* Hoeffding Inequalities
  1. Single Slot Machine

不難從形式中找到對應意義的符號：

只要從 開始推導：

代入由所推得的以及其他參數可以得到：

* 1. All Slot Machines

Hoeffding Inequalities告訴我們，對所有的m：

由於M跟t均大於1，因此：

令*：*

所以代回式子可以得到：

所以反過來就可以推得：

* 1. 抽抽樂

要有某些數字全都是綠色，可以知道AB這兩種獎券不可以同時被抽到，他們顏色的情形是互補的；CD也是同樣道理。因此可以知道，這五抽裡面的獎券種類最多只可以包含兩種，並且只可以是，AC、AD、BC、BD這四種組合，所以獎券總共有以下的取法數量：

但是要注意，AC跟AD的取法中，他們都包含「全都是A」的取法，所以會多算一次，要扣除；BC跟BD、AC跟BC還有AD跟BD也是同理：

最後再除以全部的取法就可以得到機率：

* 1. 抽抽樂-續

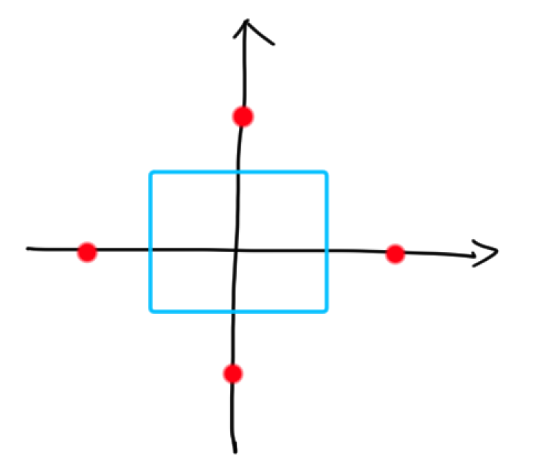
如果要五張券裡面的數字2都是綠色的，那這五張券只能是B或D這兩種獎券：

可以發現，，BAD Data 發生的機率有所不同。

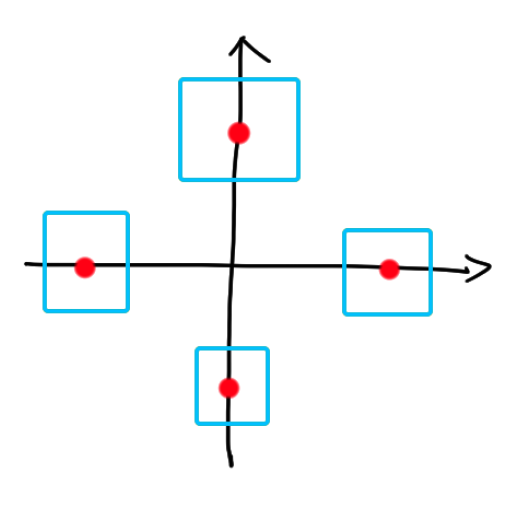
* VC Dimension

1. negative rectangle

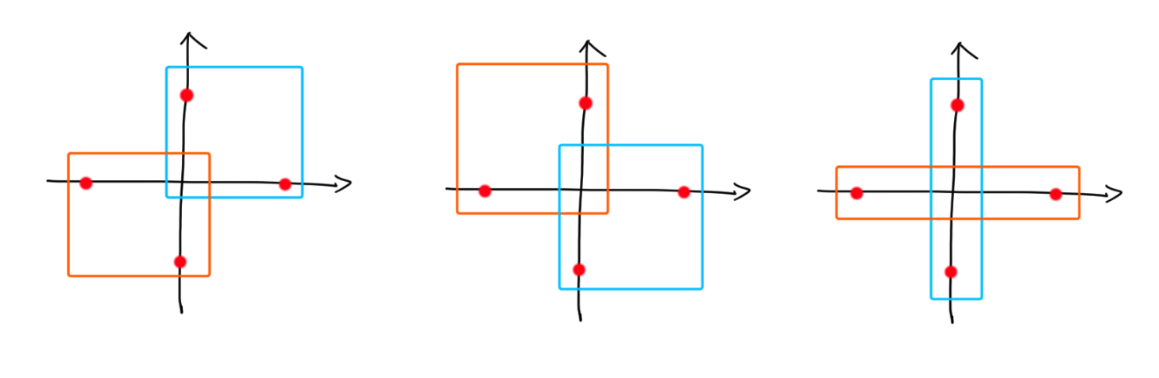
下面的圖中，紅色點是4個input vectors，而每一個藍色框框或橘色框框都是一個hypothesis，他們都屬於negative rectangle 這個hypothesis set。



