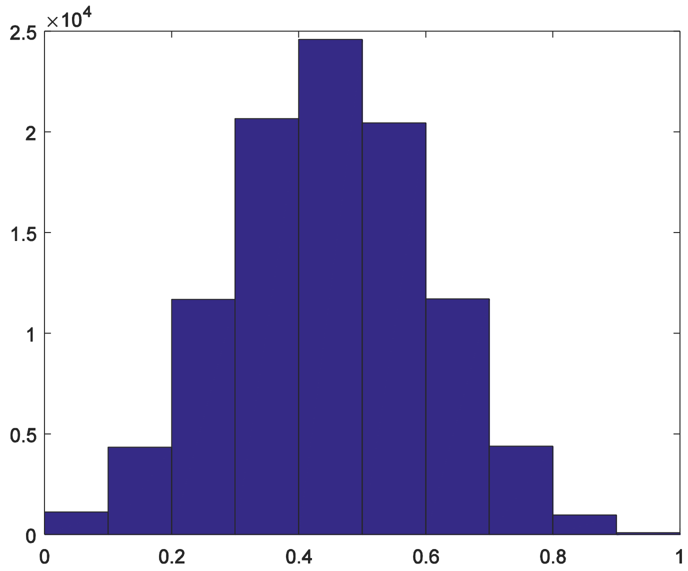
Homework1

(I discussed this homework with Liuke ,Xiaoxingxu and TAs.)

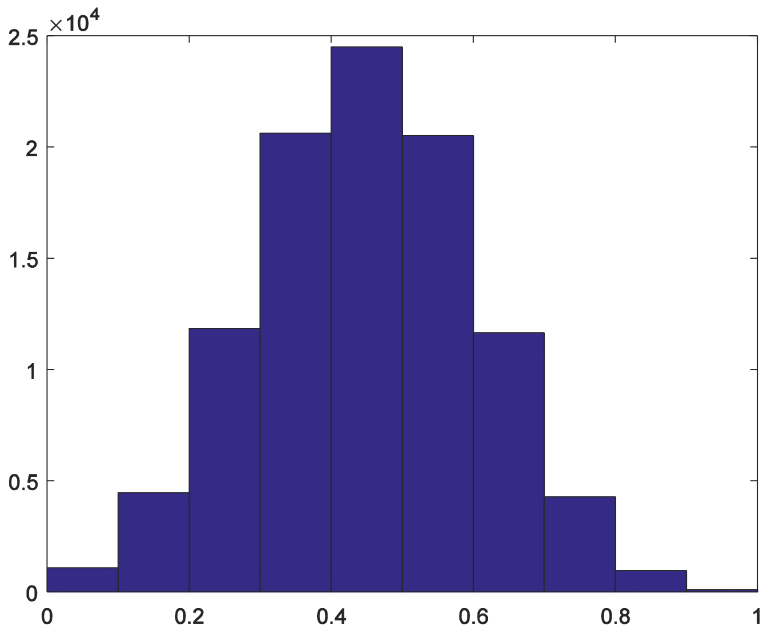
1.(a)u=sum(sum(round(rand(10,1000))))/(10*1000)

$$\mu = 0.5$$

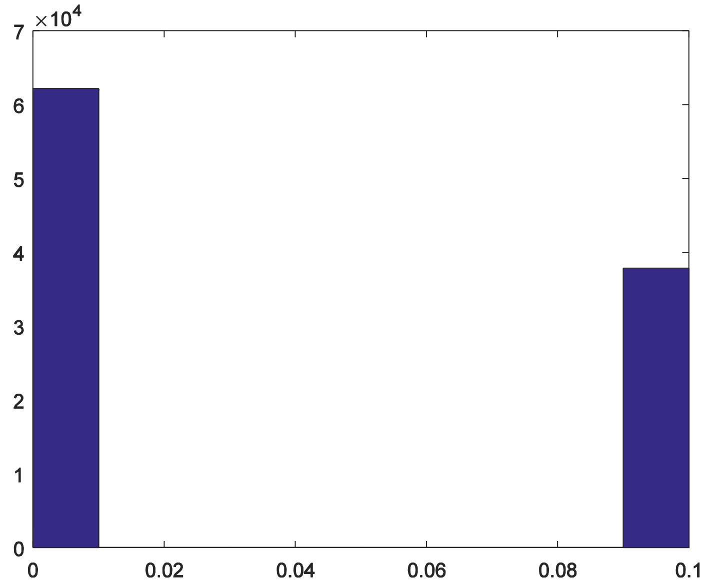
（b）for the first coin:The X axis is v. And Y axis is the time of experiments.



