CS156a Problem Set #4

1. We have

mH(2N) \* e-N/8ℇ^2

Where

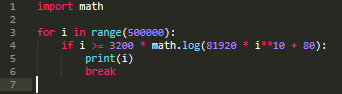
and dvc = 10

Solve for N and plug in:

And:

MH Nd\_vc+ 1 for the simple approximate bound

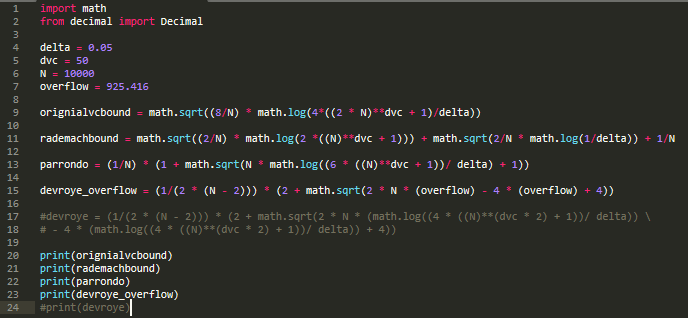
Thus, we can plug into a calculator to solve:





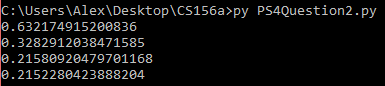
So, the closest answer is [d].

2. We can use the quadratic formula to isolate ℇ and then plug in values:



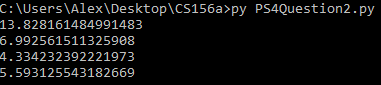
I calculated a value that was very large as “overflow” since it created a large value in the process. I used that for large values of N, like N = 10000.

Giving the output:



Thus, Devroye is our smallest bound, so our answer is [d].

3. Using the same code for Question 2 except with N = 5 and using “devroye” instead of “devroye\_overflow” (values in my code):



Thus, Parrondo and Van den Broek is our smallest bound, so our answer is [c].

4 - 6. The hypothesis that minimizes the mean squared error on the examples is:

Ein = 2

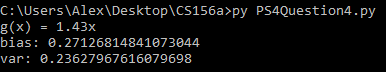
Where the sum is from i = 1 to N where N = 2

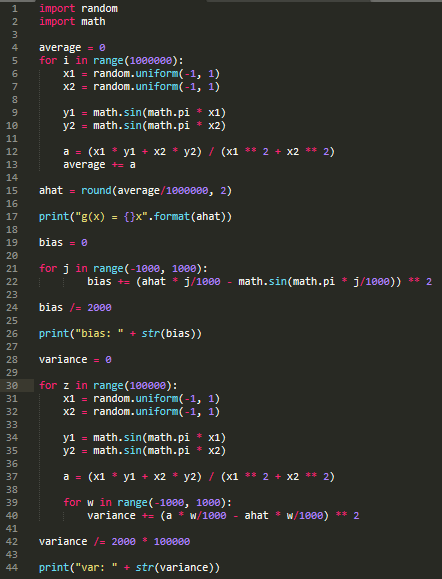
We want to minimize it, so differentiate and set equal to zero:

x1(y1 – ax1) + x2(y2 – ax2) = 0

a = (x1y1 + x2y2)/ (x12 + x22)

Thus, we can calculate that value for a large sample, which we can use to calculate bias and variance:





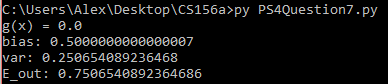
Thus, this gives the answers to Questions 4 – 6.

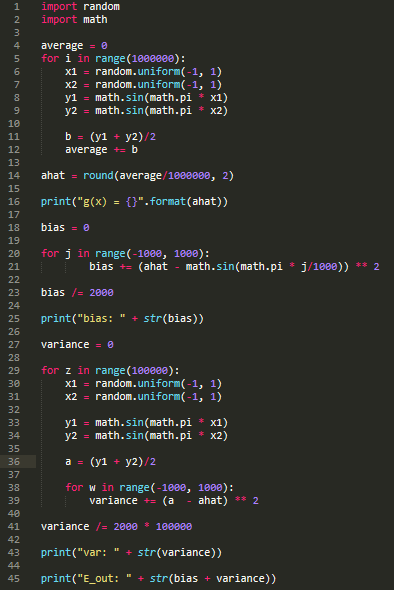
4. We got g(x) = 1.43x, so the answer is [e]

5. We got bias = 0.27, so the answer is closest to [b].

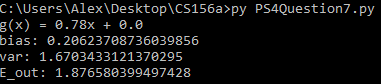
6. We got the variance = 0.24, so the answer is closest to [a].

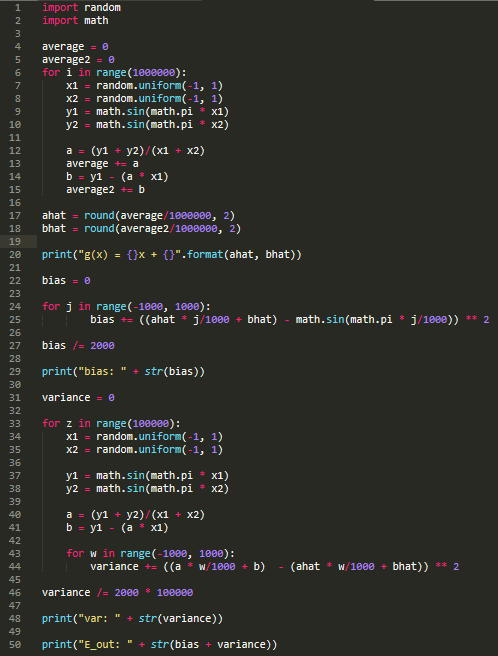
7. Like the derivation from the Ein formula above, we can do this for all five learning models. Then, we can calculate bias and variance and add them to get Eout:

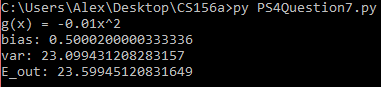
a. 