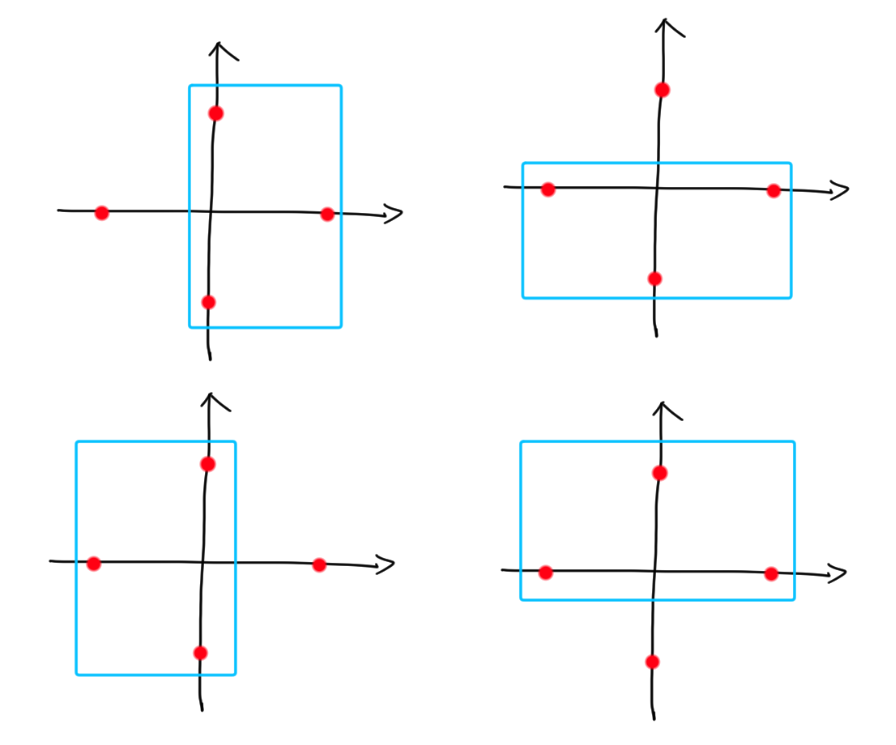
上圖是0個點回傳，框框沒有包含任何點，全部點回傳。



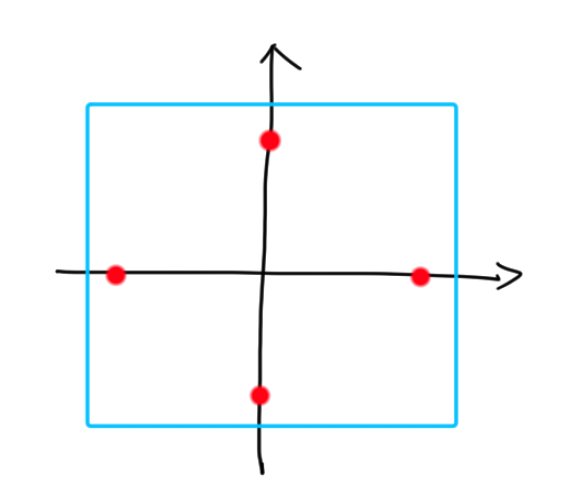
上圖是1個點回傳，圖中的四個藍色框框代表四個hypothesis。對於一個藍色框框來說，框框內的點回傳，另外三個點回傳。



