Homework 4

Problem 1

d is the correct answer.

From the givens of the problem, we have $\delta = 0.05$ and we want ϵ to be as close to 0.05 as possible. After plugging in 0.05 for δ and $N^{d_{vc}}$ as m_h, We have the equation:

$$\epsilon = \sqrt{\frac{8}{N} \ln \frac{4(2N)^{10}}{0.05}}$$

Plugging in the given options for N yields that 460,000 for N yields the closet ϵ to 0.05 (0.496). Thus, we get d as our answer.

Problem 2

d is the correct answer.

I plugged in the given values for a and b and solved for ϵ in parts c and d. The bounds were 0.632, 0.342, 0.224, and 0.215 respectively, so the smallest bound corresponds to answer d.

Problem 3

c is the correct answer.

I conducted a very similar approach to that of problem 2 by plugging in N = 5, and instead of $N^{d_{vc}}$ for m_h, we plug in 2^N because N is much less than $d_v c$. The smallest bound arose from answer choice c.

Problem 4

e is the correct answer.

See attached code. I used 100,000 runs and for each run, generated two random points and returned the best slope that minimized the mean squared error. The average best slope over the runs was 1.42, not one of the answer choices.

Problem 5

b is the correct answer.

See attached code. I took points from -1 to 1 in increments of 0.01 and calculated the bias for each point. The average bias is 0.27, closest to answer choice b.

Problem 6

a is the correct answer.

I made a vector of g^Ds and calculated the average hypothesis. Then I generated points from -1 to 1 and calculated the variance for each and then got the average variance to be 0.27, closest to answer choice a.

Problem 7

b is the correct answer.

We can simply compare the expected values for the out-of-sample errors for all the options. In lecture, choice a had the error of 0.75. From problems 5 and 6, choice b had an error of 0.54. From lecture, choice c had an error of 1.90. See attached code for choices d and e which had errors of 1.2995 and 102.13, respectively. The best expected value of out-of-sample error is thus choice b.

Problem 8

c is the correct answer.

Here, we see that as long as $\binom{N}{q} = 0$, we have that $m_h(\text{N+1}) = 2m_h(\text{N}) = 2^{N+1}$. If the VC dimension is q, the above is satisfied as $m_h(\text{q}) = 2m_h(\text{q-1}) + 0 = 2^q$ which satisfies the definition of a VC dimension.

Problem 9

b is the correct answer.

The intersection of all sets cannot have more points than that of the hypothesis set containing the minimum VC dimension. Thus, our upper bound is the hypothesis with the minimum VC dimension. If the intersection is an empty set, then the VC dimension is 0 so that's our lower bound. Thus, we get answer choice b.

Problem 10

e is the correct answer.

Just getting the lower bound for this allows for the elimination of answer choices a,b, and c. Taking the union of all hypothesis sets means we take elements of the hypothesis that has the max. VC dimension, hence our lower bound. For the upper bound, there is the scenario where the dichotomy from each individual hypothesis set is independent of the dichotomies from every other hypothesis set, in which case, we would take a sum of the VC dimensions consisting of the VC dimensions for each hypothesis set. Looking at the forum led to the differentiation between choices d and e (the diagram with the circles). The (k-1) term allows for scenarios where one hypothesis covers dichotomies missing from other hypothesis.