6.041/6.431 Spring 2008 Quiz 1
Wednesday, March 12, 7:30 - 9:30 PM.

SOLUTIONS

Name: ____________________

Recitation Instructor: ____________________

TA: ____________________

Question	Part	Score	Out of
0			3
1	all		40
2	a		5
	b		5
	c		6
	d		6
3	a		5
	b		6
	c		6
	d		6
	e		6
	f		6
	g		10
6.041 Total			100
6.431 Total			110

- 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 one two-sided 8.5 by 11 formula sheet plus a calculator.
- You have 120 minutes to complete the quiz.
- Be neat! You will not get credit if we can't read it.
- Graded quizzes will be returned in recitation on Tuesday 3/18.

Problem 0: (3 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.

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
Multiple Choice Questions: **CLEARLY** circle the appropriate choice. Scratch paper is available if needed, though **NO** partial credit will be given for the Multiple Choice.

a. (4 pts) One of the following statements is NOT true:

   (i) $\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B)$

   (ii) $\mathbf{P}(A \cup (B \backslash A)) = \mathbf{P}(A) + \mathbf{P}(B \backslash A)$

   (iii) $\boxed{\mathbf{P}(A \cap (B \backslash A)) \text{ can be greater than } 0}$

   (iv) if $A \cap B$ is empty then $\mathbf{P}(A|B) \neq \mathbf{P}(A)$ (assume $\mathbf{P}(B) \neq 0$ and $\mathbf{P}(A) \neq 0$)

   Solution: The intersection $A \cap (B \backslash A)$ is always empty, so $\mathbf{P}(A \cap (B \backslash A))$ is always zero.

b. (4 pts) Two people each independently pick one integer between 1 and 5, inclusive. All choices are equally likely. Denote these numbers by X and Y. Then $\mathbf{P}(|X - Y| > 2 \mid X > 2)$ is equal to:

   (i) 6/25

   (ii) $\boxed{1/5}$

   (iii) 3/25

   (iv) 2/5

   Solution: There are $15 (= 3 \times 5)$ ways of picking X and Y given $X > 2$. Only three out of them satisfy $|X - Y| > 2$:

$$X = 4, Y = 1,$$
$$X = 5, Y = 1,$$
$$X = 5, Y = 2.$$

   It follows that $\mathbf{P}(|X - Y| > 2 \mid X > 2) = \frac{3}{15} = \frac{1}{5}$.

c. (4 pts) There are n chairs placed in a perfect circle. The chairs are indistinguishable, and thus there is no identifiable start or stop to the circle. The number of ways n people can be placed in the chairs is

   (i) $n!$

   (ii) $\boxed{(n-1)!}$

   (iii) n

   (iv) $n - 1$

   Solution: Since the chairs are placed in a circle, it doesn't make any difference where the first person chooses to sit. Then there are $n - 1$ people left to take $n - 1$ chairs. So the total number of ways n people can be placed in a circle of chairs is $1 \cdot (n - 1)!$.

d. (4 pts) When your friend Mike is asked if he smokes, he flips a fair coin. If the outcome is heads, he says no. If the outcome is tails, he tells the truth. Assume the probability that Mike smokes is p. Given Mike just said he doesn't smoke (he said no), the probability that the outcome of the coin flip is heads is

(i) p

(ii) $1/(1+p)$

(iii) $1/(1-p)$

(iv) $\boxed{1/(2-p)}$

Solution: The conditional probability of head given Mike said no is written as

$$\begin{aligned}\mathbf{P}(\text{head} \mid \text{Mike said no}) &= \frac{\mathbf{P}(\text{head, Mike said no})}{\mathbf{P}(\text{Mike said no})} \\ &= \frac{\mathbf{P}(\text{The coin flip is head and Mike said no})}{\mathbf{P}(\text{Head, Mike said no}) + \mathbf{P}(\text{Tail, Mike said no})} \\ &= \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}(1-p)} \\ &= \frac{1}{2-p}.\end{aligned}$$

e. (4 pts) Assume that X and Y are independent random variables. Let $Z = X^2 + Y^3$. Then