For the random coin:The X axis is v. And Y axis is the time of experiments.

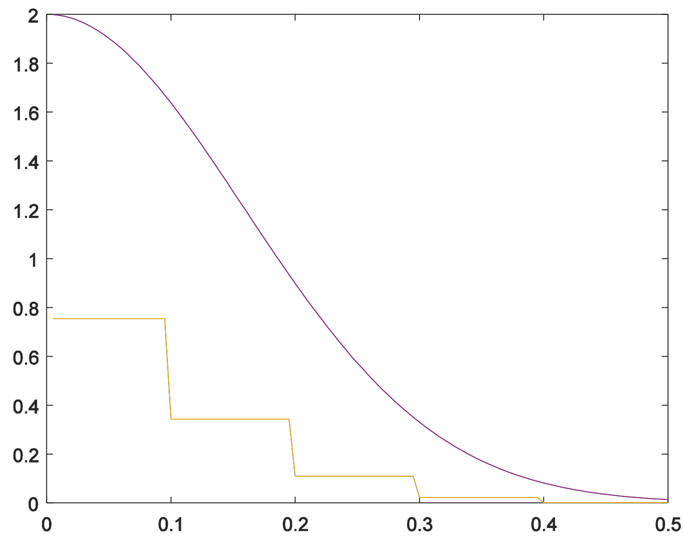


The min coin: The X axis is v. And Y axis is the time of experiments.



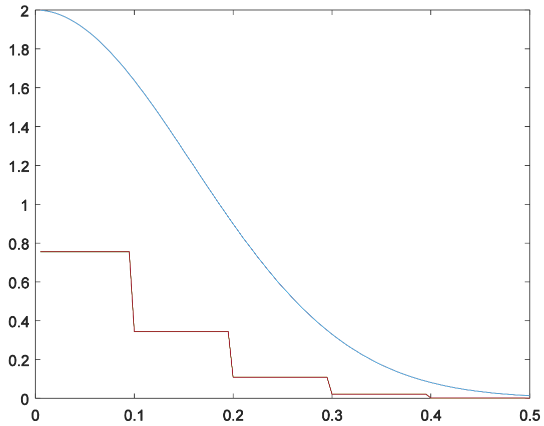
(c):For the first coin: The X axis is ε. And Y axis is P.

The purple one is the hoeffding bound. And the orange one is the P of the first coin



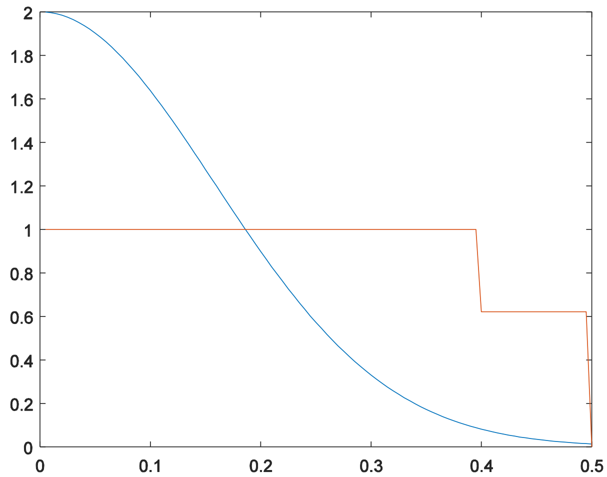
For a random coin: The X axis is ε. And Y axis is P.

The blue one is hoeffding bound. The brown one is P of the random coin.



For the min coin: The X axis is ε. And Y axis is P.

The blue one is the hoeffding bound and the red one is the P of the min coin.



