	E(M,N,S) = \(\frac{1}{2N} \left(\frac{\mathbb{PM}}{\mathbb{Page No.}} \) YOUVA
A.1>	a) M=1, E < 0.05, S=0.03
	we have given forequation
	N = 839.941
	b) M = 100 N = 1760.9750
	c) $M = 1000$ N = 2682.00908

2.2

For N=4, we can pick points: (1,3),(2,4),(3,1),(4,2). It's easy to see that these points are shattered by positive rectangles. So mH(4)=24.

The idea is that for any two points, if we draw a rectangle using them as diagnoal points, the rectangle should NOT contain any other point. Otherwise, whenever the two diagnoal points have values 1, the middle point will have value 1 as well, which excludes the possibility of having -1.

For N=5, if we draw horizontal and vertical lines through each of the four points above, the plane is divided into grids. The four points enclusing a 9-grid area. It's clear that the fifth point can't lie within the 9-grid area. Otherwise, there'll always a rectangle (constructed by two points) contains the fifth point.

In the same way, if we place the fifth point outside the 9-grid area, it's easy to see that the point will always lie below or above at least two points (in either x or y direction). These three points construct a rectangle which contains a point in it. This shows that mH(5) < 25.

We have the VC dimension dVC(H)=4, and $mH(N) \le \sum_{i=0}^{\infty} i=04(Ni)$.

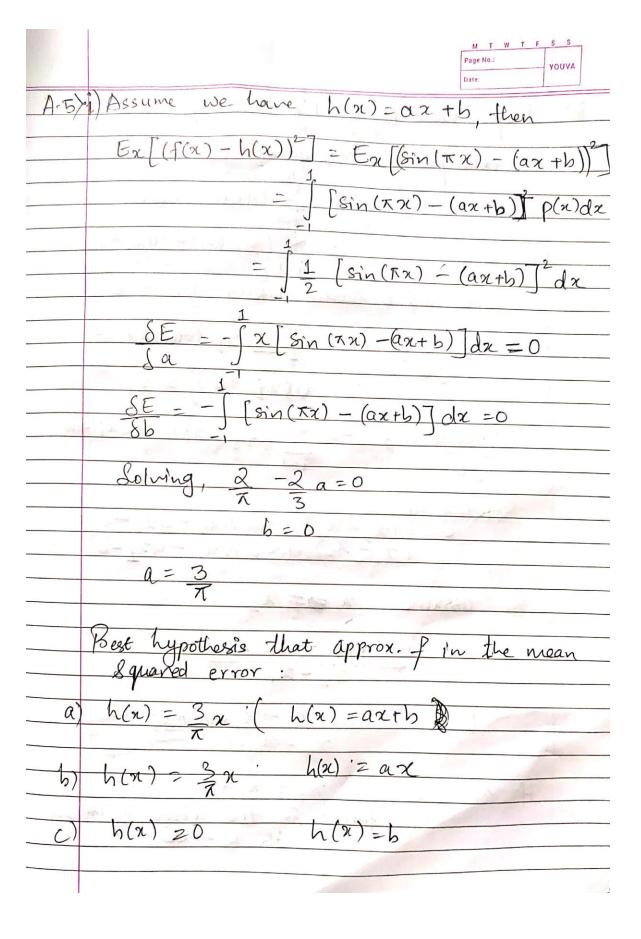
	Vc bound = 8x ln (4x growth (2x)) W T F
AZ	N=100
- 9	the former of the state of the
	vc bound = 0.848159
	N = 10000
	C - 1 - 1 by
	Vc bound = 0.104278
	(32) 34 -
	8 + m 2 2 2 2 2 2 4 1

