

Fundamentos de Aprendizaje Automático y Reconocimiento de Patrones

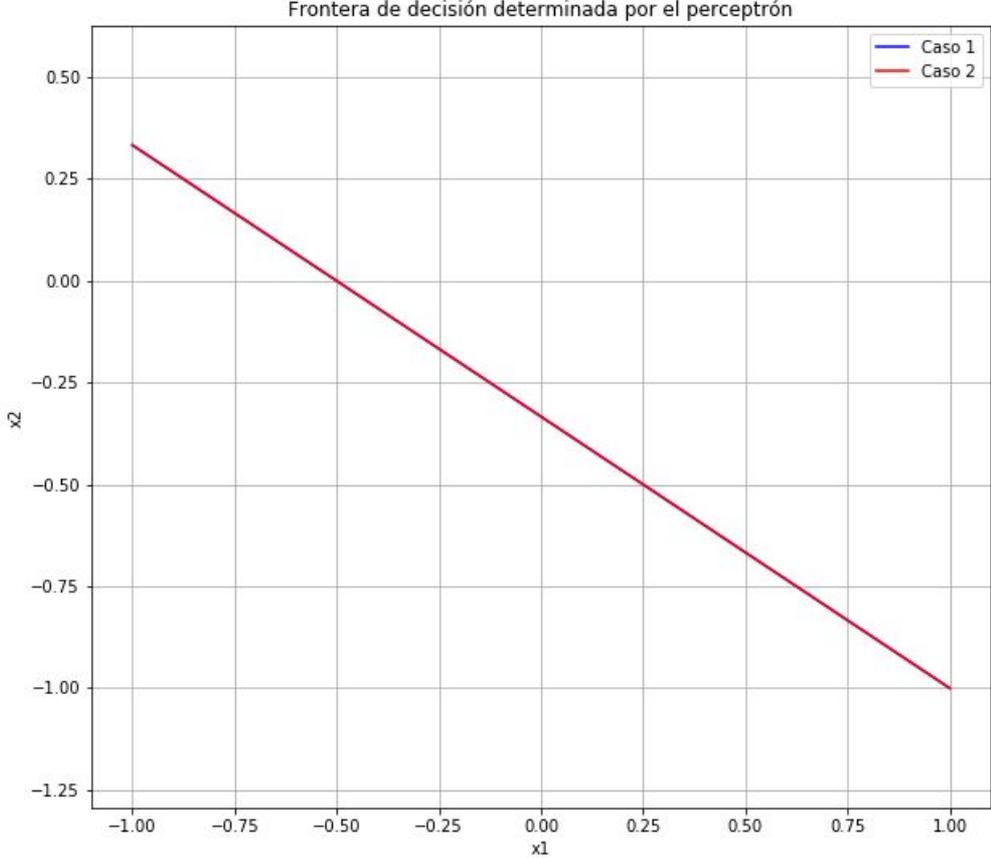
Práctico 01

2024

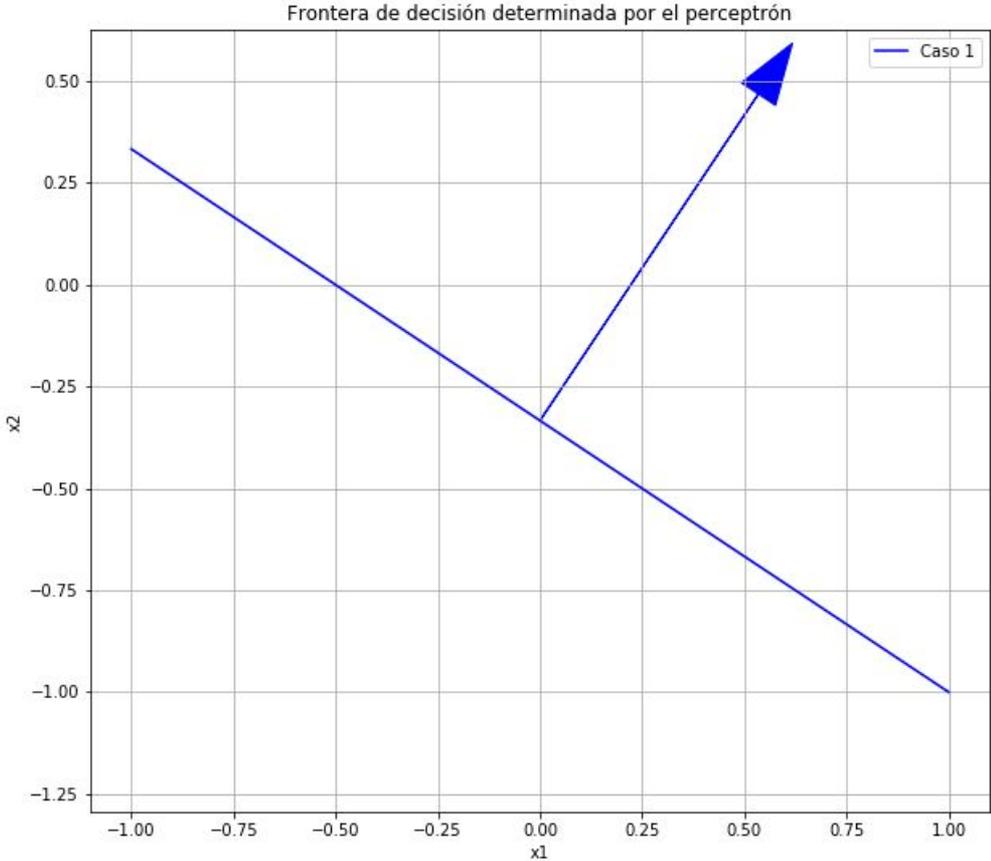
Tabla de contenido

- El perceptrón
 - Ejercicio 2
 - Ejercicio 3
- Desigualdad de Hoeffding
 - Ejercicio 4
- 9:45: Cuestionario práctico 1

