6.041/6.431 Spring 2008 Final Exam
Wednesday, May 21, 9:00AM - 12:00PM

DO NOT TURN THIS PAGE OVER UNTIL YOU ARE TOLD TO DO SO

Name: ____________________

Recitation Instructor: ____________________

TA: ____________________

6.041/6.431: ____________________

Question	Part	Score	Out of
0			0
1	all		30
2	a		5
	b		6
	c		5
	d		8
	e		5
	f		5
	g		5
3	a		4
	b		5
	c		5
	d		5
	e		6
	f		6
Total			100

- Write your solutions in this quiz packet, only solutions in the quiz packet will be graded.
- Question one, multiple choice questions, will receive no partial credit. Partial credit for question two and three will be awarded.
- You are allowed 3 two-sided 8.5 by 11 formula sheet plus a calculator.
- You have 180 minutes to complete the exam.
- Be neat! You will not get credit if we can't read it.
- We will send out an email with more information on how to gain access to your graded final exam.
- **Good Luck!**

Problem 0: (0 pts) Write your name, your assigned recitation instructor's name, and assigned TA's name on the cover of the quiz booklet. The Instructor/TA pairing is listed below. Also write the class you are registered for: 6.041 or 6.431.

Recitation Instructor	TA	Recitation Time
Vivek Goyal	Natasa Blitvic	10 & 11 AM
Michael Collins	Danielle Hinton	10 & 11 AM
Shivani Agarwal	Stavros Valavanis	12 & 1 PM
Dimitri Bertsekas (6.431)	Aman Chawla (6.431)	1 & 2 PM

Question 1: (30 pts) Multiple choice questions. **CLEARLY** circle the best answer for each question below. Each question is worth 3 points each, with no partial credit given.

a. n balls are randomly thrown into m urns. The expected number of empty urns is:

   (i) $n - m$
   (ii) $\boxed{m(1 - \frac{1}{m})^n}$
   (iii) $n(1 - \frac{1}{n})^m$
   (iv) $\binom{n}{m}(\frac{1}{m})^n$

   Solution: Define random variable I_i which equals 1 if the ith bin is empty, and 0 if it is not empty. The expected value of the number of empty urns is the sum of $\mathrm{E}[I_1 + \cdots + I_m]$. The probability that the ith urn is empty is the probability that each of the m balls does not fall into the ith bin, which is $(1 - \frac{1}{m})^n$. The expected value of I_i is thus $(1 - \frac{1}{m})^n$, and the expected value of $\mathrm{E}[I_1 + \cdots + I_m] = m(1 - \frac{1}{m})^n$.

b. Assume the failure time for any laptop is exponentially distributed with parameter λ. Suppose we have 100 laptops, all of which are started simultaneously and all of which are independent. Then the expected time until the $2nd$ failure is:

   (i) $\frac{2}{\lambda}$
   (ii) $\frac{1}{\lambda} + \frac{1}{2\lambda}$
   (iii) $\boxed{\frac{1}{100\lambda} + \frac{1}{99\lambda}}$
   (iv) $\frac{2}{100\lambda}$

   Solution: Recall from Quiz 2, the distribution on the minimum of 100 competing exponentials all with rate λ is exponential with rate 100λ. After the first laptop fails, there are 99 laptops still competing. By the Memoryless Property of exponentials, the expected time until the next laptop failure is $\frac{1}{99\lambda}$ and independent of the time of the first laptop failure. Consequently, the expected time until the second laptop failure is $\frac{1}{100\lambda} + \frac{1}{99\lambda}$.

c. The number of people waiting at a bank machine, N, is modeled as a Poisson random variable with parameter λ. Assume $H_0 : \lambda = 1$, and $H_1 : \lambda = 0.5$. Based on a single observation, $N = n$, we accept H_0 if and only if $n \geq 2$. The probability of false acceptance of H_0 is given by:

   (i) $1 - 1.5e^{-1}$
   (ii) $0.5(1 - e^{-0.5})$
   (iii) $\boxed{1 - 1.5e^{-0.5}}$
   (iv) $(1 - e^{-0.5})$

   Solution: The false acceptance probability is $P(n \geq 2; H_0) = 1 - P(n = 0; H_0) - P(n = 1; H_0)$. As N is Poisson, this probability is given in (iii).

d. Let $X_1, X_2, \ldots, X_n$ be independent random variables distributed uniformly over the interval $[0, 1]$. Define the random variable $R_n = \min(X_1, X_2, \ldots, X_n)$. Then:

(i) $\lim_{n\to\infty} \mathbf{E}[R_n] = c$, for some $c > 0$

(ii) $\boxed{\lim_{n\to\infty} \mathbf{E}[R_n] = 0}$

(iii) $\mathbf{E}[R_n] = 0.5 \ \ \forall\ n$.

