1. $P[|E_{in}(g) - E_{out}(g)| > \varepsilon] \leq 2Me^{-2\varepsilon^2 N}$

$\varepsilon = 0.05 \rightarrow$ want $2Me^{-2(0.05)^2 \cdot N} \leq 0.03$

$M=1$: $e^{-2(0.05)^2 N} \leq \frac{0.03}{2}$

$-2 \cdot 0.05^2 N \leq \ln(0.015)$

$N \geq -\frac{\ln(0.015)}{2 \cdot 0.05^2}$

$N \geq 839.94$

$\boxed{\text{the least number of examples among the choices is 1000 [b]}}$

2. for $M = 10$: $2(10)e^{-2(0.05)^2 \cdot N} \le 0.03$

$$e^{-2(0.05)^2 \cdot N} \le \frac{0.03}{20}$$

$$-2(0.05)^2 N \le \ln(0.03/20)$$

$$N \ge -\frac{\ln(0.03/20)}{2(0.05)^2}$$

$$N \ge 1300.5$$

$\boxed{\text{the least number of examples required of the given choices is } 1500 \Rightarrow [c]}$

3. for $M = 100$: (same eq. as 1 and 2)

$$N \geq -\frac{\ln(0.03/200)}{2(0.05)^2}$$

$$N \geq 1761.0$$

the least number of examples required of the given choices is 2000 => [d]

4. The break point in $\mathbb{R}^2$ is 4, because there is always a line that separates any 3 points. In $\mathbb{R}^3$, this means that there is always a plane that separates any 4 points, because a plane in $\mathbb{R}^3$ is similar to a line in $\mathbb{R}^2$. So, adding one more point does not guaruntee that a plane will separate the 5 points. For example, if there are 4 points on a plane and a 5th one that is not on that plane, then there is no plane that can separate all possible labelings of the 5 points. answer: [b]

5. From lecture 6, we know that $m_H(N)$ is polynomial in N, and from lecture 5, we know that $m_H(N) \le 2^N$ and $m_H(N)$ is not polynomial in N only when $m_H(N)$ does not have a break point. So, the growth formulas that are polynomial in N or equal 2^N are valid.

(i) $1+N$ ✓ this is possible because it is linear in N and $\le 2^N$. In fact, we know that this is the growth function for positive rays.

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

✓ this is a possible growth function too because it is polynomial in N.

(iii) $\sum_{i=1}^{\lfloor\sqrt{N}\rfloor}\binom{N}{i}$ → This is not polynomial in N and not equal to 2^N.

This is not a valid growth function because it is not polynomial in N and not equal to 2^N (for a H with no break points).

$m_H(N) \le \sum_{i=0}^{\lfloor\sqrt{N}\rfloor}\binom{N}{i}$ this inequality will not hold for large N and small k.

(iv) $2^{\lfloor N/2 \rfloor}$ this growth function is not polynomial in N, so it is not valid. X

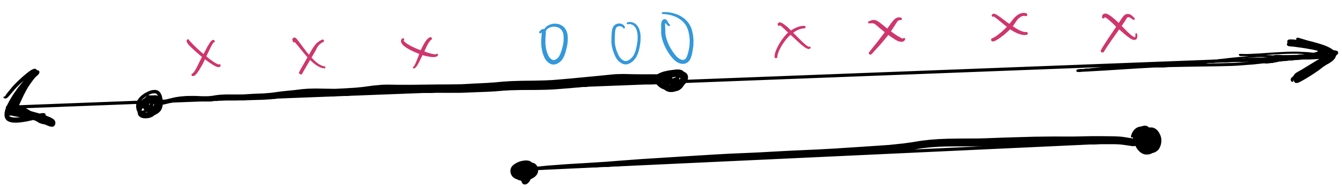
(v) 2^N this is the growth function for a a hypothesis with no break point, so this is a valid growth function. ✓

i, ii, and v are valid => $\boxed{[b]}$

6. $h: \mathbb{R} \to \{-1, +1\}$

$h(x) = 1$ if it is within two randomly chosen intervals

$h(x) = -1$ otherwise



we want to know the smallest number of points where there are no two intervals that can shatter any labeling of the points

try 3:

$o = +1$

$x = -1$

x	x	o	✓
x	o	x	✓
o	x	x	✓
x	o	o	✓
o	o	x	✓
o	x	o	✓
x	x	x	✓
o	o	o	✓

intervals outside of points

at least one interval covers all points

4 points also works because all arrangements will be similar to 3 points other than:

x o x o

In which case each o can correspond to one interval.

