Stat 4382 Stochastic Processes

Syllabus

Stat 4382 Course Information

Instructor:	Dr. Larry P. Ammann
Office hours:	Monday 2:30-3:30 pm, others by appt.
Email:	ammann@utdallas.edu
Office:	FO 2.402C
Phone:	(972) 883-2162
Text:	Introduction to Probability Models, 11^{th} Edition
Author:	Sheldon Ross

Topics	**Chapters**
Review of probability	class notes, text: 1,2
Conditional probability and expectation	text: 3
Markov chains	text: 4.1-4.3
Asymptotic properties of Markov chains	text: 4.4-4.6
Bernoulli and Poisson processes	class notes, text: 5
Renewal processes	text: 7.1-7.3
Queueing processes	text 8.1-8.3
Brownian motion and Gaussian processes	text: 10.1,10.2,10.6,10.7

Grading Policy

Course grade will be based on two exams and homework:
Exam 1, 1/3
Exam 2, 1/3
Homework, 1/3.
Note: the complete syllabus is available here:
`http://www.utdallas.edu/~ammann/stat4382_syllabus.pdf`

Supplemental Notes

Markov chain example

Analysis of list replacement policies.

Suppose N items are contained in a stack in which the cost of retrieving an item is an increasing function of its position in the stack. Suppose also that the cost of pushing items down one position in the stack is negligible. Then we can reduce retrieval costs by replacing selected items in such a way that more frequently requested items are near the top of the stack. One policy for doing this is to take the requested item out of the stack, push items that were above it down one position, and then replace requested item at the top of the stack. We will analyze this policy by tracking the position of a special item after each replacement. Let X_n denote the position of the special item after n replacements. Let p denote the probability that the special item is requested and assume each of the other items has the same chance of being requested. Denote that probability by

$$r = \frac{1-p}{N-1}.$$

Then the special item is returned to the top if and only if it is requested. So,

$$P(X_1 = 1|X_0 = j) = p,\ 1 \leq j \leq N$$

It will stay at any other position as long as the item selected was above it. Then

$$P(X_1 = j|X_0 = j) = (j-1)r$$

The special item will move down one position if any item below it is selected. Then

$$P(X_1 = j+1|X_0 = j) = (N-j)r.$$

To simplify, suppose $N = 5$. Then the TPM is given by

$$P = \begin{bmatrix} p & 1-p & 0 & 0 & 0 \\ p & r & 3r & 0 & 0 \\ p & 0 & 2r & 2r & 0 \\ p & 0 & 0 & 3q & r \\ p & 0 & 0 & 0 & 1-p \end{bmatrix}$$

Statiionary distribution for state 1 is given by

$$\pi_1 = p\pi_1 + p\pi_2 + p\pi_3 + p\pi_4 + p\pi_5 = p\left(\sum_{i=1}^{5}\pi_i\right) = p.$$

The represents the long-term proportion of times the special item is at the top of the stack.

Now suppose we swap positions of a selected item with the item just above it if the selected item is not at the top, and return it to the top if that is where it was. Then the special item will remain at the same position as long as the item just below it is not selected. If the special item is at the bottom, it will move up one position if it is selected, otherwise it stays there. Then the TPM for this policy is given by

$$P = \begin{bmatrix} 1-r & r & 0 & 0 & 0 \\ p & 3r & r & 0 & 0 \\ 0 & p & 3r & r & 0 \\ 0 & 0 & p & 3r & r \\ 0 & 0 & 0 & p & 1-p \end{bmatrix}$$

Stationary distribution for this MC is obtained as follows.

$$\begin{aligned}
\pi_1 &= (1-r)pi_1 + p\pi_2 \quad \Rightarrow \pi_2 = \frac{r}{p}\pi_1 \\
\pi_2 &= r\pi_1 + 3r\pi_2 + p\pi_3 \quad \Rightarrow p\pi_3 = (p+r)\pi_2 - q\pi_1 \Rightarrow \pi_3 = \frac{r^2}{p^2}\pi_1 \\
\pi_3 &= r\pi_2 + 3r\pi_3 + p\pi_4 \quad \Rightarrow \pi_4 = \frac{r^3}{p^3}\pi_1 \\
\pi_5 &= r\pi_4 + (1-p)\pi_5 \quad \Rightarrow \pi_5 = \frac{r}{p}\pi_4 = \frac{r^4}{p^4}\pi_1.
\end{aligned}$$

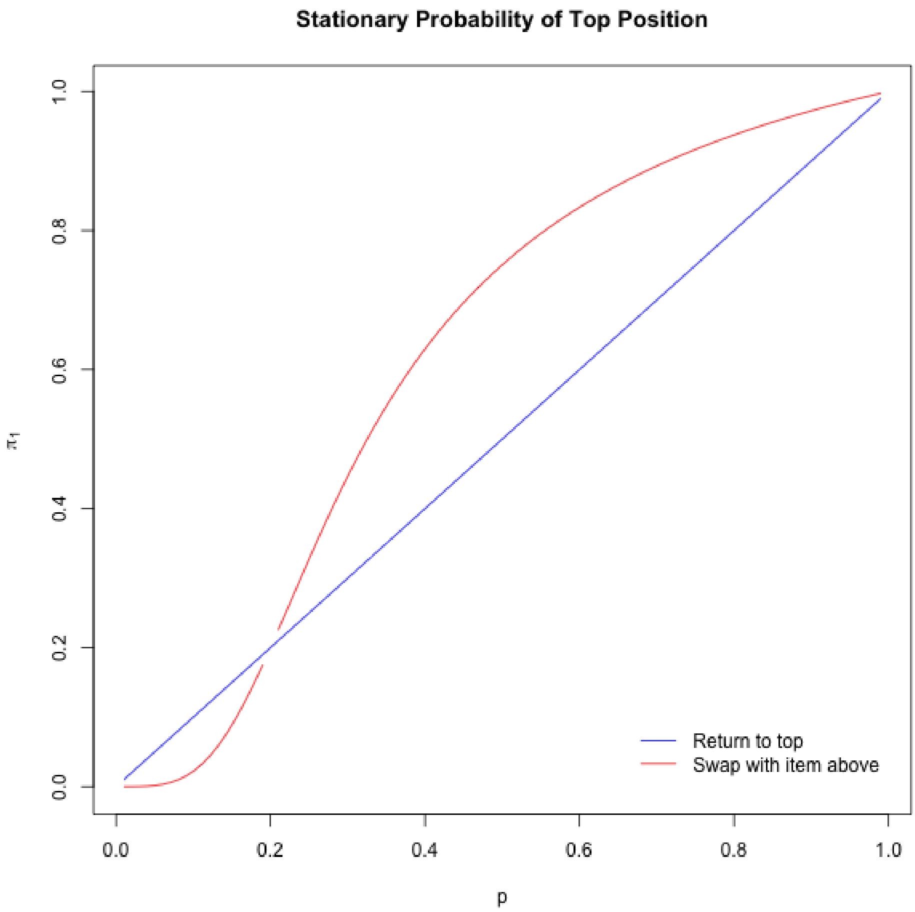
Since $\sum \pi_k = 1$, then

$$1 = \pi_1 \sum_{k=0}^{4} \left(\frac{r}{p}\right)^k.$$

If $r = p$ ($p = 1/5$), then $1 = 5\pi_1$, and so $\pi_1 = 1/5$. If $r \neq p$, then

$$\pi_1 = \left[\sum_{k=0}^{4} \left(\frac{r}{p}\right)^k\right]^{-1} = \frac{1 - r/p}{1-(r/p)^5}.$$

A comparison of these two policies is given by the following plot. It shows that the swap with item above is better than replace at top.



