

## Solution to HW 4, Problem 1

### COS 340 - Spring 2012

Consider two hypothetical students Alpha and Beta. Suppose each of them takes 36 courses.  $X_i$  is an indicator random variable that is 1 if Alpha gets an A in the  $i_{th}$  course. Similarly,  $Y_i$  is an indicator random variable that is 1 if Beta gets an A in the  $i_{th}$  course. Let  $X = \sum_{i=1}^{36} X_i$  and  $Y = \sum_{i=1}^{36} Y_i$ . Suppose that  $P[X_i = 1] = .8$  and  $P[Y_i = 1] = .6$  for all  $i$ . Further, assume that all the  $X_i$  and  $Y_i$  variables are independent of each other.

1. Compute  $E[X - Y]$  and  $Var[X - Y]$ .

$$\begin{aligned} E[X - Y] &= E[X] - E[Y] = E\left[\sum_{i=1}^{36} X_i\right] - E\left[\sum_{i=1}^{36} Y_i\right] = \sum_{i=1}^{36} E[X_i] - \sum_{i=1}^{36} E[Y_i] = \\ &= \sum_{i=1}^{36} E[X_1] - \sum_{i=1}^{36} E[Y_1] = 36 \times .8 - 36 \times .6 = 7.2. \end{aligned}$$

by linearity of expectation and because all the  $X_i$  are identically distributed and the  $Y_i$  are identically distributed.

Note that  $E[XY] = E[X]E[Y]$  because  $X$  and  $Y$  are independent of each other, and that  $Var\left[\sum_{i=1}^{36} X_i\right] = \sum_{i=1}^{36} Var[X_i]$  because the  $X_i$  are independent of each other. The same goes for the variance of the  $Y_i$ . Finally, note that, because the  $X_i$  are identically distributed and the  $Y_i$  are identically distributed, that  $Var[X_i] = Var[X_1]$  for all  $i$  for both the  $X_i$  and the  $Y_i$ .

$$\begin{aligned} Var[X - Y] &= E[(X - Y)^2] - E[X - Y]^2 = \\ &= E[X^2 - 2XY + Y^2] - E[X]^2 + 2E[X]E[Y] - E[Y]^2 = \\ &= E[X^2] - 2E[XY] + E[Y^2] - E[X]^2 + 2E[X]E[Y] - E[Y]^2 = \\ &= E[X^2] - 2E[X]E[Y] + E[Y^2] - E[X]^2 + 2E[X]E[Y] - E[Y]^2 = \\ &= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 = Var[X] + Var[Y] = Var\left[\sum_{i=1}^{36} X_i\right] + Var\left[\sum_{i=1}^{36} Y_i\right] = \\ &= \sum_{i=1}^{36} Var[X_i] + \sum_{i=1}^{36} Var[Y_i] = 36 \times Var[X_1] + 36 \times Var[Y_1] = \\ &= 36 \times (E[X_1^2] - E[X_1]^2 + E[Y_1^2] - E[Y_1]^2) = \\ &= 36 \times ((.8 \times 1^2) - (.8 \times 1)^2 + (.6 \times 1^2) - (.6 \times 1)^2) = 14.4 \end{aligned}$$

Use the Chebyshev inequality to give an upper bound on  $P[X \leq Y]$ . (i.e. the probability that Alpha gets at most as many A's as Beta gets.)

Let  $W \sim X - Y$  and let  $\sigma = \sqrt{Var[W]}$ . Then, by Chebyshev, we have:

$$\begin{aligned} P[|W - E[W]| \geq (E[W]/\sigma)\sigma] &\leq (\sigma/E[W])^2 \\ \rightarrow P[W - E[W] \geq (E[W]/\sigma)\sigma] + P[E[W] - W \geq (E[W]/\sigma)\sigma] &\leq (\sigma/E[W])^2 \\ \rightarrow P[E[W] - W \geq E[W]] &\leq (\sigma/E[W])^2 - P[E[W] - W \geq (E[W]/\sigma)\sigma] \\ \rightarrow P[-W \geq 0] &\leq (\sigma/E[W])^2 - P[E[W] - W \geq (E[W]/\sigma)\sigma] \\ \rightarrow P[Y - X \geq 0] &\leq (\sigma/E[W])^2 - P[E[W] - W \geq (E[W]/\sigma)\sigma] \end{aligned}$$