(i) $\boxed{\text{var}(Z) = \text{var}(X^2) + \text{var}(Y^3)}$

(ii) $\text{var}(Z) = (\text{var}(X))^2 + (\text{var}(Y))^3$

(iii) $\text{var}(Z) = \mathbf{E}[X^2] + \mathbf{E}[Y^3]$

(iv) $\text{var}(Z) = \text{var}(X^2) + \text{var}(Y^3) + 2\mathbf{E}[X^2]E(Y^3)$

Solution: The independence of two random variables implies the independence of functions of the two. Thus we are able to sum the variances.

f. (4 pts) Assume you roll a fair 6-sided die twice[1]. Define X to be the outcome of the first roll and Y the outcome of the second roll. Then

(i) $\boxed{\mathbf{E}[XY] = (21/6)^2}$

(ii) $\mathbf{E}[XY] = 0$

(iii) $\mathbf{E}[XY] = 91/36$

(iv) $\mathbf{E}[XY] = 91/6$

Solution:) Since the die is fair, the rolls are independent. Thus X and Y are independent. Therefore $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$. The outcomes of a roll range from 1 through 6, with the probability of each outcome being $\frac{1}{6}$. Thus, $\mathbf{E}[X] = \mathbf{E}[Y] = \frac{1}{6}(1+2+3+4+5+6) = \frac{21}{6}$. Thus, the answer follows.

g. (4 pts) You roll a fair 6-sided die 100 times. Let X be the number of times a 6 appears. Then

(i) $\mathbf{E}[X^2] = 100/6 - 100/36$

[1] A fair 6 sided die is labeled with the integers 1 through 6, with each side being equally likely to appear. Each roll of the die is independent of all other rolls.

(ii) $\mathbf{E}[X^2] = 100/6$

(iii) $\boxed{\mathbf{E}[X^2] = 100/6 - 100/36 + (100/6)^2}$

(iv) $\mathbf{E}[X^2] = 100/6 - (100/6)^2$

Solution:) Define X_i to be an indicator random variable that takes the value 1 if a 6 appears on the ith throw and 0 otherwise. Since the die is fair the X_i's are independent. Building on this we write $X = \sum_{i=1}^{100} X_i$. Thus

$$\begin{aligned}
\mathbf{E}[X^2] &= \text{var}(X) + (\mathbf{E}[X])^2 \\
&= 100\ \text{var}(X_i) + (100\ \mathbf{E}[X_i])^2 \\
&= 100\ [\mathbf{E}(X_i^2) - (\mathbf{E}[X_i])^2] + (100/6)^2 \\
&= 100\ [1/6 - 1/36] + (100/6)^2 \\
&= 100/6 - 100/36 + (100/6)^2
\end{aligned}$$

h. (4 pts) X refers to the time of the first arrival of a head in a coin flipping sequence (i.e. $X = 4$ means the first head was observed on the $4th$ flip). Assume the probability of a heads on any coin flip is p, and all flips are independent of each other. If Y is a random variable indicating the time of the second arrival of a head, then

(i) $\mathbf{E}[Y] = 1/p$

(ii) $\mathbf{E}[Y] = 1 + 1/p$

(iii) $\mathbf{E}[Y] = 1 + p$

(iv) $\boxed{\mathbf{E}[Y] = 2/p}$

Solution: iv) The time of the second arrival is the sum of two independent random variables: the time of the first arrival and the time between the first and the second arrivals. Both of these are geometric random variables with parameter p. Thus the expected value of their sum is the sum of their expected values and therefore $2/p$.

i. (4 pts) If X is a binomial random variable where the probability of a success is p and the number of trials is denoted by n, then

(i) $\mathbf{P}(X \geq 2) = (1-p)^n - np^{n-1}$

(ii) $\mathbf{P}(X \geq 2) = 1 - p - np(1-p)$

(iii) $\boxed{\mathbf{P}(X \geq 2) = 1 - (1-p)^n - np(1-p)^{n-1}}$

(iv) $\mathbf{P}(X \geq 2) = p^n + np^{n-1}$

Solution: We know that $\mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Note $\mathbf{P}(X \geq 2) = 1 - \mathbf{P}(X < 2) = 1 - (1-p)^n - np(1-p)^{n-1}$.

