HW3:

Q1: $2Me^{-2\epsilon^2 N} \leq 0.03$

M=1, ϵ=0.05; Find N:

$$2(0.05)^2 N \geq -\ln\left(\frac{0.03}{2}\right)$$

$$N \geq \frac{-\ln\left(\frac{0.03}{2}\right)}{2(0.05)^2} \geq 840$$

Find closest answer

[a] 500; wrong because less than 840;
[b] 1000 − 840 = 160
[c] 1500 − 840 = 660
[d] 2000 − 840 = 1160
[e] More examples needed; wrong bcz we need N.

Hence correct answer is **b** because it is the closest to our N.

Q2: $2Me^{-2\epsilon^2 N} \leq 0.03$

M=10; ϵ=0.05; Find N:

$$2(0.05)^2 N \geq -\ln\left(\frac{0.03}{2(10)}\right)$$

$$N \geq \frac{-\ln\left(\frac{0.03}{2(10)}\right)}{2(0.05)^2} \approx 1300.$$

Hence according to the warm above the correct answer is **c** bcz it is the closest.

Q3: $2Me^{-2\epsilon^2 N} \leq 0.03$

M=100; ϵ=0.05; Find N:

$$2(0.05)^2 N \geq -\ln\left(\frac{0.03}{2(100)}\right)$$

$$N \geq \frac{-\ln\left(\frac{0.03}{2(100)}\right)}{2(0.05)^2} \approx 1761$$

Hence according to the above warm, the correct answer is **d** because it is the closest.

Q4: The correct answer is **b** bcz in the case M seperating planes only a single point needs to be added in $\mathbb{R}^3$ to create a dichotomy that cannot be shattered hence making 5 the break point in $\mathbb{R}^3$ for this model.

Q5:

(i) $1+N$ is a possible growth function because it is a first degree polynomial and the growth function of positive rays as mentioned in lecture 5 slide 11.

(ii) $1 + N + \binom{N}{2} = 1 + N + \frac{N!}{2!(N-2)!}$,

$= 1 + N + \frac{N(N-1)(N-2)!}{2(N-2)!} = 1 + N + \frac{N^2 - N}{2}$ hence making it a second degree polynomial hence marking it a possible growth function.

(iii) $\sum_{i=1}^{1126} \binom{N}{i}$ is not in the form 2^N and has a maximum power N^{1126} hence making it not a polynomial hence marking it not a possible growth function.

(iv) 2^{1126} is not in the form 2^N and is not a polynomial hence it is not a possible growth function.

(v) 2^N it is a possible growth function because it is in the form 2^N.

Hence **b** is the correct answer as it is the only one that matches the above analysis.

(7) 5 is breakpoint from question b hence $m_H(N) = 2^N$

[a] Wrong because function does not equal to 2^N for $N \leq 5$:

N=1 should output 2 but

$\binom{1+1}{1} = 0$ hence wrong.

[b] Wrong because function does not equal to 2^N for $N \leq 5$:

N=3 should output 8 but

$\binom{4}{2} + 1 = 7$ hence wrong.

[c] Correct bcz for all $N \leq 5$ it is equal to 2^N. Also, it is correct it accounts for all disjoint two interval $\left(\binom{N+1}{2}\right)$ and for all overlapping interval $\left(\binom{N+1}{2}\right)$. Furthermore, the +1 accounts for 1 interval where no points are selected hence making this the correct growth function for five intervals.

[d] Wrong because function does not equal to 2^N for $N \leq 5$:

N=1 should output 2 but

$\binom{2}{2} + \binom{2}{1} + \binom{2}{0} + 1 = 4 \neq 2$

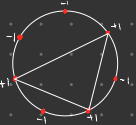
hence wrong.

[e] Wrong because we found the correct choice to be C.

(8) Start with largest number 9 and place them all on a circle(s) as we are looking at convex sets and classify the points:

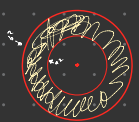
For 9 points we get four intervals, however, when triangle only has 3 intervals/sides hence No H can shatter 9 points.

For 7 points.



Here we have 3 intervals and hence we can shatter 7 points. as it can be depicted in three intervals. Even in the case of $\{+1, -1, +1, -1, +1, -1, +1\}$ we still only 3 intervals because as seen by the model of the points on the circle above, the last and first points are the same hence part of the same interval as the circle wraps around. Hence for 7 points we will never have more than 3 many ± 1 intervals thus shattering 7 points. As 9 cannot be shattered and 7 is the largest number that can be shattered the correct answer is **d**.

(10)



We know the equation of a circle is in the form $r^2 = x^2 + y^2$ however we only care about the distance of these points from the center and any point $x, y, z \in \mathbb{R}^3$ plane must be the following distance from the center; $a^2 \leq r^2 \leq b^2$ the yellow region. Now this can be simplified $\sqrt{a^2} \leq \sqrt{r^2} \leq \sqrt{b^2}$ $\rightarrow$ to $\rightarrow$ $a \leq r \leq b$ which is a single interval learning model and from lecture 5 slide 17 we know that the growth function of this model is $\binom{N+1}{2} + 1$ hence the correct answer is **b**.

(6)

6. In the two interval learning model we want a hypothesis that covers two intervals thus a break point would be when the distribution is covered in more than two intervals. As there are no ways to classify the following set of five points, {+1, -1, +1, -1, +1}, into two intervals, only 3 intervals, hence 5 points are our break point. Hence for 5 points there exists such a dichotomy that cannot be shattered by our hypothesis hence making 5 points the break point. Looking at all possible distributions of four points, there is no way to create three intervals as only zero, one, or two intervals can be created with four points then four points can be shattered hence so can three points hence making 5 the least number of breakpoints.

[a] Wrong because of the above analysis.
[b] Wrong because of the above analysis.
[c] Correct because of the above analysis.
[d] Wrong because of the above analysis.
[e] Wrong because of the above analysis.

(8)

8. In order to find the smallest breakpoint for the case of M intervals we need to find a set that requires M + 1 intervals as we cannot shatter a distribution that requires M + 1 intervals. Similar to question six we want to find how many points create such an interval. If we were to create a distribution with M many +1's each directly followed by a -1 then we would have created a set with 2M points as we have M many +1 points and M many -1 points. Also, with this distribution of 2M points we would have created M intervals. Thus to create M + 1 intervals we would just need to add one extra point that will create a new interval which would be +1 in this case as the last point in our sequence was -1 hence adding the extra interval hence serving as a breakpoint. As we only had to add one more point to generate that extra interval then the smallest break point is at 2M + 1. This model works as seen in the case of M = 2 in question six because
(2*2)+1 = 5 which was the breakpoint.

[a] Wrong because of the above analysis.
[b] Wrong because of the above analysis.
[c] Wrong because of the above analysis.
[d] Correct because of the above analysis.
[e] Wrong because of the above analysis.