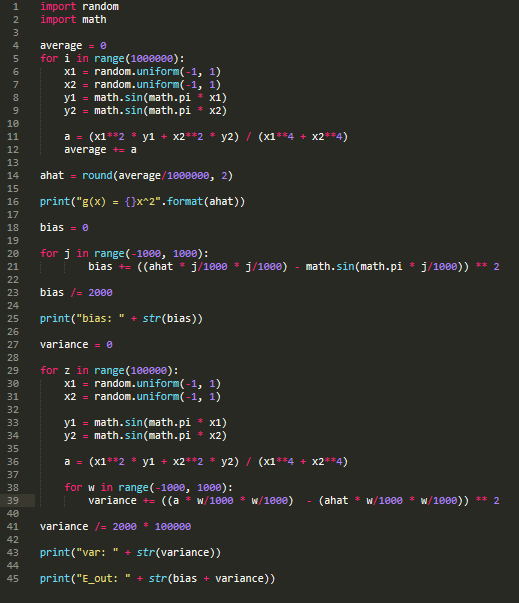


b. Sum Question 5 and 6: 0.271 + 0.236 = 0.508

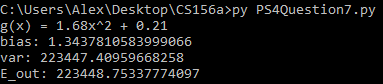
c. 

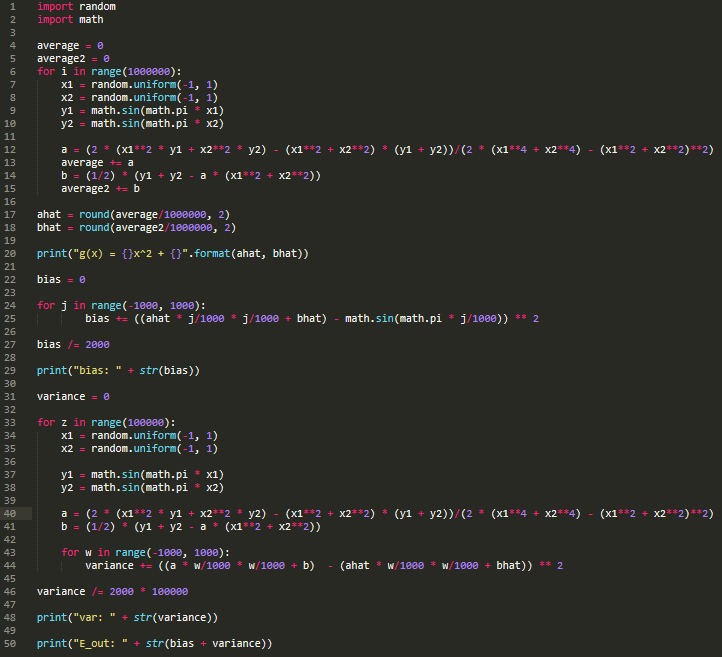


d. 



e.





So, the answer is [b].

8. We can write out the growth function:

Since this should work for any q 1 and N 1, I will pick q = 3 and N = 3. We can then choose some values for d\_vc to see how the growth function behaves:

d\_vc = 1:

5 = 2(4) – 1 -> 5 != 7

d\_vc = 2:

11 = 2(7) – 1 -> 11 != 13

d\_vc = 3:

15 = 2(8) – 1 -> 15 = 15

d\_vc = 4:

16 = 2(8) – 1 -> 16 != 15

d\_vc = 5:

16 = 2(8) – 1 -> 16 != 15

As shown, as d\_vc < q, the left side of the equation is smaller than the left. When d\_vc = q, the equation is fulfilled. When d\_vc > q, the left side of the equation is bigger than the right.

Therefore, the VC dimension that fulfills the growth function is d\_vc = q. The answer is [c].

9. Since we are taking the intersection of same breaking points within different VC dimension, we know that the smallest value for an intersection will be the empty set. If nothing intersects between the sets, we will have the empty set which is equal to zero. This is our lower bound. For the upper bound, we must consider the scenario where the sets are the same. This will narrow our options down to b and c since the max and mix will be equal to the intersection since the sets are all the same. Thus, the min = max = dvc. Next, we will consider the scenario where the intersection is not equal to the entire set. In this case, the breaking point will be the same, but the VC dimension will be different. However, the intersection will be the set of the lower dimension, since the values of the lower dimension will be contained within the higher dimension. Thus, the upper bound will be the minimum of the dimensions of the sets. The answer is [b].

|  |  |  |
| --- | --- | --- |
| X | X | X |
| O | X | X |
| X | O | X |
| X | X | O |

|  |  |  |
| --- | --- | --- |
| O | O | O |
| X | O | O |
| O | X | O |
| O | O | X |

10. Since we are taking the union, the union of a set of a smaller dimension and a bigger dimension will be the bigger dimension since the set of the smaller dimension will be contained within the bigger dimension. Thus, the lower bound will be the maximum of the dimensions being intersected. For the upper bound, we can get to a value bigger than the sum of the dimensions. Take for example these sets:

Both of sets have a dimension of one and a breaking point of two since we can only change one value. When combined, the dimension becomes three since we can go from {X, X, X} to {O, O, O}. The dimension is equal to 2 – 1 + sum(dimensions) = 1 + 2 = 3.

Thus, the answer is [e].