Bernoulli, Poisson, and Renewal Processes

An important family of Markov chains is generated by sums of i.i.d. random variables. Let X_i, $i \geq 1$, be i.i.d. discrete random variables assumed WLOG to be integer valued with pmf,

$$P(X_i = k) = p_k, \quad \sum_k p_k = 1.$$

Define

$$S_0 = 0, \quad S_n = \sum_{i=1}^{n} X_i, \; n \geq 1.$$

Then it is easily seen that $\{S_n\}$ is a homogeneous Markov chain with transition probabilities

$$\begin{aligned} P[i,j] &= P(S_{n+1} = j | S_n = i) \\ &= P(S_n + X_{n+1} = j | S_n = i) \\ &= P(X_{n+1} = j - i | S_n = i) \\ &= p_{j-i}. \end{aligned}$$

Note that the last equality holds since X_{n+1} and X_k, $1 \leq k \leq n$ are independent. In most cases we are not much interested in the MC behavior of such processes, but there are other properties we will examine. These properties will be discussed for some special examples and extensions.

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. It is the counting process representation 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. 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

$0 \le n_0 < n_1 < \cdots < n_r < \infty$, then $S_{n_0}, S_{n_1} - S_{n_0}, \cdots, S_{n_r} - S_{n_{r-1}}$ are independent random variables. If a stochastic process satisfies this property for every $r \ge 1$, then we say the 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,

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

   This implies that

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

   which does not depend on m. Also,

$$Cor(S_m, S_n) = \frac{Cov(S_m, S_n)}{\sqrt{mpq}\sqrt{npq}} = \sqrt{\frac{n}{m}}$$

   which does not depend on 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 counts the number of failures before the first 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:

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

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

Another important class of stochastic processes is the Poisson process. This process has many important applications in addition to providing a 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 non-negative 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 under conditions 1,2 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 $0 \le t_0 < t_1 < \cdots < t_{n+1}$ 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

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

Hence,

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

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, $1 \le i \le r$. 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.

Example 5.21, p. 334. Suppose insurance claims arrive according to a Poisson process with intensity λ and the claim amounts associated with those claims are i.i.d. with mean μ and are independent of the claim arrival times (reasonable under non-disaster conditions). Also suppose the value covered by the policy is depreciated over time. That is, the claim amount C is reduced to

$$Ce^{-\alpha T}$$

for a claim made at time T. Then the total cost of all claims up to time t is

$$D(t) = \sum_{i=1}^{N(t)} C_i e^{-\alpha S_i},$$

where S_i is the arrival time of the *i-th* claim, C_i is the amount of that claim, and *N(t)* is the number of claims up to time t. Since this is a random sum of random variables, we can obtain its expected value by conditioning on *N(t)*.

$$E[D(t)] = \sum_{n=0}^{\infty} E[D(t)|N(t) = n]P(N(t) = n) = \sum_{n=0}^{\infty} E[D(t)|N(t) = n]\frac{(\lambda t)^n}{n!}e^{-\lambda t}.$$

Recall that conditioned on $N(t) = n$, the times of the n arrivals are distributed as the ordered values of n i.i.d. uniform r.v.'s on $(0, t)$. Let $U_1, \cdots, U_n$ denote the i.i.d. uniform r.v.'s and let $U_{(1)}, \cdots, U_{(n)}$ denote their ordered values. Then

$$\begin{aligned}
E[D(t)|N(t) = n] &= E\left[\sum_{i=1}^{n} C_i e^{-\alpha U_{(i)}}\right] \\
&= \sum_{i=1}^{n} E\left[C_i e^{-\alpha U_{(i)}}\right] \\
&= \sum_{i=1}^{n} E[C_i] E\left[e^{-\alpha U_{(i)}}\right] \quad \text{(independence of claim amounts and times)} \\
&= \mu \sum_{i=1}^{n} E\left[e^{-\alpha U_{(i)}}\right] \\
&= \mu E\left[\sum_{i=1}^{n} e^{-\alpha U_{(i)}}\right] \\
&= \mu E\left[\sum_{i=1}^{n} e^{-\alpha U_i}\right] = n\mu E\left[e^{-\alpha U}\right],
\end{aligned}$$

where the last step follows from the fact that a sum of ordered values is the same as the sum of the unordered values. Since U is uniformly distributed, then

$$E\left[e^{-\alpha U}\right] = \frac{1}{t}\int_0^t e^{-\alpha s} ds = \frac{1}{\alpha t}(1 - e^{-\alpha t}).$$

Therefore,

$$E[D(t)|N(t)] = N(t)\frac{\mu}{\alpha t}(1 - e^{-\alpha t}),$$

and so

$$E[D(t)] = \lambda t \frac{\mu}{\alpha t}(1 - e^{-\alpha t}) = \lambda \frac{\mu}{\alpha}(1 - e^{-\alpha t}).$$

Markov Shot Noise Process. A similar problem is a model for noise which arrives at discrete times but whose value decays over time. Specifically, assume noise events arrive according to a Poisson process $N(t)$ with intensity λ. The initial magnitude of the i^{th} noise event is a r.v. X_i and the noise magnitude at time t is

$$X_i e^{-\alpha(t-S_i)},$$

where S_i is the arrival time of the i^{th} noise event and α is a constant. We assume $\{X_i,\ i \geq 1\}$ are i.i.d., are independent of $N(t)$, $t > 0$, and are independent of the arrival process. Then the total noise at time t is given by

$$X(t) = \sum_{i=1}^{N(t)} X_i e^{-\alpha(t-S_i)},$$

for $N(t) > 0$, and $X(t) = 0$ when $N(t) = 0$. This model is referred to as a *shot noise* process.