上圖是2個點回傳，圖中的每個藍色跟橘色框框各自代表1個hypothesis。對於一個框框來說，框框內的兩個點回傳，另外兩個點回傳。



上圖是3個點回傳，圖中的每個藍色框框代表1個hypothesis。框框內的三個點回傳，另外一個點回傳。



最後一張圖是4個點回傳，也就是每個點都回傳。

1. Multiple intervals

個參數，就是有M個intervals可以用，我們先討論positive interval。首先，可以從課堂的one positive interval對N個點的公式：

推廣到如果是M個positive intervals對N個點的公式：

這時候如果回憶高中所學：

白話來說就是偶數項加起來等於。

所以我們可以知道，根據M個positive intervals的公式，如果我們令：

也就是說，M個positive intervals可以shatter 2M個點。

但如果是個點：

會發現少了最後一項，如果我們幫他補上去：

也就代表M個positive intervals無法shatter任何個點。

不過當我們舉一個實際的例子來看，例如的時候，上面的推論可以知道我們可以shatter 個點，無法shatter任何 個點，那為甚麼positive intervals無法shatter 個點，或者說剛好只差1個情形做不到？答案就是這樣的分布情形，因為我們只有2個positive intervals，但是想要弄出需要3個。

這時候就是negative interval上場的時候了，上面的這種情形其實就是2個negative interval，所以如果我們連同negative interval也考慮進來，那麼M個intervals就可以shatter 個點了。

那如果是個點呢？我們一樣先從positive interval開始討論：

會發現少了最後一項，如果我們幫他補上去：

從上面推導的過程可以知道，positive interval想要shatter 個點還差種情形做不到，我們一樣以上面2個positive interval來做舉例，對於，哪7種情形做不到？其實就跟剛剛很類似，需要3個positive interval的情形我們就無法辦到：

除了上面3種外還有另外4種。那這7種我們有辦法靠negative interval做到嗎？可以發現有一些可以，例如：

但是，有一些不行：

這種的他要嘛需要3個positive interval，要嘛3個negative interval。

如果推廣到 個點，我們無法透過 個interval，去擊敗那種要嘛需要 個positive interval，要嘛 個negative interval的情形；因此我們無法shatter任何 個點。

因此我們最終可以知道，個interval的Hypothesis Set，我們可以shatter 個點，但無法shatter任何 個點，也就是說VC dimension 是。

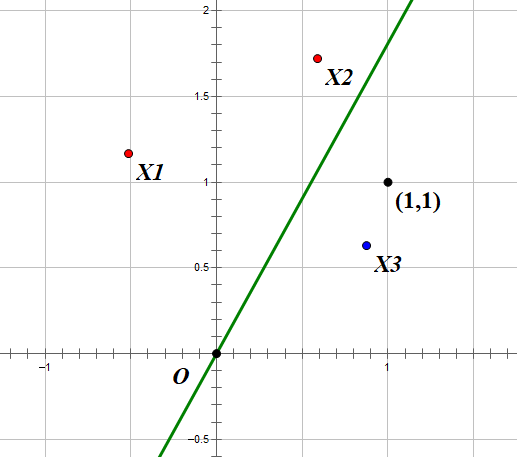
1. origin-passing perceptrons

由於perceptron得通過原點，就很像他被釘在了一個地方不可以任意移動只能旋轉。而在二維平面中，一條只能在原點旋轉的線，對於個散佈在平面上的點，就等同把這些點壓到剩下一個維度，並在這僅剩的一個維度做「切一刀」的功能，也就是在線的一邊回傳，另一邊回傳。那麼其效果就跟課堂上提到的positive and negative ray效果是一樣的，growth function 是。

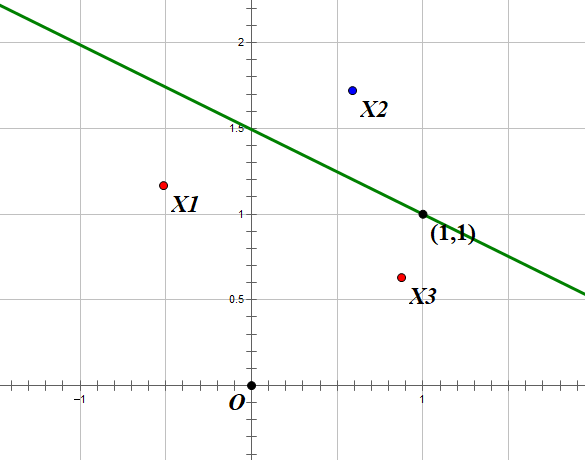
1. Union hypothesis set

上一題的origin-passing perceptrons可以知道VC dimension是2，同理我們也可以知道(1,1)-passing perceptrons的VC dimension也是2，當我們把這兩個Hypothesis set聯集起來，可以分別對跟的資料做討論。

