1. Exercise 1.10

(a) μ is 0.5.
code: coin=randi([0:1], [1000,10]);

(b)

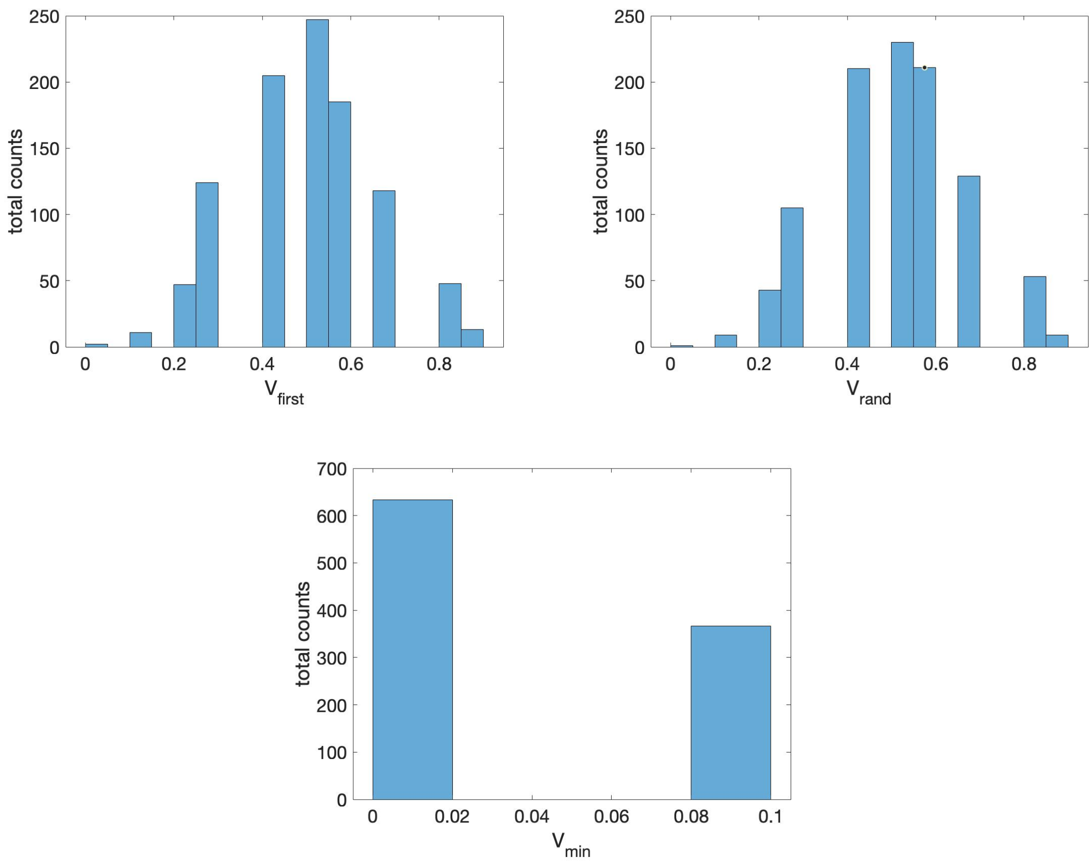
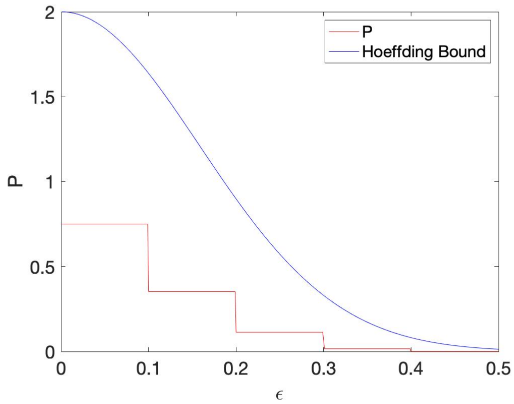
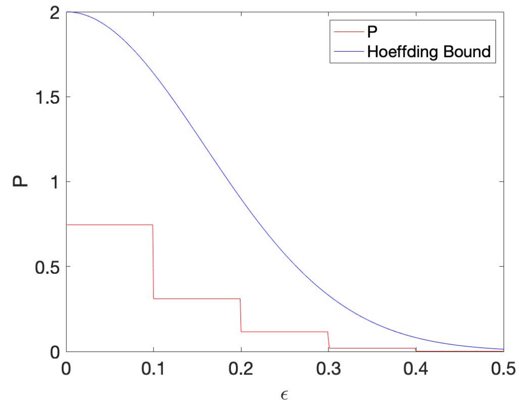