(iv) $\lim_{n\to\infty} \mathbf{E}[R_n]$ is not defined

Solution: To find $\mathrm{E}[R_n]$, we first find $f_{R_n}(r)$ by derived distribution. $P(R_n \le r) = 1 - P(X_i \ge r, \ldots, X_n \ge r)$, which is equal to $1-(1-r)^n$ for $0 \le r \le 1$ and either 0 or 1 outside of this range. Differentiating the CDF to get the pdf, $f_{R_n}(r) = n(1-r)^{n-1}$ for $0 \le r \le 1$. To find the mean, we solve $\mathrm{E}[R_n] = \int_0^1 rn(1-r)^{n-1}dr$ via integration by parts. This yields $\mathrm{E}[R_n] = \frac{1}{n+1}$. In the limit as n tends to ∞, $\mathrm{E}[R_n]$ goes to zero.

e. A rubber coin changes its probability of a head depending on the outcome of the previous coin flip. If the previous flip is a Head, then the probability of a Head is equal to 0.8. If the previous flip is a tail, the probability of a Head is 0.2. After a very large number of flips, the probability of a Head is approximately

(i) $\boxed{0.5}$

(ii) 0.8

(iii) 0.2 after an even number of flips and 0.8 after an odd number of flips.

(iv) It cannot be determined as it depends on the outcome of the initial throw.

Solution: The rubber coin experiment can be modeled by a 2 state Markov Chain, S = {0,1}. Let state '0' correspond to a tail, and let state '1' correspond to a head. $P(X_{n+1} = 1|X_n = 1) = 0.8$, $P(X_{n+1} = 0|X_n = 0) = 0.2$, $P(X_{n+1} = 1|X_n = 0) = 0.8$, and $P(X_{n+1} = 0|X_n = 1) = 0.2$. Since this Markov chain has a single recurrent class, the steady state probabilities exist, and describe the probability of being in each state after a large number of flips. Noting the symmetry, or solving balance equations gives $\pi_1 = 0.5$.

f. You have one dollar and your friend has two dollars. You decide to play several games of chess where the loser gives the winner a dollar. You stop playing when either person has zero dollars. If the probability of you winning is 0.6, then the probability of the game terminating with you having 3 dollars is given by:

(i) 16/76

(ii) $\boxed{36/76}$

(iii) 0.36

(iv) 0.6

Solution: This experiment can be modeled by a 4 state Markov Chain S = {0,1, 2, 3}, where each state denotes the amount of money you have left. States 0 and 3 are recurrent, and states 1 and 2 are transient. The transitions probabilities between the states are $p_{00} = 1, p_{10} = .4, p_{12} = .6, p_{21} = .4, p_{23} = .6, p_{33} = 1$. We need to calculate the probability of being absorbed into state 3 given that $X_0 = 1$. Setting $a_3 = 1$, $a_0 = 0$, and solving the absorption equations $a_1 = .4a_0 + .6a_2$, and $a_2 = .4a_1 + .6a_3$ yields $a_1 = 36/76$.

g. Your email messages get classified through a spam filter. This filter sends a message to your inbox with probability p; otherwise, the message goes to your spam folder. Assume you have n total messages, and each message is independent of any other message. Let X denote the number of messages in your inbox and Y the number of messages in your spam folder. Then $\text{var}(X - Y)$ is given by:

(i) $4np(1-p) + n^2$

(ii) $2np(1-p)$

(iii) $\boxed{4np(1-p)}$

(iv) $np(1-p)$

Solution: Var(X−Y) = Var(X − (n − X)) = 4Var(X). Since X is a binomial random variable, 4Var(X) = $4np(1-p)$.