當 時，令某組資料可以被origin-passing perceptrons分出6種的情形，如果這時候有(1,1)-passing perceptrons成員的幫助，其效果就好像是把原本的origin-passing perceptrons平移了一些距離，因此可以「換到另一個位置」去「切一刀」，也就shatter了該組資料。



例如在上面的圖中，資料中的三個點為，綠色的線是origin-passing perceptrons的其中一個成員，可以看到他在圖片中將分別分成了「紅，紅，藍」的情形，在這裡將紅色代表藍色代表；我們可以很明顯的知道，origin-passing perceptrons沒有一個成員可以把 分成「藍，紅，藍」或「紅，藍，紅」，但是(1,1)-passing perceptrons的某些成員可以，例如：



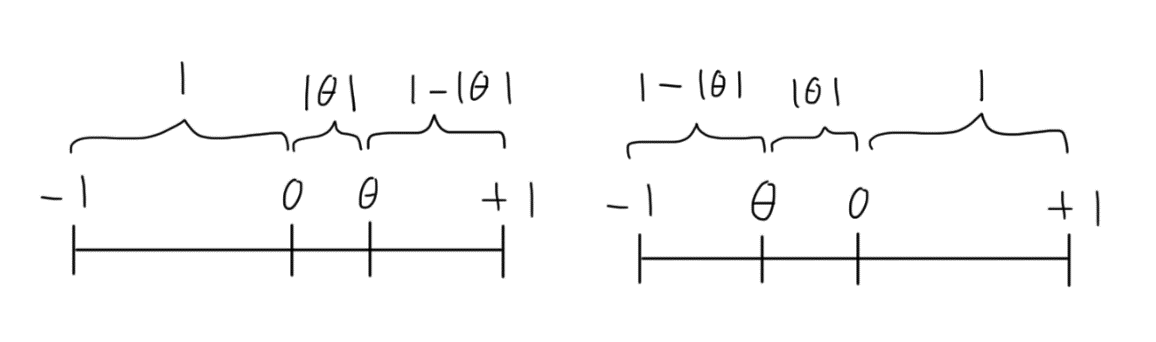
因此我們可以知道，兩個Hypothesis Set聯集之後，某種角度來說就好像是讓origin-passing perceptrons有了「局部移動」的能力，可以shatter 一組 的資料。

但是當時，就算origin-passing perceptrons獲得了一點點平移的能力，依舊不能改變他們「身為一個2D perceptrons」的極限： 2維平面的perceptrons的VC dimension是3。因此兩個Hypothesis Set聯集之後無法shatter任何一組 的資料。以另一個直覺的角度來說，想像有四個點排成一直線，想要透過perceptrons切出「」是不可能的。