The moment generating function of this process can be obtained by conditioning on $N(t)$ and using the property that the conditional distribution of the arrival times given $N(t) = n$ is the distribution of the ordered values of n i.i.d. uniform r.v.'s. Then

$$E[\exp\{\theta X(t)\}|N(t) = n] = E\left[\exp\left(\theta \sum_{i=1}^{n} X_i e^{-\alpha(t-U_{(i)})}\right)\right],$$

where $U_1, \ldots U_n$ are i.i.d uniform r.v.s on $(0, t)$, and $U_{(1)}, \ldots U_{(n)}$ are their ordered values. Since X_i are i.i.d., U_i are i.i.d. and are independent of X_i, then

$$\sum_{i=1}^{n} X_i e^{-\alpha(t-U_{(i)})}$$

and

$$\sum_{i=1}^{n} X_i e^{-\alpha(t-U_i)}$$

have the same distribution. Therefore,

$$E[\exp\{\theta X(t)\}|N(t) = n] = \left(E\left[X_i e^{-\alpha(t-U_i)}\right]\right)^n.$$

Let $M_X(\theta)$ denote the m.g.f. of X_i and note that the expectation in the right-hand-side of the previous equation refers to the joint distribution of X_i, U_i which are independent. Hence,

$$\begin{aligned}
E\exp\{\theta X e^{-\alpha(t-U)}\} &= E\left[\frac{1}{t}\int_0^t \exp\{\theta X e^{-\alpha(t-u)}\}du\right] \\
&= \frac{1}{t}\int_0^t E\left[\exp\{\theta X e^{-\alpha y}\}\right] dy \quad (y = t - u) \\
&= \frac{1}{t}\int_0^t M_X\left(\theta e^{-\alpha y}\right).
\end{aligned}$$

Denote this by B. Then

$$\begin{aligned}
E\exp\{\theta X(t)\} &= \sum_{n=0}^{\infty} B^n \frac{(\lambda t)^n}{n!} \\
&= e^{\lambda t(B-1)} \\
&= \exp\left(\lambda \int_0^t [M_X(\theta e^{-\alpha y}) - 1]dy\right)
\end{aligned}$$

Moments can be obtained by differentiating the MGF. So for example, if $E(X) = \mu$ and $Var(X) = \sigma^2$, then

$$E[X(t)] = \frac{\lambda\mu}{\alpha}(1 - e^{-\alpha t}),$$

$$Var[X(t)] = \frac{\lambda(\mu^2 + \sigma^2)}{2\alpha}(1 - e^{-2\alpha t}),$$

$$
\begin{aligned}
Cov(X(t), X(t+s)) &= e^{-\alpha s} Var[X(t)] \\
&= e^{-\alpha s} \frac{\lambda(\mu^2+\sigma^2)}{2\alpha}(1-e^{-2\alpha t}).
\end{aligned}
$$

Also, we can obtain the limiting distribution of $X(t)$ by letting $t \to \infty$ in the MGF. This gives

$$
\lim_{t\to\infty} E[\exp\{\theta X(t)\}] = \exp\left\{\lambda \int_0^\infty [M_X(\theta e^{-\alpha y}) - 1]dy\right\}.
$$

Now consider the special case in which the X_i are exponentially distributed with rate η. Then

$$
M_X(\theta) = \frac{\eta}{\eta - \theta}
$$

and so

$$
\begin{aligned}
\lim_{t\to\infty} E[\exp\{\theta X(t)\}] &= \exp\left\{\lambda \int_0^\infty \left(\frac{\eta}{\eta - \theta e^{-\alpha y}} - 1\right) dy\right\} \\
&= \exp\left\{\frac{\lambda}{\alpha} \int_0^\theta \frac{1}{\eta - x} dx\right\} \quad (x = \theta e^{-\alpha y}) \\
&= \left(\frac{\eta}{\eta-\theta}\right)^{\lambda/\alpha}.
\end{aligned}
$$

This is the MGF of a gamma r.v. with shape λ/α and rate η.

Renewal Processes

The renewal property of Poisson processes provides a third definition of a Poisson process: a counting process with interarrival times that are independent and have the same exponential distribution with rate λ. We can generalize this definition to define other counting processes whose interarrival times are i.i.d r.v.'s, $X_k,\ k \geq 1$. The distribution of the interarrival times is defined over $[0, \infty)$ with

$$
P(X_k > 0) > 0.
$$

Such processes are called *renewal processes*. For example, every other arrival in a Poisson process is a renewal process, but not a Poisson process, because the interarrival times for those arrivals have a *Gamma* distribution with shape *2*.

Obviously renewal processes satisfy the renewal property when the process is reset at the time of an arrival. However, the Poisson process is the only renewal process in which time to the first arrival after resetting at an arbitrary time t has the same exponential distribution as all the other interarrival times. To allow for resetting at arbitrary times, the first arrival time, X_1, is allowed to have a different distribution, but $X_k,\ k \geq 2$, are i.i.d. We will only consider renewal processes in which all interarrival times have the same distribution. Let F denote the d.f. of the interarrival times. Let

$$\mu = E(X_n),\ n \geq 1.$$

Since

$$F(0) = P(X_k = 0) < 1,$$

then $\mu > 0$. We will assume that $\mu < \infty$.

The arrival or renewal times are defined by

$$S_0 = 0, \quad S_n = \sum_{k=1}^{n} X_k,\ n \geq 1.$$

The counting process version is given by

$$N(t) = \max\{n:\ S_n \leq t\}.$$

The Strong Law of Large Numbers implies that

$$\frac{S_n}{n} \to \mu \text{ as } n \to \infty$$

with probability 1. This also implies that $N(t) < \infty$ for $0 < t < \infty$ and that

$$N(\infty) = \lim_{t \to \infty} N(t) = \infty$$

with probability 1.

Let X, Y be independent r.v.'s with distribution functions F_X, F_Y, respectively, and assume they have density functions f_X, f_Y. The convolution of F_X, F_Y is defined to be the distribution of the sum of the r.v.s and is denoted by $F_X * F_Y$,

$$\begin{aligned}
F_X * F_Y(t) &= P(X + Y \leq t) \\
&= \int_0^\infty P(X \leq t - s | Y = s) f_Y(s) ds \\
&= \int_0^\infty F_X(t - s) f_Y(s) ds \\
&= \int_0^\infty \left[\int_0^{t-s} f_X(u) du\right] f_Y(s) ds.
\end{aligned}$$

The density function of the convolution is

$$f_{X+Y}(t) = \int_0^\infty f_X(t - s) f_Y(s) ds.$$

If F is the d.f. of the interarrival times of a renewal process, then the distribution of

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

is

$$F_n(t) = F^{*n}(t) = F * \cdots * F(t).$$

Since

$$\{N(t) \geq n\} \Leftrightarrow \{S_n \leq t\},$$

then

$$P(N(t) = n) = P(S_n \leq t) - P(S_{n+1} \leq t) = F_n(t) - F_{n+1}(t).$$

An alternative expression for the p.m.f. of *N(t)* can be obtained by conditioning on S_n.

$$\begin{aligned} P(N(t) = n) &= \int_0^\infty P(N(t) = n | S_n = y) f_n(y) dy \\ &= \int_0^t P(X_{n+1} > t - y | S_n = y) f_n(y) dy \\ &= \int_0^t (1 - F(t - y)) f_n(y) dy. \end{aligned}$$

In the case of a Poisson process, F is the exponential distribution with rate λ and F_n is the gamma distribution with shape n and rate λ. Substitution of these distributions into the above integral gives the Poisson p.m.f. for *N(t)*.