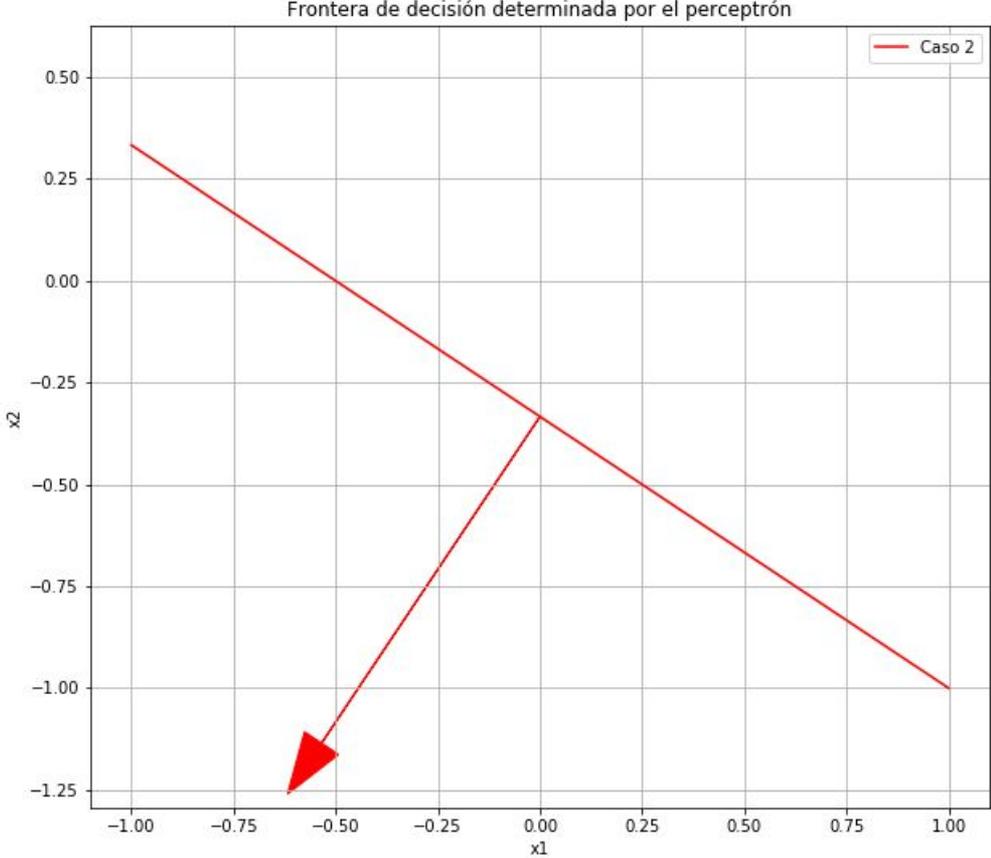
Superficie de decisión



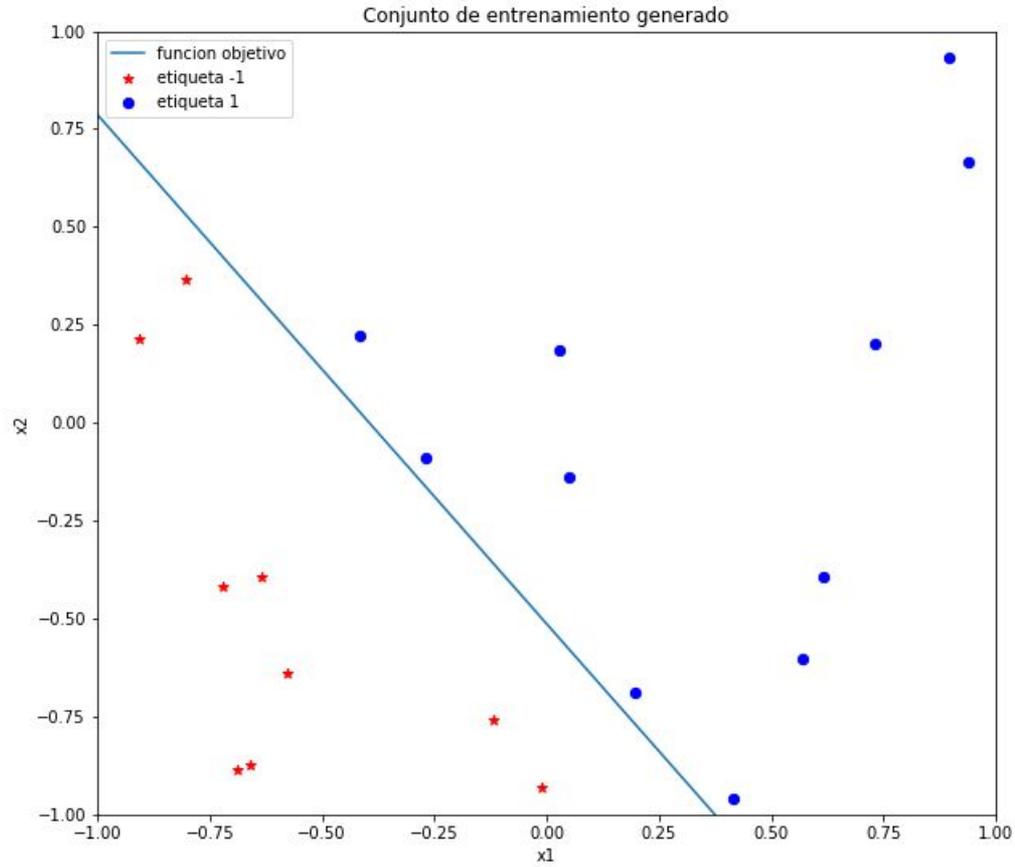
Superficie de decisión



Superficie de decisión



Ejercicio 3: Generar conjunto de entrenamiento



Ejercicio 3: funciones de numpy de utilidad

- `random.rand()`,
- `ones()`, `zeros()`,
- `concatenate()`,
- `where()`,
- `hstack()`,
- `dot()`,
- `reshape()`,
- `sign()`

