UC Berkeley
Department of Electrical Engineering and Computer Science

EE 126: PROBABLITY AND RANDOM PROCESSES

Solutions 9
Spring 2006

Exam format

- The second midterm will be held on Wednesday May 17; CHECK the final exam schedule. You are permitted to bring a calculator, and three sheets of hand-written notes on $8.5 \times 11''$ paper.
- Questions on the exam can be based on any material from Chapters 1 through to the END of chapter 7, as discussed in lectures #1 through #30, covered in homeworks #1 through #11, and all discussion sections.

Review problems

Problem 9.1

Let X and Y be independent random variables that are uniformly distributed on $[0, 1]$.

(a) Find the mean and variance of $X - 2Y$.

(b) Let A be the event $\{X \leq Y\}$. Find the conditional PDF of X given that A occurred. (A fully labeled sketch will suffice.)

Solution:

(a) We have that

$$E[X - 2Y] = E[X] - 2E[Y] = \frac{1}{2} - 2\frac{1}{2} = -\frac{1}{2}.$$

and

$$\text{Var}[X - 2Y] = \text{Var}[X] + 4\text{Var}[Y] = \frac{1}{12} + 4\frac{1}{12} = \frac{5}{12}.$$

(b) We are asked to compute

$$\begin{aligned}
f_{X|X\leq Y}(x) &= \frac{d}{dx}\left[P_{X|X\leq Y}(X \leq x | X \leq Y)\right] \\
&= \frac{d}{dx}\left[\frac{P(X \leq x, X \leq Y)}{P(X \leq Y)}\right] \\
&= \frac{d}{dx}\left[\frac{x - \frac{x^2}{2}}{\frac{1}{2}} 1[0 \leq x \leq 1]\right] \\
&= 2(1 - x)1[0 \leq x \leq 1]
\end{aligned}$$

Problem 9.2

Male and female patients arrive at an emergency room according to independent Poisson processes, with each process having a rate of 3 per hour. Let $M_{t,t'}$ be the number of male arrivals between t and t'. Let $F_{t,t'}$ be the number of female arrivals between time t and t'.

(a) Write down the PMF of $M_{3,5} + F_{4,6}$.

(b) Calculate the variance of $M_{3,5} + M_{4,6}$.

(c) What is the expectation of the arrival time of the *last* patient to arrive *before* 4 p.m.?

(d) Starting from a particular time, what is the expected time until there is an arrival of at least one male and at least one female patient?

Solution:

(a) Since $M_{3,5}$ and $F_{4,6}$ are independent Poisson random variables with parameter $\lambda t = 6$, $M_{3,5} + F_{4,6}$ is Poisson with parameter 12

(b) $M_{3,5} + M_{4,6} = M_{3,4} + 2M_{4,5} + M_{5,6} = X + 2Y + Z$, where X,Y and Z are independent random variables with parameters 3. So, $\text{Var}[M_{3,5}+M_{4,6}] = \text{Var}[X]+4\text{Var}[Y]+\text{Var}[Z] = 3(1 + 4 + 1) = 18$.

(c) We are asked to compute the expected time of the first arrival in the merged process going backwards in time. The merged process has parameter 6, so the expectation of the arrival time of the *last* patient to arrive *before* 4 p.m. is 4 p.m. - $\frac{1}{6}$ hours.

(d) Let X_m and X_f be the interarrival times of male and female patients processes. X_m and X_f are independent r.vs exponentially distributed with parameter 6. We are asked to compute $E[Z] = E[\max X_m, X_f] = \frac{1}{2}$. The calculation is done in more generality in Problem 9.8 (d).

Problem 9.3

Consider a factory that produces $X_n \geq 0$ gadgets on day n. The X_n are independent identically distributed random variables with mean 5, variance 9, $\mathbb{E}[X_n^3] = 412$, and $\mathbb{E}[X_n^4] < \infty$. Furthermore, we are told that $\mathbb{P}(X_n = 0) > 0$.

(a) Find an approximation to the probability that the total number of gadgets produced in 100 days is less than 440.

(b) Find (approximately) the largest value of n such that $\mathbb{P}(X_1+\cdots+X_n \geq 200+5n) \leq 0.05$.