One of the major results for renewal processes concerns the mean function,

$$\begin{aligned} m(t) &= E(N(t)) \\ &= \sum_{n=1}^{\infty} P(N(t) \geq n) \\ &= \sum_{n=1}^{\infty} P(S_n \leq t) \\ &= \sum_{n=1}^{\infty} F_n(t) \\ &= \sum_{n=1}^{\infty} F^{*n}(t). \end{aligned}$$

This function is referred to as the renewal function and it uniquely determines the interarrival time distribution and, hence, the renewal process.

The renewal function satisfies an integral equation,

$$m(t) = F(t) + \int_0^t m(t - s) f(s) ds,$$

called the renewal equation. This equation is obtained by conditioning on the time of the first renewal. The renewal property implies that for $0 < s < t$,

$$E[N(t) | X_1 = s] = 1 + E[N(t - s)]$$

and

$$E[N(t)|X_1 = s] = 0,\ s > t.$$

This gives

$$\begin{aligned} m(t) = E[N(t)] &= \int_0^\infty E[N(t)|X_1 = s] f(s) ds \\ &= \int_0^t [1 + m(t-s)] f(s) ds \\ &= F(t) + \int_0^\infty m(t-s) f(s) ds. \end{aligned}$$

This equation can be solved in general for only a few cases. Of main interest is the limiting behavior of the renewal function.

First note that

$$S_{N(t)} \leq t < S_{N(t)+1},$$

and so

$$\frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} < \frac{S_{N(t)+1}}{N(t)}.$$

The Strong Law of Large Numbers implies that

$$\frac{S_{N(t)}}{N(t)} \to \mu$$

and

$$\frac{S_{N(t)+1}}{N(t)} \to \mu$$

with probability 1. Therefore,

$$\frac{N(t)}{t} \to \frac{1}{\mu}$$

with probability 1 as $t \to \infty$. The *Elementary Renewal Theorem* states that this limit also holds for the mean function,

$$\lim_{t\to\infty} \frac{m(t)}{t} = \frac{1}{\mu}.$$

Homework Assignments

Homework 1

Due date: 1/30/2017

1. Problem 8, p. 14.
   If $P(E) = 0.9$ and $P(F) = 0.8$, show that $P(E \cap F) \geq 0.7$. More generally, show that
   $$P(E \cap F) \geq P(E) + P(F) - 1.$$

2. Problem 10, p. 14.
   Show that
   $$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i).$$
   Hint: use mathematical induction on n, or else express $\cup E_i$ as a disjoint union.

3. Problem 20, p. 15.
   Three dice are tossed. What is the probability the same number appears on exactly two of the three dice?

4. Problem 21, p. 15.
   Suppose that 5% of men and 0.25% of women are color-blind in a population that consists of an equal number of males and females. If a randomly chosen person is colorblind, what is the probability this person is male?

5. Problem 22, p. 15.
   Suppose A and B play a sequence of independent games, each worth 1 point, until one has 2 more points than the other. If the probability A wins a game is p, what is the probability the game ends after $2n$ games? What is the probability that A wins?

6. Problem 29, p. 16.
   Suppose that $P(E) = 0.6$. What can you say about $P(E|F)$ when
   (a) E and F are mutually exclusive?
   (b) $E \subset F$?
   (c) $F \subset E$?

7. Problem 42, p. 18.
   There are three coins in a box. One is a two-headed coin, one is a fair coin, and the third is a biased coin that comes up heads 75% of the time. Suppose one of those coins is selected at random and then flipped. If that coin comes up heads, what is the proability that it was the two-headed coin that was flipped?

Solutions for Homework 1

1. Problem 8, p. 14.
   From the Theorem of Total Probability,

$$\begin{aligned} P(E) &= P(E \cap F) + P(E \cap F^c) \\ &\leq P(E \cap F) + P(F^c) \\ &= P(E \cap F) + 1 - P(F). \end{aligned}$$

   Therefore,

$$P(E \cap F) \geq P(E) + P(F) - 1.$$

2. Problem 10, p. 14.
   Proof by induction: the result holds trivially for $n = 1$, so now suppose it holds for n, that is, suppose

$$P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i).$$

   It remains to show that it holds for $n + 1$. Since

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B),$$

   then

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} E_i\right) &= P\left(\bigcup_{i=1}^{n} E_i \cup E_{n+1}\right) \\ &\leq P\left(\bigcup_{i=1}^{n} E_i\right) + P(E_{n+1}) \\ &\leq \sum_{i=1}^{n} P(E_i) + P(E_{n+1}) \\ &= \sum_{i=1}^{n+1} P(E_i). \end{aligned}$$

   Alternative proof: let $A_1 = E_1$ and define

$$A_k = E_k \bigcap \left(\bigcup_{i=1}^{k-1} E_i\right)^c, \; k \geq 2.$$

Note that A_i, $i \geq 1$, are disjoint sets, and

$$\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{n} A_i, \; n \geq 1.$$

Since $A_i \subset E_i$, then $P(A_i) \leq P(E_i)$, and so

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} E_i\right) &= P\left(\bigcup_{i=1}^{n} A_i\right) \\ &= \sum_{i=1}^{n} P(A_i) \\ &\leq \sum_{i=1}^{n} P(E_i). \end{aligned}$$

3. Problem 20, p. 15.
   This is an experiment with equally likely outcomes, so we need to use counting methods to find this probability. Label the dice 1,2,3. There are 6 ways to choose the number that appears twice, there are 5 ways to choose the number that appears once, and there are 3 ways to choose which die has the number that appears once. So the numerator of this probability is $6 * 5 * 3 = 90$. The denominator is 6*6*6=216, so this probability is $90/216 = 5/12$.

4. Problem 21, p. 15.
   This is an application of Bayes Theorem.

$$P(\text{male} \mid \text{colorblind}) = \frac{P(\text{male} \cap \text{colorblind})}{P(\text{colorblind})},$$

$$\begin{aligned} P(\text{colorblind}) &= P(\text{colorblind} \mid \text{male})P(\text{male}) + P(\text{colorblind} \mid \text{female})P(\text{female}) \\ &= (.05)(.5) + (.0025)(.5) \\ &= 0.02625. \end{aligned}$$

   Therefore,

$$P(\text{male} \mid \text{colorblind}) = \frac{.025}{.0265} = 0.952381.$$

5. Problem 22, p. 15.
   Let A_k denote the event that A wins game k and let B_k denote the event that B wins game k. Note that play ends after two games if

$$(A_1 \cap A_2) \cup (B_1 \cap B_2)$$

and it will continue otherwise. Play ends after 4 games if

$$[(A_1 \cap B_2) \cup (B_1 \cap A_2)] \cap [(A_3 \cap A_4) \cup (B_3 \cap B_4)].$$

Note that

$$[(A_1 \cap A_2) \cup (B_1 \cap B_2)] = [(A_1 \cap B_2) \cup (B_1 \cap A_2)]^c.$$

Continuing in the same way, let

$$E_n = (A_{2n-1} \cap B_{2n}) \cup (B_{2n-1} \cap A_{2n}).$$

Since

$$P(E_i) = 2p(1-p),$$

then the probability that play ends after $2n$ games is given by

$$P\left(\left[\bigcap_{i=1}^{n-1} E_i\right] \bigcap E_n^c\right) = [2p(1-p)]^{n-1}(1-2p(1-p)),$$