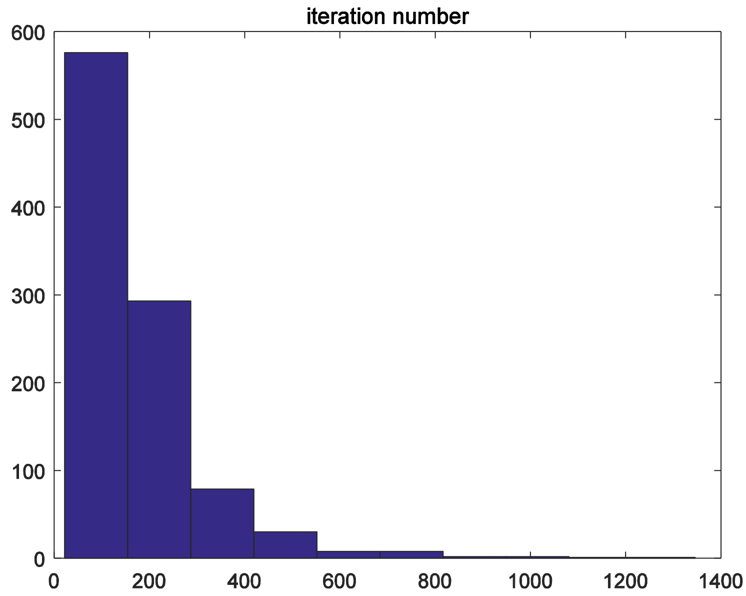
(d)The first coin and the random coin obey the Hoeffting bound. And the the min coin does not obey it. Because the distribution of the min coin is not random.

(e) Those bins are not obey the Hoeffding bound. Because the probability of get red marbles and green marbles is not random.

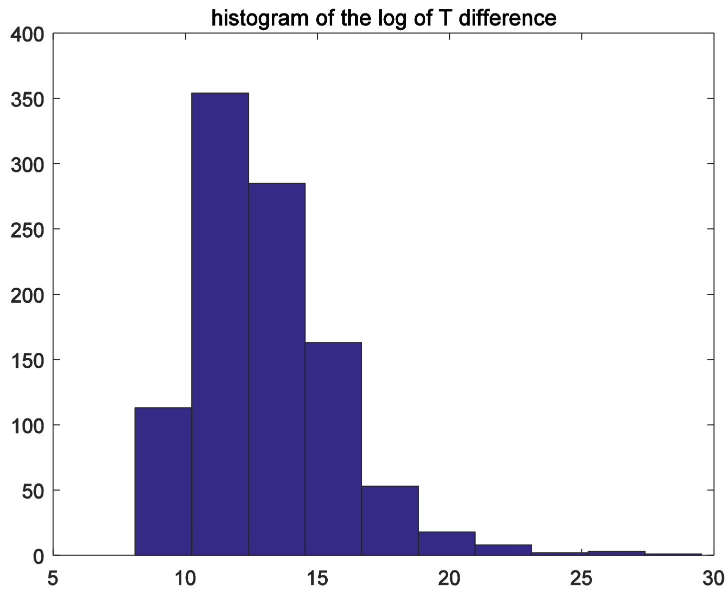
2.See attachment files.

In this case, inputs of perceptron_experiment() is N = 100; d = 11; num_samples = 1000;

The X axis is the number of iteration and the Y axis is the time of experiments.



The X axis is the log of T difference and the Y axis is the time of experiments



3.(a)When $\mu = 0.05$

When N =10*1, we have: $\mathrm{P} = 1 - P_{(all\ \ tail)} = 0.05^0 * (1 - 0.05)^{10} = 0.5987$

When N=10*1,000 , we have : $\mathrm{P} = 1 - (0.5987)^{1000} = 1$

When N=10*1,000,000, we have $\mathrm{P} = 1 - (0.5987)^{1000000} = 1$

When $\mu = 0.8$

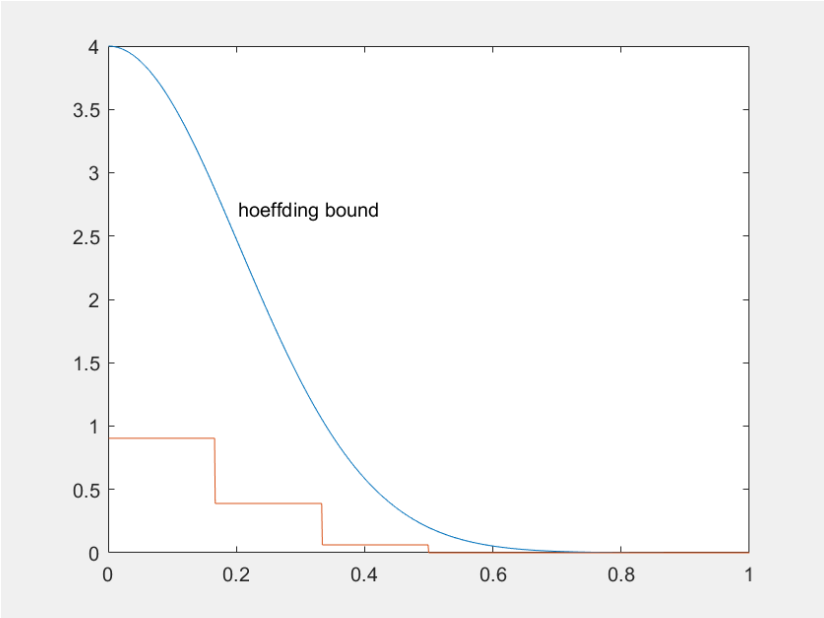
When N =10*1, we have: $P = 1 - P_{(all\ \ tail)} = 0.8^0 * (1 - 0.8)^{10} = 1.024 * 10^{-7}$.