h. X_n is a Bernoulli process with probability of success equal to 0.5. The probability that the 5th success occurs in the 10th time slot is given by:

(i) $\boxed{\binom{9}{4}(0.5)^{10}}$

(ii) $\binom{9}{4}(0.5)^{9}$

(iii) $\binom{10}{4}(0.5)^{10}$

(iv) $\binom{10}{4}(0.5)^{9}$

Solution: X_n is a Pascal random variable, and we're looking for $P(X_5 = 10)$,which is given by (i).

i. X is a Gaussian random variable with mean 1 and variance 1. We observe the random variable $Y = \Theta X$ where Θ takes on the values of 1 (hypothesis H_0) and -1 (hypothesis H_1) with probabilities p and $1-p$ respectively. If we observe one value y, then the acceptance region for hypothesis H_0 that minimizes the probability of error is all y such that:

(i) $\boxed{y \geq \frac{1}{2}\log\left(\frac{1-p}{p}\right)}$

(ii) $y \geq -\frac{1}{2}\log\left(\frac{1-p}{p}\right)$

(iii) $y \geq \frac{1}{2}\log\left(\frac{p}{1-p}\right)$

(iv) $y \geq \frac{1}{4}\log\left(\frac{1-p}{p}\right)$

Solution: In this Bayesian Hypothesis Testing framework, we use the MAP decision rule to minimize the probability of error. Using Bayes Rule, the comparison of the posterior densities simplifies to comparing the ratio of the observation model under H_0 $f(y|H_0) = \frac{1}{\sqrt{2\pi}}e^{\frac{-(x+1)^2}{2}}$ and the observation model under H_1 $f(y|H_1) = \frac{1}{\sqrt{2\pi}}e^{\frac{-(x-1)^2}{2}}$ to the ratio of their priors $\frac{P(H_1)}{P(H_0)} = \frac{1-p}{p}$. Simplifying using the log function yields the decision rule which accepts observations of y for the decision boundary in (i).

j. Let $X_1, X_2, \ldots, X_{16}$ and $Y_1, Y_2, \ldots, Y_{16}$ be independent random variables, uniformly distributed on the interval $[0, 1]$. Let:

$$W = \frac{(X_1 + X_2 + \ldots X_{16}) - (Y_1 + Y_2 + \ldots Y_{16})}{16}$$

The best approximation to the quantity $\mathbf{P}(|W - \mathbf{E}[W]| < 0.001)$ is:

(i) $\Phi\left(\frac{0.001}{4\sqrt{6}}\right) - \Phi\left(\frac{-0.001}{4\sqrt{6}}\right)$

(ii) $\Phi\left(\frac{0.001}{\sqrt{6}}\right) - \Phi\left(\frac{-0.001}{\sqrt{6}}\right)$

(iii) $\boxed{\Phi\left(0.004\sqrt{6}\right) - \Phi\left(-0.004\sqrt{6}\right)}$

(iv) $\Phi\left(0.001\sqrt{6}\right) - \Phi\left(-0.001\sqrt{6}\right)$

Solution: To use the central limit theorem, the random variable W must be the sum of iid random variables, and be normalized to have unit variance. Define $Z_i = X_i - Y_i$. Var(Z) = Var(X) + Vay(Y) = $\frac{2}{12}$. Since Var(W) = Var(Z)/16, we must normalize W by $\frac{\sqrt{\frac{1}{6}}}{4}$. This normalization is shown in (iii).

Problem 2 (39 pts) Consider a Bernoulli process X_1, X_2, X_3, ... with unknown probability of success q. As usual, define the kth inter-arrival time T_k as

$$T_1 = Y_1, \qquad T_k = Y_k - Y_{k-1}, \quad k = 2, 3, \ldots$$

where Y_k is the time of the kth success. This problem explores estimation of q from observed inter-arrival times $\{t_1, t_2, t_3, \ldots\}$. Parts (a) - (d) focus on Bayesian estimation, while parts (e) - (g) focus on classical estimation.