(c) For each definition of Z_n given below, state whether the sequence Z_n converges in probability. (Answer yes or no, only.)

  (a) $Z_n = \frac{X_1+\cdots+X_n}{n}$.

  (b) $Z_n = \frac{X_1+\cdots+X_n-5n}{\sqrt{n}}$.

  (c) $Z_n = \frac{X_1^2+\cdots+X_n^2}{n}$.

(d) $Z_n = X_1 X_2 \cdots X_n$.

Solution:

(a) $\Phi\left(\frac{439.5-500}{30}\right) = 1 - \Phi\left(\frac{500-439.5}{30}\right) = 0.0217$

(b) $\mathbf{P}\left(\frac{X_1+..+X_n-5n}{3\sqrt{n}} > \frac{199.5}{3\sqrt{n}}\right) = 1 - \Phi\left(\frac{199.5}{3\sqrt{n}}\right) \leq 0.05 \quad \Rightarrow \quad n \leq 1624$

(c) $\mathbf{P}(N \geq 220) = \mathbf{P}(S_{219} \leq 1000) = \mathbf{P}\left(\frac{S_{219}-5(219)}{3\sqrt{219}} \leq \frac{1000.5-5(219)}{3\sqrt{219}}\right) = \Phi\left(\frac{-94.5}{3\sqrt{219}}\right) =$ 0.0166

1. (a) Yes, to $\mathbf{E}[X]$ (SLLN).
   (b) No.
   (c) Yes, to $\mathbf{E}[X^2]$ (SLLN).
   (d) Yes, to 0.

Problem 9.4

Consider the Markov chain specified by the following state transition diagram. Let X_n be the value of the state at time n.

(a) For all states $i = 0, 1, \ldots, 6$, find μ_i the expected time to absorption starting from state i, (i.e. entering a recurrent state.)

(b) Find $\lim_{n\to\infty} \mathbb{P}(X_n = 2 | X_0 = 3)$.

Solution:

(a) By definition, $\mu_i = 0$ for $i = 0, 1, 2, 5, 6$. In fact, states 0,1,2,5,6 are recurrent. Finally, u_3 and u_4 can be obtained by solving the system of equations

$$\begin{aligned} u_3 &= 1 + \frac{1}{2} u_4 \\ u_4 &= 1 + \frac{2}{10} u_4 \end{aligned}$$

from which $u_3 = \frac{5}{3}$ and $u_4 = \frac{4}{3}$.

(b) Fix $0, 1, 2$ as absorbing states, consider the absorption probability a_i that $0, 1, 2$ are eventually reached starting from state i.

We have that $\lim_{n\to\infty} \mathbb{P}(X_n = 2 | X_0 = 3) = \pi_2 a_3$.

where π_2 is the steady state distribution of the recurrent class $0, 1, 2$. Solving the system of equations

$$\begin{aligned} \pi_1 &= \frac{1}{2} \pi_0 \\ \pi_2 &= \frac{1}{2} \pi_1 \\ \pi_0 + \pi_1 + \pi_2 &= 1 \end{aligned}$$

we get $\pi_2 = \frac{1}{7}$. And a_3 can be obtained solving this system of equations

$$\begin{aligned} a_3 &= \frac{1}{2}a_4 + \frac{1}{2} \\ a_4 &= \frac{2}{10}a_3 \end{aligned}$$

from which, $a_3 = \frac{5}{9}$. Hence, $\lim_{n\to\infty} \mathbb{P}(X_n = 2|X_0 = 3) = \frac{1}{7}\frac{5}{9}$

Problem 9.5

The random variable X is distributed as a binomial random variable with parameters $n > 1$ and $0 < p < 1$. The pair of random variables (Y, Z) takes values on the set

$$\{(0,1),(1,0),(0,-1),(-1,0)\}$$

with equal probability if $X \leq n/2$. Similarly, (Y, Z) takes values on the set

$$\{(1,1),(1,-1),(-1,1),(-1,-1)\}$$

with equal probability if $X > n/2$.

(a) Are X and Y independent?

(b) Are Y and Z independent?

(c) Conditioned on X being even, are Y and Z independent?

(d) Conditioned on $Y = 1$, are X and Z independent?

(e) Conditioned on $Z = 0$, are X and Y independent?

