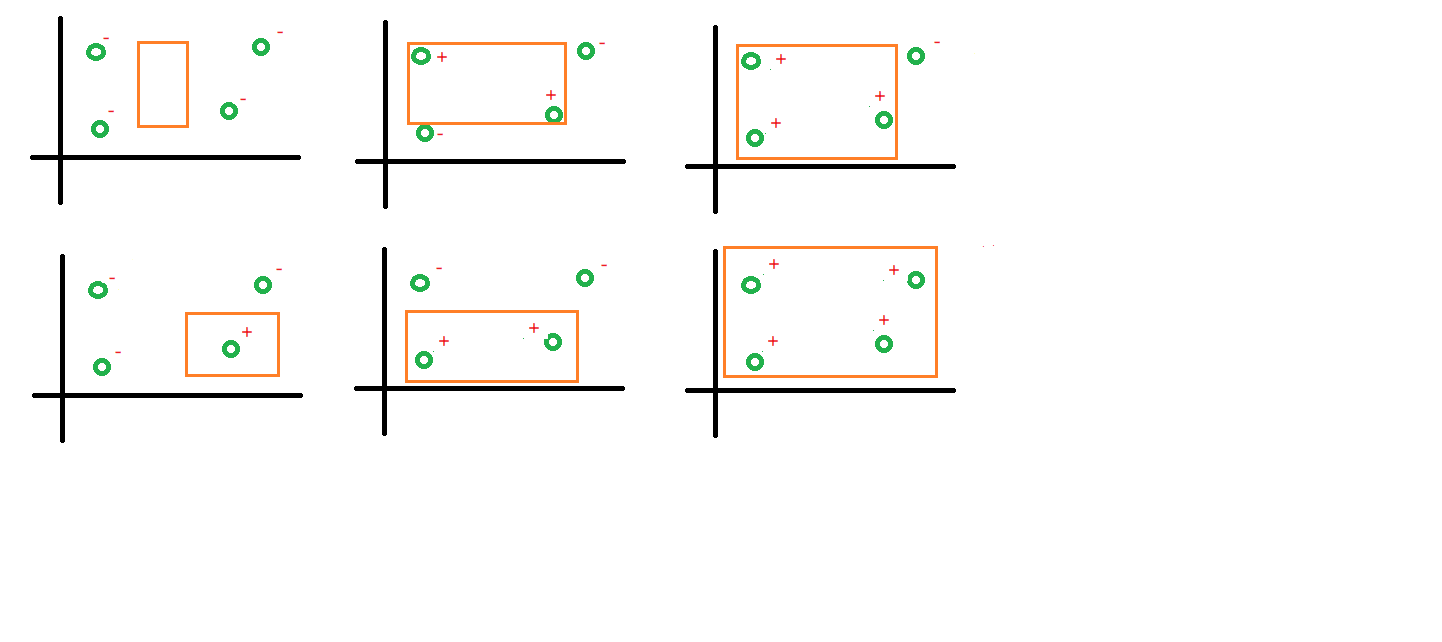
1. 
2. An example 4-point set is shown below with all typical labels and the

corresponding realization.   
There exists a 4-point set shattered by the concept set, so we have VC-dim≥ 4.



1. H = {hα| hα(x) = sign( |αx) mod 4 - 2| - 1), α∈ R}  
   We can easily find that the output range of the hypothesis H set is bounded to {-1,1}

Let Y ∈{+1,-1}^N be the output set and X∈R be the input set  
For each y∈Y, we can always find a corresponding input set and an adjusting theαof h∈H that hα(x) = y

∵ {+1,-1}^N combinations can be constructed  
∴ H can shatter any N inputs. That means the VC-dimension of H is ∞.