因此我們可以知道聯集之後的VC dimension是3。

* Decision Stumps

首先我們圖像化：



上圖畫出了兩種的情形。

可以知道不管是正的或負地，都可以劃分出三個區域：。接著分成 兩種情況探討，並且是noise的機率，此題中為10%：

：

可以知道在這三個區間內，「」跟「」這兩個區域有的比例會犯錯，「」這個區域則是 的比例會犯錯，所以可以列出犯錯的式子：

但是不要忘記，我們的區是從到，所以要記得除2才會是錯誤率：

：

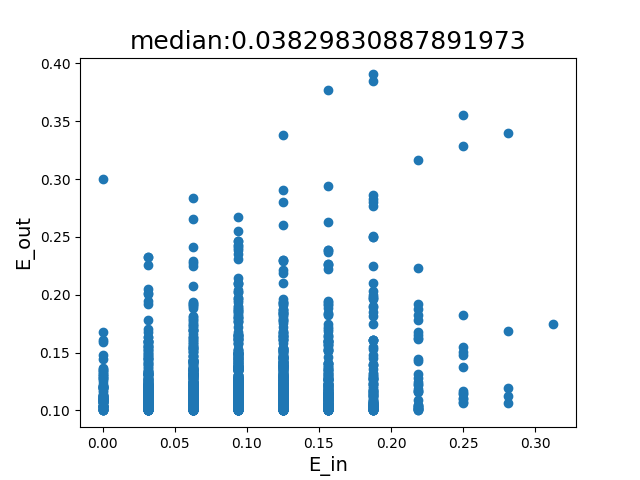
流程跟上面一樣。「」跟「」這兩個區域有的比例會犯錯，「」這個區域則是 的比例會犯錯，所以可以列出犯錯的式子：

一樣不要忘記除以2：

最後觀察的結論，可以統整成：

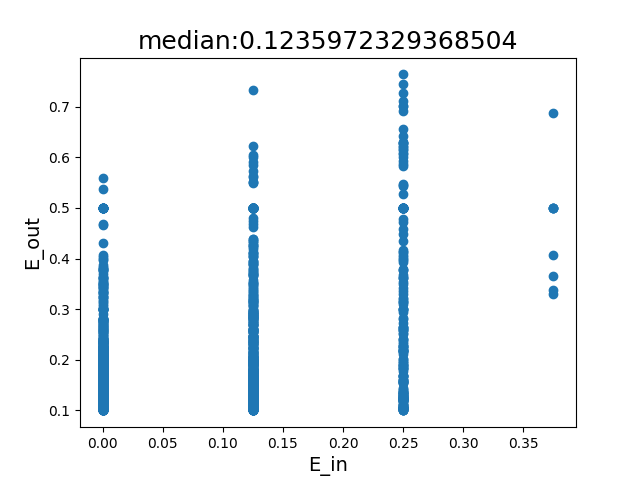
只要將代入就可以得到題目的公式了：

1. Artificial data with size 32



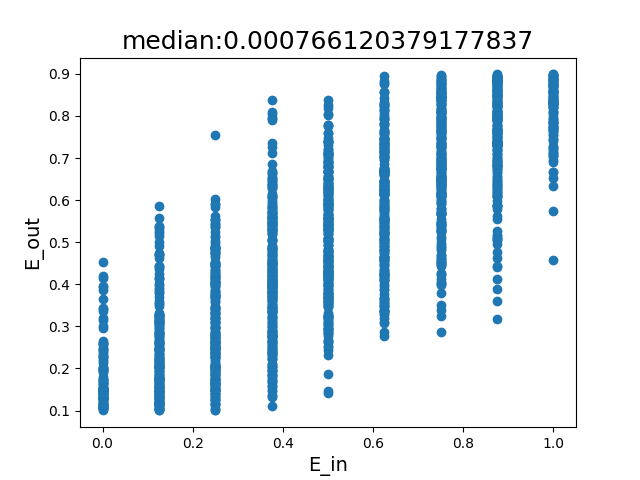
的median大約是0.038。

1. Artificial data with size 8



的median大約是0.12。可以發現 的median變高了，並且大於0.5的情形也變多了。回顧上面的公式，可以知道要大於0.5，必須是，也就是說資料量比較小時，noise帶來的影響比較大，會讓Decision Stump選擇來達到低的。

1. Artificial data with size 8 and random picking



的median大約是0.0007。可以發現 的median變超低。首先因為均勻分布選取所產生的影響，圖形呈現斜線上升會隨著上升，而的上升代表的取值逐漸以為主，才可以逐漸拉高；可以發現在0到1的範圍，雖然不好確定是否數量差不多，但是可以發現大致都有一定的分布。

* Bonus: Perceptrons that Pass Special Points

1. Cover’s Theorem

就如同題目提供的pdf檔中，作者所提及的一段話：「Now, by forcing the hyperplane to pass through a certain fixed point, we are in fact moving the problem to one in dimensions, instead of.」所以如果我們要求perceptrons通過個「錨點anchor points」，其效果就好像維度從變成了，因此公式就會被改寫為：