Solution:

(a) No. Suppose n even.

$$P(Y=1) = P(Y=1|X<n/2)P(X<n/2)+P(Y=1|X\geq n/2)P(X\geq n/2) = \frac{1}{2}\frac{1}{4}+\frac{1}{2}\frac{1}{2} = \frac{3}{8}$$

But $P(Y = 1|X = 1) = \frac{1}{4}$.

(b) No. Suppose n even.

$$P(Z=1) = P(Z=1|X<n/2)P(X<n/2)+P(Z=1|X\geq n/2)P(X\geq n/2) = \frac{1}{2}\frac{1}{4}+\frac{1}{2}\frac{1}{2} = \frac{3}{8}$$

But $P(Z = 1|Y = 1) = \frac{1}{2}$.

(c) No. Suppose n even, such that $P(X < n/2|X \text{ even}) = P(X \geq n/2|X \text{ even}) = 1/2$. Then,

$$\begin{aligned}
P(Y = 1|X \text{ even}) &= P(Y = 1|X < n/2, X \text{ even})P(X < n/2|X \text{ even}) \\
&\quad + P(Y = 1|X \geq n/2, X \text{ even})P(X \geq n/2|X \text{ even}) \\
&= P(Y = 1|X < n/2)P(X < n/2|X \text{ even}) \\
&\quad + P(Y = 1|X \geq n/2)P(X \geq n/2|X \text{ even}) \\
&= \frac{1}{2}\frac{1}{4} + \frac{1}{2}\frac{1}{2} = \frac{3}{8}
\end{aligned}$$

But $P(Y = 1|X \text{ even}, Z = 1) = \frac{1}{2}$.

(d) No.

$$P(X > n/2|Y = 1) = \frac{P(Y = 1|X > n/2)P(X > n/2)}{P(Y = 1)} = \frac{2}{3}$$

However, $P(X > n/2|Y = 1, Z = 0) = 0$.

(e) No. $P(Y = 1|Z = 0, X > n/2) = 0$. However, $P(Y = 1|Z = 0) > 0$.

Problem 9.6

Consider $N+1$ independent Poisson arrival processes, such that the i^{th} process has arrival rate $\lambda_i = i\lambda$, for $i = 1, 2, \ldots, (N + 1)$. Let N be a binomial random variable, independent of all the Poisson arrival processes; $\mathbb{E}[N] = \mu$ and $\text{var}(N) = v$.

(a) Assume that N represents the number of successes in n independent Bernoulli trials, each with a probability p of success. Find n and p in terms of μ and v.

(b) Find, in terms of λ, μ and v, the mean of the number of total arrivals from the sum of the processes, in a time interval of length t.

Solution:

(a) Since $\mu = np$ amd $v = mp(1 - p)$, we get $n = \frac{\mu^2}{\mu - v}$ and $p = \frac{\mu - v}{\mu}$.

(b) The total arrivals from the sum of the processes can be written ad $Z = \sum_{i=1}^{N+1} Y_i$, where Y_i is a Poisson distribution with parameter $\lambda i t$. From the law of total expectation, we have

$$\begin{aligned}
E[Z] &= E[E[Z|N]] \\
&= \sum_{j=0}^{n} P(N=j) E\left[\sum_{i=1}^{j+1} Y_i\right] \\
&= \sum_{j=0}^{n} P(N=j) \sum_{i=1}^{j+1} E[Y_i] \\
&= \sum_{j=0}^{n} \binom{n}{j} p^j (i-p)^{n-j} \sum_{i=1}^{j+1} \lambda i t \\
&= \lambda t \sum_{j=0}^{n} \binom{n}{j} p^j (i-p)^{n-j} \frac{(j+1)(j+2)}{2} \\
&= \lambda t \sum_{j=0}^{n} \binom{n}{j} p^j (i-p)^{n-j} \frac{j^2+3j+2}{2} \\
&= \lambda t \left[\frac{E[N^2]}{2} + \frac{3E[N]}{2} + 2\right] \\
&= \lambda t \left[\frac{np(1-p+np)}{2} + \frac{3np}{2} + 2\right]
\end{aligned}$$

Problem 9.7
Let $X_1, X_2, \ldots$ be an i.i.d. sequence of normal random variables with mean μ and variance σ^2. Further, let $Y_n = \frac{\sum_{i=1}^{n} X_i}{\sqrt{n}}$ for $n = 1, 2, \ldots$.

