Assignment 3

2023-12675 박지호

4.2

- Let

  $$X_i = \begin{cases} 0 & \text{(if the result of } i\text{-th dice roll is 6)} \\ 1 & \text{(otherwise)} \end{cases}$$

  - $$\mathbf{E}[X_i] = \frac{1}{6}$$
  - $$\mathbf{Var}[X_i] = \mathbf{E}[X_i^2] - (\mathbf{E}[X_i])^2 = \frac{1}{6} - \frac{1}{36} = \frac{5}{36}$$
- $$X = \sum_{i=1}^{n} X_i$$
  - $$\mathbf{E}[X] = n \times \frac{1}{6}$$
  - $$\mathbf{Var}[X] = n \times \frac{5}{36}$$
  - $$\sigma[X] = \sqrt{\frac{5n}{36}}$$
- Markov's inequality:
  - $$p = \Pr\left[X \geq \frac{n}{4}\right] \leq \frac{\mathbf{E}[X]}{\frac{n}{4}} = \frac{2}{3}$$
  - The bound obtained with Markov's inequality is constant with respect to n, and isn't too practical
- Chevyshev's inequality:
  - $$p = \Pr\left[X \geq \frac{n}{4}\right] = \Pr\left[\left|X - \frac{n}{4}\right| \geq \frac{n}{12}\right] \leq \frac{\mathbf{Var}[X]}{\left(\frac{n}{12}\right)^2} = \frac{20}{n}$$
  - The bound obtained by Chebyshev's inequality gets smaller as n grows, and for a n larger than 30, is smaller than the bound from Markov's inequality
- Chernoff bound (Using the formula for the sum of Poisson trials)
  - $$p = \Pr\left[X \geq \frac{n}{4}\right] = \Pr\left[X \geq \left(1 + \frac{1}{2}\right)\mathbf{E}[X]\right] \leq \left(\frac{e^{\frac{1}{2}}}{\left(1 + \frac{1}{2}\right)^{1+\frac{1}{2}}}\right)^{\frac{n}{6}} = \left(\frac{8e}{27}\right)^{\frac{n}{12}}$$
  - The Chernoff bound gets exponentially smaller as n grows, and for a large enough n it is smaller than both Markov's inequality bound and Chebyshev's inequality bound.
- Chernoff bound (General)
  - We'll use the moment generating function from 4.3.(a)

- $$p = \Pr\left[X \geq \frac{n}{4}\right] \leq \min_{t>0} \left(\frac{1}{6}e^t + \frac{5}{6}\right)^n e^{-\frac{1}{4}nt}$$
- $$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{1}{6}e^t + \frac{5}{6}\right)^n e^{-\frac{1}{4}nt}\right) = 0 \\ &\Leftrightarrow \frac{1}{6}ne^t\left(\frac{1}{6}e^t + \frac{5}{6}\right)^{n-1} e^{-\frac{1}{4}nt} - \frac{1}{4}n\left(\frac{1}{6}e^t + \frac{5}{6}\right)^n e^{-\frac{1}{4}nt} = 0 \\ &\Leftrightarrow \frac{1}{6}e^t - \frac{1}{4}\left(\frac{1}{6}e^t + \frac{5}{6}\right) = 0 \\ &\Leftrightarrow \frac{1}{8}e^t - \frac{5}{24} = 0 \\ &\Leftrightarrow e^t = \frac{5}{3} \\ &\Leftrightarrow t = \ln\frac{5}{3} > 0 \end{aligned}$$
  - For this t, we get
    $$\Pr\left[X \geq \frac{n}{4}\right] \leq \left(\frac{2^4 \times 5^3}{3^7}\right)^{\frac{n}{4}}$$
- $$\left(\frac{8e}{27}\right)^{\frac{1}{12}} \approx 0.9821, \left(\frac{2^4 \times 5^3}{3^7}\right)^{\frac{1}{4}} \approx 0.9779$$
- This bound also gets exponentially smaller as n grows, and is better than the Poisson-specific version.