$$\rightarrow P[X - Y \leq 0] \leq (\sigma/E[W])^2 - P[E[W] - W \geq (E[W]/\sigma)\sigma]$$

$$\rightarrow P[X \leq Y] \leq (\sigma/E[W])^2 - P[E[W] - W \geq (E[W]/\sigma)\sigma]$$

$$\rightarrow P[X \leq Y] \leq (\sigma/E[W])^2 = (\sqrt{\text{Var}[X - Y]}/E[X - Y])^2 = (14.4/51.84) = .278$$

$$\rightarrow P[X \leq Y] \leq .278.$$

2. Use Chernoff bounds for  $X$  to give an upper bound on  $P[X \leq .6 \times 36]$ . (i.e. the probability that Alpha gets an A in at most 60% of the 36 courses.)

$$P[X \leq (1 - \delta)E[X]] \leq e^{-\delta^2 E[X]/2}$$

$$.6 \times 36 = (1 - \delta)E[X] \rightarrow \delta = 1 - .6 \times 36/E[X] = 1 - .6/.8 = .25$$

$$\rightarrow P[X \leq .6 \times 36] \leq e^{-.25^2 \times .8 \times 36/2} = .407$$

$$\rightarrow P[X \leq .6 \times 36] \leq .407.$$

3. Let  $Z_i = (1 - X_i)$  and let  $Z = \sum_{i=1}^{36} Z_i$ . Apply Chernoff bounds to get an upper bound on  $P[Z \geq .4 \times 36]$ .

$$P[Z \geq c \times E[Z]] \leq e^{-(c \ln c - c + 1) \times E[Z]}$$

$$.4 \times 36 = c \times E[Z] \rightarrow c = .4 \times 36/E[Z].$$

$$E[Z] = E\left[\sum_{i=1}^{36} (1 - X_i)\right] = E\left[36 - \sum_{i=1}^{36} X_i\right] = 36 - E\left[\sum_{i=1}^{36} X_i\right] =$$

$$36 - \sum_{i=1}^{36} E[X_i] = 36 - 36 \times .8 = 7.2.$$

$$P[Z \geq .4 \times 36] \leq e^{-(2 \ln 2 - 2 + 1) \times E[Z]} = .062 \rightarrow P[Z \geq .4 \times 36] \leq .062$$

## Solution to HW 4, Problem 2

### COS 340 - Spring 2012

$n$  men go to a party and check their hats in. When they return, the hats are returned to them in random order. Let  $X$  be the number of men who get their original hats back.

1. Compute  $E[X]$  and  $Var[X]$ .

Let  $X_i$  be the event that the  $i_{th}$  man gets his hat back. Specifically, let  $X_i$  be an indicator random variable that is 1 if the  $i_{th}$  man gets his hat back and 0 otherwise. Then, by linearity of expectation, for  $E[X]$  we have:

$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = n \times 1/n = 1.$$

For variance, we do a similar calculation:

$$Var[X] = Var\left[\sum_{i=1}^n X_i\right] = E\left[\left(\sum_{i=1}^n X_i\right)^2\right] - E\left[\sum_{i=1}^n X_i\right]^2$$

Note that  $\left(\sum_{i=1}^n X_i\right)^2 = \sum_{i=1}^n \sum_{j=1}^n X_i \times X_j$  and that, therefore:

$$E\left[\left(\sum_{i=1}^n X_i\right)^2\right] = E\left[\sum_{i=1}^n \sum_{j=1}^n X_i \times X_j\right] = \sum_{i=1}^n \sum_{j=1}^n E[X_i \times X_j]$$

by linearity of expectation.

