Massachusetts Institute of Technology
6.042J/18.062J: Mathematics for Computer Science
Professors Mauricio Karchmer and Ron Rivest

Handout 29
May 12, 1995

Problem Set 10 Solutions

Problem 10-1. **(a)** Let X and Y be two random variables defined on the same probability space. Show that:

$$E(X) = \sum_k \Pr(Y = k) E(X \mid Y = k)$$

Proof. We know that

$$P(X = i) = \sum_k P(X = i \mid Y = k) P(Y = k)$$

and so

$$\begin{aligned}
E(X) &= \sum_i i P(X = i) \\
&= \sum_i i \left[\sum_k P(X = i \mid Y = k) P(Y = k) \right] \\
&= \sum_i \sum_k i P(X = i \mid Y = k) P(Y = k) \\
&= \sum_k \sum_i i P(X = i \mid Y = k) P(Y = k) \\
&= \sum_k P(Y = k) \sum_i i P(X = i \mid Y = k) \\
&= \sum_k P(Y = k) E(X \mid Y = k) \\
&= \sum_k E(X \mid Y = k) P(Y = k)
\end{aligned}$$

■

(b) Consider the following random experiment. We have n (fair) dice in a bin. We shake the bin and tip it over. Some of the dice roll out (it is equally likely that any number from 1 to n will roll out). We then sum up all of the values on the top faces of the dice that roll out. What is the expected value of this sum?

We will use the result proven in part (a). Let random variable Y be the number of dice that roll out and random variable X be the sum. We know that $E(X \mid$

$Y = k) = 3.5k$. (Since the expected value of one roll is 3.5 and expectation is linear). Thus

$$\begin{aligned} E(X) &= \sum_{k=1}^{n} E(X \mid Y = k) P(Y = k) \\ &= \sum_{k=1}^{n} 3.5k \frac{1}{n} \\ &= 3.5 \frac{1}{n} \sum_{k=1}^{n} k \\ &= \frac{3.5n(n+1)}{2n} \\ &= 1.75(n+1) \end{aligned}$$

Problem 10-2. Suppose we toss a fair coin n times. We define the random variables X_i $(i = 1, \ldots, n)$ by:

$$X_i = \begin{cases} 1 & \text{if the } i\text{th toss is heads} \\ 0 & \text{otherwise} \end{cases}$$

Let $X = \sum_{i=1}^{n} X_i$. Compute the following (where $a, b \in \{0, 1\}$ and $0 \le k \le n$):

(a) $\Pr(X_i = a \;\&\; X_j = b \mid X = k)$

We can solve this using the definition of conditional probability. We wish to compute

$$\frac{P(X_i = a \;\&\; X_j = b \;\&\; X = k)}{P(X = k)}$$

The denominator is always $\binom{n}{k}/2^n$ (since to get k 1s we need to choose which positions get the value 1) so we just work on the numerator.

Let $c = a + b$. To get k tosses to be 1 when we fix the value of tosses i and j to sum up to c, we need to get $k - c$ 1s out of the remaining $n - 2$ positions. We can do this in $\binom{n-2}{k-c}$ ways. So the conditional probability is

$$\frac{\binom{n-2}{k-c}}{\binom{n}{k}}$$

(b) $\Pr(X = k \mid X_i = a \;\&\; X_j = b)$

We solve this in a similar way. We first note that we want to find

$$\frac{P(X_i = a \;\&\; X_j = b \;\&\; X = k)}{P(X_i = a \;\&\; X_j = b)}$$

Note that the denominator is always 1/4. From part (a) the numerator is

$$\frac{\binom{n-2}{k-c}}{2^n}$$

and so the answer is

$$\frac{\binom{n-2}{k-c}}{2^{n-2}}$$

Problem 10-3. You are given a wheel (like a roulette wheel) which contains the numbers from 1 to $2n$ arranged in some way. The wheel has a spinner and on each spin you get the two numbers at the opposite ends of the spinner. How would you arrange the numbers on the wheel to maximize the expected value of:

(a) The sum of the numbers chosen?