The probability that play ends after $2n$ games and A wins is

$$P\left(\left[\bigcap_{i=1}^{n-1} E_i\right] \bigcap A_{2n-1} \bigcap A_{2n}\right) = [2p(1-p)]^{n-1}p^2.$$

Therefore, from the Theorem of Total Probability the probability A wins is given by

$$\begin{aligned}
P(A \text{ wins}) &= \sum_{n=1}^{\infty} P(\{A \text{ wins}\} \cap \{\text{game ends after } 2n \text{ games}\}) \\
&= \sum_{n=1}^{\infty} [2p(1-p)]^{n-1}p^2 \\
&= \frac{p^2}{1-2p(1-p)} \\
&= \frac{p^2}{p^2+(1-p)^2}.
\end{aligned}$$

6. Problem 29, p. 16.

(a) $$P(E|F) = \frac{P(E \cap F)}{P(F)} = 0.$$

(b) $$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(E)}{P(F)}.$$

(c) $$P(E|F) = \frac{P(E \cap F)}{P(F)} = \frac{P(F)}{P(F)} = 1.$$

7. Problem 42, p. 18.
   This is an application of Bayes Theorem. Label the coins c_1, c_2, c_3. Then

$$P(H \mid c_1) = 1,\ P(H \mid c_2) = 0.50,\ P(H \mid c_3) = 0.75,$$

   and so, by the Theorem of Total Probability,

$$\begin{aligned} P(H) &= P(H \mid c_1) * \frac{1}{3} + P(H \mid c_2) * \frac{1}{3} + P(H \mid c_2) * \frac{1}{3} \\ &= \frac{1}{3}(1 + 0.50 + 0.75) \\ &= 0.75. \end{aligned}$$

   Therefore,

$$P(c_1 \mid H) = \frac{1/3}{0.75} = 4/9 = 0.444.$$

Homework 2

Due date: 2/13/2017

1. Problem 25, p. 82 (10^{th} ed: p. 88)
2. Problem 26, p. 82 (10^{th} ed: p. 88)
3. Problem 28, p. 82 (10^{th} ed: p. 88)
4. Problem 37, p. 169 (10^{th} ed: p. 178)
5. Problem 40, p. 170 (10^{th} ed: p. 179)
6. Problem 46, p. 171 (10^{th} ed: p. 180)
7. Suppose that your company receives a large shipment of computer memory cards and that the defective rate for these cards is p. If the proportion of defectives in the shipment is too high, you want to reject the shipment. To test the shipment, you randomly select 10 cards from the shipment (assume sampling with replacement, although the result would be essentially the same for sampling without replacement since the sample size would be much smaller than the population size). The shipment will be rejected if at least 2 defective cards are found in the sample. Plot the probability of rejecting the shipment as a function of p. What is the value of p for which the probability of rejecting the shipment is 0.80?

Solutions for Homework 2

1. Problem 25, p. 82 (10^{th} ed: p. 88). 7 games will be played iff each team has won 3 games after 6 games have been played. Each such outcome has probability $p^3(1-p)^3$. The number of such outcomes is the number of subsets of size 3 taken from (1,2,3,4,5,6) since that represents the games won by A. This gives

   $$P(N=7) = \binom{6}{3} p^3(1-p)^3 = 20[p(1-p)]^3.$$

   Maximum of $p(1-p)$ is obtained by setting first derivative equal to 0,

   $$1 - 2p = 0$$

   and then noting that the second derivative (-2) is negative. This implies that $[p(1-p)]^3$ also is maximized at $p = 1/2$, and therefore so is $P(N=7)$.

2. Problem 26, p. 82 (10^{th} ed: p. 88). Let N_i denote the number of games played. Then $i \leq N_i \leq 2i - 1$. For $i = 2$,

   $$\begin{aligned} P(N_2 = 2) &= P(AA) + P(BB) = p^2 + (1-p)^2 \\ P(N_2 = 3) &= 1 - P(N = 2) = 2p(1-p), \end{aligned}$$

   so

   $$E(N_2) = 2[p^2 + (1-p)^2] + 3[2p(1-p)] = 2[p^2 + 2p(1-p) + (1-p)^2] + 2p(1-p) = 2 + 2p(1 -$$

   Similarly, for N_3, note that $N_3 = 3$ when one team wins first 3 games; $N_3 = 4$ when A wins 2 out of first 3 games and then wins game 4 or B wins 2 games of first 3 and then wins game 4; $N_3 = 5$ when each team wins 2 of first 4 games.

   $$\begin{aligned} P(N_3 = 3) &= P(AAA) + P(BBB) = p^3 + (1-p)^3 \\ P(N_3 = 4) &= \binom{3}{1} p^2(1-p)p + \binom{3}{1}(1-p)^2 p(1-p) \\ &= 3p^3(1-p) + 3p(1-p)^3 \\ P(N_3 = 5) &= \binom{4}{2} p^2(1-p)^2. \end{aligned}$$

   As a check, note that these probabilities sum to 1. Then

   $$\begin{aligned} E(N_3) &= 3p^3 + 3(1-p)^3 + 12p^3(1-p) + 12p(1-p)^3 + 30p^2(1-p)^2 \\ &= 3 + 3p^3(1-p) + 3p(1-p)^3 + 12p^2(1-p)^2 \\ &= 3 + 3p(1-p)[p^2 + (1-p)^2 + 4p(1-p)] \\ &= 3 + 3p(1-p) + 6p^2(1-p)^2. \end{aligned}$$

3. Problem 28, p. 82 (10^{th} ed: p. 88). Let $A = \{HH \cup TT\}$ denote the event that the process continues after two flips, and let N denote the number of flips until the last two flips are different. Note for example, that

$$\begin{aligned}
P(X=1, N=0) &= P(HT) = p(1-p), \\
P(X=1, N=2) &= P(A, HT) = (p^2 + (1-p)^2)p(1-p), \\
P(X=1, N=4) &= P(AA, HT) = [p^2 + (1-p)^2]^2 p(1-p)
\end{aligned}$$

So for $k \geq 0$,

$$P(X=1, N=2k) = [p^2 + (1-p)^2]^k p(1-p),$$

and

$$\begin{aligned}
P(X=1) &= \sum_{k=0}^{\infty} [p^2 + (1-p)^2]^k p(1-p) \\
&= p(1-p) \frac{1}{1 - p^2 - (1-p)^2} \\
&= \frac{p(1-p)}{2p(1-p)} \\
&= 1/2.
\end{aligned}$$

In the alternative experiment to get $X = 1$ we must have all heads followed by a tail and the first flip must be heads. So in this case,

$$\begin{aligned}
P(X=1) &= \sum_{k=1}^{\infty} p^k (1-p) \\
&= (1-p) \frac{p}{1-p} \\
&= p.
\end{aligned}$$

4. Problem 37, p. 169 (10^{th} ed: p. 178).

$$\begin{aligned}
E(X) &= E(X|A)P(A) + E(X|B)P(B) + E(X|C)P(C) \\
&= 2.6/3 + 3/3 + 3.4/3 = 3.0
\end{aligned}$$

Note that the 2nd moment of a Poisson r.v. with mean μ is

$$E(N^2) = \text{Var}(N) + \mu^2 = \mu + \mu^2.$$

Therefore,

$$
\begin{aligned}
E(X^2) &= E(X^2|A)P(A) + E(X^2|B)P(B) + E(X^2|C)P(C) \\
&= [(2.6 + 2.6^2) + (3 + 3^2) + (3.4 + 3.4^2)]/3 \\
&= 36.32/3 = 12.11
\end{aligned}
$$

and so $\text{Var}(X) = 12.11 - 9 = 3.11$.