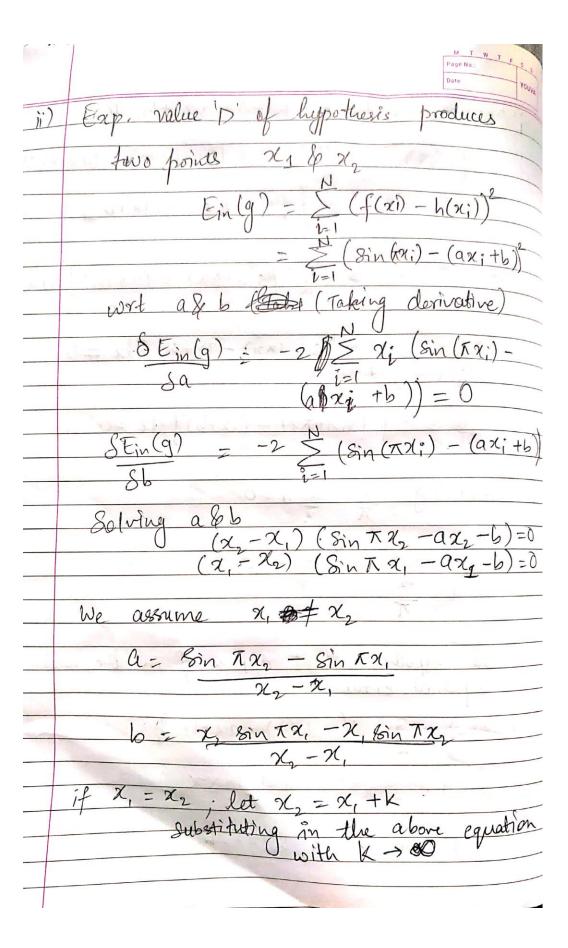
2.4

```
#!/usr/bin/env python3
2
 3
      import numpy as np
4
5
6
     def get sample size(dvc, deviation, epsilon, n):
7
          max iterations = 1000
8
          tolerance = 10
9
10
          for _ in range(max_iterations):
               rhs = 8 * np.log(4 * (pow(2 * n, dvc) + 1) / deviation)
11
               rhs = rhs / epsilon / epsilon
12
13
14
              result = n - rhs
15
               if np.abs(result) <= tolerance:</pre>
16
                  break
17
18
               n = int(rhs)
19
          return n
20
21
     if name == " main ":
22
          dvc = 10
23
24
          deviation = 0.05
25
          epsilon = 0.05
26
27
          # Starting value for N
28
          n = 1000
29
30
          sample data = get sample size(dvc, deviation, epsilon, n)
          print(f"Sample Size: {sample data}")
31
32
```