It makes no difference how we arrange the numbers. The expected value is the same. Why? If we label the ends of the spinner we can define 2 random variables X and Y. X is the number pointed to by one end of the spinner and Y is the other number. We wish to maximize $E(X+Y)$. But since $E(X+Y) = E(X) + E(Y)$ and $E(X) = E(Y) = (2n+1)/2$, all arangements give the same expected value of $2n+1$.

(b) The product of the numbers chosen?

The best way to do this is to ensure that 1 is opposite 2, 3 is opposite 4, ..., and $2n-1$ is opposite $2n$. Let's call this arrangement A^*. This will maximize the expected value of the product.

Consider any arrangement that isn't the one described above. It therefore must contain an odd i that isn't paired up with $i+1$. Pick the smallest such i and call it k. Say k is opposite p and $k+1$ is opposite q. We know that p and q are bigger than k (by choosing k to be the smallest such number we are assured that all numbers smaller than k are "properly" paired up). Now consider the new arrangement where k is paired with $k+1$ and p is paired with q. It has a better expected product than the previous arrangement.

To prove this note that since all pairs are equally likely when we turn the spinner, it suffices to show that the sum of the products of the numbers (denoted by S) is larger in the second arrangement.

$$\begin{aligned}
S_2 - S_1 &= k(k+1) + pq - kp - (k+1)q \\
&= k^2 + k + pq - kp - kq - q \\
&= k^2 + (1-p-q)k + (p-1)q \\
&= (k-q)(k-(p-1))
\end{aligned}$$

Thus $S_2 - S_1$ is greater than zero if k is smaller than both q and $p-1$. But this is true because of our choice of k. Thus the second arrangement is better.

Now we can transform any arrangement that isn't equal to A^* by repeatedly applying the following operation:

"Find the smallest odd i that isn't paired up with $i+1$. If i is paired with p and $i+1$ with q then pair i with $i+1$ and p with q"

Each time we perform this operation, we increase the expected value of the product. Thus A^* must be the best arrangement since it is the only arrangement where we cannot perform this operation.

Problem 10-4. Let S be a sample space. Say random variable $X : S \to \mathbb{N}$ has distribution

$$X = \begin{cases} 2 & \text{with probability } 1/4, \\ 3 & \text{with probability } 1/4, \\ 5 & \text{with probability } 1/2; \end{cases}$$

and random variable $Y : S \to \mathbb{N}$ has distribution

$$Y = \begin{cases} 1 & \text{with probability } 2/3, \\ 4 & \text{with probability } 1/3; \end{cases}$$

where X and Y are independent.

Let random variable $Z : S \to \mathbb{N}$ be given by $Z = \max(X, Y)$.

(a) What is the expected value of Z?

We first compute the values of Z along with their associated probabilities. Note that $P(X = x \;\&\; Y = y) = P(X = x)P(Y = y)$ since X and Y are independent.

X	Y	Z	Z^2	Pr
2	1	2	4	$\frac{1}{6}$
3	1	3	9	$\frac{1}{6}$
5	1	5	25	$\frac{1}{3}$
2	4	4	16	$\frac{1}{12}$
3	4	4	16	$\frac{1}{12}$
5	4	5	25	$\frac{1}{6}$

Now

$$E(Z) = 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 5 \cdot \frac{1}{3} + 4 \cdot \frac{1}{12} + 4 \cdot \frac{1}{12} + 5 \cdot \frac{1}{6} = 4$$

(b) What is the variance of Z? Recall that $\mathrm{Var} Z = E(Z^2) - E(Z)^2$. So we first compute $E(Z^2)$ using the data from the above table

$$E(Z^2) = 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 25 \cdot \frac{1}{3} + 16 \cdot \frac{1}{12} + 16 \cdot \frac{1}{12} + 15 \cdot \frac{1}{6} = \frac{52}{3}$$

Finally $\mathrm{Var}(Z) = 52/3 - 16 = \frac{4}{3}$.

Problem 10-5. We are given a random vector of n distinct numbers. We then determine the maximum of these numbers by the following procedure:

Pick the first number. Call it the "current maximum". Go through the rest of the vector (in order) and each time we find a number x that exceeds our "current maximum", we update the "current maximum" with x.