Figure 1: V_1, V_{rand} *and* V_{min}

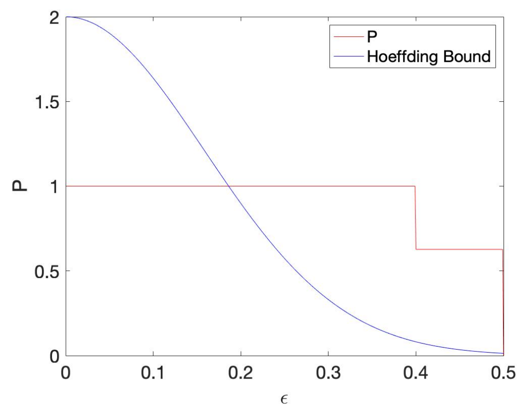
(c)



(a) v_{first}



(b) v_{rand}



(c) v_{min}

Figure 2: P with V_1, V_{rand} *and* V_{min}

(d)

First coin and random coin obey the Hoeffding bound, but min coin do not since we will use a fixed hypothesis which will have $v = 0.5$ if we choose one coin or every time randomly. In this case, we select the min every time which like we change our hypothesis to a smaller v after generating the data set, thus, the Hoeffding bound doesn't hold.

(e)

It's like every time we choose the sample with the least red ball chosen from the bin that have half green balls and red balls, which is the same as we choose some balls from the bin that have very few red balls. So the bin's balls changed, the hypothesis changed, the Hoeffding doesn't hold.

2.

