MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Spring 2009)

Problem Set 4 Solutions
Due: March 4, 2009

1. Since the X_i are iid, for any i and j, we have:

$$E[X_i|X_1 + X_2 + \cdots + X_n = x] = E[X_j|X_1 + X_2 + \cdots + X_n = x].$$

But since we also know that:

$$E[X_1 + \cdots + X_n|X_1 + \cdots + X_n = x] = x$$

by the linearity of expectation we find that:

$$E[X_1|X_1 + X_2 + \cdots + X_n = x] = \frac{x}{n}.$$

2. (a) Clearly, it is not possible that the number of buses exceed the number of minutes you wait. Thus, the support of the PMF is $\{(n, t)|t \geq n \geq 1\}$.

(b) What does it mean to have been at the bus stop exactly t minutes and have n buses arrive? It means that at time t, the n^{th} bus came and you got on. And then in the previous $t - 1$ minutes, $n - 1$ buses came and you didn't get on. Thus,

$$p_{N,T}(n, t) = pq * \binom{t-1}{n-1} [p(1-q)]^{n-1}(1-p)^{t-n} \quad t \geq n \geq 1. \tag{1}$$

(c) To get the marginals, we sum the joint.

$$\begin{aligned} p_T(t) &= pq \sum_{n=1}^{t} \binom{t-1}{n-1} [p(1-q)]^{n-1}(1-p)^{t-n} && (2) \\ &= pq[p(1-q) + 1 - p]^{t-1} \quad \text{(binomial theorem)} && (3) \\ &= pq(1-pq)^{t-1} \quad t = 1, 2, \ldots && (4) \end{aligned}$$

Does this make intuitive sense? Consider an alternative way of getting this. T is the time we wait at the bus stop, ie. the time until we get on a bus and leave. Consider a trial occuring every minute: you get on the bus with probability pq and you don't with probability $1 - pq$. Then the time until you get on is just a geometric random variable. For N, does the above thinking give you a hint which will allow us to avoid doing a nasty sum? N is the number of buses which come until we get on a bus and leave. Consider trials occuring every time a bus comes: with probability q, you get on a bus, and with probability $1 - q$, you don't get on a bus. Again we have a geometric. If you want to do the sum,

$$\begin{aligned} p_N(n) &= p^n q(1-q)^{n-1} \sum_{t=n}^{\infty} \binom{t-1}{n-1} (1-p)^{t-n} && (5) \\ &= p^n q(1-q)^{n-1} p^{-n} && (6) \\ &= q(1-q)^{n-1} \quad n = 1, 2, \ldots && (7) \end{aligned}$$

The second equality above can be done by looking at small n (1,2,3..), computing the sum (look at the summand, write it as a derivative, interchange the sum and derivative operator, do the sum, then take the derivative of the result), and finding a pattern.

(d) To find the conditional, think Bayes.

$$
\begin{aligned}
p_{N|T}(n|t) &= \frac{p_N(n,t)}{p_T(t)} && (8)\\
&= \frac{\binom{t-1}{n-1}[p(1-q)]^{n-1}(1-p)^{t-n}}{(1-pq)^{t-1}} && (9)\\
&= \binom{t-1}{n-1}\left(\frac{p(1-q)}{p(1-q)+1-p}\right)^{n-1}\left(\frac{1-p}{p(1-q)+1-p}\right)^{t-n} \quad t \geq n \geq 1. && (10)
\end{aligned}
$$

$$
p_{T,N}(t|n) = \frac{p_{T,N}(n,t)}{p_N(n)} = p^n\binom{t-1}{n-1}(1-p)^{t-n} \quad t \geq n \geq 1. \tag{11}
$$

3. Recall that a random variable is a function which assigns a value of the random variable to each sample point in the sample space of an experiment (Drake page 42). For this experiment, we throw two fair six-sided dice. There are a total of 36 possible outcomes, or sample points, each of which are equally likely. Therefore the probability of any particular outcome is 1/36. For each sample point, the four random variables S, T, M, and N are assigned values. The table below lists these values for each sample point.

S T / M N	1		2		3		4		5		6	
1	2	1	3	2	4	3	5	4	6	5	7	6
	1	1	2	1	3	1	4	1	5	1	6	1
2	3	2	4	4	5	6	6	8	7	10	8	12
	2	1	2	2	3	2	4	3	5	2	6	2
3	4	3	5	6	6	9	7	12	8	15	9	18
	3	1	3	2	3	3	4	3	5	3	6	3
4	5	4	6	8	7	12	8	16	9	20	10	24
	4	1	4	2	4	3	4	4	5	4	6	4
5	6	5	7	10	8	15	9	20	10	25	11	30
	5	1	5	2	5	3	5	4	5	5	6	5
6	7	6	8	12	9	18	10	24	11	30	12	36
	6	1	6	2	6	3	6	4	6	5	6	6

The rows correspond to the value of the first die, and the columns correspond to the values of the second die. For example, if the result of the experiment is a 2 on the first die and a 6 on the second, which we can denote as $(2,6)$, we consult row 2 and column 6 to find these values for the random variables: $S = 8$, $T = 12$, $M = 6$, and $N = 2$.