What is the expected number of times we update the "current maximum"?

Let's fix the n numbers we are given. We can assume that we are given a random permutation of the n numbers.

Let's define random variables $X_i (1 \leq i \leq n)$ by

$$X_i = \begin{cases} 1 & \text{if element } i \text{ is the maximum} \\ & \text{of } V[1], \ldots, V[i] \\ 0 & \text{otherwise} \end{cases}$$

Note that the number of times we update the current maximum is precisely $X_1 + \ldots + X_n$. Since expectation is a linear operator, we can compute $E(X_1 + \ldots + X_n)$ buy finding $E(X_i)$ for all i and summing them up. Note also that since X_i is an "indicator" variable, we only have to find $P(X_i = 1)$. $X_i = 1$ if element i is the largest of $V[1], \ldots, V[i]$. In a random permutation, this happens with probability $\frac{1}{i}$. Why? If you take i distinct numbers and randomly permute them, the probability that the largest one occupies the last (or any) position is $1/i$. If we note that

$$P(X_i = 1) = \sum_{n_1, \ldots, n_i} P(X_i = 1 \mid \text{first } i \text{ elements are } n_1, \ldots, n_i) \cdot P(\text{first } i \text{ elements are } n_1, \ldots, n_i)$$

and further observe that the first term is $1/i$ we can conclude that $P(X_i = 1) = 1/i$. Thus

$$\begin{aligned} E(X) &= \sum_i P(X_i = 1) \\ &= \sum_{i=1}^{n} 1/i \\ &= H_n \approx \ln n \end{aligned}$$

Recall that H_n is the nth Harmonic number.

Problem 10-6. **(a)** Show that if random variables X and Y are independent then $\mathrm{Cov}(X, Y) = 0$

First we show that if X and Y are independent, then $E(XY) = E(X)E(Y)$.

$$E(XY) = \sum_{x,y} xyP(X = x \;\&\; Y = y)$$

$$\begin{aligned}
&= \sum_x \sum_y xyP(X = x \;\&\; Y = y) \\
&= \sum_x \sum_y xyP(X = x)P(Y = y) \\
&= \sum_x xP(X = x) \sum_y yP(Y = y) \\
&= \sum_x xP(X = x)E(Y) \\
&= E(Y) \sum_x xP(X = x) \\
&= E(X)E(Y)
\end{aligned}$$

Since $\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$, $\mathrm{Cov}(X, Y) = 0$.

(b) Show that having $\mathrm{Cov}(X, Y) = 0$ does not necessarily mean that X and Y are independent.

Let (X, Y) have joint probability given by the table below:

X	Y	P
-1	1	$\frac{1}{3}$
0	0	$\frac{1}{3}$
1	1	$\frac{1}{3}$

Note that X and Y are not independent. $P(X = 1 \;\&\; Y = 1) = \frac{1}{3} \neq \frac{2}{9} = P(X = 1)P(Y = 1)$. But $E(X) = 0$ and $E(XY) = 0$ (since $XY = X$). Thus $\mathrm{Cov}(X, Y) = 0$.

(c) Let $X_1, \ldots, X_n$ be random variables. Prove that

$$\mathrm{Var}(X_1 + \ldots + X_n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

$$\begin{aligned}
\mathrm{Var}(X_1 + \ldots + X_n) &= E\left[X_1 + \ldots + X_n)^2\right] - E\left[X_1 + \ldots + X_n)\right]^2 \\
&= E\left[\sum_i X_i^2 + (\sum_{i<j} 2X_iX_j)\right] - \left[\sum_i E(X_i)^2 + \sum_{i<j} 2E(X_i)E(X_j)\right] \\
&= \sum_i E(X_i^2) + \sum_{i<j} 2E(X_iX_j) - \sum_i E(X_i)^2 - \sum_{i<j} 2E(X_i)E(X_j) \\
&= \sum_i E(X_i^2) - E(X_i)^2 + \sum_{i<j} 2(E(X_iX_j) - E(X_i)E(X_j)) \\
&= \sum_i \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)
\end{aligned}$$