Machine Learning 10-701
http://alex.smola.org/teaching/cmu2013-10-701x/
Carnegie Mellon University

Homework 3 Solutions

---

1 Probability Inequalities [Jing; 25 pts]

(a)

$$\begin{aligned} P(\bar{X}_n \geq 0.5) &\leq \frac{E[\bar{X}_n]}{0.5} \\ &\leq \frac{0.2}{0.5} \\ &\leq \frac{2}{5} \end{aligned}$$

(b) First, calculate the variance of a bernoulli random variable: $Var(X) = p(1-p)$.

$$\begin{aligned} Var(X_i) &= 0.2(1-0.2) \\ Var(X_i) &= 0.2(0.8) \\ Var(X_i) &= 0.16 = \frac{4}{25} \end{aligned}$$

Now, find the bound.

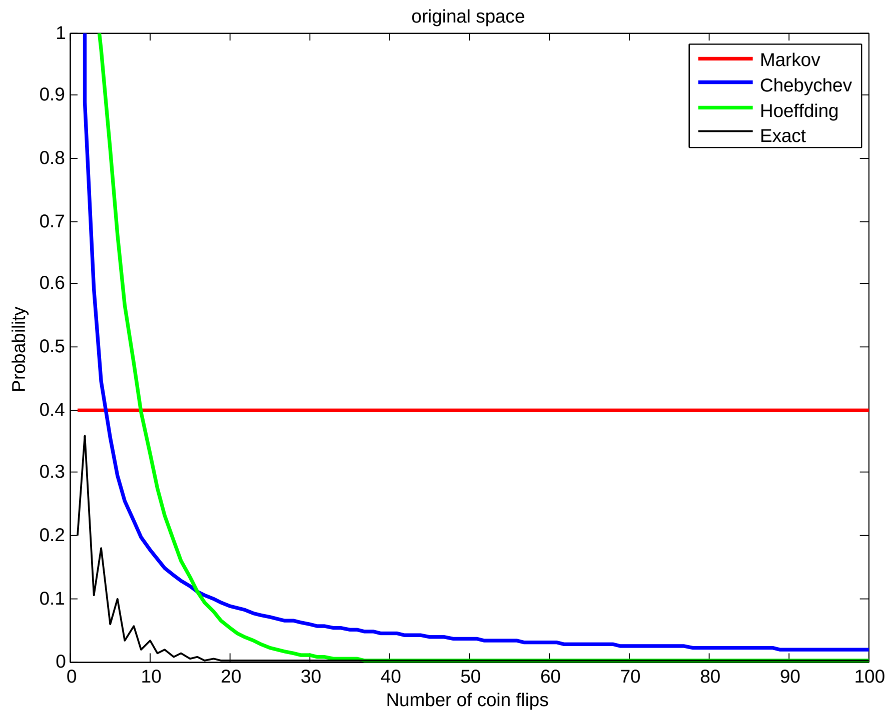
$$\begin{aligned} P(\left|\bar{X}_n - E[\bar{X}_n]\right| \geq k) &\leq \frac{Var[\bar{X}_n]}{k^2} \\ P(\left|\bar{X}_n - 0.2\right| \geq 0.3) &\leq \frac{Var[\frac{1}{n}\sum_{i=1}^{n} X_i]}{(0.3)^2} \\ P(\bar{X}_n \geq 0.5) &\leq \frac{\left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var(X_i)}{0.09} \\ P(\bar{X}_n \geq 0.5) &\leq \frac{\left(\frac{1}{n}\right)^2 n(0.16)}{0.09} \\ P(\bar{X}_n \geq 0.5) &\leq \frac{100 \cdot 4}{9n \cdot 25} \\ P(\bar{X}_n \geq 0.5) &\leq \frac{16}{9n} \end{aligned}$$

(c)