j. (4 pts) You are in possession of n balls and s bins, where each bin is labeled 1 through s. You throw each ball into a bin at random, where each bin is equally likely and each throw is independent of any other throw. Let the random variable Y_i be defined as follows

$$Y_i = \begin{cases} 1, & \text{if bin } i \text{ is EMPTY} \\ 0, & \text{otherwise} \end{cases}$$

Let $Y = \sum_{i=1}^{s} Y_i$ be the total number of EMPTY bins. Then

(i) $\boxed{\mathbf{E}[Y] = \frac{(s-1)^n}{s^{n-1}}}$

(ii) $\mathbf{E}[Y] = \frac{1}{s^n}$

(iii) $\mathbf{E}[Y] = 1$

(iv) $\mathbf{E}[Y] = \frac{n(s-1)^n}{s^n}$

Solution: Note that although each Y_i is dependent of each other, we still have $\mathbf{E}[Y] = \Sigma_i^s \mathbf{E}[Y_i]$. The probability for the ith bin to be empty is $(\frac{s-1}{s})^n$, so $\mathbf{E}[Y_i]$ becomes

$$\mathbf{E}[Y_i] = 1 \cdot (\frac{s-1}{s})^n + 0 \cdot (1 - (\frac{s-1}{s})^n) = (\frac{s-1}{s})^n.$$

It follows that $\mathbf{E}[Y] = s \cdot (\frac{s-1}{s})^n = \frac{(s-1)^n}{s^{n-1}}$.

Question 2: Many of the parts in Question 2 can be answered independent of other parts.

Assume the following joint PMF for the discrete random variables X and Y.

y=3	1/6	1/9	0	1/9
y=2	0	0	2/9	0
y=1	1/6	0	1/9	1/9
	x=1	x=2	x=3	x=4

$p_{X,Y}(x,y)$, the joint PMF of X and Y.

a. (5 pts) Find the marginal PMF of X, $p_X(x)$.

Solution: To find the marginal distribution of X you can easily sum over Y from the joint, $p_X(x) = \sum_{y=1}^{3} p_{X,Y}(x,y)$. Given the above table, we simply sum the columns to find $p_X(x)$. This leads to

$$p_X(x) = \begin{cases} 1/3 & \text{for } x = 1 \\ 1/9 & \text{for } x = 2 \\ 1/3 & \text{for } x = 3 \\ 2/9 & \text{for } x = 4 \\ 0 & \text{otherwise.} \end{cases}$$

Assume the following events:

$$\begin{aligned} A &: \{Y < 3\} \\ B &: \{X = 1\} \\ C &: \{X = 4\} \end{aligned}$$

b. (5 pts) Define the event D as $D = ((A \cap B^c) \cup C)^c$. Find $\mathbf{P}(D)$.

Solution: Converting the Boolean expressions we find

$$\begin{aligned} A \cap B^c &= \{(2,1),(2,2),(3,1),(3,2),(4,1),(4,2)\} \\ \{A \cap B^c\} \cup C &= \{(2,1),(2,2),(3,1),(3,2),(4,1),(4,2),(4,3)\} \\ ((A \cap B^c) \cup C)^c &= \{(1,1),(1,2),(1,3),(2,3),(3,3)\} \end{aligned}$$

Summing these we find $\mathbf{P}(D) = 4/9$.

c. (6 pts) Find $\mathbf{E}[Y|A]$.

Solution: First we find the marginal PMF of Y,

$$p_Y(y) = \begin{cases} 7/18, & y = 1, \\ 2/9, & y = 2, \\ 7/18, & y = 3, \\ 0, & \text{otherwise,} \end{cases}$$

and the conditional PMF of Y,

$$p_{Y|A}(y) = \begin{cases} 7/11, & y = 1, \\ 4/11, & y = 2, \\ 0, & \text{otherwise.} \end{cases}$$

Now we have

$$\mathbf{E}[Y \mid A] = 2 \cdot p_{Y|A}(2) + 1 \cdot p_{Y|A}(1) = 8/11 + 7/11 = 15/11.$$

d. (6 pts) Define the event E as $E = (B \cup C)$. Conditioned on E are X and Y independent? Give a mathematical justification for your answer.