You may find the following integral useful: For any non-negative integers k and m,

$$\int_0^1 q^k (1-q)^m dq = \frac{k!\, m!}{(k+m+1)!}$$

For parts (a) - (c) assume q is sampled from the random variable Q which is uniformly distributed over $[0, 1]$.

a. (5 pts) Compute the PMF of T_1, $p_{T_1}(t_1)$

   Solution: Using the theorem of total probability, we have

$$p_{T_1}(t) = \int_0^1 p_{T_1|Q}(t, q) f_Q(q) dq = \int_0^1 (1-q)^{t-1} q dq = \frac{1}{(t+1)t}$$

b. (6 pts) Compute the least squares estimate (LSE) of Q from the first recording T_1.

   Solution: The lease squares estimate coincides with the conditional expectation of Q given $N1$, which is derived as

$$\begin{aligned}
\mathbf{E}[Q \mid T_1 = t] &= \int_0^1 p_{Q|T_1}(q \mid t) q dq \\
&= \int_0^1 \frac{p_{T_1|Q}(t \mid q) f_Q(q)}{p_{T_1}(t)} q dq \\
&= \int_0^1 t(t+1) q (1-q)^{t-1} q dq \\
&= \int_0^1 t(t+1) q^2 (1-q)^{t-1} dq \\
&= t(t+1) \frac{2(t-1)!}{(t+2)!} \\
&= \frac{2}{t+2}
\end{aligned}$$

c. (5 pts) Compute the maximum a posteriori (MAP) estimate of Q given the k recordings, $T_1 = t_1, \ldots, T_k = t_k$.

Solution: We write the posterior probability distribution of Q given $T_1 = t_1, \ldots, T_k = t_k$

$$\begin{aligned} f_{Q,T_1,\ldots,T_k}(q \mid t_1,\ldots,t_k) &= \frac{f_Q(q)\prod_i^k P_{T_i}(T_i = t_i \mid Q = q)}{\int_0^1 f_Q(q)\prod_i^k P_{T_i}(T_i = t_i \mid Q = q)dq} \\ &= \frac{q^k(1-q)^{\sum_i^k t_i - k}}{c} \\ &= \frac{1}{c} q^k (1-q)^{\sum_i^k t_i - k}, \end{aligned}$$

where the denominator integrates out q so it could be viewed as a constant scaler c.

To maximize the above probability we set its derivative with respect to q to zero

$$kq^{k-1}(1-q)^{\sum_i^k t_i - k} - (\sum_i^k t_i - k) q^k (1-q)^{\sum_i^k t_i - k - 1} = 0,$$

or equivalently

$$k(1-q) - (\sum_i^k t_i - k)q = 0,$$

which yields the MAP estimate

$$\hat{q} = \frac{k}{\sum_{i=1}^k t_i}.$$

For this part only assume q is sampled from the random variable Q which is now uniformly distributed over $[0.5, 1]$

d. (8 pts) Find the linear least squares estimate (LLSE) of the second inter-arrival time (T_2), from the observed first arrival time ($T_1 = t_1$).

Solution: The LLSE of T_1 given T_2 is

$$\hat{T}_2 = \mathbf{E}[T_2] + \frac{cov(T_1, T_2)}{\text{var}(T_1)}(T_1 - \mathbf{E}[T_1]),$$

where the coefficients are

$$\mathbf{E}[T_1] = \mathbf{E}[T_2] = \int_{0.5}^1 2 * 1/q dq = 2\ln 2,$$

and

$$\begin{aligned} \text{var}(T_1) = \text{var}(T_2) &= \mathbf{E}\left[\text{var}(T_1 \mid Q)\right] + \text{var}\left[\mathbf{E}(T_1 \mid Q)\right] \\ &= \mathbf{E}\left[\frac{1-Q}{Q^2}\right] + \text{var}\left[\frac{1}{Q}\right] \\ &= \mathbf{E}[1/Q^2] - \mathbf{E}[1/Q] + \mathbf{E}[1/Q^2] - \mathbf{E}[1/Q]^2 \\ &= 2 - 2\ln 2 + 2 - (2\ln 2)^2 \\ &= 4 - 2\ln 2 - (2\ln 2)^2, \end{aligned}$$

and their covariance

$$\begin{aligned}\text{cov}(T_1, T_2) &= \mathbf{E}[T_1 T_2] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= \mathbf{E}\left[\mathbf{E}[T_1 T_2 \mid Q]\right] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= \mathbf{E}\left[\mathbf{E}[T_1 \mid Q]\mathbf{E}[T_2 \mid Q]\right] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= \mathbf{E}\left[1/Q^2\right] - \mathbf{E}[T_1]\mathbf{E}[T_2] \\ &= 2 - 4(\ln 2)^2\end{aligned}$$