(a) Using the Chebyshev inequality, give an upper bound for $\mathbb{P}(|Y_n - \mathbb{E}[Y_n]| \geq \epsilon)$.

(b) Evaluate $\mathbb{P}(|Y_n - \mathbb{E}[Y_n]| \geq \epsilon)$ exactly in terms of Φ, the CDF of the standard normal random variable.

(c) Hence compute $\mathbb{P}(|Y_n - \mathbb{E}[Y_n]| \geq \epsilon)$ and its Chebyshev upper bound for $\epsilon/\sigma = 0.5$, $\epsilon/\sigma = 1.0$, and $\epsilon/\sigma = 2.0$.

(d) For $n > k$, find the linear least squares estimate of Y_n given $Y_k = y$ and its mean squared error.

Solution:

(a) We have

$$\text{var}(Y_n) = \frac{\text{var}\left(\sum_{i=1}^{n} X_i\right)}{n} = \frac{n\sigma^2}{n} = \sigma^2.$$

Hence, by the Chebyshev inequality,

$$P(|Y_n - E[Y_n]| \geq \epsilon) \leq \frac{\text{var}(Y_n)}{\epsilon^2} = \frac{\sigma^2}{\epsilon^2}.$$

(b) Since Y_n is a linear function of independent normal random variables, it is normal. Hence

$$P(|Y_n - E[Y_n]| \geq \epsilon) = P\left(\frac{|Y_n - E[Y_n]|}{\sqrt{\mathrm{var}(Y_n)}} \geq \frac{\epsilon}{\sigma}\right) = 2\left(1 - \Phi\left(\frac{\epsilon}{\sigma}\right)\right).$$

(c)

ϵ/σ	$P(\|Y_n - E[Y_n]\| \geq \epsilon)$	Chebyshev bound
0.5	$2(1-\Phi(0.5)) = 2(1-0.6915) = 0.6170$	4.0
1.0	$2(1-\Phi(1.0)) = 2(1-0.8413) = 0.3174$	1.0
2.0	$2(1-\Phi(2.0)) = 2(1-0.9772) = 0.0456$	0.25

(d) We have

$$Y_n = \frac{\sqrt{k}Y_k + \sum_{i=k+1}^{n} X_i}{\sqrt{n}}.$$

Therefore, the least squares estimate of Y_n given $Y_k = y$ is given by

$$E[Y_n|Y_k = y] = E\left[\left.\frac{\sqrt{k}Y_k + \sum_{i=k+1}^{n} X_i}{\sqrt{n}}\right| Y_k = y\right] = \frac{\sqrt{k}y + (n-k)\mu}{\sqrt{n}},$$

which, being a linear function of y, must be equal to the linear least squares estimate. Alternatively, we compute

$$\mathrm{cov}(Y_n, Y_k) = \mathrm{cov}\left(\sqrt{\frac{k}{n}}Y_k, Y_k\right) + \mathrm{cov}\left(\frac{\sum_{i=k+1}^{n} X_i}{\sqrt{n}}, Y_k\right) = \mathrm{cov}\left(\sqrt{\frac{k}{n}}Y_k, Y_k\right).$$

Now,

$$\mathrm{cov}\left(\sqrt{\frac{k}{n}}Y_k, Y_k\right) = E\left[\left(\sqrt{\frac{k}{n}}Y_k - E\left[\sqrt{\frac{k}{n}}Y_k\right]\right)(Y_k - E[Y_k])\right] = \sqrt{\frac{k}{n}}\,\mathrm{var}(Y_k).$$

Hence the linear least squares estimate of Y_n given $Y_k = y$ is given by

$$E[Y_n] + \frac{\mathrm{cov}(Y_n, Y_k)}{\mathrm{var}(Y_k)}(y - E[Y_k]) = \sqrt{n}\mu + \sqrt{\frac{k}{n}}(y - \sqrt{k}\mu) = \frac{\sqrt{k}y + (n-k)\mu}{\sqrt{n}}.$$