Solution: Yes they are independent. We write the joint conditional distribution

y=3	3/10	1/5
y=2	0	0
y=1	3/10	1/5
	x=1	x=4

$p_{X,Y|E}(x, y)$, the conditional joint PMF of X and Y.

and the marginal conditional distribution of X and Y given E,

$$p_{X|E}(x) = \begin{cases} 3/5, & x = 1, \\ 2/5, & x = 4, \\ 0, & \text{otherwise,} \end{cases}$$

$$p_{Y|E}(y) = \begin{cases} 1/2, & y = 1, \\ 0, & y = 2, \\ 1/2, & y = 3, \\ 0, & \text{otherwise.} \end{cases}$$

Now we obtain that $p_{X,Y|E}(x, y) = p_{X|E}(x)p_{Y|E}(y)$, which proves the conditional independence of X and Y given E.

Question 3: Many parts of Question 3 can be answered independent of other parts. Part (g) is required for 6.431 students only.

As a member of the elite 6.041/6.431 student body you are entitled to shop at the local probabilistic grocery store. This is considered a great privilege. Not unexpectedly choosing fruit is far from deterministic. To obtain a piece of fruit you approach the fruit manager and ask him for a piece of fruit. He will give you an apple with probability 2/3 or an orange with probability 1/3. You always receive a piece of fruit after each request, and each request is independent of all other requests.

a. (5 pts) What is the probability you receive exactly 3 apples in your first 10 requests?

Solution: The number of apples received is a binomial random variable, so the probability is $\mathbf{P}(3 \text{ apples received}) = \binom{10}{3} \left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^7$.

b. (6 pts) Let X be the number of requests you make until you receive your first orange, i.e. you receive your first orange on request number Z. What is $\mathbf{E}[X]$?

Solution: X is a geometric random variable, thus $\mathbf{E}[X] = \dfrac{1}{p} = 3$

c. (6 pts) Given you received 3 apples and 7 oranges in 10 requests, what is the probability that you received exactly 2 apples within the first 5 requests? *For full credit your final answer should contain no summations.*

Solution: The probability of any particular ordering is the same, similar to coin tossing problems. Thus we can use counting to find the desired probability. For the denominator count the number of ways you can choose 3 apples from 10 pieces of fruit $=\binom{10}{3}$. For the numerator count the total number of ways you can choose the 2 apples in the first 5, and then one apple in the last 5 $=\binom{5}{2}\binom{5}{1}$. Thus

$$\frac{\binom{5}{2}\binom{5}{1}}{\binom{10}{3}} = \frac{5}{12}$$

Note you could also count oranges instead of apples with the same result.

You proceed to the check out to pay for your fruit. Prices are as follows: each apple is \$3, each orange is \$15. Let T be your total bill in dollars, assuming you have 10 pieces of fruit.

d. (6 pts) Find $\mathbf{E}[T]$

Solution: Let W_i be the cost of the ith piece of fruit. Note $T = \sum_{i=1}^{10} W_i$. Using the linearity of expectation, and the identical distribution of W_i $\forall$ i we find

$$\mathbf{E}[T] = \sum_{i=1}^{10} \mathbf{E}[W_i] = 10\ \mathbf{E}[W] = 10\ (3\frac{2}{3} + 15\frac{1}{3}) = 10\ (2+5) = 70$$

e. (6 pts) Find var(T)

Solution: We now include the independence of the W_i's to the previous statements to claim the variance of a sum of W_i's is equal to the sum of variances.

$$\text{var}(T) = \text{var}(\sum_{i=1}^{10} W_i) = \sum_{i=1}^{10} \text{var}(W_i) = 10 \ \text{var}(W)$$