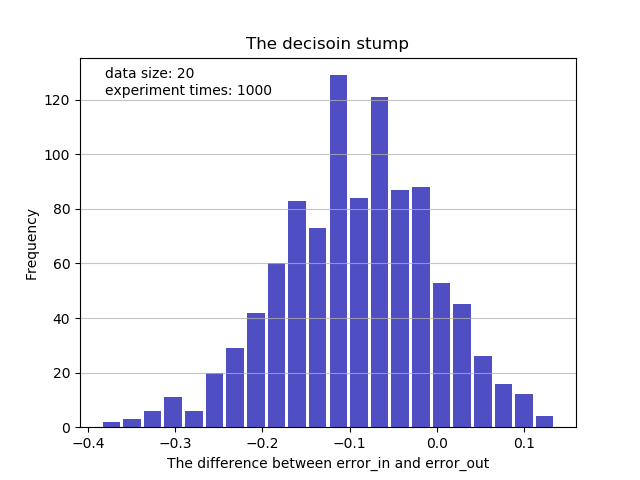
1. Prove that dvc(H1∩H2)≦dvc(H1) where H1 and H2 that come with non-empty intersection  
   First, ∀n < dvc(H1∩H2) that means any n inputs can be shattered by H1∩H2  
   ∵H1∩H2 ⊆ H1 and H1∩H2 ∉{∅}  
   ∴Any n inputs can be shattered by H1. That means dvc(H1∩H2) ≦ dvc(H1)
2. H1 as the positive-ray hypothesis set

H2 as the negative-ray hypothesis set

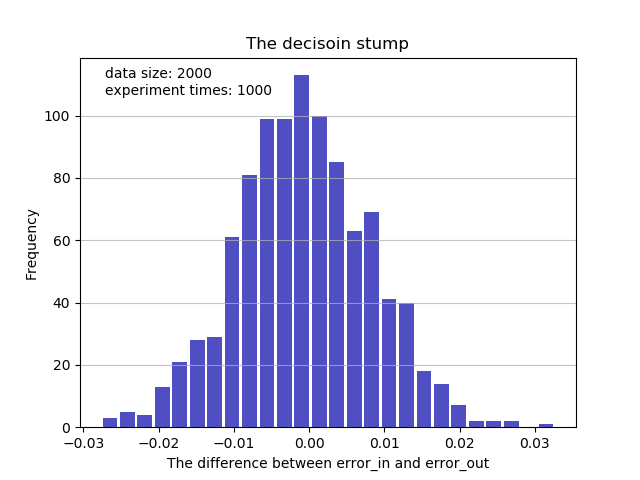
mH1 (N) = N + 1 = mH2 (N)，mH1∪H2 (N) = 2(N+1)-2 = 2N

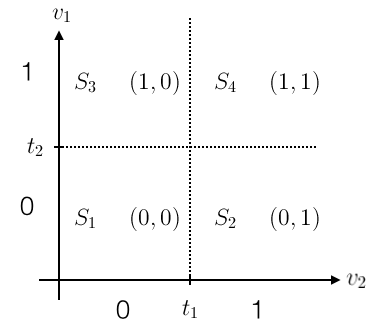
When N=3 , mH1∪H2 (N) = 2N≠2^N. So the dvc(H1∪H2) is 2.

1. hs,θ(x) = s‧sign(x-θ)，with θ∈ [-1,1]   
   Given a target function f that P(x|y) = 0.8 where f(x)=y and P(x|y) =0.2 where f(x)≠y   
   Assume that h error rate is x  
   ∵ When s=1, x = abs(θ/2) otherwise x = 1-abs(θ/2)  
   ∴ s=1, Eout=0.8\* abs(θ/2)+0.2(1- abs(θ/2)) = 0.2 + 0.3 abs(θ)  
    s=-1, Eout=0.2\* abs(θ/2)+0.8(1- abs(θ/2)) = 0.8-0.3 abs(θ)  
   Eout( hs,θ) = 0.5+0.3s(abs(θ)|-1)
2. I find that the smaller data size and the larger difference between average error in and average error out. The data size affects the hoeffding inequality and that means the bad sample happened with higher probability.



1. I find that the difference between average error in and error out gets closer with the increasing of the data size. That proves the correctness of the hoeffding inequality.



1. H = {ht,s | ht,s(x) = 2[[v ∈ s]] -1, where vi = [xi>ti], S a collection of vectors in {0,1}^d, t ∈ R^d}  
   What is the VC-dimension of the "simplified decision trees'' hypothesis set?  
     
   我們將二維度上的threshold值ti與各向量轉換到二維平面上，上面二维的例子中，**simplified decision trees**的Dvc 與hyper-rectangular regions的數量相等。D維向量可以用D條直線最多分割出2^D 個hyper-rectangular regions，代表說可以shatter掉2^d個點。  
   ∴VC-dimension of the "simplified decision trees'' is 2^d.