To determine the probability mass function (PMF) $p_S(s) = P(S = s)$, we first examine the table and conclude that S takes on possible experimental values between 2 and 12. Then, for each possible value of S, we count the number of outcomes that correspond to that value and multiply this number by the probability of each outcome (1/36). For instance, to find $p_S(4)$, we count three outcomes, $(1,3)$, $(2,2)$, and $(3,1)$. Therefore $p_S(4) = 3*(1/36) = 1/12$. The result of applying this method to every possible value of s is sketched on page 2. We determine probability mass functions $p_T(t)$, $p_M(m)$, and $p_N(n)$ in a similar fashion.

For part (e), the conditional PMF is given by

$$p_{S|M}(s|m) = \frac{p_{S,M}(s,m)}{p_M(m)}$$

which is defined only when $p_M(m)$ is non-zero. To find the compound PMF $p_{S,M}(s,m)$, we count the number of outcomes satisfying $S = s$ *and* $M = m$. The value of the compound PMF is obtained by multiplying this number by 1/36. The value of the conditional PMF is then obtained by dividing this value by $p_M(m)$. For example, to find $p_{S|M}(3|2)$, we first find $p_{S,M}(3,2)$ by counting two outcomes, $(1,2)$ and $(2,1)$, where $S = 3$ and $M = 2$. Therefore $p_{S,M}(3,2) = 2 * (1/36) = 1/18$. Since $p_M(2) = 3/36$, we have that $p_{S|M}(3|2) = 2/3$. The conditional PMF is plotted on page 3.

For part (f), we find the compound PMF $p_{M,N}(m,n)$ by counting the number of outcomes satisfying $M = m$ *and* $N = n$ and multiplying by the probability of each outcome (1/36). The compound PMF is plotted on page 3.

4. Let random variable X be the number of trials you need to open the door, and define K_i to be the event that the ith key selected opens the door.

   (a) We have

$$\begin{aligned} p_X(1) &= P(K_1) = \frac{1}{5} \\ p_X(2) &= P(K_1^c)P(K_2 \mid K_1^c) = \left(\frac{4}{5}\right)\left(\frac{1}{4}\right) = \frac{1}{5} \\ p_X(3) &= P(K_1^c)P(K_2^c \mid K_1^c)P(K_3 \mid K_1^c \cap K_2^c) = \left(\frac{4}{5}\right)\left(\frac{3}{4}\right)\left(\frac{1}{3}\right) = \frac{1}{5} \end{aligned}$$

   Proceeding as such, we see that the PMF for X is

$$p_X(x) = \begin{cases} \frac{1}{5} & x = 1,2,3,4,5 \\ 0 & \text{otherwise} \end{cases}$$

   We can also view the problem as ordering the keys in advance and then trying them in succession, in which case the probability of any of the five keys being correct is 1/5. Since we have a discrete uniform distribution, the mean and variance can be readily determined:

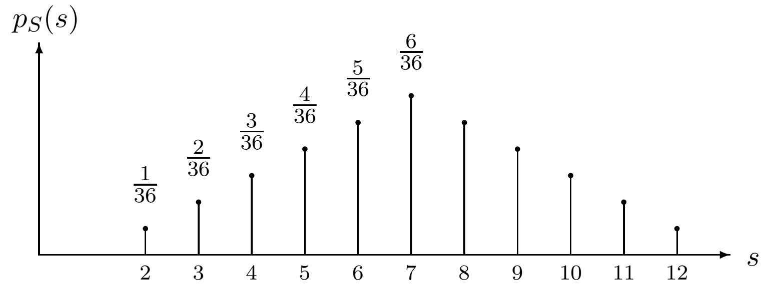
$$\mathbf{E}[X] = \frac{a+b}{2} = \frac{1+5}{2} = \boxed{3}, \qquad \text{var}(X) = \frac{(b-a)(b-a+2)}{12} = \frac{(5-1)(5-1+2)}{12} = \boxed{2}.$$

   (b) In this case, X is a geometric random variable with $p = 1/5$. The mean and variance are then:

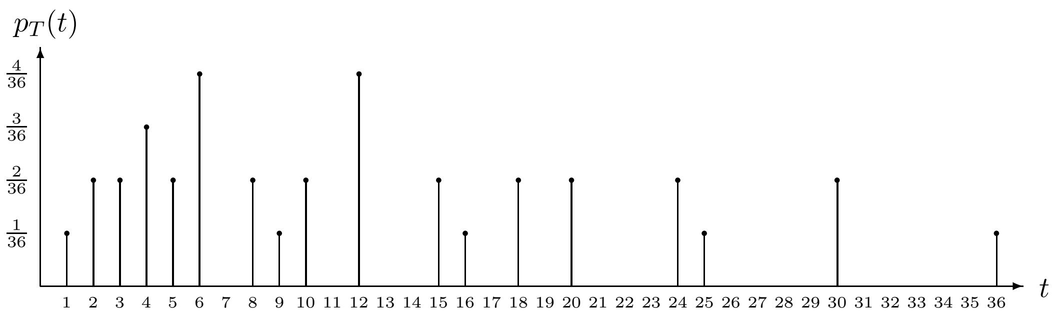
$$\mathbf{E}[X] = \frac{1}{p} = \frac{1}{1/5} = \boxed{5}, \qquad \text{var}(X) = \frac{1-p}{p^2} = \frac{1-1/5}{(1/5)^2} = \boxed{20}.$$

5. The key observation is that either Y or Z must be constant. When $Y = X^2$ is constant, i.e. $X \in \{a, -a\}$ for some $a \geq 0$, the assumptions $\mathbf{E}[X] = 0$ and $\text{var}(X) = 1$ imply $a = 1$, hence

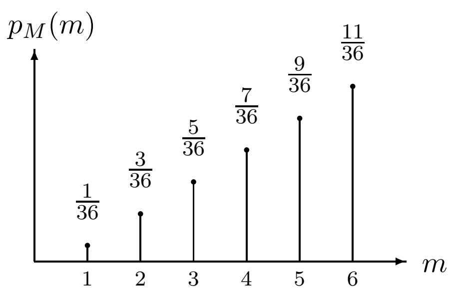
$$p_X(x) = \begin{cases} 0.5, & x \in \{1, -1\}, \\ 0, & \text{otherwise.} \end{cases}$$


$p_S(s)$
$\frac{1}{36}$ $\frac{2}{36}$ $\frac{3}{36}$ $\frac{4}{36}$ $\frac{5}{36}$ $\frac{6}{36}$
2 3 4 5 6 7 8 9 10 11 12
s

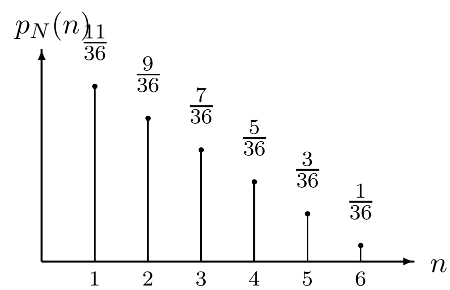
(a)


$p_T(t)$
$\frac{4}{36}$ $\frac{3}{36}$ $\frac{2}{36}$ $\frac{1}{36}$
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
t

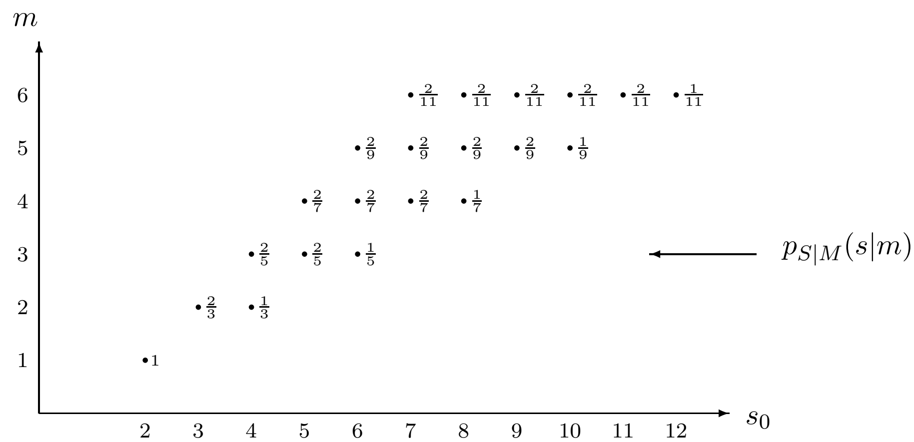
(b)


$p_M(m)$
$\frac{1}{36}$ $\frac{3}{36}$ $\frac{5}{36}$ $\frac{7}{36}$ $\frac{9}{36}$ $\frac{11}{36}$
1 2 3 4 5 6
m