The mean square error of the estimate is

$$E\left[\left.\left(\frac{\sqrt{k}Y_k + \sum_{i=k+1}^{n} X_i}{\sqrt{n}} - \frac{\sqrt{k}y + (n-k)\mu}{\sqrt{n}}\right)^2\right| Y_k = y\right] = \frac{\sum_{i=k+1}^{n}\mathrm{var}(X_i)}{n} = \left(1 - \frac{k}{n}\right)\sigma^2.$$

Problem 9.8

The BART is broken, so that only two kinds of vehicles go from Berkeley to San Francisco: taxis and buses. The interarrival time of taxis, in minutes, is an independent exponential random variable with parameter λ_1, i.e. its PDF is $f_{I_T}(t) = \lambda_1 e^{-\lambda_1 t}$ for $t \geq 0$, while the interarrival time of buses, in minutes, is an independent exponential random variable with parameter λ_2, i.e., its PDF is $f_{I_B}(t) = \lambda_2 e^{-\lambda_2 t}$ for $t \geq 0$.

Suppose Joe and Harry arrive at Berkeley at 7:00 a.m.

(a) What is the average time before they see the first vehicle?

(b) What is the probability that the first vehicle they see is a bus, and what is the probability that the first vehicle they see is a taxi?

In a taxi, the travel time to San Francisco, in minutes, is an independent exponential random variable with parameter μ_1, i.e., its PDF is $f_{D_T}(t) = \mu_1 e^{-\mu_1 t}$ for $t \geq 0$. On the other hand, in a bus, the travel time to San Francisco, in minutes, is an an independent exponential random variable with parameter μ_2, i.e., its PDF is $f_{D_B}(t) = \mu_2 e^{-\mu_2 t}$ for $t \geq 0$.

(c) Suppose Joe and Harry arrive at downtown Berkeley at 7:00 a.m., take the first vehicle that passes, and arrive in San Francisco X minutes later. Find the transform of X.

(d) Suppose that a taxi and a bus arrive simultaneously, and Joe takes the taxi while Harry takes the bus. Let Y be the number of minutes from their departure from Berkeley till they meet again in San Francisco. Find $Exs[Y]$.

There are, in fact, two different kinds of buses: fast buses and slow buses. For any bus that arrives at Berkeley downtown, it is a fast bus with probability p, and it is a slow bus with probability $1-p$. Whether a bus is fast or slow is independent of everything else.

(e) If they stay at Berkeley for l minutes, how many fast buses will they see on average?

(f) If they stay indefinitely, what is the probability that they will see k fast buses before they see k slow buses?

Solution:

(a) Vehicles arrive according to a Poisson process with rate $\lambda_1 + \lambda_2$. Owing to the memorylessness of the Poisson process, the time until the first vehicle after 7:00 a.m. is distributed as an exponential random variable with parameter $\lambda_1 + \lambda_2$. Therefore, the average time before they see the first vehicle is $1/(\lambda_1 + \lambda_2)$.

(b) By the properties of merged Poisson processes, we have

$$P(\text{first vehicle is a taxi}) = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

and

$$P(\text{first vehicle is a bus}) = \frac{\lambda_2}{\lambda_1 + \lambda_2}.$$

(c) We have $X = W + D$, where W is the time until the arrival of the first vehicle, and D is the travel time. Now,

$$M_W(s) = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 - s},$$

since, as established in part (a), W is an exponential random variable with parameter $\lambda_1+\lambda_2$. As for D, we have

$$\begin{aligned}
M_D(s) &= E\left[e^{sD}\right] \\
&= E\left[e^{sD}\middle|\,\text{first vehicle is a taxi}\right]\frac{\lambda_1}{\lambda_1+\lambda_2} + E\left[e^{sD}\middle|\,\text{first vehicle is a bus}\right]\frac{\lambda_2}{\lambda_1+\lambda_2} \\
&= M_{D_T}(s)\frac{\lambda_1}{\lambda_1+\lambda_2} + M_{D_B}(s)\frac{\lambda_2}{\lambda_1+\lambda_2} = \frac{\mu_1}{\mu_1-s}\frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\mu_2}{\mu_2-s}\frac{\lambda_2}{\lambda_1+\lambda_2}.
\end{aligned}$$