Now, we compute  $E[X_i \times X_j]$  for when  $i = j$  and for when  $i \neq j$ :

$$E[X_i \times X_i] = 1/n \times (1)(1) + (n-1)/n \times (0)(0) = 1/n$$

$$E[X_i \times X_j] = \frac{1}{n(n-1)} \times (1)(1) + 2 \times \frac{n-2}{n(n-1)} (1)(0) + \frac{n^2-3n+n^2}{n(n-1)} \times (0)(0) = \frac{1}{n(n-1)}$$

So, computing our sum over  $E[X_i \times X_j]$ , we have:

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^n E[X_i \times X_j] &= n \times E[X_i \times X_i] + (n^2 - n) \times E[X_i \times X_j] \text{ for } i \neq j \\ &= n(1/n) + (n(n-1))/(n(n-1)) = 2, \end{aligned}$$

which brings us to our final expression:

$$Var[X] = \sum_{i=1}^n \sum_{j=1}^n E[X_i \times X_j] - E\left[\sum_{i=1}^n X_i\right]^2 = 2 - 1^2 = 1.$$

2. Use the Chebyshev inequality to give an upper bound on  $P[X \geq k]$ .

Let  $\sigma = \sqrt{Var[X]}$ :

$$P[|X - E[X]| \geq c \times \sigma] \leq 1/c^2$$

$$\rightarrow P[X - E[X] \geq c \times \sigma] + P[E[X] - X \geq c \times \sigma] \leq 1/c^2$$

$$\rightarrow P[X \geq c \times \sigma + E[X]] \leq 1/c^2 - P[E[X] - X \geq c \times \sigma]$$

$$\rightarrow P[X \geq c \times \sigma + E[X]] \leq 1/c^2.$$

Solving for  $c$  in terms of  $k$  we get:

$$k = c \times \sigma + E[X] \rightarrow c = (k - E[X])/\sigma.$$

Bringing us to the final answer:

$$P[X \geq k] \leq Var[X]/(k - E[X])^2 = 1/(k-1)^2$$

## Solution to HW 4, Problem 3

### COS 340 - Spring 2012

A scientist is running an experiment on growing red and yellow spores in a petri dish. At any time moment a yellow spore may divide with an equal probability into either (2 yellow and 2 red) or (3 yellow and 1 red) spores. A red spore may divide with an equal probability into either (5 red) or (4 yellow and 1 red) spores. The experiment starts with 5000 red and 3000 yellow spores in the dish.

The scientist terminates the experiment by killing all the spores in the dish once the red spore count reaches 25000. What is the expected number of yellow spores just before the experiment is terminated?

In order to figure this out, we must first formalize this process as a Martingale process,  $X_t$ .

Let  $X_t$  = (number of red spores) - (number of yellow spores) after the  $t_{th}$  spore division. Note that  $t$  here does not represent time directly, but rather represents the number of divisions since the process began.

In order to show that  $X_t$  is a Martingale process, we need to show that, for all  $t$ ,

$E[X_{t+1}|X_0, \dots, X_t] = X_t$ . Recall that  $X_t$  denotes the state of the system just after the  $t_{th}$  spore division.

Two events can occur just after  $X_t$ : either a red spore can choose to divide, or a yellow spore can choose to divide. These are the only possibilities and so, by conditioning, we get:

$$\begin{aligned} E[X_{t+1}|X_0, \dots, X_t] &= E[X_{t+1}|X_0, \dots, X_t, \text{yellow spore divides just after } t]P[\text{yellow spore divides just after } t] \\ &+ E[X_{t+1}|X_0, \dots, X_t, \text{red spore divides just after } t]P[\text{red spore divides just after } t]. \end{aligned}$$

But, we can note a symmetry in the expectations.