To find the variance of W we exploit $\text{var}(W) = \mathbf{E}[W^2] - \mathbf{E}[W]^2$. We have $\mathbf{E}[W] = 7$ from the previous part, thus we only require $\mathbf{E}[W^2]$ which is easily computed as

$$\begin{aligned} \mathbf{E}[W^2] &= 3^2 \ (\frac{2}{3}) + 15^2 \ (\frac{1}{3}) \\ &= 6 + 75 = 81 \end{aligned}$$

Putting it all together we find

$$\begin{aligned} \text{var}(T) &= 10 \ \text{var}(W) \\ &= 10 \ (81 - 49) = 320 \end{aligned}$$

You leave the grocery store and get on a bus with your friend Mais. The bus has k empty seats in a single row, where k is an integer greater than 2. You and Mais each choose a seat at random. Assume only one person can sit in a seat.

f. (6 pts) What is the probability you and Mais choose adjacent seats, i.e. you sit next to each other?

Solution: One approach is to use counting. The number of ways that you and Mais choose adjacent seats is $2(k-1)$, and the number of all ways of seating is $2\binom{k}{2}$. Therefore

$$\mathbf{P}(\text{You and Mais choose adjacent seats}) = \frac{2(k-1)}{2\binom{k}{2}} = \frac{2}{k}.$$

Another approach is to use sequential probability. Denote the seat you choose as A and the seat Mais choose as B. There are three ways for Mais and you to sit together: you choose the 1st seat and Mais chooses the 2nd seat; you choose the last seat and Mais chooses the second last seat; you choose a seat which is not at either end and Mais chooses the seat at the left or right to you. We have

$$\begin{aligned} \mathbf{P}(A, B \text{ are adjacent seats}) &= \sum_{i=2}^{k-1} \mathbf{P}(A = i) \ \ (\ \mathbf{P}(B = i+1) + \mathbf{P}(B = i-1) \ ) \\ &\quad + \mathbf{P}(A = 1)\mathbf{P}(B = 2) + \mathbf{P}(A = k)\mathbf{P}(B = k-1) \\ &= (k-2)(\frac{1}{k} \cdot \frac{1}{k-1}) + \frac{1}{k} \cdot \frac{1}{k-1} + \frac{1}{k} \cdot \frac{1}{k-1} \\ &= 2/k. \end{aligned}$$

Part (g) is required for 6.431 students only. 6.041 students may attempt part (g), but will only receive extra credit if they have the correct answer. i.e. there is no partial credit for 6.041 students on part (g).

On any given day there are N shopping carts in the store, where N is a discrete uniform random variable from 1 to 100. All carts are labeled with sequential integers from 1 to N. Let event A be that immediately upon entering the store you see cart number 30. Also assume that the carts are randomly placed in the store such that you are equally likely to see any of the N carts when you enter the store.

g. (10 pts) Find the PMF of N conditioned on event A, $p_{N|A}(n)$.

Solution: By Bayes' rule we have

$$p_{N|A}(n) = \frac{\mathbf{P}(A \mid N = n) p_N(n)}{\mathbf{P}(A)},$$

where $\mathbf{P}(A \mid N = n)$ is the probability to see cart number 30 when there are $N = n$ carts, $p_N(n)$ is the prior probability of N equaling to n, and $\mathbf{P}(A)$ is the probability of A when N is distributed uniformly between 0 and 100. So it follows that

$$\begin{aligned} p_{N|A}(n) &= \frac{\mathbf{P}(A \mid N = n) p_N(n)}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}(A \mid N = n) p_N(n)}{\sum_{i=1}^{100} \mathbf{P}(A \mid N = i) p_N(i)} \\ &= \frac{\frac{1}{n} \cdot \frac{1}{100}}{\sum_{i=30}^{100} \frac{1}{i} \cdot \frac{1}{100}} \\ &= \frac{\frac{1}{n}}{\sum_{i=30}^{100} \frac{1}{i}}, \quad \text{for} \quad n = 30, 31, 32, \ldots, 100, \end{aligned}$$

and $p_{N|A}(n) = 0$ for $n = 1, 2, \ldots, 29$.