4.3

(a)

- Let X be the binomial random variable with the distribution of $\mathbf{B}(n,p)$
  - We can represent X as
    $$X = \sum_{i=1}^{n} X_i$$
  - Where X_is are the independent random variables with the distribution of $\mathbf{Bernoulli}(\boldsymbol{p})$
- $$\begin{aligned} M_{X_i}(t) &= \mathbf{E}[e^{tX_i}] \\ &= p \times e^t + (1-p) \times e^0 \end{aligned}$$
  - $$M_{X(t)} = \prod_{i=1}^{n} M_{X_i}(t) = (1 - p + pe^t)^n$$

(b)

- $M_{X+Y}(t) = M_X(t) \times M_Y(t) = (1-p+pe^t)^n \times (1-p+pe^t)^m = (1-p+pe^t)^{m+n}$

(c)

- Let Z be the binomial random variable with the distribution of $\mathbf{B}(m+n,p)$
  - $M_Z(t) = M_{X+Y}(t)$
- $X+Y$ has the same distribution as Z

4.4

- Let X be the number of heads obtained after n flips
- Explicit calculation:
  - $$\Pr(X \geq k) = \sum_{i=k}^{n} \binom{n}{i} \left(\frac{1}{2}\right)^{i} \left(\frac{1}{2}\right)^{n-i} = \frac{1}{2^n} \sum_{i=k}^{n} \binom{n}{i}$$
  - Calculated using the python code in Listing 1

```python
import math

n = 100 # The number of coin flips
k = 55 # The amount of heads to get
cases = 0

for i in range(k, n + 1):
    cases += math.comb(n, i)

# Dividing at the end to avoid floating point problems
print(cases / (2 ** n))
```

Listing 1: Python code used to calculate the probability explicitly

- Chernoff bound (Using the formula for the sum of Poisson trials)
  - $\mathbf{E}[X] = \frac{n}{2}$ (trivial)
  - $$\Pr[X \geq k] \leq \Pr\left[X \geq \left(1 + \left(\frac{2k}{n} - 1\right)\right) \times \frac{n}{2}\right] = \left(\frac{e^{\frac{2k}{n}-1}}{\left(\frac{2k}{n}\right)^{\frac{2k}{n}}}\right)^{\frac{n}{2}}$$
- Chernoff bound (General)
  - We'll use the moment generating function from 4.3.(a)
  - $$\Pr[X \geq k] \leq \min_{t>0} \left(\frac{1}{2}e^t + \frac{1}{2}\right)^n e^{-kt}$$

- $\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{1}{2}e^t + \frac{1}{2}\right)^n e^{-kt}\right) = 0$

  $\Leftrightarrow \frac{1}{2}n\left(\frac{1}{2}e^t + \frac{1}{2}\right)^{n-1} e^{-kt} - k\left(\frac{1}{2}e^t + \frac{1}{2}\right)^n e^{-kt} = 0$

  $\Leftrightarrow \frac{1}{2}ne^t - k\left(\frac{1}{2}e^t + \frac{1}{2}\right) = 0$

  $\Leftrightarrow e^t = \frac{k}{n-k}$

  $\Leftrightarrow t = \ln \frac{k}{n-k} > 0 \quad (\text{For given } n, k)$

- Getting 55 or more heads from 100 flips
  - Explicit calculation: 0.184100
  - Chernoff bound (Using the formula for the sum of Poisson trials): $\Pr[X \geq k] \leq 0.785009$
  - Chernoff bould (General): $\Pr[X \geq k] \leq 0.606023$
  - The Chernoff bound is way larger than the actual probability
- Getting 550 or more heads from 1000 flips:
  - Explicit calculation: 8.65268×10^{-4}
  - Chernoff bound (Using the formula for the sum of Poisson trials): $\Pr[X \geq k] \leq 8.88684 \times 10^{-2}$
  - Chernoff bould (General): $\Pr[X \geq k] \leq 6.68181 \times 10^{-3}$
  - The Chernoff bound is still way larger than the actual probability, and the scale difference is much larger

4.9

(a)

