6.041/6.431 Spring 2006 Quiz 1
Wednesday, March 15, 7:30-9:30 PM.

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

Name: ______________________

Recitation Instructor: ______________________

TA: ______________________

Question	Score	Out of
0		3
1		30
2		40
3		27
Your Grade		100

- You have 120 minutes to complete the quiz.
- At the end of the quiz period you will turn in this quiz packet, and 2 blue books. Question 1 will be answered in the quiz packet, while question 2 and 3 will be answered in their own respective blue books.
- Question 1 is True False, no partial credit is given.
- For questions 2 and 3, you should concisely indicate your reasoning and show all relevant work. Grades will be based on our judgment of your level of understanding as reflected by what you have written.
- You may give an answer in the form of an arithmetic expression (sums, products, ratios, factorials) of <u>numbers</u> that could be evaluated using a calculator. Expressions like $\binom{8}{3}$ or $\sum_{k=0}^{5}(1/2)^k$ are also fine.
- This is a closed-book exam except for one double-sided, handwritten, 8.5 by 11 formula sheet.
- Calculators are allowed.
- Be neat! If we can't read it, we can't grade it.

Massachusetts Institute of Technology

Department of Electrical Engineering & Computer Science

6.041/6.431: Probabilistic Systems Analysis

(Spring 2006)

Problem 0: (3 points)

Write your name, your recitation instructor's name, and TA's name on the cover of the quiz booklet and your 2 blue books. Include the question number on the cover of the blue book.

Problem 1: (30 points)
Each of the following statements is either True or False. There will be **no partial credit** given for the True False questions, thus any explanations will not be graded. Please **clearly** indicate True or False in the below, ambiguous marks will receive zero credit. All questions have equal weight.

Consider a probabilistic model with a sample space Ω, a collection of events that are subsets of Ω, and a probability law $\mathbf{P}()$ defined on the collection of events—all exactly as usual. Let A, B and C be events.

(a) If $\mathbf{P}(A) \leq \mathbf{P}(B)$, then $A \subseteq B$. True False

(b) Assuming $\mathbf{P}(B) > 0$, $\mathbf{P}(A \mid B)$ is at least as large as $\mathbf{P}(A)$. True False

Now let X and Y be random variables defined on the same probability space Ω as above.

(c) If $\mathbf{E}[X] > \mathbf{E}[Y]$, then $\mathbf{E}[X^2] \geq \mathbf{E}[Y^2]$. True False

(d) Suppose $\mathbf{P}(A) > 0$. Then $\mathbf{E}[X] = \mathbf{E}[X \mid A] + \mathbf{E}[X \mid A^C]$. True False

(e) If X and Y are independent and $\mathbf{P}(C) > 0$,
then $p_{X,Y|C}(x, y) = p_{X|C}(x)\, p_{Y|C}(y)$. True False

(f) If for some constant c we have $\mathbf{P}(\{X > c\}) = \frac{1}{2}$, then $\mathbf{E}[X] > \frac{c}{2}$. True False

In a simple game involving flips of a fair coin, you win a dollar every time you get a head. Suppose that the maximum number of flips is 10, however, the game terminates as soon as you get a tail.

(g) The expected gain from this game is 1. True False

Let X be a uniformly distributed continuous random variable over some interval $[a, b]$.

(h) We can uniquely describe $f_X(x)$ from knowing its mean and variance. True False

Let X be an exponentially distributed random variable with a probability density function $f_X(x) = e^{-x}$.

(i) Then $P(\{0 \leq X \leq 3\} \cup \{2 \leq X \leq 4\}) = 1 - e^{-4}$ True False

Let X be a normal random variable with mean 1 and variance 4. Let Y be a normal random variable with mean 1 and variance 1.

(j) $\mathbf{P}(X < 0) < \mathbf{P}(Y < 0)$. True False

Problem 2: (40 points)
Please write all work for Problem 2 in your **first blue book**. No work recorded below will be graded. All questions have approximately the same weight.

Borders Book store has been in business for 10 years, and over that period, the store has collected transaction data on all of its customers. Various marketing teams have been busy using the data to classify customers in hopes of better understanding customer spending habits.

Marketing Team A has determined that out of their customers, 1/4 are *low frequency buyers* (i.e., they don't come to the store very often). They have also found that out of the *low frequency buyers*, 1/3 are *high spenders* (i.e., they spend a significant amount of money in the store), whereas out of the *high frequency buyers* only 1/10 are *high spenders*. Assume each customer is either a *low* or *high frequency buyer*, and each customer is either a *high* or *low spender*.