Therefore we have derived the linear least squares estimator

$$\hat{T}_2 = 2\ln 2 + \frac{2 - 4(\ln 2)^2}{4 - 2\ln 2 - (2\ln 2)^2}(T_1 - 2\ln 2) \approx 1.543 + 0.113 T_1.$$

For the remaining parts assume q is an unknown parameter in the interval $(0, 1]$. Denote the true parameter by q^*. For the remaining parts denote by $\widehat{Q}_k$ the maximum likelihood estimate (MLE) of Q given k recordings, $T_1 = t_1, \ldots, T_k = t_k$.

e. (5 pts) Compute $\widehat{Q}_k$. Is this different from your MAP estimate of part (e)?

Solution: The likelihood function is $\prod_{i=1}^{k} P_{T_i}(T_i = t_i \mid Q = q) = q^k(1-q)^{\sum_i^k t_i - k}$, which can be maximized using the same procedure as in (c) and yields $\widehat{Q}_k = \frac{k}{\sum_{i=1}^k t_i}$. This is not different from the MAP estimate of part (e). Since the MAP estimate of part (e) is calculated using a uniform prior, the likelihood function is a 'scaled' version of posterior probability and they can be maximized at the same value of q.

f. (5 pts) Show that for all $\epsilon > 0$

$$\lim_{k\to\infty} \mathbf{P}\left(\left|\frac{1}{\widehat{Q}_k} - \frac{1}{q^*}\right| > \epsilon\right) = 0$$

Solution: Since $\dfrac{1}{\widehat{Q}_k} = \dfrac{\sum_{i=1}^k T_i}{k}$, and that each T_i is independent identically distributed, it follows that $\dfrac{1}{\widehat{Q}_k}$ is actually a sample mean estimator. The weak law of large numbers says that, when the number of samples increases to infinity, the sample mean estimator converges to the actual mean, which is $\frac{1}{q^*}$ in this case. So we can write the limit of probability as

$$\lim_{k\to\infty} \mathbf{P}\left(\left|\frac{1}{\widehat{Q}_k} - \frac{1}{q^*}\right| > \epsilon\right) = \lim_{k\to\infty} \mathbf{P}\left(\left|\frac{\sum_{i=1}^k T_i}{k} - \mathbf{E}[T_1]\right| > \epsilon\right) = 0.$$

g. (5 pts) Assume $q^* \geq 0.5$. Give a lower bound on k such that

$$\mathbf{P}\left(\left|\frac{1}{\widehat{Q}_k} - \frac{1}{q^*}\right| \leq 0.1\right) \geq 0.95$$

Solution: Chebyshev inequality states that

$$\mathbf{P}\left(\left|\frac{\sum_{i=1}^{k} T_i}{k} - \mathbf{E}[T_1]\right| \geq \epsilon\right) \leq \frac{\text{var}(T_1)}{k\epsilon^2}.$$

So we have

$$\mathbf{P}\left(\left|\frac{1}{\widehat{Q}_k} - \frac{1}{q^*}\right| \leq 0.1\right) = \mathbf{P}\left(\left|\frac{\sum_{i=1}^{k} T_i}{k} - \frac{1}{q^*}\right| \leq 0.1\right) = 1 - \mathbf{P}\left(\left|\frac{\sum_{i=1}^{k} T_i}{k} - \mathbf{E}[T_1]\right| \geq 0.1\right) \geq 1 - \frac{\text{var}(T_1)}{k * 0.1^2}$$

