

## Problem 1

**b** is the correct answer.

We want the probability bound  $2Me^{2\epsilon^2 N}$  to be at most .03. So, we can solve for

$$2Me^{-2\epsilon^2 N} = .03$$

$$2e^{-2(.05)^2 N} = .03$$

$$e^{-2(.05)^2 N} = .015$$

$$N = 839.94$$

So for any  $N$  less than this value, our probability bound will be greater than .03. So the answer is 1000. We can check this by plugging in 500 and 1000 into  $2Me^{2\epsilon^2 N}$ . At 500, we get .164, which is too big. At 1000, we get .013, which is below the desired bound. Thus 1000 is the correct answer.

## Problem 2

**c** is the correct answer.

We want the probability bound  $2Me^{2\epsilon^2 N}$  to be at most .03. So, we can solve for

$$2Me^{-2\epsilon^2 N} = .03$$

$$20e^{-2(.05)^2 N} = .03$$

$$e^{-2(.05)^2 N} = .0015$$

$$N = 1300.46$$

So for any  $N$  less than this value, our probability bound will be greater than .03. So the answer is 1500. We can check this by plugging in 1000 and 1500 into  $2Me^{2\epsilon^2 N}$ . At 1000, we get .135, which is too big. At 1500, we get .011, which is below the desired bound. Thus 1500 is the correct answer.

## Problem 3

**d** is the correct answer.

We want the probability bound  $2Me^{2\epsilon^2 N}$  to be at most .03. So, we can solve for

$$2Me^{-2\epsilon^2 N} = .03$$

$$200e^{-2(.05)^2 N} = .03$$

$$e^{-2(.05)^2 N} = .00015$$

$$N = 1760.98$$

So for any  $N$  less than this value, our probability bound will be greater than .03. So the answer is 2000. We can check this by plugging in 1500 and 2000 into  $2Me^{2\epsilon^2 N}$ . At 1500, we get .111, which is too big. At 2000, we get .009, which is below the desired bound. Thus 2000 is the correct answer.

## Problem 4

**b** is the correct answer.

First of all, we know this is true because of the VC dimension + 1 theory. Secondly, for 5 points in  $\mathbb{R}^3$ , we get one of the following situations. One situation is if all 5 points are on the same plane. Then we obviously cannot shatter this, because we cannot even shatter 4 points in the  $\mathbb{R}^2$  case. Another situation is if 4 points are on the same plane and one is on a different plane. We obviously cannot shatter this either, since we could not shatter 4 points in  $\mathbb{R}^2$ . In our last situation, we can form a plane with 3 points that separates the 2 other points. Then, we have that the dichotomy where the 3 points in that plane are +1, the 1 point on one side of that plane is -1, and the 1 point on the other side of the plane is also -1 cannot be achieved with the Perceptron Model.

## Problem 5

**b** is the correct answer.

We have that if there is no break point,  $m_{\mathcal{H}}(N) = 2^N$ , and that if there is any break point,  $m_{\mathcal{H}}(N)$  is polynomial in  $N$ . So, we know that (i) and (ii) are actual growth functions, since they are the growth functions for positive rays and positive intervals and are of the form  $\sum_{i=0}^{k-1} \binom{N}{i}$  (polynomial). And we have that (v) is just  $2^N$ . Then we have that (iii) and (iv) are not polynomial in  $N$  and not  $2^N$ . So we have our answer.

## Problem 6

**c** is the correct answer.

We have that with the "2-intervals" learning model, we can have at most 2 distinct sets of positive points. This is enough to shatter 3 and 4 points, as it is impossible to have 3 distinct sets of positive points with just 3 or 4 points. But with 5 points, it is possible to have this: +1, -1, +1, -1, +1, and we cannot achieve this dichotomy with the "2-intervals" learning model. So the answer is 5.

## Problem 7

**c** is the correct answer.

We get to this answer the following way. There are  $\binom{N+1}{4}$  ways to place 2 distinct intervals (choosing 4 bounds in total), and  $\binom{N+1}{2}$  ways to place 1 distinct interval (choosing 2 bounds in total). Then there is just 1 dichotomy in which all the intervals are placed together so that all the points are negative.

## Problem 8

**d** is the correct answer.

**d** is the only option that is consistent with the "1-interval" and "2-interval" learning models - for the "1-interval" learning model we had a break point of 3, and for the "2-interval" learning model we had a break point of 5. This also makes sense in general, as given  $M$  intervals to play with, we can have at most  $M$  distinct sets of positive points. So we are going to break once we reach  $2M + 1$  points, because at this point we can have  $M + 1$  distinct sets of positive points as a possible dichotomy.

## Problem 9

**d** is the correct answer.

Basically solved this problem by doing the brute force way, trying each answer one by one by drawing it out. 1 point can clearly be shattered. 3 points can also clearly be shattered because a triangle is more powerful than a line. If we arrange all the points in a circle, it is clear that 5 and 7 points can be shattered as well. However, trying to shatter 9 points arranged in a circle fails (alternate +1/-1). So we have our answer.

## Problem 10

**e** is the correct answer.

Here, consider the number of dichotomies we can get on  $N$  points when we put the  $N$  points on a line. When we do this, the concentric circles act as intervals. It's tempting to say that the answer is **b**, since it seems the same as the "1-interval" model. But we can see that with concentric circles, we can shatter 3 points. So the breakpoint for concentric circles is greater than 3. So, let us now go through the options and see if any of them are viable. **a** has a breakpoint of 2, so that doesn't work. As we mentioned above, **b** has a breakpoint of 3, so that doesn't work. **c** also has a breakpoint less than 3, and **d** is not less than or equal to  $2^N$  for all values of  $N$ . So we have our answer.