_n))$$

That is, the conditional distribution of the future given the present and the past depends only on the present. This is the Markov property for continuous-time processes.

8. If $N_1, \cdots, N_r$ are independent Poisson processes with intensities $\lambda_1, \cdots, \lambda_r$, respectively, then $N = \sum_1^r N_k$ is a Poisson process with intensity $\lambda = \sum_1^r \lambda_k$. The proof of this property follows from the result derived earlier that a sum of independent Poisson random variables has a Poisson distribution. This process of summing a collection of stocastic processes to form a new process is referred to as the **superposition** of the processes.

9. Suppose that the Poisson process N represents arrivals of customers to a service center that contains r queues and that there is a monitor that decides which queue a new arrival will enter. Suppose also that this decision is made independently of the arrival process and independently of the queue assignments of other arrivals such that the probability that an arrival is assigned to queue k is p_k, where $\sum p_k = 1$. Let $N_k(t)$ denote the number of arrivals to queue k up to time t, $1 \leq i \leq r$. Then $N_1, \cdots, N_r$ are independent Poisson processes with intensities $\lambda_k = p_k \lambda$. This process of splitting the arrivals of a Poisson process is referred to as **thinning**.

   We can represent thinning as follows. Let $X_i,\ i \geq 1$ be independent random variables that are also independent of the arrival process N with

$$P(X_i = 1) = p_k,\ 1 \leq k \leq r.$$

   Then $N_k(t) = 0$ if $N(t) = 0$; otherwise

$$N_k(t) = \sum_{i=1}^{N(t)} I\{X_i = k\}.$$

We can think of the arrivals as marked by the X_i's, and these marks indicate into which queue the arrival is sent.

This representation of thinning give a generalization of Poisson processes called compound Poisson processess. Suppose $N(t),\ t \geq 0$, is a Poisson process with intensity λ and $X_i,\ i \geq 1$ are i.i.d. non-negative random variables that are also independent of the arrival process N. Define $X_0 = 0$, and let

$$X(t) = \sum_{k=0}^{N(t)} X_k,\ t \geq 0.$$

Then $X(t),\ t \geq 0$ is a compound Poisson process. We can think of this as a Poisson arrival process, $N(t)$, in which the k^{th} arrival has associated with it a value X_k, and $X(t)$ represents the total value of all arrivals up to time t. Note that this is a random sum of random variables, so

$$\begin{aligned} E[X(t)] &= E[N(t)]E(X_1) = \mu\lambda t, \\ Var[X(t)] &= \sigma^2\lambda t + \mu^2\lambda t = (\sigma^2 + \mu^2)\lambda t, \end{aligned}$$

where $E(X_1) = \mu$ and $Var(X_1) = \sigma^2$. More generally, suppose that M_X is the MGF of X_1. Then

$$\begin{aligned} E[\exp\{\theta X(t)\}|N(t) = n] &= E[\exp\{\theta \sum_{k=0}^{n} X_k\}] \\ &= E[\prod_{k=0}^{n} \exp\{\theta X_k\}] \\ &= \prod_{k=0}^{n} E[\exp\{\theta X_k\}] \\ &= [M_X(\theta)]^n. \end{aligned}$$

This implies that

$$\begin{aligned} E[\exp\{\theta X(t)\}|N(t)] &= [M_X(\theta)]^{N(t)} \\ &= \exp\{N(t) \log(M_X(\theta))\}. \end{aligned}$$

And so,

$$\begin{aligned} E[\exp\{\theta X(t)\}] &= E(E[\exp\{\theta X(t)\}|N(t)]) \\ &= E[\exp\{N(t) \log(M_X(\theta))\}] \\ &= \exp\{\lambda t[M_X(\theta) - 1]\}. \end{aligned}$$

We can perform other thinning-related operations. For example, let A denote a subset of $[0, \infty)$, and define

$$\begin{aligned} N_A(t) &= \sum_{k=0}^{N(t)} I(X_k \in A) \\ X_A(t) &= \sum_{k=0}^{N(t)} X_k I(X_k \in A). \end{aligned}$$

Then $N_A(t)$ represents the number of arrivals up to time t whose values are in A, and $X_A(t)$ represents the total value of those arrivals.

Homework Assignment 4

1. Let $\{N_t,\ t \geq 0\}$ be a Poisson process with intensity λ. Find
   (a) $P(N_s = 1 | N_t = 1)$, for any $0 < s < t < \infty$.
   (b) $P(N_s = 1 | N_t = 1)$, for any $0 < t < s < \infty$.
   (c) $E(N_s | N_t = 1)$, for any $0 < s < t < \infty$.
   (d) $E(N_s | N_t = 1)$, for any $0 < t < s < \infty$.

2. Suppose that the errors in the first draft of code produced by the software group of a large corporation occur randomly and independently such that the likelihood that a line contains an error is 0.03. Also suppose that multiple errors on the same line are just counted once. What is the probability that the first 40 lines of code will be error-free? What is the expected line number of the first error? Given that the second error occurred at line number 30, what is the probability that the first error occurred after line number 10?

3. In a large network of computer servers and clients, interruptions due to server breakdowns occur according to a Poisson process with a mean rate of 1.5 per 50 days.
   (a) Suppose you have a critical job that will require 10 days to run. What is the probability that this job will not be interrupted by server breakdowns?
   (b) To improve system reliability, you decide to perform regularly scheduled maintenance of the servers. How often should this maintenance be performed to ensure that there is no more than a 10% chance that a server breakdown will occur before the next scheduled maintenance?
   (c) Given that 2 breakdowns have occurred during the first 30 days, what is the probability that both breakdowns occurred during the first 15 days?
   (d) What is the probability that the time of the second breakdown is more than 60 days after startup?

4. Customers arrive at a pizza restaurant starting at 5:00 pm according to a Poisson process with intensity 25 per hour. Suppose that there is a special promotion at this restaurant so that 20% of the customers have a half-price coupon. Assume that the event that a customer has a coupon is independent of the arrival process.

   (a) What is the expected number of customers with coupons who arrive during the first hour?

   (b) Find the mean waiting time to the first full-paying customer.

   (c) Suppose that the owner of the restaurant gives a coupon to every third full-paying customer. Let $M(t)$ denote the number of full-paying customers who arrive up to time t and who receive a coupon. Is this a Poisson process? Why or why not? What is the expected time between arrivals of such customers?

5. Let $N(t)$ be a Poisson process with intensity λ, and let T_k be the waiting time to the k^{th} arrival.

   (a) Find $E(N(t)|T_3 = s)$, for $t > s$.

   (b) Find $E(T_4|N(t) = 2)$.

   (c) Find $E(N(t)|N(s) = r)$ for $t > s$.

6. Suppose that hits on a faculty web site can come from inside UTD or outside UTD. Hits from inside UTD arrive according to a Poisson process with mean 40/day and hits from outside UTD arrive according to a Poisson process with mean 20/day, independent of the number of hits from inside UTD. Also, 30% of hits from inside UTD access problem solutions, but only 10% of hits from outside UTD access problem solutions.

   (a) What is the mean time between hits that access problem solutions.

   (b) What proportion of hits that access problem solutions come from inside UTD?

   (c) Given that that were 50 hits over 2 days from inside UTD that accessed problem solutions, what is the expected total number of hits over 2 days?

   (d) If a hit did not access problem solutions, what is the probability that it came from outside UTD?