Problem 11

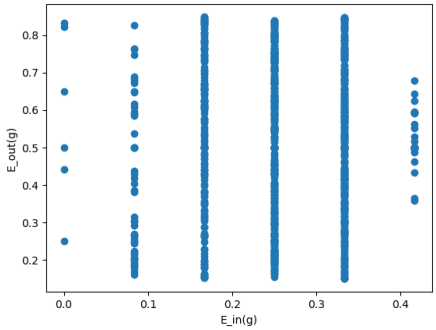
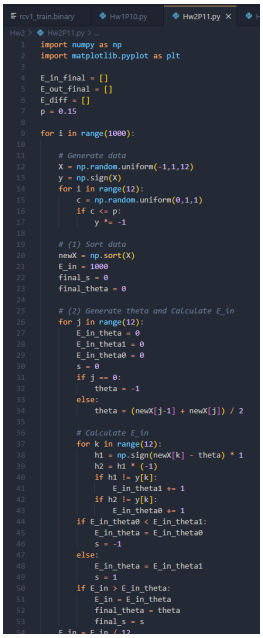


Figure 1: scatter plot

```
The median of E_out - E_in is  0.25
```

Figure 2: median



```python
import numpy as np
import matplotlib.pyplot as plt

E_in_final = []
E_out_final = []
E_diff = []
p = 0.15

for i in range(1000):

    # Generate data
    X = np.random.uniform(-1,1,12)
    y = np.sign(X)
    for i in range(12):
        c = np.random.uniform(0,1,1)
        if c <= p:
            y *= -1

    # (1) Sort data
    newX = np.sort(X)
    E_in = 1000
    final_s = 0
    final_theta = 0

    # (2) Generate theta and Calculate E_in
    for j in range(12):
        E_in_theta = 0
        E_in_theta1 = 0
        E_in_theta0 = 0
        s = 0
        if j == 0:
            theta = -1
        else:
            theta = (newX[j-1] + newX[j]) / 2

        # Calculate E_in
        for k in range(12):
            h1 = np.sign(newX[k] - theta) * 1
            h2 = h1 * (-1)
            if h1 != y[k]:
                E_in_theta1 += 1
            if h2 != y[k]:
                E_in_theta0 += 1
        if E_in_theta0 < E_in_theta1:
            E_in_theta = E_in_theta0
            s = -1
        else:
            E_in_theta = E_in_theta1
            s = 1
        if E_in > E_in_theta:
            E_in = E_in_theta
            final_theta = theta
            final_s = s
    E_in = E_in / 12
```

Figure 3: snapshot

Problem 12

For the P11, we choose θ, s to minimize E_{in}. And now, we choose θ, s uniformly. From the median, we can easily observe that the $E_{out} - E_{in}$ is almost 0. The conclusion is that if we choose θ, s to minimize E_{in}, we will have more $E_{out} - E_{in}$, while if we choose θ, s uniformly, we can minimize $E_{out} - E_{in}$ but we will have more E_{in}.

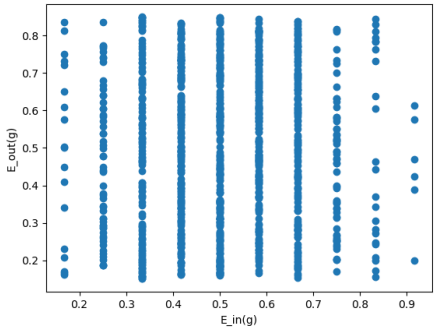


Figure 4: scatter plot

```
The median of E_out - E_in is  -0.014203931743953596
```

Figure 5: median