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

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) $\frac{1}{100\lambda} + \frac{1}{99\lambda}$
   (iv) $\frac{2}{100\lambda}$

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) $1 - 1.5e^{-0.5}$
   (iv) $(1 - e^{-0.5})$

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) $\lim_{n\to\infty} \mathbf{E}[R_n] = 0$
   (iii) $\mathbf{E}[R_n] = 0.5 \ \ \forall\, n.$
   (iv) $\lim_{n\to\infty} \mathbf{E}[R_n]$ is not defined

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) 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.

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) 36/76
(iii) 0.36
(iv) 0.6

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) $4np(1-p)$
(iv) $np(1-p)$

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

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) $y \geq \frac{1}{2}\log\left(\frac{1-p}{p}\right)$
(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) $y \geq \frac{1}{4}\log\left(\frac{1-p}{p}\right)$

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

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

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, $p_{T_1}(t_1)$

b. (6 pts) Compute the least squares estimate (LSE) of Q from the first recording $T_1 = t_1$.

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

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 ($T_1 = t_1$).

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 (c)?

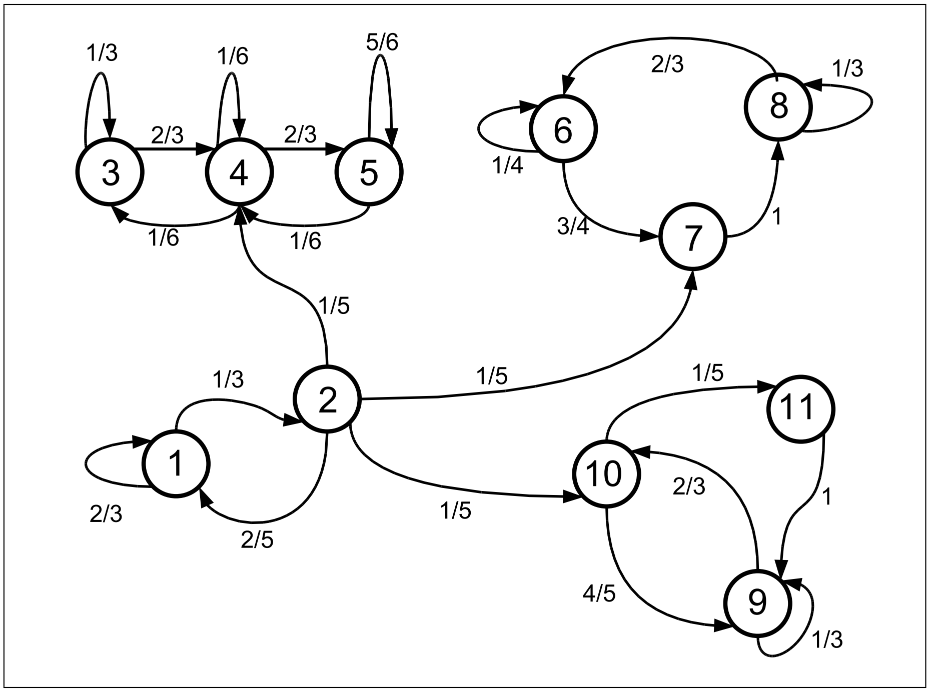
f. (5 pts) Show that for all $\epsilon > 0$

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

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

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

Problem 3 (31 pts) Let $X_0, X_1, X_2, X_3 \dots$ 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.

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.

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 $\text{var}(Y)$.

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 $\mathbf{E}[Z_4]$.

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.

This question does not rely on the above figure. A friend offers you a game of chance: You flip a fair 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 $\mathbf{E}[W]$.