To ensure the above probability to be greater than 0.95, we need that

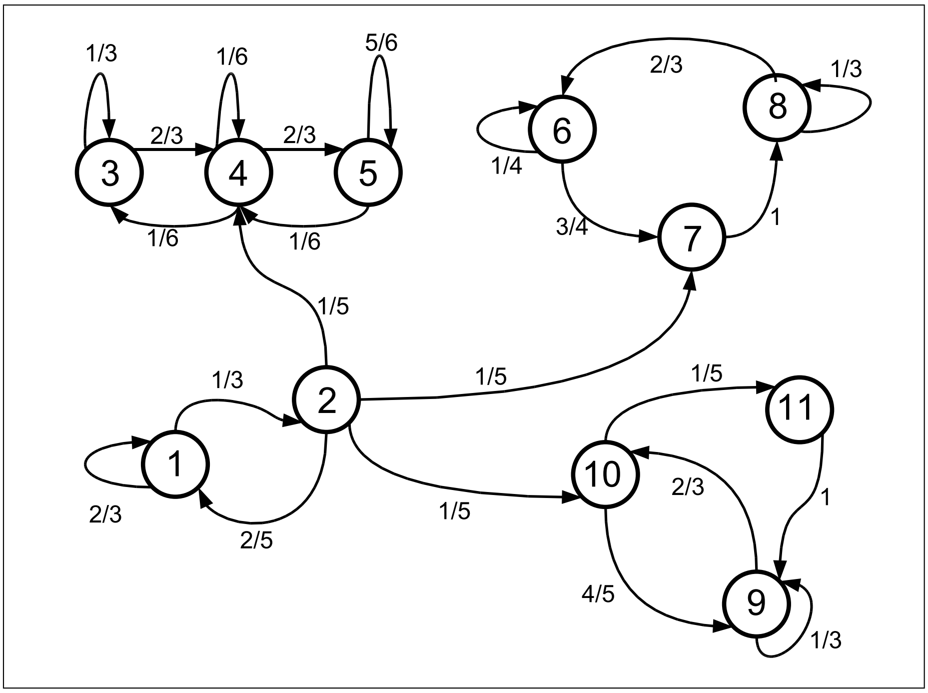
$$1 - \frac{\text{var}(T_1)}{k * 0.1^2} = 1 - \frac{\frac{1-q}{q^2}}{k * 0.1^2} \geq 0.95,$$

or

$$k \geq 2000\frac{1-q}{q^2} \geq 4000,$$

where the last inequality holds for all q ranging from 0.5 to 1.

Problem 3 (31 pts) Let $X_0, X_1, X_2, X_3 \ldots$ be the consecutive states of the Markov chain shown in the figure below. All questions may be answered independently.



a. (4 pts) Identify the set of transient states and all recurrent classes.

Solution: There are two transient states: 1 and 2, and three recurrent classes: (3,4,5),(6,7,8),(9,10,11).

b. (5 pts) Assume $X_0 = 3$. Find a numerical answer for $\mathbf{P}(X_{n-1} = 4 \mid X_n = 3)$, where n is very large.

Solution: We have

$$\begin{aligned}\mathbf{P}(X_{n-1} = 4 \mid X_n = 3) &= \frac{\mathbf{P}(X_{n-1} = 4, X_n = 3)}{\mathbf{P}(X_n = 3)} \\ &= \frac{\mathbf{P}(X_n = 3 \mid X_{n-1} = 4)\mathbf{P}(X_{n-1} = 4)}{\mathbf{P}(X_n = 3)} \\ &= \frac{1}{6} * \frac{\mathbf{P}(X_{n-1} = 4)}{\mathbf{P}(X_n = 3)}.\end{aligned}$$

Since n is very large, the effect of initial state vanishes. So we approximate $\mathbf{P}(X_{n-1} = 4)$ and $\mathbf{P}(X_n = 3)$ by their steady distribution probabilities π_4 and π_3, noting that this is a Birth Death Chain. They satisfy

$$\pi_3 = 1/3\pi_3 + 1/6\pi_4,$$

or equivalently

$$\frac{\pi_4}{\pi_3} = 4.$$

Now we have

$$\mathbf{P}(X_{n-1} = 4 \mid X_n = 3) = \frac{1}{6} * \frac{\mathbf{P}(X_{n-1} = 4)}{\mathbf{P}(X_n = 3)} = \frac{1}{6} * \pi_4/\pi_3 = \frac{2}{3}$$

c. (5 pts) Assume $X_0 = 7$. Let Y be the total number of times the process enters state 6 or state 8 before the 10th visit to state 7. Count X_0 as the first visit to state 7. Find $\mathbf{E}[Y]$ and $\text{var}(Y)$.