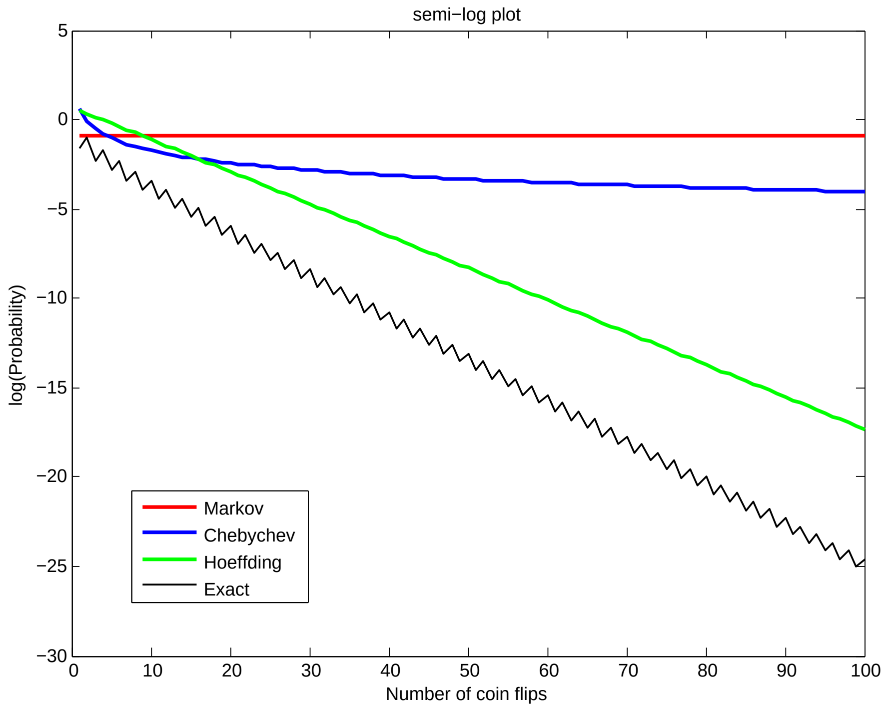
$$\begin{aligned} P(\left|\bar{X}_n - E[\bar{X}_n]\right| \geq \epsilon) &\leq 2e^{-2n\epsilon^2/(b-a)^2}c \\ P(\left|\bar{X}_n - 0.2\right| \geq 0.3) &\leq 2e^{-2n(0.3)^2/(1-0)^2} \\ P(\bar{X}_n \geq 0.5) &\leq 2e^{\frac{-2n \cdot 9}{100}} \\ P(\bar{X}_n \geq 0.5) &\leq 2e^{\frac{-9n}{50}} \end{aligned}$$

Homework 3 Solutions

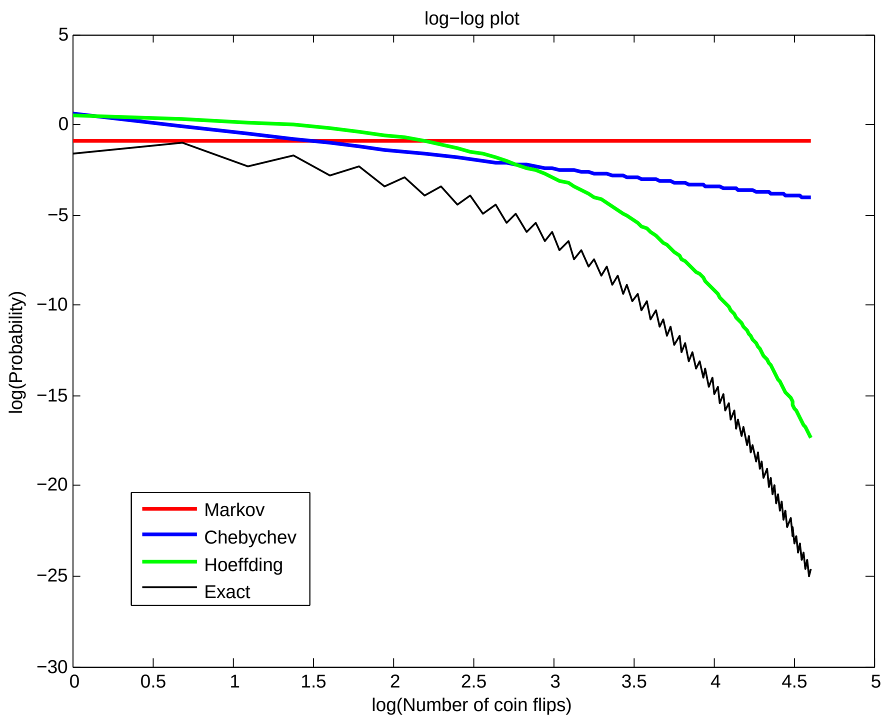
(d) and (e)



(f)


semi−log plot
log(Probability)
Number of coin flips
Markov
Chebychev
Hoeffding
Exact
5
0
−5
−10
−15
−20
−25
−30
0
10
20
30
40
50
60
70
80
90
100

(g)



(h)

Original Space

- Markov is linear, has a $y = b$ relationship, $y = \frac{2}{5}$.

Semi-log Space

- Markov is linear, has $y = log(b)$, which is still a straight line.
- Hoeffding is linear, because taking the log of an exponential will get you a linear $y = ax$ relationship, where $a = -9/50$.

log-log Space

- Markov is still linear.
- Chebychev is linear, because $y = x^b$ is a line with slope $b = -1$