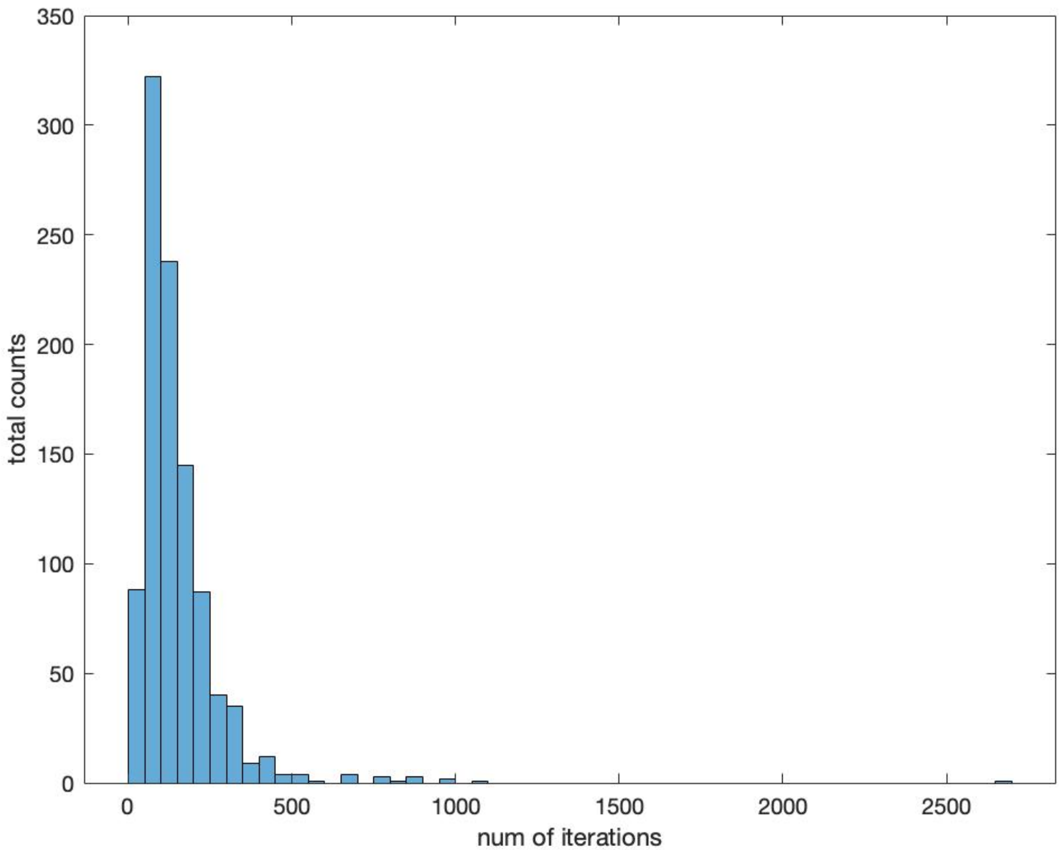


Figure 3: number of iteration

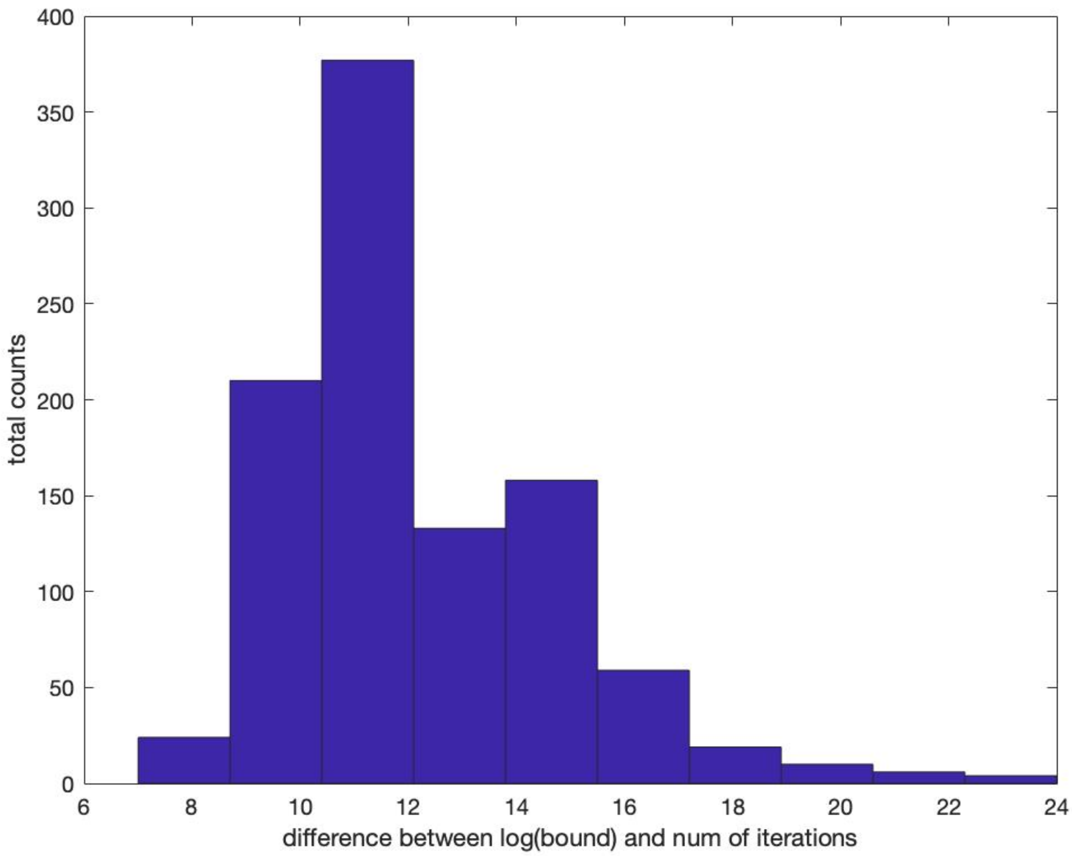


Figure 4: difference between log bound and number of iteration

From this experiment, number of iteration is far below the theoretical bond.

3. LFD 1.7

(a)

For 1 coin:

$$P_0 = (1-\mu)^{10}$$
$$\Rightarrow P_0 \approx 0.6,\ if\ \mu = 0.05$$
$$P_0 \approx 1 \times 10^{-7},\ if\ \mu = 0.8$$

For 1000 coins:

$$P_0 = 1 - (1 - (1 - \mu)^{10})^{1000}$$
$$\Rightarrow P_0 \approx 1,\ if\ \mu = 0.05$$
$$P_0 \approx 1.024 \times 10^{-4},\ if\ \mu = 0.8$$

For 1,000,000 coin:

$$P_0 = 1 - (1 - (1 - \mu)^{10})^{1000000}$$
$$\Rightarrow P_0 \approx 1,\ if\ \mu = 0.05$$
$$P_0 \approx 0.0973,\ if\ \mu = 0.8$$