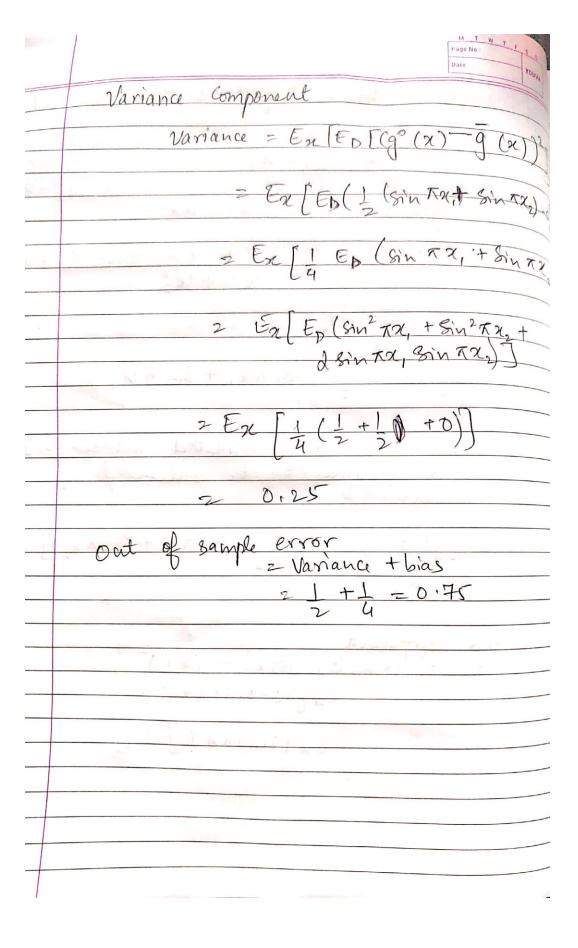
Sample Size: 45290

2.5





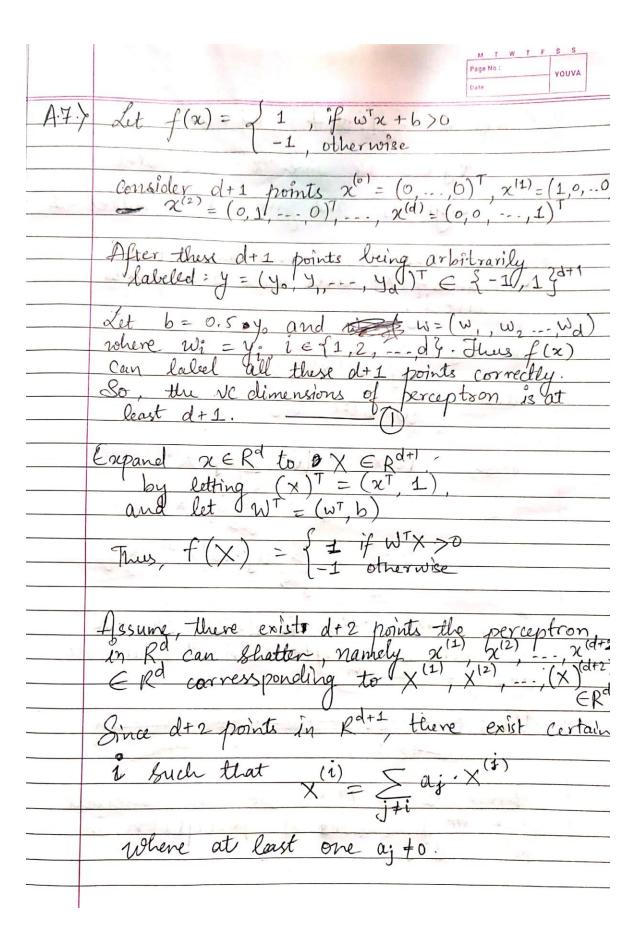
M T W T F S S Page No.: Date: YOUVA
we can obtain hypothesis langential tof
$g^{p}(x) = \sin \left(\frac{\pi x}{2} - \sin \frac{\pi x}{2} + \frac{x_{2} \sin \frac{\pi x_{1}}{2}}{x_{1} \sin \frac{\pi x_{2}}{2}} \right)$
for hypothesis set $h(x) = ax$ $x_2 - x$,
$g'(x) = \chi_r \sin \pi \chi_r + \chi_2 \sin \pi \chi_2 \cdot \chi$ $\chi_r^2 + \chi_r^2$
for h(x) = 6
$g^{\circ}(\pi) = 1 \left(\sin \pi \pi_{1} + \sin \pi \pi_{2} \right)$
bias, variance can be calculated by $h(x) = b$
$E_{D}[g^{D}(x)] = E_{D}[\sum_{z} (\sin x x_{1} + \sin x x_{2})]$
z 0
bias = $\mathbb{E}_{x} [g(x) - f(x)]^{2}$ $= \mathbb{E}_{x} [(G-\sin(\pi x))^{2}]$
= Ex ((sin (Tx))) 2 1 2



To show that k is a break point for H, we need to show H cannot shatter any set of k points x1, xk.

- If k is a break point, then mH (k) < 2k,
- In general, it is easier to find a break point for H than to compute the full growth function for that H.
- So the correct option is (d)

2.7



	Page No.: Page No.: Youv
	Let $S = Sj j \neq i$, $aj \neq 0$.
	tj∈S, we give x (j) the label
-	Sign (aj), and give x(i) a label-1
	By our assumption, there exists W that makes $f(x)$ label those $d+2$ points
	correctly
	So, $\forall j \in S$, we have $aj \cdot W^{T} \times^{(j)} > 0$
_	and $W^{T} \times (i) < 0$
	Also, $W^{T} \times^{(i)} = W^{T} \left(\sum_{j \neq i} a_{j} \cdot X^{(j)} \right)$ $= \emptyset W^{T} \left(\sum_{j \in S} a_{j} \cdot X^{(j)} \right)$
	jes dj. W. X. (i)
-	So our assumption is false. The VC dimension of perceptron lin Rd is at most d+1.
	Booms Stude from the above, we can
	From 1 & 2 we can conclude
	that VC dimensions of perceptron in
+	

	M T W T Page No.:	F S S
Ta) Eout(g) < Ein(g) + Inlinem		
E = 0.05 Jest boundary		
- lest bound = I ln(2)	Experience	
=	2 (20.05)	0.096
Jrain bound M = 1000 N = 400 train bound = 1 ln(2M) E)	
= lu (2*100)		5-115
Ein (g) = 0.115 Etest(g) = 0 So Flues, error bar on in sim is higher than the error test error	.096 ple er	ror

more samples then the sam 6) on