#### 1) If a yellow spore divides, we have:

$$\begin{aligned} E[X_{t+1}|X_0, \dots, X_t, \text{yellow spore divides just after } t] &= E[X_{t+1}|X_t, \dots] \text{ because } X_{t+1} \text{ is clearly independent of } X_0, \dots, X_{t-1} \text{ given } X_t \\ &= X_t + E[\text{red spores generated from yellow division} - \text{yellow spores generated from yellow division}] \\ &= X_t + E[\text{red spores generated from a yellow division}] - E[\text{yellow spores generated from a yellow division}]. \end{aligned}$$

$$E[\text{number of spores after yellow division}] = 1/2(1Y + 2R) + 1/2(2Y + 1R) = 1.5Y + 1.5R$$

$$\rightarrow E[\text{yellow spores generated}] = 1.5, E[\text{red spores generated}] = 1.5.$$

$$\begin{aligned} &\rightarrow E[X_{t+1}|X_0, \dots, X_t, \text{yellow spore divides just after } t] \\ &= X_t + E[\text{red spores generated from a yellow division}] - E[\text{yellow spores generated from a yellow division}] \\ &= X_t + 1.5 - 1.5 = X_t. \end{aligned}$$

#### 2) If a red spore divides, we have:

$$\begin{aligned} E[X_{t+1}|X_0, \dots, X_t, \text{red spore divides just after } t] &= E[X_{t+1}|X_t, \dots] \text{ because } X_{t+1} \text{ is clearly independent of } X_0, \dots, X_{t-1} \text{ given } X_t \\ &= X_t + E[\text{red spores generated from red division} - \text{yellow spores generated from red division}] \\ &= X_t + E[\text{red spores generated from a red division}] - E[\text{yellow spores generated from a red division}]. \end{aligned}$$

$$E[\text{number of spores after red division}] = 1/2(4Y + 0R) + 1/2(0Y + 4R) = 2Y + 2R$$

$$\rightarrow E[\text{yellow spores generated}] = 2, E[\text{red spores generated}] = 2.$$

$$\begin{aligned} &\rightarrow E[X_{t+1}|X_0, \dots, X_t, \text{red spore divides just after } t] \\ &= X_t + E[\text{red spores generated from a red division}] - E[\text{yellow spores generated from a red division}] \end{aligned}$$

$$= X_t + 2 - 2 = X_t.$$

**Plugging these into our original equation:**

$$\begin{aligned} & E[X_{t+1} | X_0, \dots, X_t] \\ &= E[X_{t+1} | X_0, \dots, X_t, \text{yellow spore divides just after } t] P[\text{yellow spore divides just after } t] \\ &+ E[X_{t+1} | X_0, \dots, X_t, \text{red spore divides just after } t] P[\text{red spore divides just after } t] \end{aligned}$$

$$\begin{aligned} &= X_t P[\text{yellow spore divides just after } t] + X_t P[\text{red spore divides just after } t] \\ &= X_t \text{ because } P[\text{yellow}] + P[\text{red}] \text{ sum to one.} \end{aligned}$$

Now, finally having shown that this is indeed a Martingale, we choose our stopping criterion  $T$  to be the point at which the number of red spores is equal to 25000. By the optimal stopping theorem, we have:

$$\begin{aligned} & E[X_T] = X_0 = 2000 \text{ and combine this with the definition of } X_T \\ & \rightarrow E[X_T] = E[\# \text{ red spores at interval } T] - E[\# \text{ yellow spores at interval } T] \\ &= 25000 - E[\# \text{ yellow spores at } T] = 2000 \\ & \rightarrow E[\# \text{ yellow spores at } T] = 25000 - 2000 = 23000. \end{aligned}$$

## Solution to HW 4, Problem 4

### COS 340 - Spring 2012

In order to implement a counter capable of counting from 1 to  $N$ , at least  $\log N$  bits are needed. In some cases, allowing this number of bits is too expensive. In this question, you will see how to implement a “probabilistic counter” - a counter that counts with small errors, and uses only  $\theta \log \log N$  bits.

The counter is defined as follows:

- Initially the counter is zero.
- An increment operation is performed as follows: If the current value of the counter is  $j$ , increase the counter value by 1 with probability  $1/2^j$ , and leave the counter in its current state with probability  $1 - 1/2^j$ .

Let  $C_k$  be a random variable denoting the value of the counter after  $k$  increment operations have occurred.