Solution: Let S_i be number of times the process enters state 6 or state 8, starting from the ith visit to state 7 to the $i+1$th visit to state 7. Then $Y = S_1 \ldots + S_9$. Using the memoryless property it follows that S_i are independent identically distributed random variables.

Each S_i is the sum of two independent geometric random variables, number of times entering 6, denoted by M, and number of times entering 8, denoted by N. Their probabilities of a succuss are 2/3 and 3/4 respectively. Note that this is a second arrival of a Bernoulli process. Now the expectation and variance of S_1 become

$$\mathbf{E}[S_1] = \mathbf{E}[M] + \mathbf{E}[N] = \frac{1}{2/3} + \frac{1}{3/4} = 17/6$$
$$\text{var}[S_1] = \text{var}(M) + \text{var}(N) = \frac{1-2/3}{(2/3)^2} + \frac{1-3/4}{(3/4)^2} = 43/36$$

It follows that

$$\mathbf{E}[Y] = 9\mathbf{E}[S_1] = 51/2, \qquad \text{var}[Y] = 9\text{var}[S_1] = 43/4.$$

d. (5 pts) Assume $X_0 = 11$. Let Z_j be equal to the number of transitions up to and including the jth time the process enters state 9. Find $\mathbf{E}[Z_4]$.

Solution: Using the memoryless property it follows that, $Z_4 - Z_3$, $Z_3 - Z_2$, $Z_2 - Z_1$ and Z_1 are independent, and that the first three of them are identically distributed. Therefore

$$\mathbf{E}[Z_4] = t_{11} + 3t_9^*,$$

where t_{11} is the mean first passage times to reach state 9 from state 11, and t_9^* is the mean recurrence time of state 9. We write the system of equations

$$\begin{aligned} t_9 &= 0, \\ t_{10} &= 1 + 4/5t_9 + 1/5\mu_{11}, \\ t_{11} &= 1 + 1t_9, \\ t_9^* &= 1 + 1/3t_9 + 2/3t_{10}, \end{aligned}$$

which yields

$$t_9 = 0, \quad t_{10} = 6/5, \quad t_{11} = 1, \quad t_9^* = 9/5,$$

so

$$\mathbf{E}[Z_4] = t_{11} + 3t_9^* = 32/5.$$

An alternate solution is as follows: Let N_i be the number of transitions between the $(i-1)$th and the ith trip to state 9. $N_1 = 1$, since from state 11, you must go to state 1. N_2, N_3 and N_4 are identically distributed, and can only be 1 with probability 1/3 (a self loop at state 9), or 2 with probability 8/15 (the path from 9 to 10, then 10 to 9), or 3 with probability (2/15) (the path from 9 to 10 to 11 then back to 9). E[N_i] is thus 1*(1/3) + 2*(8/15) + 3*(2/15) = 9/5. E[Z_4] = 1 + 3E[N_i] = 32/5.

e. (6 pts) Assume $X_0 = 1$. Given the process eventually reaches state 5, what is the expected number of transitions up to and including the first time the process enters state 4? Briefly explain your reasoning so that we can understand your approach.

Solution: Define A as the event that the recurrent class 3,4,5 is eventually reached, we may simplify the transition probabilities from state 2 by:

$$\begin{aligned}\mathbf{P}(X_{n+1} = 4 \mid X_n = 2, A) &= \frac{\mathbf{P}(X_{n+1} = 4, A \mid X_n = 2)}{\mathbf{P}(A \mid X_n = 2)} \\ &= \frac{\mathbf{P}(X_{n+1} = 4 \mid X_n = 2)}{\mathbf{P}(A \mid X_n = 2)} \\ &= \frac{p_{24}}{\frac{p_{24}}{p_{24}+p_{27}+p_{2,10}}} \\ &= \frac{1/5}{\frac{1/5}{1/5+1/5+1/5}} \\ &= 3/5,\end{aligned}$$

where $\mathbf{P}(A \mid X_n = 2)$ equals one third because all three absorption classes are equal likely to be reached. Therefore we view all three recurrent classes as one single absorption state s and $p_{2s} = 3/5$. All we need is the expected number of transitions until absorption. We write the system of equations