Any arrangement of 5 points cannot necessarily be shattered by any two intervals because: o x o x o

requires that each o corresponds to an interval, since the intervals are continuous, there cannot be a x in between two o's that are within the same interval. Thus, 5 is the smallest break point for the 2-intervals learning model. => $\boxed{[c]}$

7. $m_H(N)$ for "2-intervals"

Based on slide 11 of lecture 6,

$$m_H(N) \leq \sum_{i=0}^{k-1} \binom{N}{i} \quad \text{for } m_H(N) \text{ w/ break point } k$$

we have 3 cases for the 2-intervals:

we know from problem 6 that 4 is the maximum number of points that can be shattered by 2 intervals, so: the '_' are the possible spots

_ 1 _ 2 _ 3 _ 4 _ for the endpoints of the intervals and the numbers are the points.

<u>case 1</u>: intervals are not overlapping at all

$\Rightarrow$ 4 endpoints for total of 2 intervals:

$\binom{N+1}{4}$ $\rightarrow$ N+1 spots for N points

<u>case 2</u>: intervals overlap, so there are exactly 2 end points: $\binom{N+1}{2}$

<u>case 3</u>: all the points are outside both intervals, which is just one case (all points are -1)

final answer: $\boxed{\binom{N+1}{4} + \binom{N+1}{2} + 1} \Rightarrow [c]$

8. M-intervals case where point falls in any of M randomly chosen intervals

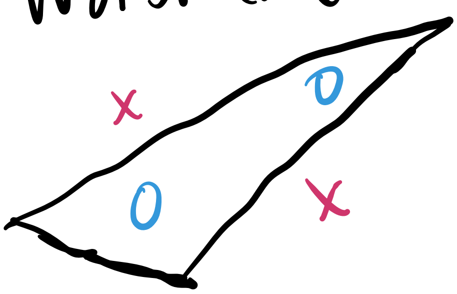
For the two intervals case, we found that the break point is 5 because if all 5 points were alternating in terms of classification (starting with +1), then it required 3 intervals. With M intervals, if we have 2M+1 points that are alternating in classification, (starting with +1), then M+1 intervals are required to shatter these points. Again, since intervals are continuous, two o's with a x in between must be in two <u>distinct</u> intervals. Thus, the smallest number of break points for M intervals is $\boxed{2M+1 \Rightarrow [d]}$

9. The largest number of points that can be shattered by this hypothesis set where +1 is inside a randomly chosen triangle.

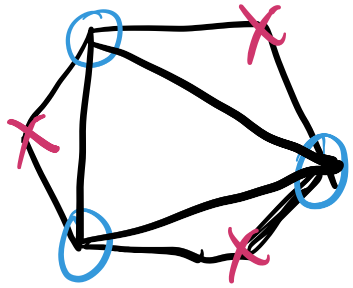
1, 3, 5, 7, 9 are options for max that can be shattered

<u>2 points</u> → The triangle can always enclose 0, 1, or 2 of the points because one side of the triangle can divide any two of the points (a line can always split 2 points) or it can be made large or small enough to have both/no points. So, 1 is not the maximum.

<u>4 points</u> → A triangle can also cut across any two sides of the quadrilateral formed by the 4 points to omit any arrangement of points. This means that 3 is also not the maximum. worst case scenario for 4 points (the ones that are supposed to be inside are diametrically opposite.