Therefore,

$$M_X(s) = M_W(s)M_D(s) = \frac{\lambda_1+\lambda_2}{\lambda_1+\lambda_2-s}\left(\frac{\mu_1}{\mu_1-s}\frac{\lambda_1}{\lambda_1+\lambda_2} + \frac{\mu_2}{\mu_2-s}\frac{\lambda_2}{\lambda_1+\lambda_2}\right).$$

(d) We have $Y=\max(D_T,D_B)$. We write $Y=Y_1+Y_2$, where $Y_1=\min(D_T,D_B)$ denotes the time till the arrival of the first vehicle, and $Y_2=Y-Y_1$ denotes the time from the arrival of the first vehicle till the arrival of the second vehicle. First, we have

$$E[Y_1] = \frac{1}{\mu_1+\mu_2}.$$

Then

$$\begin{aligned}
E[Y_2] &= E[Y_2|D_T<D_B]P(D_T<D_B) + E[Y_2|D_T\ge D_B]P(D_T\ge D_B) \\
&= \frac{1}{\mu_2}\frac{\mu_1}{\mu_1+\mu_2} + \frac{1}{\mu_1}\frac{\mu_2}{\mu_1+\mu_2}.
\end{aligned}$$

So

$$E[Y] = \frac{1}{\mu_1+\mu_2}\left(1+\frac{\mu_1}{\mu_2}+\frac{\mu_2}{\mu_1}\right).$$

(e) Fast buses arrive according to a Poisson process with rate $p\lambda_2$. Hence the average number of fast buses they will see is $lp\lambda_2$.

(f) The probability that exactly i slow buses arrive before the arrival of the kth fast bus is the probability that the k fast bus is the $(k+i)$th bus, which, using the Pascal PMF, is given by

$$\binom{k+i-1}{k-1}p^k(1-p)^i.$$

The probability that they see k fast buses before they see k slow buses is the probability that strictly less than k slow buses arrive before the arrival of the kth fast bus, which is therefore given by

$$\sum_{i=0}^{k-1}\binom{k+i-1}{k-1}p^k(1-p)^i.$$

Alternatively, the probability of seeing k fast buses before k slow buses is equivalent to the probability of seeing k or more fast buses in the first $2k-1$ buses, which is given by

$$\sum_{i=k}^{2k-1}\binom{2k-1}{i}p^i(1-p)^{2k-1-i}.$$

Problem 9.9

A hungry mouse is trapped in a cage with three doors. At each "turn", the mouse gets to open one of the three doors and eat a piece of cheese behind the door if one is present. Each door is chosen w ith equal probability on each turn, regardless of whether a piece of cheese was found on the previous turn. If no cheese was found on the previous turn, there is a probability of 3/4 that cheese will be found behind each door on the current turn. If cheese was found on the previous turn and the same door is chosen on the current turn, then there is a probability of 0 that cheese will be found; whilst if cheese was found on the previous turn and a different door is chosen on the current turn, then there is a probability of 1 that cheese will be found.

(a) If you observe the mouse's behavior over 1000 turns, in approximately what fraction of turns do you expect the mouse to eat a piece of cheese?

(b) Suppose no cheese was found on the previous turn. What is the expected number of turns before the mouse eats a piece of cheese?

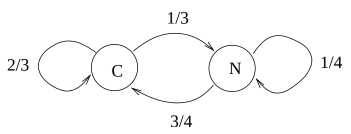
(c) Suppose no cheese was found on the previous turn. What is the expected number of turns before the mouse eats n pieces of cheese?

(d) Suppose cheese was found on the previous turn. Using the Central Limit Theorem, approximate the probability that the number of turns before the mouse eats 100 pieces of cheese exceeds 152.

(e) You look into the cage and observe the mouse eating a piece of cheese from behind door number 1. What is the probability that, if you observe the mouse three turns later, it will again be eating a piece of cheese from behind door number 1?

Solution:

(a) We model the problem with the following two-state Markov chain, where C represents the state where cheese is found, and N represents the state where no cheese is found.



Therefore, the steady-state probabilities satisfy

$$\frac{1}{3}\pi_C = \frac{3}{4}\pi_N$$

and

$$\pi_C + \pi_N = 1,$$

whence we conclude that $\pi_N = 4/13$ and $\pi_C = 9/13$. Thus, over 1000 turns, we expect that the mouse eats a piece of cheese in approximately 9/13 of them.