**Part A** Suppose that  $k$  increment operations have been performed. Our first goal is to show that the actual number of increment operations is well approximated by  $2^{C_k}$ , at least in expectation. We will prove:  $E[2^{C_k}] = k + 1$ .

**Step 1:** Define, for  $i > 0$ , the random variable  $X_i := 2^{C_i} - 2^{C_{i-1}}$ . Show that for any value of  $t$ :  $E[X_i | C_{i-1} = t] = 1$ .

Note that given  $C_{i-1} = t$ ,  $C_i$  can only take on two values. Namely, if an increase occurs when incrementing the previous value (happens with probability  $1/2^t$ ), then  $C_i = t + 1$ . Otherwise, if an increase does not occur (happens with probability  $1 - 1/2^t$ ), then  $C_i = t$ . We can now compute the expectation directly.

$$E[X_i | C_{i-1} = t] = (2^{t+1} - 2^t)P[C_i = t + 1 | C_{i-1} = t] + (2^t - 2^t)P[C_i = t | C_{i-1} = t] \\ = (2^{t+1} - 2^t)(1/2^t) + (0)(1 - 1/2^t) = (2 - 1) = 1.$$

**Step 2:** Conclude that  $E[X_i] = 1$

$$E[X_i] = \sum_{t=0}^{\infty} E[X_i | C_{i-1} = t] P[C_{i-1} = t] = 1 \sum_{t=0}^{\infty} P[C_{i-1} = t] = 1.$$

Note that  $P[C_{i-1} = t]$  is a probability distribution over  $t$  and therefore:  $\sum_{t=0}^{\infty} P[C_{i-1} = t] = 1$ .

**Step 3:** Conclude that  $E[2^{C_k}] = k + 1$ .

Note the following recursive formulation:

$$E[2^{C_k}] = E[X_k] + E[2^{C_{k-1}}] = 1 + E[2^{C_{k-1}}]$$

And, therefore, we have:

$$E[2^{C_k}] = 1 + \dots + 1 (\text{repeated } k \text{ times}) + E[2^{C_0}] = k + 1.$$

**Part B** Our next goal is to calculate the variance of  $2^{C_k}$ .

**Step 1:** In order to calculate the variance, we must first compute  $E[(2^{C_k})^2] = E[4^{C_k}]$ . Define, for  $i > 0$ , the random variable:  $Y_i = 4^{C_i} - 4^{C_{i-1}}$ . Show that  $E[Y_i | C_{i-1} = t] = 3 \cdot 2^t$ .

Again, note that the value of  $C_i$  conditioned on  $C_{i-1} = t$  can be either  $t + 1$  with probability  $1/2^t$  or  $t$  with probability  $1 - 1/2^t$ . Thus, computing the expectation we get:

$$E[Y_i | C_{i-1} = t] = (4^{t+1} - 4^t)(1/2^t) + (4^t - 4^t)(1 - 1/2^t) = (2^{2t+2} - 2^{2t})(1/2^t) + 0$$

$$= 2^{t+2} - 2^t = 4 \cdot 2^t - 1 \cdot 2^t = 3 \cdot 2^t.$$

**Step 2:** Conclude that  $E[Y_i] = 3i$ .

Using our result from step 3 in part A above, we have:

$$E[Y_i] = \sum_{t=0}^{\infty} E[Y_i | C_{i-1} = t] P[C_{i-1} = t] = 3 \cdot \sum_{t=0}^{\infty} 2^t P[C_{i-1} = t] = 3 \cdot E[2^{C_{i-1}}] = 3 \cdot i.$$