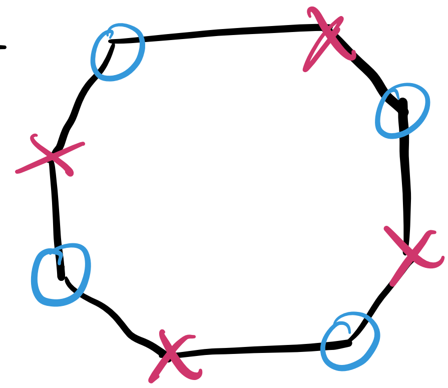


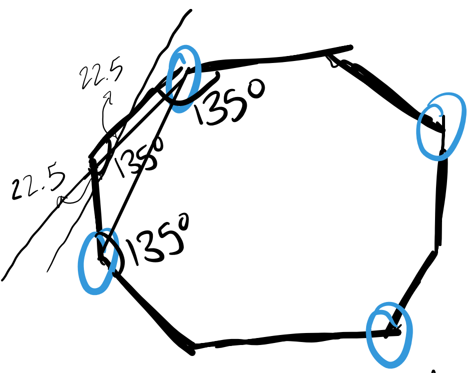
<u>6 points</u> → Again, we can show that the worst-case scenario for 6 points where the labels are alternating around a hexagon can still be shattered by a triangle. In this



case, the 3 +1 points can form the endpoints of the triangle.

8 points → Now, we attempt to show that alternating labels around an octogon can also be shattered by a triangle. There is no triangle that can enclose the O's here. In the worst-case scenario, where the

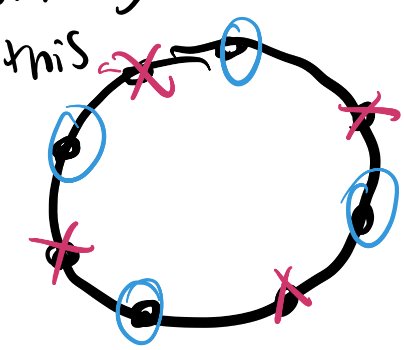




8 points are equidistant around a circle and alternatively labeled +1 and -1, the convex properties of a triangle do not allow for a valid triangle to only enclose the O's, it must enclose an adjacent X to do so, which means that our hypothesis set cannot shatter the hypothesis.

$x^2+y^2=r^2$

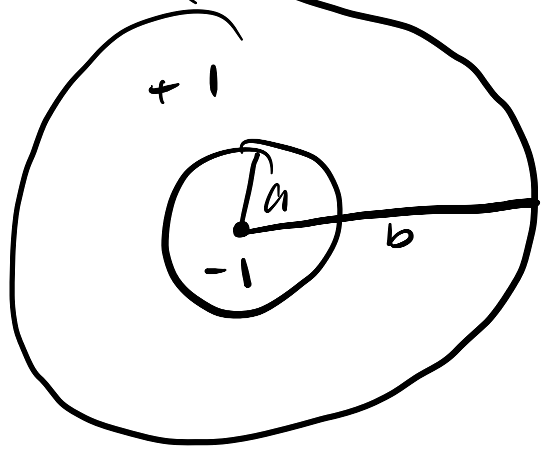
all 8 points satisfy this



So, 8 points cannot be shattered, but 6 points can, so we know the answer is

$\boxed{[d] \Rightarrow 7}$

10. $a^2 \le x_1^2 + x_2^2 \le b^2$



This is essentially labeling all the points inside the larger circle but outside the smaller circle as +1, and -1 otherwise. This is the same as labeling points inside the interval [a,b] as +1, and outside the interval as -1. The selection of a and b are the only constraints that matter, because this selected interval is just rotated at one point to create a circle. Thus, we can calculate the growth function in a very similar manner to that of problem 7. In this case, there are only two scenarios for the interval: the two end points are amongst the N+1 gaps between N points $\binom{N+1}{2}$ or all the points are outside the interval (1).

Final answer: $\boxed{\binom{N+1}{2} + 1 \Rightarrow [b]}$