When N=10*1,000 , we have : $P = 1 - (1.024 * 10^{-7})^{1000} = 1.0239 * 10^{-4}$

When N=10*1,000,000, we have $P = 1 - (1.024 * 10^{-7})^{1000000} = 0.0973$

(b)The plot is showing below: the orange one is probability.

X axis is ε and Y axis is P.



4.(a)

$$\alpha P(t \geq \alpha) = \int_{\alpha}^{\infty} \alpha P(t)dt$$

because $\alpha \geq 0$ and $t \geq \alpha$, we have

$$\alpha P(t \geq \alpha) = \int_{\alpha}^{\infty} \alpha P(t)dt \leq \int_{0}^{\infty} \alpha P(t)dt \leq \int_{0}^{\infty} tP(t)dt = E(t)$$

Thus, we have $\alpha P(t \geq \alpha) \leq E(t)$

$$P(t \geq \alpha) \leq \frac{E(t)}{\alpha}$$

(b)

$$\sigma^2{}_u = E[((u - E(u))^2)]$$

$$= E[(u-\mu)^2]$$
$$= \sigma^2$$

From the problem(a), we have: $\mathrm{P}(\mathrm{t} \geq \alpha) \leq \frac{E(t)}{\alpha}$

$$\text{Thus, } \mathrm{P}[(\mathrm{u}-\mu)^2 \geq \alpha] \leq \frac{E[(\mathrm{u}-\mu)^2]}{\alpha} = \frac{\sigma^2}{\alpha}$$

(c)

$$\text{Because } \mathrm{u} = \frac{\sum_{i=0}^{N} u_n}{N}$$

$$\mathrm{Var(u)} = \mathrm{Var}\left(\frac{\sum_{i=0}^{N} u_n}{N}\right)$$
$$= \frac{1}{N^2} Var\left(\sum_{i=0}^{N} u_n\right)$$
$$= \frac{1}{N^2} \sum_{i=0}^{N} Var(u_n)$$
$$= \frac{1}{N^2} * \mathrm{N} * \sigma^2$$
$$= \frac{\sigma^2}{N}$$

From the conclusion of problem(b), we have

$$\mathrm{P}(\mathrm{t} \geq \alpha) \leq \frac{\sigma^2}{N\alpha}$$

5.(a)

$$\mathrm{E} = \sum_{n=1}^{N} (h - y_n)^2$$

$$\mathrm{E}' = 2N^2 h_{mean} - 2N h_{mean} \sum_{N=1}^{N} y_n$$

When E is minimum, E=0, we have

$$2N^2 h_{mean} = 2N h_{mean} \sum_{N=1}^{N} y_n$$

Thus,

$$h_{mean} = \frac{\sum_{n=1}^{N} y_n}{N}$$

(b)Set when n=i, E is minimum:

$$\mathrm{E} = \sum_{n=1}^{N} |h - y_n|$$

$$\text{Thus}, \mathrm{E}' = \sum_{n=1}^{i} 1 - \sum_{i}^{N} 1 \quad or \quad \mathrm{E}' = -\sum_{n=1}^{i} 1 + \sum_{i}^{N} 1$$

Let E' =0, we have

$$\sum_{n=1}^{i} 1 = \sum_{i}^{N} 1$$

Which means i should be the median of N.

Thus, the half points are at most h_{med} and half points are at least h_{med}.

(c)hen $\mathrm{y} \to \infty$, $\sum_{n=1}^{N} y_n \to \infty, Thus\ h_{mean} \to \infty$.

But the h_{med} $does\ not\ change.\ \ Because\ not\ wonder\ how\ y_n\ changes,$ $it\ has\ no\ connection\ to\ the\ value\ of\ the\ h_{med}.$