(b) We wish to know the mean first passage time t_N to reach state C from state N. We have

$$t_N = 1 + \frac{1}{4} t_N,$$

which implies that $t_N = 4/3$.

(c) We wish to know the mean first passage time to reach state C from state N followed by $(n-1)$ recurrence times of state C. The mean recurrence time t_C^* of state C is given by

$$t_C^* = 1 + \frac{1}{3} t_N = 1 + \frac{4}{9} = \frac{13}{9}.$$

Hence the expected number of turns until the mouse eats n pieces of cheese is $4/3 + (n-1)13/9$.

(d) Let Y_{100} be the number of turns before the mouse eats 100 pieces of cheese. Then, $Y_{100} = \sum_{i=1}^{100} R_i$, where R_i denotes the number of turns from eating the $(i-1)$th piece of cheese till eating the ith piece of cheese. The random variable R_i is equal to 1 with probability 2/3 (transition from C to C) and equal to $1 + S$ with probability 1/3 (transition from C to T), where S is a geometric random variable with parameter 3/4. We have already established that $E[R_i] = 13/9$. Therefore,

$$\begin{aligned} E[R_i^2] &= 1 \cdot \frac{2}{3} + E[(1+S)^2] \cdot \frac{1}{3} = \frac{2}{3} + \frac{1}{3}(1 + 2E[S] + E[S^2]) \\ &= \frac{2}{3} + \frac{1}{3}(1 + 2E[S] + \mathrm{var}(S) + (E[S])^2) = \frac{2}{3} + \frac{1}{3}\left(1 + \frac{8}{3} + \frac{1/4}{(3/4)^2} + \frac{16}{9}\right) = \frac{71}{27}. \end{aligned}$$

Hence

$$\mathrm{var}(R_i) = \frac{71}{27} - \left(\frac{13}{9}\right)^2 = \frac{44}{81}.$$

Since it is clear that $R_1, R_2, \ldots, R_{100}$ is an i.i.d. sequence, it follows by the Central Limit Theorem that

$$\begin{aligned} P(Y_{100} > 152) &= P\left(\frac{Y_{100} - E[Y_{100}]}{\sqrt{\mathrm{var}(Y_{100})}} > \frac{152 - 100 \cdot 13/9}{10 \cdot \sqrt{44/81}}\right) = P\left(\frac{Y_{100} - E[Y_{100}]}{\sqrt{\mathrm{var}(Y_{100})}} > \sqrt{\frac{289}{275}}\right) \\ &\approx 1 - \Phi(1.03) = 1 - 0.8485 = 0.1515. \end{aligned}$$

(e) We now model the problem with a four-state Markov chain. The states are $C1$, $C2$, $C3$, and N, where $C1$, $C2$, and $C3$ represent the states where cheese is found behind the first, second, and third doors, respectively, and N represents the state where no cheese is found. The transition probabilities are

$$p_{CiCj} = \begin{cases} 1/3 & \text{if } j \neq i, \\ 0 & \text{otherwise}, \end{cases} \qquad \text{for all } i, j = 1, 2, 3,$$

$$p_{CiN} = 1/3, \qquad \text{for all } i = 1, 2, 3,$$

$$p_{NCi} = 1/4, \qquad \text{for all } i = 1, 2, 3,$$

and

$$p_{NN} = 1/4.$$

There are seven three-transition paths from state $C1$ to state $C1$, which are

(i) $C1 \to C2 \to C3 \to C1$,
(ii) $C1 \to C2 \to N \to C1$,
(iii) $C1 \to C3 \to C2 \to C1$,
(iv) $C1 \to C3 \to N \to C1$,
(v) $C1 \to N \to C2 \to C1$,
(vi) $C1 \to N \to C3 \to C1$, and
(vii) $C1 \to N \to N \to C1$.

Paths (i) and (iii) occur with probability 1/27, paths (ii), (iv), (v), and (vi) occur with probability 1/36, and path (vii) occurs with probability 1/48. Therefore, the probability of going from state $C1$ to state $C1$ in three transitions is

$$\frac{2}{27} + \frac{4}{36} + \frac{1}{48} = \frac{89}{432}.$$