Ejercicio 3: código generar conjunto de entrenamiento

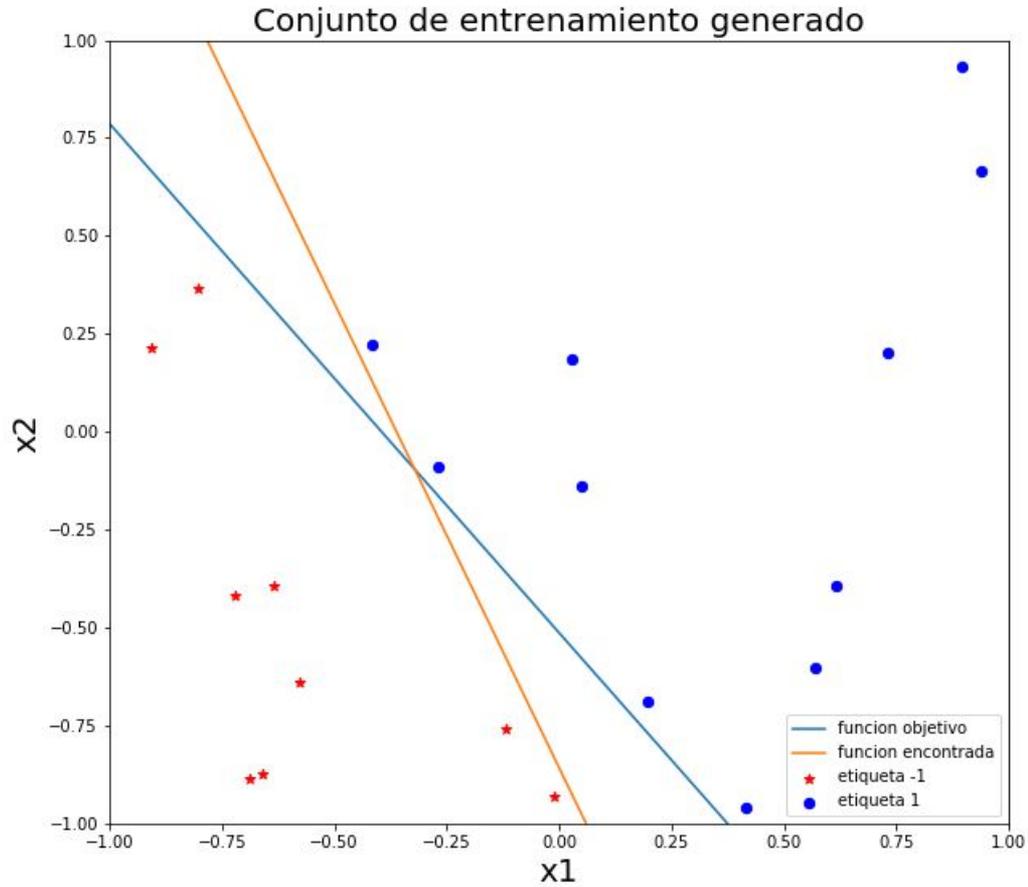
```
1 def generar_conjunto_entrenamiento(cantidad_muestras, d=2, semilla=None):
2     """Esta función genera un conjunto de entrenamiento de tamaño cantidad_muestras
3     Entrada:
4         cantidad_muestras: cantidad de muestras que se devuelven
5         d: dimensión en la que viven las muestras generadas
6         semilla: valor que se le asigna al método random.seed()
7     Salida:
8         w_rand: vector de tamaño d+1 que contiene los parámetros
9                 de la función objetivo elegida aleatoriamente
10        X: matriz de tamaño (cantidad_muestras x (d+1)) que contiene las
11           muestras de entrenamiento generadas
12        y: vector de tamaño cantidad_muestras que contiene las etiquetas de las
13           muestras generadas
14        """
```