(a) Compute the probability that a randomly chosen customer is a *high spender*.

(b) Compute the probability that a randomly chosen customer is a *high frequency buyer* given that he/she is a *low spender*.

You are told that the only products Borders sells are books, CDs, and DVDs. You are introduced to Marketing Team B which has identified 3 customer groupings. These groups are collectively exhaustive and mutually exclusive. They have also determined that each customer is equally likely to be in any group, customers are *i.i.d* (independent and identically distributed), and each customer buys only one item per day. They refer to the groupings as C_1, C_2, and C_3, and have determined the following conditional probabilities:

$$\begin{aligned}
\mathbf{P}(\text{purchases a book} \mid \text{customer in } C_1) &= 1/2 \\
\mathbf{P}(\text{purchases a CD} \mid \text{customer in } C_1) &= 1/4 \\
\mathbf{P}(\text{purchases a DVD} \mid \text{customer in } C_1) &= 1/4 \\
\mathbf{P}(\text{purchases a book} \mid \text{customer in } C_2) &= 1/2 \\
\mathbf{P}(\text{purchases a CD} \mid \text{customer in } C_2) &= 0 \\
\mathbf{P}(\text{purchases a DVD} \mid \text{customer in } C_2) &= 1/2 \\
\mathbf{P}(\text{purchases a book} \mid \text{customer in } C_3) &= 1/3 \\
\mathbf{P}(\text{purchases a CD} \mid \text{customer in } C_3) &= 1/3 \\
\mathbf{P}(\text{purchases a DVD} \mid \text{customer in } C_3) &= 1/3
\end{aligned}$$

(c) Compute the probability that a customer purchases a book or a CD.

(d) Compute the probability that a customer is in group C_2 or C_3 given he/she purchased a book.

Now in addition to the data from Marketing Team B, you are told that each book costs \$15, each CD costs \$10, and each DVD costs \$15.

(e) Compute the PMF, expected value and variance of the revenue (in dollars) Borders collects from a single item purchase of one customer?

(f) Suppose that n customers shop on a given day. Compute the expected value and variance of the revenue Borders makes from these n customers.

The following question is required for 6.431 Students
(6.041 Students may attempt it for 5 points of Extra Credit)

Skipper is very abnormal, not that there's anything wrong with that. He doesn't fit into any of the marketing teams' models. Every day Skipper wakes up and walks to Borders Bookstore. There he flips a fair coin repeatedly until he flips his second tails. He then goes to the counter and buys 1 DVD for each head he flipped. Let R be the revenue Borders makes from Skipper each day.

(g) What's the daily expected revenue from Skipper? What's the variance of the daily revenue from Skipper?

Problem 3: (27 points)
Please write all work for Problem 3 in your **second blue book**. No work recorded below will be graded. All questions have approximately the same weight.

We have s urns and n balls, where $n \geq s$. Consider an experiment where each ball is placed in an urn at random (i.e., each ball has equal probability of being placed in any of the urns). Assume each ball placement is independent of other placements, and each urn can fit any number of balls. Define the following random variables:

For each $i = 1, 2, \ldots, s$, let X_i be the number of balls in urn i.
For each $k = 0, 1, \ldots, n$, let Y_k be the number of urns that have exactly k balls.

Note: Be sure to include ranges of variables where appropriate.

(a) Are the X_i's independent? Yes or No? Please explain your answer.

(b) Find the PMF, mean, and variance of X_i.
i.e. compute $p_{X_i}(k)$, $\mathbf{E}[X_i]$, and $\text{var}(X_i)$.

(c) For this question only assume $n = 10$ and $s = 3$. Find the probability that the first urn has 3 balls, the second has 2, and the third has 5.
i.e. compute $\mathbf{P}(X_1 = 3 \cap X_2 = 2 \cap X_3 = 5)$

(d) Compute $\mathbf{E}[Y_k]$.

(e) Compute $\text{var}(Y_k)$. You may assume $n \geq 2k$.

(f) **This problem is required for 6.431 Students**
(6.041 Students may attempt it for 5 points of Extra Credit)
What is the probability that no urn is empty?
i.e. compute $P(X_1 > 0 \cap X_2 > 0 \cap \ldots \cap X_s > 0)$