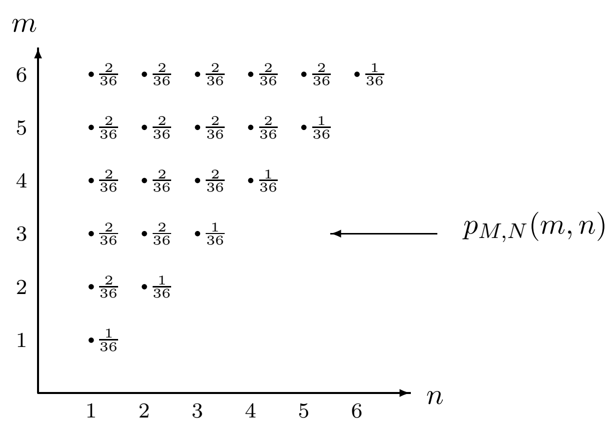
(c)


$p_N(n)$
$\frac{11}{36}$ $\frac{9}{36}$ $\frac{7}{36}$ $\frac{5}{36}$ $\frac{3}{36}$ $\frac{1}{36}$
1 2 3 4 5 6
n

(d)



(e)



(f)

When $Z = (X+1)^2$ is constant, i.e. $X \in \{b, -b-2\}$ for some $b \geq -1$, the assumptions $\mathbf{E}[X] = 0$ and $\mathrm{var}(X) = 1$ imply $b = \sqrt{2} - 1$, hence

$$p_X(x) = \begin{cases} 0.25(2+\sqrt{2}), & x = \sqrt{2} - 1, \\ 0.25(2-\sqrt{2}), & x = -\sqrt{2} - 1, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

To show that either Y or Z is constant, assume the contrary. Then $p_X(a) > 0$ imples $p_X(-a) > 0$ for every a (otherwise $Y = a^2$ means $X = a$ and $Z = (a+1)^2$, which contradicts independence of Y and Z). Similarly, $p_X(a) > 0$ implies $p_X(-a-2) > 0$. Combining the two observations shows that $p_X(a) > 0$ implies $p_X(a+2) > 0$, which means that the range of X is infinite.

On the other hand, when the value of $Y = X^2$ is fixed, X (and hence $Z = (X+1)^2$) can take at most two different values. Since Y and Z are independent, Z takes at most two values, unconditionally. Hence X takes not more than four different values. The contradiction proves that either Y or Z are constant.

G1†. (a) Let's associate a Bernoulli random variable X_k to urn k, where "success" corresponds to extracting a white ball. Then, clearly X_k has the probability mass function:

$$X_k = \begin{cases} 1 & \text{with probability } \frac{k}{n+1} \\ 0 & \text{with probability } 1 - \frac{k}{n+1} \end{cases}.$$

Since X_k is a Bernoulli random variable (or by direct calculation), it follows that

$$E[X_k] = \frac{k}{n+1}, \qquad \mathrm{var}[X_k] = \frac{k}{n+1}\frac{n+1-k}{n+1}.$$

The total number of white balls extracted is just

$$X = X_1 + \cdots + X_n.$$

This implies

$$E[X] = \sum_{k=1}^{n} E[X_k] = \sum_{k=1}^{n} \frac{k}{n+1} = \frac{n}{2}.$$

This can also be easily obtained from the complete symmetry of the problem between white and black balls.

For the variance, since all the X_i are independent, if follows that

$$\mathrm{var}[X] = \sum_{k=1}^{n} \mathrm{var}[X_k] = \sum_{k=1}^{n} \frac{k}{n+1}\frac{n+1-k}{n+1} = \frac{1}{6}\frac{n(n+2)}{n+1}.$$

(b) The probability of obtaining all white balls, except a black one from urn k is equal to

$$\frac{n+1-k}{n+1}\prod_{j\neq k}\frac{j}{n+1},$$

since the first term corresponds to the probability of "failure" in urn k, and "success" in all the other urns. This expression can be simplified (by multiplying and dividing by the missing term) to

$$\frac{n+1-k}{n+1}\frac{n+1}{k}\prod_{j=1}^{n}\frac{j}{n+1} = \frac{n+1-k}{k}\frac{n!}{(n+1)^n}.$$

(c) Define the events $B = \{\text{only one black ball was extracted}\}$ and $B_k = \{\text{the ball extracted from urn } k \text{ is b}$
Using the results of the previous exercise (and canceling common terms), we have

$$P(B_k|B) = \frac{P(B\cap B_k)}{P(B)} = \frac{P(B\cap B_k)}{\sum_{j=1}^{n} P(B\cap B_j)} = \frac{\frac{n+1-k}{k}}{\sum_{j=1}^{n}\frac{n+1-j}{j}}.$$