5. Problem 40, p. 170 (10^{th} ed: p. 179). This problem can be solved by representing N as a random sum of random variables. Let M denote the number of door selections before selecting door 3, let X_j be i.i.d. r.v.'s that are independent of M with

$$
\begin{aligned}
P(X_j = 2) &= P(\text{select door 1}|\text{not 3}) = \frac{.5}{.8} = 5/8 \\
P(X_j = 3) &= P(\text{select door 2}|\text{not 3}) = \frac{.3}{.8} = 3/8.
\end{aligned}
$$

Note that we are using the conditional distribution of the door selection given that door 3 is not selected for this p.m.f. Then

$$
N = \sum_{j=0}^{M} X_j,
$$

where $X_0 = 0$. So N is a random sum of random variables and M has a geometric distribution with success probability 0.2. Also,

$$
\begin{aligned}
E(M) &= \frac{1-0.2}{0.2} = 4, \\
Var(M) &= \frac{1-0.2}{0.04} = 20, \\
E(X_j) &= \frac{19}{8}, \\
Var(X_j) &= \frac{47}{8} - \left(\frac{19}{8}\right)^2 = 15/64 = 0.234375.
\end{aligned}
$$

This gives

$$
\begin{aligned}
E(N) &= E(M)E(X_j) = 9.5, \\
Var(N) &= Var(X_j)E(M) + [E(X_j)]^2 Var(M) = 113.75.
\end{aligned}
$$

For part (b), we can enumerate all of the possibilities.

Selection	Travel	Probability
3	0	1/3
13	2	1/6
23	3	1/6
123, 213	5	1/3

This gives,

$$\begin{aligned} E(N) &= 5/2 \\ Var(N) &= E(N^2) - (E(N))^2 = 10.5 - 6.25 = 4.25, \end{aligned}$$

which is considerably smaller than the variance for part (a).

6. Problem 46, p. 171 (10^{th} ed: p. 180).
(a)

$$\begin{aligned} Cov(X, E(Y|X)) &= E[XE(Y|X)] - E(X)E[E(Y|X)] \\ &= E[E(XY|X)] - E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \\ &= Cov(X, Y). \end{aligned}$$

(b) If $E(Y|X) = a + bX$, then from part (a),

$$\begin{aligned} Cov(X, Y) &= Cov(X, E(Y|X)) \\ &= Cov(X, a + bX) \\ &= bCov(X, X) \\ &= bVar(X). \end{aligned}$$

Therefore,

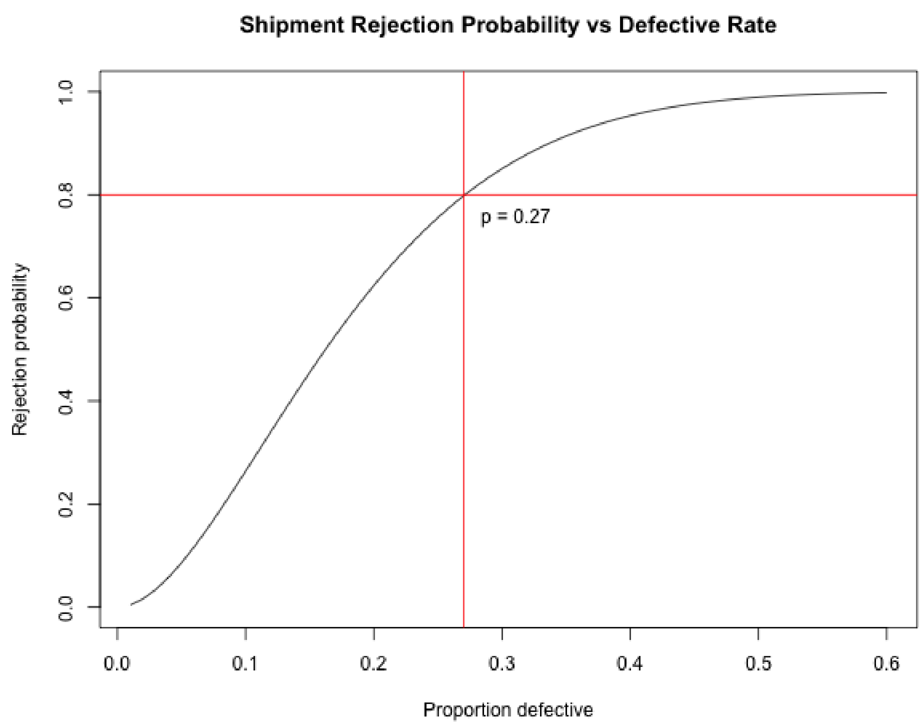
$$b = \frac{Cov(X, Y)}{Var(X)}.$$

7. In this test the number of defectives would have a Binomial distribution with $n = 10$ and so the probability the shipment will be rejected is

$$P(\text{reject shipment}) = 1 - (1-p)^{10} - 10p(1-p)^9.$$

A plot of the rejection probability is shown below.

The probability of rejecting a shipment will be at least 0.8 when the defective rate is at least 0.27.


Shipment Rejection Probability vs Defective Rate
Rejection probability
Proportion defective
p = 0.27
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.1
0.2
0.3
0.4
0.5
0.6

Homework 3

Due date: 3/1/2017

1. 4.5, p. 261 (4.5, p. 276, 10^{th} ed.)
2. 4.8, p. 262 (4.8, p. 276, 10^{th} ed.)
3. 4.14, p. 262 (4.14, p. 277, 10^{th} ed.)
4. 4.23, p. 264 (4.23, p. 278, 10^{th} ed.)
5. Let $X_k, k \geq 0$ be a Markov chain with states $0, 1$ and transition probability matrix

$$P = \begin{bmatrix} 0.99 & 0.01 \\ 0.12 & 0.88 \end{bmatrix}$$

(a) Find $P(X_4 = 1 | X_1 = 1)$.
(b) If $P(X_0 = 0) = .1$, find $P(X_4 = 1)$.
(c) Find the stationary distribution of this Markov chain.