**Step 3:** Conclude that the variance of  $2^{C_k}$  is equal to  $3/2 \cdot k \cdot (k+1) + 1 - (k+1)^2$ .

Again, note the following recursive relation; note we use the formula  $\sum_{i=1}^n i = n(n+1)/2$ :

$$E[4^{C_k}] = E[Y_k] + E[4^{C_{k-1}}] = 3i + 3(i-1) + \dots + 3(2) + 3(1) + E[4^{C_0}]$$

$$\rightarrow E[(2^{C_k})^2] = \frac{3k(k+1)}{2} + 1$$

And recall that  $E[2^{C_k}] = (k+1)$ . So, as our final answer, we have:

$$Var[2^{C_k}] = E[(2^{C_k})^2] - E[2^{C_k}]^2 = \frac{3k(k+1)}{2} + 1 - (k+1)^2.$$

**Part C:** The last goal is to show concentration around the mean. Suppose we have 10 such counters. Let  $T$  be the average value of  $2^{C_k}$  over all counters. Show that with probability at least 80%,  $T$  is in the range  $(k+1) \cdot (1 \pm .5)$ .

Hint: Compute the variance  $T$ .

First, compute expected value and variance of  $T$ . Let  $V_i$  denote the value of the  $i_{th}$  counter after  $k$  iterations. We will assume the counters are independent.

$$E[T] = E[\sum_{i=1}^{10} V_i / 10] = 1/10 E[\sum_{i=1}^{10} V_i] = 1/10 \sum_{i=1}^{10} E[V_i] = 1/10 \cdot 10 \cdot E[V_1] = E[V_1] = k+1. \quad Var[T] =$$

$$Var[\sum_{i=1}^{10} V_i / 10] = \frac{1}{100} \sum_{i=1}^{10} Var[V_i] = \frac{1}{10} (3/2 \cdot k \cdot (k+1) + 1 - (k+1)^2)$$

$$= \frac{1}{10} (3/2 k^2 + 3/2 k + 1 - k^2 - 2k - 1) = \frac{1}{10} (1/2 k^2 - 1/2 k) = (k^2 - k)/20.$$

Note the use of the independence of the counters and the fact that  $Var[aX] = a^2 Var[X]$  above. Now, armed with the expected value and the variance, we can apply Chebyshev's inequality.

$$P[|T - E[T]| \geq c \cdot \sigma] \leq 1/c^2$$

$$\rightarrow P[|T - (k+1)| \geq c \cdot \sigma] \leq 1/c^2$$

$$\rightarrow P[T - (k+1) \geq c \cdot \sigma] + P[(k+1) - T \geq c \cdot \sigma] \leq 1/c^2$$

$$\rightarrow P[T \geq c \cdot \sigma + (k+1)] + P[T \leq (k+1) - c \cdot \sigma] \leq 1/c^2.$$

Now we must pick the value of  $c$  such that the right-hand sides of the inequalities enclosed in the probabilities are  $(k+1)(1.5)$  and  $(k+1)(.5)$ . We can do this by setting them equal and solving for  $c$ :

For the upper bound:

$$c \cdot \sigma + (k+1) = (k+1)(3/2) \rightarrow c \cdot \sigma = (3/2)(k+1) - (k+1)$$

$$\rightarrow c = (k+1)/(2\sigma)$$

This value of  $c$ , happily, works for the lower bound as well:

$$(k+1) - c \cdot \sigma = (k+1)(1/2) \rightarrow -c \cdot \sigma = (-1/2)(k+1)$$

$$\rightarrow c = (k+1)/(2\sigma)$$

Thus, if  $P[|T - E[T]| \geq (k+1)/2]$  then  $T$  is within the range  $(k+1)(1 \pm .5)$  and, using the Chebyshev bound, we get:

$$P[|T - E[T]| \geq (k+1)/2] = P[T \text{ within range } (k+1)(1 \pm .5)] \leq ((2\sigma)/(k+1))^2$$

$$= (4Var[T])/(k+1)^2 = \frac{4(k^2 - k)}{20(k^2 + 2k + 1)} = \frac{1}{5} \cdot \frac{(k^2 - k)}{(k^2 + 2k + 1)} \leq \frac{1}{5}$$

$$\rightarrow P[T \text{ within range } (k+1)(1 \pm .5)] \leq \frac{1}{5}.$$

Thus, with probability at least 80%,  $T$  is within the range  $(k+1)(1 \pm .5)$ .