$$\begin{aligned}\mu_s &= 0, \\ \mu_1 &= 1 + 2/3\mu_1 + 1/3\mu_2, \\ \mu_2 &= 1 + 3/5\mu_s + 2/5\mu_1,\end{aligned}$$

which yields

$$\mu_1 = 20/3.$$

This question does not rely on the above figure. A friend offers you a game of chance: You flip a coin until you get two consecutive heads (i.e. 2 heads in a row; e.g. THTTTHTHH, HTHTTHH, etc.). For each flip of the coin your friend will pay you \$2. Let W denote your winnings.

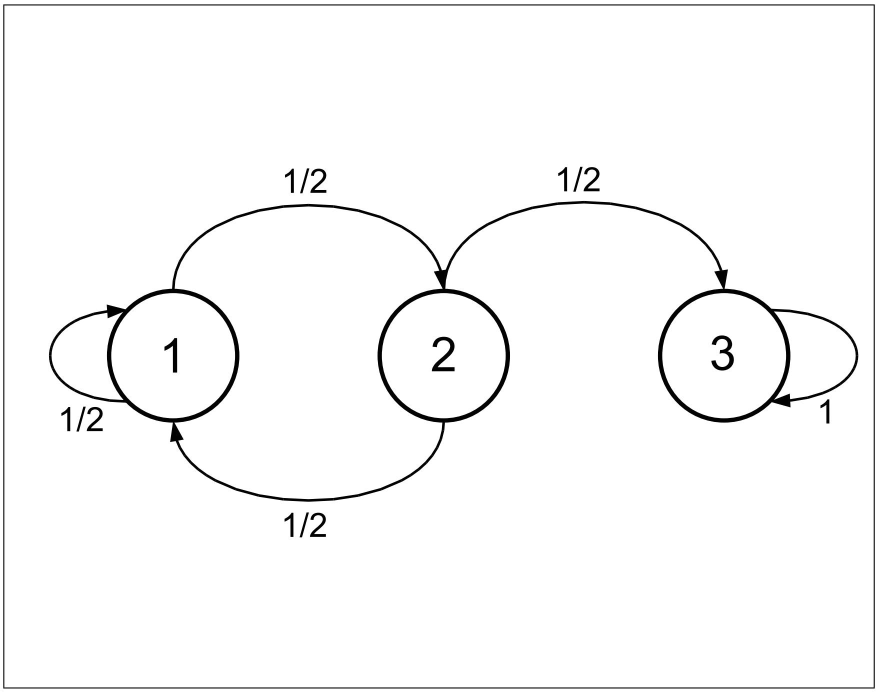
f. (6 pts) Draw a minimum state Markov chain to describe the game and find $\mathbf{E}[W]$.

Solution: Intuitively, we only need to keep count of number of consecutive heads, or in other words, to store the result of the last flip. If the last flip came up with tail, the count is simply zero; if the last flip came up with head, we add the count by one; however when the count reaches 2 the game ends, so we make this an absorption state. We design a Markov chain with three states as in the following figure

1 : Number of consecutive head is zero,

2 : Number of consecutive head is one,

3 : Two Heads in a row.



Their transition probabilities are

$$p_{11} = 0.5, p_{12} = 0.5, p_{13} = 0,$$

$$p_{21} = 0.5, p_{22} = 0, p_{23} = 0.5,$$

$$p_{31} = 0, p_{32} = 0, p_{33} = 1,$$

where states 1, 2 are transient, and state 3 is recurrent. Now $\mathbf{E}[W]$ can be interpreted as the expected time till absorption into state 3, which follows this system of equations:

$$\begin{aligned} \mu_3 &= 0, \\ \mu_1 &= 1 + 0.5\mu_1 + 0.5\mu_2, \\ \mu_2 &= 1 + 0.5\mu_1 + 0.5\mu_3, \end{aligned}$$

which gives

$$\mu_1 = 6, \mu_2 = 4, \mu_3 = 0.$$

Note that we are starting with the count being zero, i.e., we are starting with state 1. Therefore the expected winnings is $\mathbf{E}[W] = 2 * \mu_1 = 12$.