Solutions for Homework 3

1. 4.5, p. 261 (4.5, p. 276, 10^{th} ed.) The p.m.f for X_3 is given by

$$\begin{aligned}
\pi_3 &= \pi_0 P^3 = [.25, .25, .50] \begin{bmatrix} 1/2 & 1/3 & 1/6 \\ 0 & 1/3 & 2/3 \\ 1/2 & 0 & 1/2 \end{bmatrix}^3 \\
&= [.25, .25, .50] \begin{bmatrix} .361 & .204 & .435 \\ .444 & .148 & .407 \\ .417 & .222 & .361 \end{bmatrix} \\
&= [0.410, 0.199, 0.391].
\end{aligned}$$

Therefore,

$$E(X_3) = 0 * 0.410 + 1 * 0.199 + 2 * 0.391 = 0.981$$

2. 4.8, p. 262 (4.8, p. 276, 10^{th} ed.) This can be represented as a 2-state MC with states H,T and TPM

$$P = \begin{bmatrix} .7 & .3 \\ .6 & .4 \end{bmatrix}$$

First we need to obtain the initial state distribution.

$$\begin{aligned}
P(X_0 = H) &= P(X_0 = H|\text{coin } 1)P(\text{coin } 1) + P(X_0 = H|\text{coin } 2)P(\text{coin } 2) \\
&= .7(.5) + .6(.5) = .65
\end{aligned}$$

This gives $\pi_0 = [.65, .35]$. The event that the coin flipped on the third day is coin 1 is the event $\{X_2 = H\}$, and so

$$P(X_2 = H) = \pi_0 P P[, 1] = [.65, .35][.67, .66]^T = 0.6665,$$

using **R** notation for sections of a matrix. Also, we must find

$$P(X_4 = H|X_0 = H) = P[1,]P^2 P[, 1] = 0.6667.$$

3. 4.14, p. 262 (4.14, p. 277, 10^{th} ed.)
P_1: all states communicate, so they form a single recurrent class.
P_2: all states communicate, so they form a single recurrent class.
P_3: $\{1, 3\}$ is a closed, communicating class, so it is recurrent. There is a positive probability of going from state 2 to recurrent state 1 and therefore never returning to state 2, so state 2 is transient. $\{4, 5\}$ is a closed, communicating class, so it is recurrent.

Note that this is a reducible MC.
P_4: $\{1, 2\}$ is a closed, communicating class, so it is recurrent. State 3 is a closed, communicating class, so it is recurrent. There is a positive probability of going from state 4 to absorbing state 3 and therefore never returning to state 4, so state 4 is transient. Likewise, there is a positive probability of going from state 5 to recurrent state 1 and therefore never returning to state 5, so state 5 is transient.

4. 4.23, p. 264 (4.23, p. 278, 10^{th} ed.) We can represent the weather condition as a two-state MC X_n with states *good, bad*. This MC has TPM

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}$$

(a) Let N_k denote the number of storms in year k. Then

$$\begin{aligned} E(N_1|X_0 = good) &= 1 \cdot P(X_1 = good|X0 = good) + 3 \cdot P(X_1 = bad|X_0 = good) \\ &= 1/2 + 2 = 2.5 \end{aligned}$$

Since

$$P^2 = \begin{bmatrix} 5/12 & 7/12 \\ 7/18 & 11/18 \end{bmatrix}$$

then

$$\begin{aligned} E(N_2|X_0 = good) &= 1 \cdot P(X_2 = good|X0 = good) + 3 \cdot P(X_2 = bad|X_0 = good) \\ &= 5/12 + 21/12 \\ &= 13/6 = 2.17 \end{aligned}$$

5. Let $X_k, k \geq 0$ be a Markov chain with states $0, 1$ and transition probability matrix

$$P = \begin{bmatrix} 0.99 & 0.01 \\ 0.12 & 0.88 \end{bmatrix}$$

(a) Find $P(X_4 = 1|X_1 = 1)$.

$$P(X_4 = 1|X_1 = 1) = P[2,]PP[, 2] = 0.685.$$

(b) If $P(X_0 = 0) = .1$, find $P(X_4 = 1)$.

$$\begin{aligned} \pi_4 &= \pi_0 P^4 \\ &= [.1, .9] \begin{bmatrix} 0.967 & 0.033 \\ 0.394 & 0.606 \end{bmatrix} \\ &= [.452, .548] \end{aligned}$$

So $P(X_4 = 1) = .548$.

(c) Find the stationary distribution of this Markov chain. The first equation in $\pi = \pi P$ is

$$\pi_0 = 0.99\pi_0 + 0.12\pi_1$$

which gives, $\pi_0 = 12\pi_1$, and so $\pi_1 = 1/13$, $\pi_0 = 12/13$.

Homework 4

Due date: 4/10/2017

1. 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 $E(N(t)|T_2 = s)$, for $0 < s < t < \infty$.

   (b) Find $E(T_3|N(t) = 1)$.

   (c) Find $E(N(s)|N(t) = r)$ for $0 < t < s < \infty$.

2. A communications center has one link that is busy 75% of the time. Suppose the center polls this link every 10 minutes to determine whether or not the link is free, and suppose whether or not it is free at one time is independent of whether or not it is free any other time.

   (a) What is the probability that the link will be busy each of the first 4 times it is polled?

   (b) What is the probability that it will take more than 5 pollings before finding 2 free times?

   (c) Given that the link was free 2 times during the first 10 pollings, what is the probability that the link was busy at the first 5 pollings?

3. Accidents occur at a particular intersection according to a Poisson process with a mean rate of 1 per 10 days. Suppose that 60% of these accidents result in no serious injuries (injuries that require hospitalization), 20% result in 1 person who was seriously injured, 15% result in 2 people who are seriously injured, 3% result in 3 people who are hospitalized, and the rest result in 4 people who are seriously injured (note: a fatality is counted as a serious injury).

   (a) What is the mean time between accidents that have at least 1 person hospitalized?

   (b) What is the expected number of people who are seriously injured as a result of accidents at this intersection over a 30 day period?

   (c) Given that the first 10 days of a 30 day period had 1 accident with no serious injuries, and 1 accident with 2 people hospitalized, what is the expected number of people with serious injuries over the 30 day period?

4. A Markov chain with states $1, 2, 3, 4$ has the following transition probability matrix,

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.3 & 0.5 & 0.2 & 0 \\ 0 & 0.3 & 0.5 & 0.2 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}.$$

(a) Given that the chain is currently in state 2, what is the probability it has not been absorbed into state 1 after 2 transitions?

(b) Given that the chain is currently in state 4, what is the probability it is absorbed into state 1 after 3 transitions?

(c) What is the mean time to absorption given the chain starts in state 4?