Ejercicio 3: código generar conjunto de entrenamiento

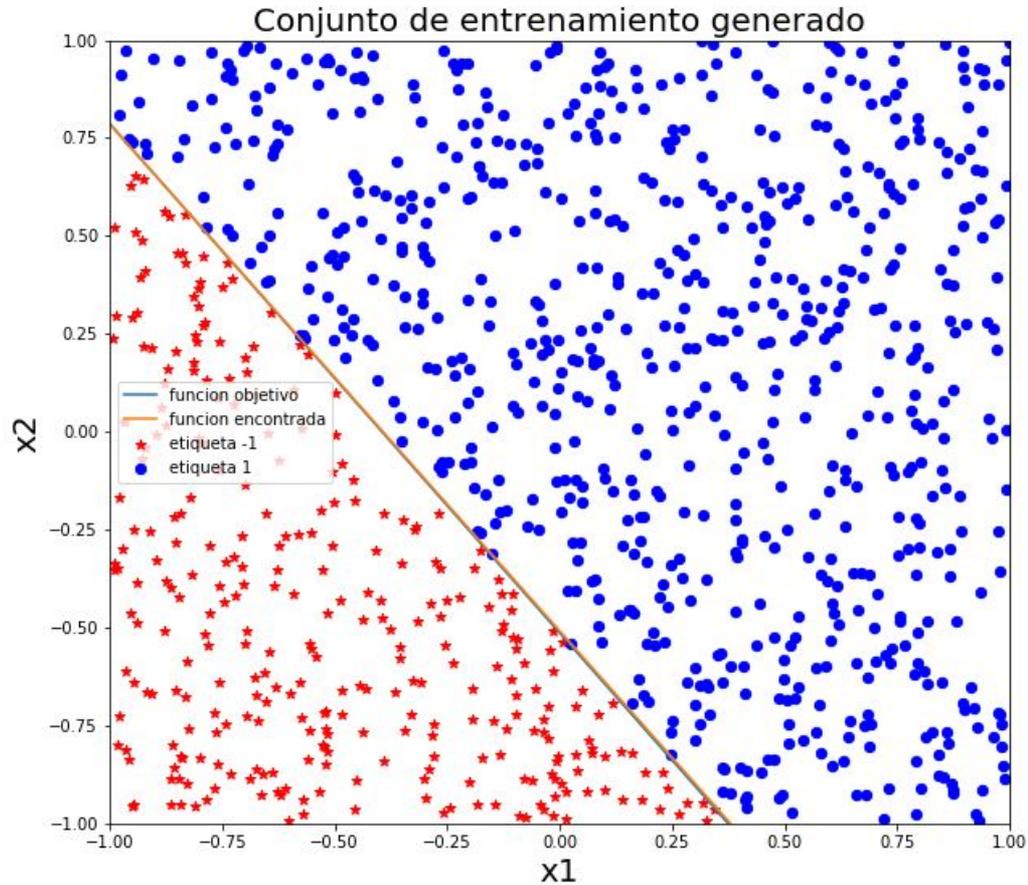
```
23 #####
24 ##### EMPIEZA ESPACIO PARA COMPLETAR CODIGO #####
25 #####
26
27 # Se elige aleatoriamente una recta del plano como función objetivo f
28
29 # w_rand =
30 w_rand = np.random.rand(d + 1)
31
32 # Se elige aleatoriamente N puntos pertenecientes a la región de interés.
33 # Se agregan los unos correspondientes a las coordenadas homogéneas.
34
35 # X =
36 X = np.random.rand(cantidad_muestras, d) * (x_max - x_min) + x_min
37 X = np.hstack((np.ones((cantidad_muestras, 1)), X))
38
39 # Se generan las salidas utilizando el modelo del perceptrón.
40
41 # y =
42 y = np.sign(np.dot(X, w_rand))
43
44 #####
45 ##### TERMINA ESPACIO PARA COMPLETAR CODIGO #####
46 #####
47
48 return w_rand, X, y
49
50 N = 20
51 w, X, y = generar_conjunto_entrenamiento(N, semilla=42)
52
53 validar_resultado("p01e03", w, X, y)
```

```
+-----+
| FuAA: validar resultado |
+-----+
El resultado es correcto. Continuar.
+-----+
```

Ejercicio 3: solución del perceptrón con $N = 20$



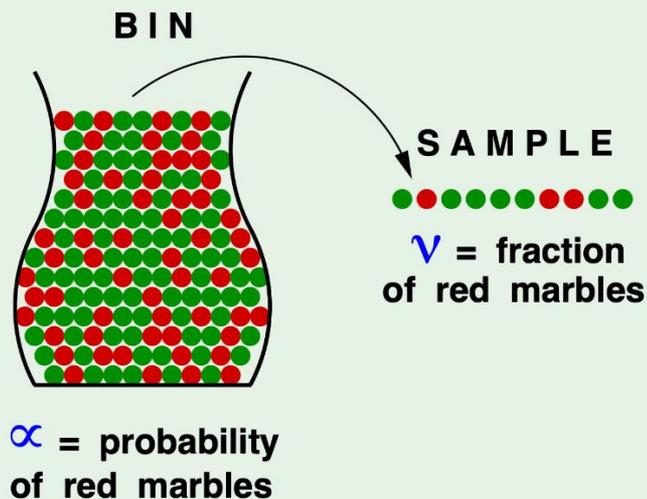
Ejercicio 3: solución del perceptrón con $N = 1000$



Desigualdad de Hoeffding

$$\mathbb{P} [|\nu - \mu| > \epsilon] \leq 2e^{-2\epsilon^2 N}$$

- Valid for all N and ϵ
- Bound does not depend on μ
- Tradeoff: N , ϵ , and the bound.
- $\nu \approx \mu \implies \mu \approx \nu$ 😊