(b)

$$\begin{aligned} P(\max_i |v_i - \mu_i| > \epsilon) &= P(|v_1 - \mu_1| > \epsilon\ or\ |v_2 - \mu_2| > \epsilon) \\ &= P(|v_1 - \mu_1| > \epsilon) + P(|v_2 - \mu_2| > \epsilon) - P(|v_1 - \mu_1| > \epsilon)P(|v_2 - \mu_2| > \epsilon) \\ &\leq 4e^{-12\epsilon^2} \end{aligned}$$

4. LFD 1.8

(a)

Assume there are N t, and have M t $\leq \alpha$, we have:

$$E(t) = \sum_{i=0}^{N} t_i P(t_i)$$
$$\alpha P(t \geq \alpha) = \sum_{j=0}^{M} \alpha P(t_j) \leq \sum_{j=0}^{M} t_j P(t_j)$$
$$Since\ t\ is\ non-negative$$
$$\leq \sum_{i=0}^{N} t_i P(t_i) = E(t)$$

Proved.

(b)

According to (a)

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

(c)

We know u is a random variable with mean μ and variance $\frac{\sigma^2}{N}$, so we replace σ^2 by $\frac{\sigma^2}{N}$, then we have

$$P[(u-\mu)^2 \geq \alpha] \leq \frac{\sigma^2}{N\alpha}$$

5. LFD 1.12

(a)

To minimize value, we have

$$\frac{\mathrm{d}E_{in}(h)}{\mathrm{d}h} = 2\sum_{n=1}^{N}(h - y_n) = 0$$
$$\Rightarrow h_{mean} = \frac{1}{N}\sum_{n=1}^{N} y_n$$

Which indeed h will be h_{mean} since

$$\frac{\mathrm{d}^2 E_{in}(h)}{\mathrm{d}^2 h} = 2N > 0$$

(b)

Denote

$$I_0 = 1, \text{ when } h > y_n$$
$$I_0 = -1, \text{ when } h \le y_n$$

To minimize value, we have

$$\frac{\mathrm{d}E_{in}(h)}{\mathrm{d}h} = \sum_{n=1}^{N}(I_0 h) = 0$$

Which means there are k $y_n < h$ and there are $N - k$ $y_n \ge h$, thus, h will be h_{med}.

(c)

In that case, $h_{mean} \to \inf$, h_{med} will still be h_{med}.

6. LFD 2.3

(a)

The positive ray will divide the region up to N + 1 regions, double it for negative ray and minus overlap 2, we have:

$$\begin{aligned} m_N(H) &= 2N \\ m_2(H) = 2^2,\ m_3(H) &= 6 < 2^3 \\ &\Rightarrow d_{vc} = 2 \end{aligned}$$

(b)

The positive ray will divide the region up to $N+1$ choose 2 +1 regions, plus negative interval which will add $N-2$ if $N > 1$, so:

$$m_N(H) = C_{N+1}^2 + 1 + N - 2 = \frac{N^2}{2} + \frac{3N}{2} - 1$$
$$m_3(H) = 2^3,\ m_4(H) = 13 < 2^4$$
$$\Rightarrow d_{vc} = 3$$

(c)

It can be seem as a positive interval, so:

$$m_N(H) = C_{N+1}^2 + 1 = \frac{N^2 + N}{2} + 1$$
$$m_2(H) = 2^2,\ m_3(H) = 7 < 2^3$$
$$\Rightarrow d_{vc} = 2$$

7 LFD 2.8

If $m_H(N)$ is 2^N, or we can find a d_{vc} that make $m_H(N)$ bounded by $N_{vc}^d + 1$, then it's a possible $m_H(N)$. So:

$$1 + N, \leq N^{d_{vc}=1} + 1 \Rightarrow Yes.$$
$$1 + N + \frac{N(N-1)}{2} \leq N^{d_{vc}=2} + 1 \Rightarrow Yes.$$
$$2^N \Rightarrow Yes.$$
$$2^{\sqrt{N}},\ d_{vc} = 1,\ \textit{but when } N = 36,\ \textit{bound } N^{d_{vc}} + 1 \textit{ not hold.}$$
$$2^{\frac{N}{2}},\ d_{vc} = 0,\ \textit{bound } N^{d_{vc}} + 1 \textit{ not hold.}$$
$$1 + N + \frac{N(N-1)(N-2)}{6},\ d_{vc} = 1,\ \textit{bound } N^{d_{vc}} + 1 \textit{ not hold.}$$