Solutions for Homework 4

1. 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 $E(N(t)|T_2 = s)$, for $0 < s < t < \infty$.
   Since $t > s$, then we can perform a renewal at time s. Let $N^*(t)$ denote the renewed process. Then

   $$N(t) = 2 + N^*(t - s),$$

   and so

   $$E(N(t)|T_2 = s) = 2 + E[N^*(t - s)] = 2 + \lambda(t - s).$$

   (b) Find $E(T_3|N(t) = 1)$.
   Since $T_3 > t$, perform a renewal at t. Then T_3 is the second arrival of the renewed process and so

   $$E(T_3|N(t) = 1) = t + E(T_2^*) = t + 2/\lambda.$$

   (c) Find $E[N(s)|N(t) = r]$ for $0 < t < s < \infty$.
   Perform a renewal at t. Then

   $$N(s) = r + N^*(s - t)$$

   and so

   $$E[N(s)|N(t) = r] = r + E[N^*(s - t)] = r + \lambda(s - t).$$

2. A communications center has one link that is busy 75% of the time. Suppose the center polls this link every 10 minutes to determine whether or not the link is free, and suppose whether or not it is free at one time is independent of whether or not it is free any other time.

   (a) What is the probability that the link will be busy each of the first 4 times it is polled?
   Let S_n denote the number of times the link is free after n pollings. Then this is a Bernoulli process with $p = .25$.

   $$P(S_4 = 0 = .75^4 = 0.3164$$

(b) What is the probability that it will take more than 5 pollings before finding 2 free times?

This event is equivalent to the event $\{S_5 \le 1\}$. Then

$$P(S_5 \le 1) = .75^5 + 5(.25)(.75^4) = 0.6328.$$

(c) Given that the link was free 2 times during the first 10 pollings, what is the probability that the link was busy at the first 5 pollings?

$$\begin{aligned}
P(S_5 = 0|S_{10} = 2) &= \frac{P(S_5 = 0,\ S_{10} = 2)}{P(S_{10} = 2)} \\
&= \frac{P(S_5 = 0,\ S_{10} - S_5 = 2)}{P(S_{10} = 2)} \\
&= \frac{P(S_5 = 0)P(S_{10} - S_5 = 2)}{P(S_{10} = 2)} \\
&= \frac{.75^5(10)(.25^2)(.75^3)}{(45)(.25^2)(.75^8)} \\
&= 2/9.
\end{aligned}$$

3. Accidents occur at a particular intersection according to a Poisson process with a mean rate of 1 per 10 days. Suppose that 60% of these accidents result in no serious injuries (injuries that require hospitalization), 20% result in 1 person who was seriously injured, 15% result in 2 people who are seriously injured, 3% result in 3 people who are hospitalized, and the rest result in 4 people who are seriously injured (note: a fatality is counted as a serious injury).

(a) What is the mean time between accidents that have at least 1 person hospitalized?

Let $N_k(t)$ denoe the number of accidents at that intersection with k people seriously injured up to time t, for $k = 0, 1, 2, 3, 4$. Then $N_k(t)$ is a Poisson process with intensity λ_k where

$$\lambda_0 = 1.8,\ \lambda_1 = 0.6,\ \lambda_2 = 0.45,\ \lambda_3 = 0.09,\ \lambda_4 = 0.06$$

per 30 days. Let $N_{1+}(t)$ denote accidents with at least one serious injury. This is a superposition of $N_1, N_2, N_3, N - 4$ and so is a Poisson process with intensity 1.2 per 30 days. Therefore, mean time between such accidents is $30/1.2 = 25$ days.

(b) What is the expected number of people who are seriously injured as a result of accidents at this intersection over a 30 day period?

$$E[X(30)] = E[N_1(30) + 2N_2(30) + 3N_3(30) + 4N_4(30)] = .6 + (2)(.45) + (3)(.09) + (4)(.0$$

(c) Given that the first 10 days of a 30 day period had 1 accident with no serious injuries, and 1 accident with 2 people hospitalized, what is the expected number

of people with serious injuries over the 30 day period?
Perform a renewal at 10 days. Then

$$E[X(30)|X(10) = 2] = 2 + E[X^*(20)] = 2 + 1.34 = 3.34.$$

4. A Markov chain with states $1, 2, 3, 4$ has the following transition probability matrix,

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.3 & 0.5 & 0.2 & 0 \\ 0 & 0.3 & 0.5 & 0.2 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}.$$

   (a) Given that the chain is currently in state 2, what is the probability it has not been absorbed into state 1 after 2 transitions?
   $P^2[2,1] = .45$, and so this probability is $1 - .45 = .55$.

   (b) Given that the chain is currently in state 4, what is the probability it is absorbed into state 1 after 3 transitions?
   $P^3[4,1] = 0.045$

   (c) What is the mean time to absorption given the chain starts in state 4?

$$Q = \begin{bmatrix} 0.5 & 0.2 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

$$S = (I - Q)^{-1} = \begin{bmatrix} 3.33 & 2.22 & 0.89 \\ 3.33 & 5.56 & 2.22 \\ 3.33 & 5.56 & 4.22 \end{bmatrix}$$

   Therefore, mean time to absorption from state 4 is 13.11.

Homework 5

1. Suppose that 98% of a very large population of SIMMs are good and you randomly select SIMMs one at a time, inspect each, and discard any SIMM that is not good.
   (a) What is the expeced number of good SIMMs you have passed when you find the first bad SIMM?
   (b) If you have just discarded your 2^{nd} bad SIMM, what is the expected number of good SIMMs that you have passed?
   (c) Given that you have inspected 20 SIMMs and found that there was 1 bad SIMM, what is the probability that the first 5 SIMMs were good?

2. Let $\{N(t),\ t \geq 0\}$ be a Poisson process with intensity λ. Let $0 < t < \infty$. Find
   (a) $P(N(s) = 1|N(t) = 1)$, for any $0 < s < \infty$.

(b) $E(N(s)|N(t) = 2)$, for any $0 < s < \infty$.
(c) $E(T(5)|N(t) = 2)$.
(d) $P(T_1 \leq s|N(t) = 2)$, for $0 < s < t$.

3. 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 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?

4. 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?

5. 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 and is independent of whether or not any other customer has a coupon.
(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 the amounts customers pay are independent and the amounts are independent of the arrival of customers. Also suppose the amount a full-paying customer pays has mean 60, s.d. 20, and the amount a customer with a coupon pays has mean 40, s.d. 10. What is the expected value and s.d. of the total amount payed by all customers who arrive during 1 hour?
(d) Suppose 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?

6. A commercial web site offers 3 products for sale on its site. During the day, hits on this

site occur as a Poisson process with mean 20/hour. 40% of hits result in a purchase. Of those that result in a purchase, 50% purchase Item1, 40% purchase Item2, the rest purchase Item3, these decisions are independent and are independent of the arrival process. Also, the prices of these items are, Item1: \$100, Item2: \$200, Item3: \$500.
(a) What is the average total sales per hour?
(b) What is the mean time between sales of Item3?
(c) Given that there were 3 purchases of Item3 during the first hour, what is the expected total sales during a 6 hour period? Note: this does not imply no Item1 or Item2 purchases were made during the first hour, just that we aren't given those values, we only are only given the number of Item3 purchases.

7. Network logs track the origin of hits on a faculty web site. They show that hits from inside UTD arrive according to a Poisson process with mean 30/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, 50% 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 there were 50 hits over 2 days from inside UTD that accessed problem solutions, what is the expected total number of hits over 2 days? If a hit did not access problem solutions, what is the probability it came from outside UTD?

Due date: 4/19/2017