Desigualdad de Hoeffding

Connection to learning

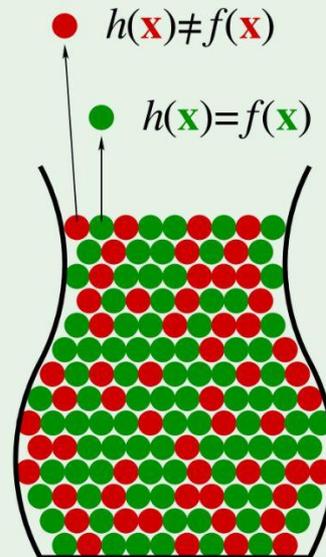
Bin: The unknown is a number μ

Learning: The unknown is a function $f : \mathcal{X} \rightarrow \mathcal{Y}$

Each marble \bullet is a point $\mathbf{x} \in \mathcal{X}$

● : Hypothesis got it **right** $h(\mathbf{x})=f(\mathbf{x})$

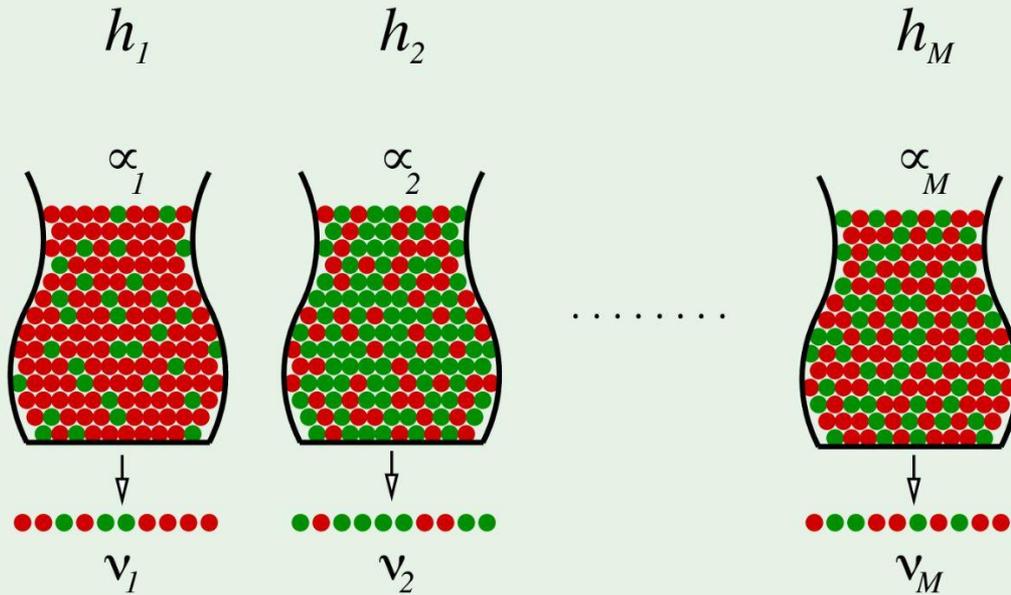
● : Hypothesis got it **wrong** $h(\mathbf{x})\neq f(\mathbf{x})$



Desigualdad de Hoeffding

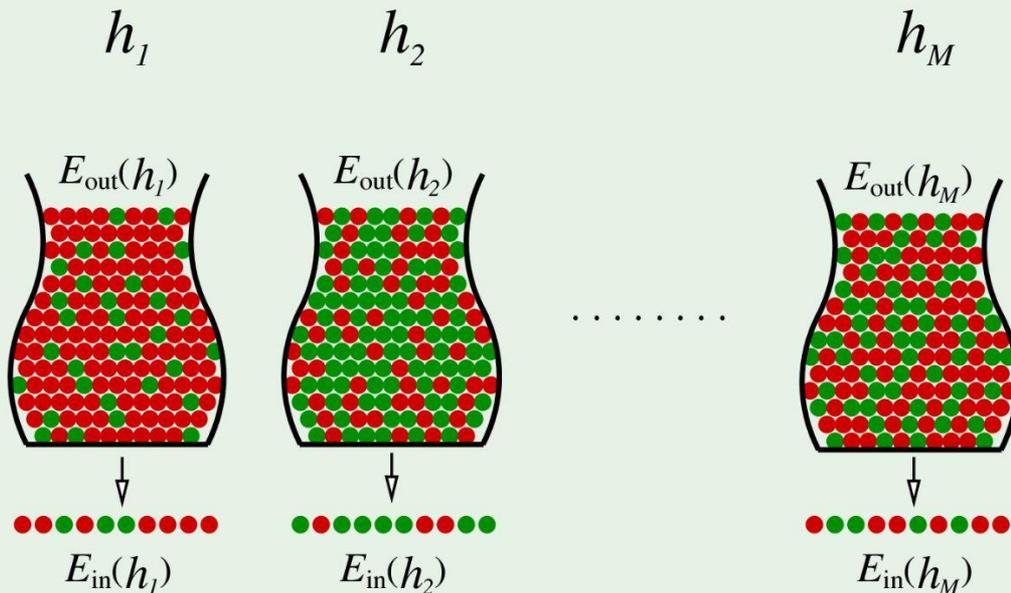
Multiple bins

Generalizing the bin model to more than one hypothesis:



Desigualdad de Hoeffding

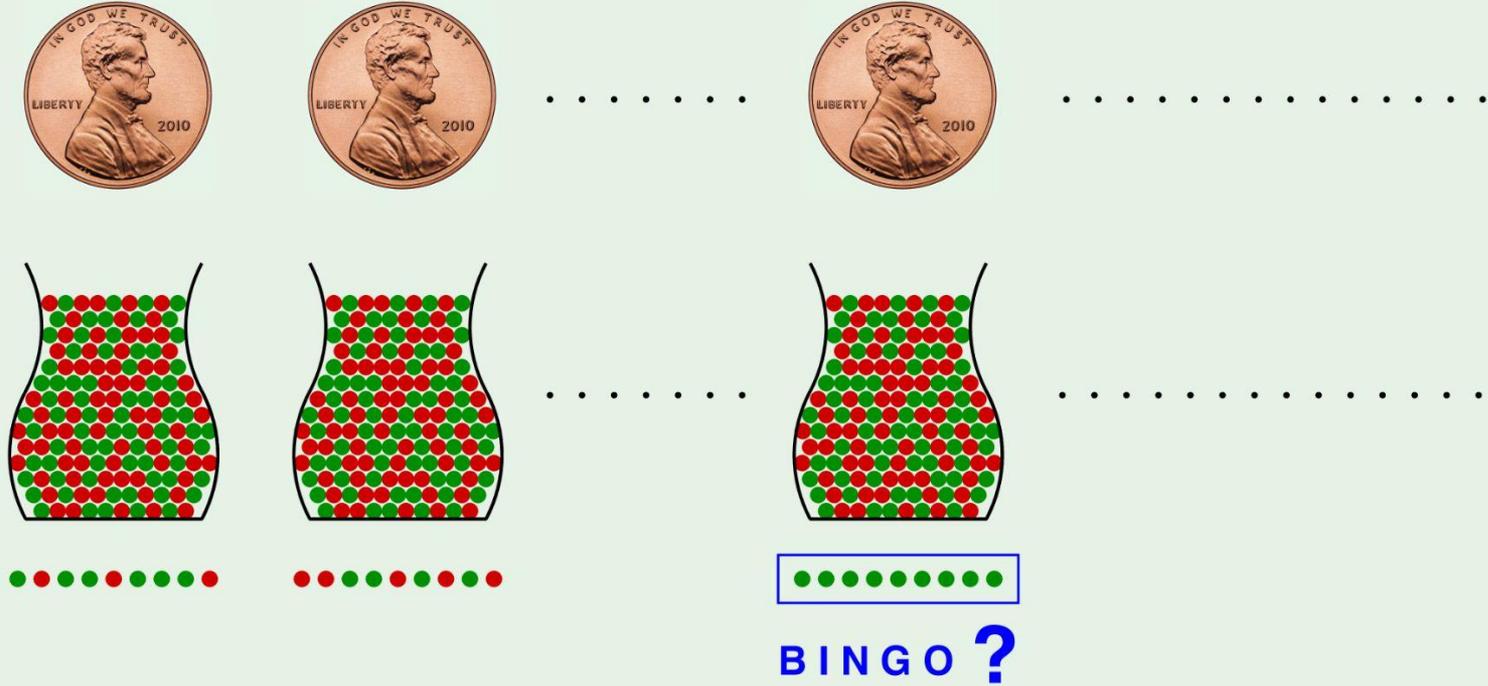
Notation with multiple bins



$$\mathbb{P} [|E_{in}(h) - E_{out}(h)| > \epsilon] \leq 2M e^{-2\epsilon^2 N}$$

Desigualdad de Hoeffding

From coins to learning



Ejercicio 5: Desigualdad de Hoeffding - Monedas al aire

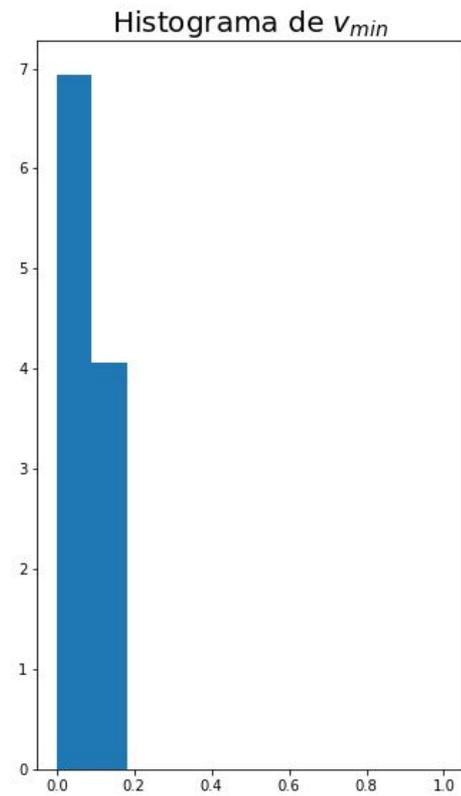
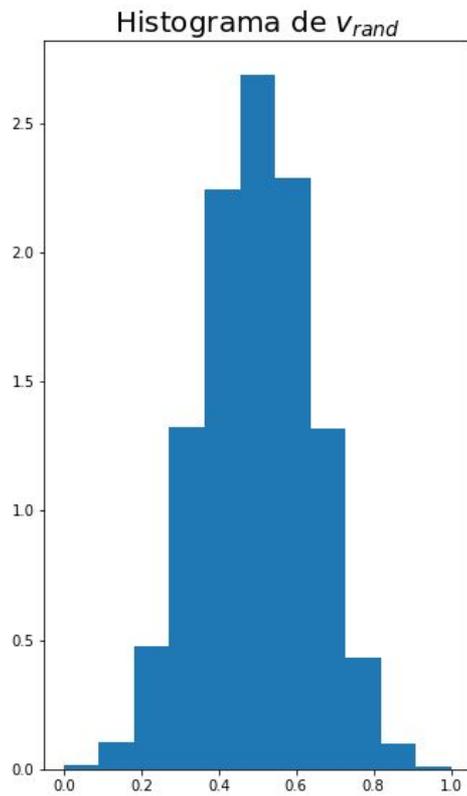
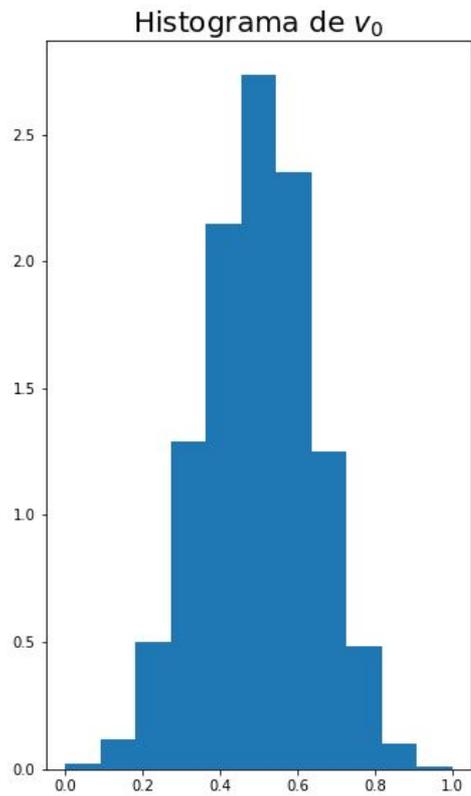
Se lanzan 1000 monedas, cada una 10 veces, de forma independiente. Considerar tres monedas de la siguiente manera: m_1 es la primera moneda lanzada, m_{rand} es una moneda elegida al azar de las 1000 y m_{min} es la moneda que tuvo la frecuencia mínima de caras. Sean v_1 , v_{rand} y v_{min} la fracción de caras obtenidas de los 10 lanzamientos.

¿Cuál(es) moneda(s) tiene(n) una distribución de v que satisface la Desigualdad de Hoeffding (un solo bolillero)? ¿Por qué?

1. sólo m_1
2. sólo m_{rand}
3. sólo m_{min}
4. m_1 y m_{rand}
5. m_{min} y m_{rand}

Ejercicio 1.10 del libro

Ejercicio 5: histogramas



Ejercicio 5: cotas y probabilidades empíricas