- Let n be the number of samples
- $\overline{X} = \frac{1}{n}\sum_{i=0}^{n} X_i$
  - $\mathbf{E}\left[\,\overline{X}\,\right] = \frac{1}{n}\sum_{i=0}^{n} \mathbf{E}[X_i] = \mathbf{E}[X]$
  - $\mathbf{Var}\left[\,\overline{X}\,\right] = \frac{1}{n^2}\sum_{i=0}^{n} \mathbf{Var}[X_i] = \frac{1}{n}\,\mathbf{Var}[X]$
- $\Pr\left[\left|\overline{X} - \mathbf{E}[X]\right| \geq \varepsilon\mathbf{E}[X]\right] \geq \frac{\frac{1}{n}\,\mathbf{Var}[X]}{(\varepsilon\mathbf{E}[X])^2} = \frac{\mathbf{Var}[X]}{(\mathbf{E}[X])^2} \times \frac{1}{n\varepsilon^2} = \frac{r^2}{n\varepsilon^2}$

- $\Pr\left[\left|\overline{X} - \mathbf{E}[X]\right| \leq \varepsilon\mathbf{E}[X]\right] > 1 - \delta \Leftrightarrow \Pr\left[\left|\overline{X} - \mathbf{E}[X]\right| \geq \varepsilon\mathbf{E}[X]\right] < \delta$

$$\Leftarrow \frac{r^2}{n\varepsilon^2} < \delta$$

$$\Leftrightarrow n > \frac{r^2}{\delta\varepsilon^2}$$

- $\mathrm{O}\left(\frac{r^2}{\delta\varepsilon^2}\right)$ samples are sufficient to solve the problem

(b)

- Applying $\delta = \frac{1}{4}$ to (a)
  - $n > 4 \times \frac{r^2}{\varepsilon^2}$
- $\mathrm{O}\left(\frac{r^2}{\varepsilon^2}\right)$ samples are are enough for this estimate

(c)

- From (b), we know that with $4 \times \frac{r^2}{\varepsilon^2}$ samples, we can ensure $\Pr\left[\left|\overline{X} - \mathbf{E}[X]\right| \leq \varepsilon\mathbf{E}[X]\right] \geq \frac{3}{4}$
- Let's say that we repeat this process of picking $4 \times \frac{r^2}{\varepsilon^2}$ samples to gain m estimate values.
  - If more than half of the estimates is within the $\varepsilon\mathbf{E}[X]$ of $\mathbf{E}[X]$, the median is within the $\varepsilon[X]$ of $\mathbf{E}[X]$
- Let Y be the number of estimates that is within the $\varepsilon\mathbf{E}[X]$ of $\mathbf{E}[X]$
  - Y follows a binomial distribution $B(m, q)$, where q is the probability that each estimate is within the $\varepsilon\mathbf{E}[X]$ of $\mathbf{E}[X]$
    - We've established that $q > \frac{3}{4}$
    - $\mathbf{E}[Y] = mq$
  - $$\Pr\left[Y \leq \frac{1}{2}m\right] \leq \min_{t<0} \frac{\left(1 - q + qe^t\right)^m}{e^{\frac{1}{2}mt}}$$
    $$= \min_{t<0} \left((1-q)e^{-\frac{1}{2}t} + qe^{\frac{1}{2}t}\right)^m$$
    - The minimum value of right hand side is $\left(2\sqrt{q(1-q)}\right)^m$
    - $\Pr\left[Y \leq \frac{1}{2}m\right] \leq \left(\frac{\sqrt{3}}{2}\right)^m (\because q > \frac{3}{4})$
  - $$\Pr\left[Y > \frac{1}{2}m\right] \geq 1 - \delta \Leftrightarrow \Pr\left[Y \leq \frac{1}{2}m\right] < \delta$$
    $$\Leftarrow \left(\frac{\sqrt{3}}{2}\right)^m < \delta$$
    $$m > \frac{\ln\frac{1}{\delta}}{\ln\left(\frac{2}{\sqrt{3}}\right)}$$
- By repeating the process to more than $\mathrm{O}\left(\ln\frac{1}{\delta}\right)$ times, we can ensure that median of estimates is within the $\varepsilon\mathbf{E}[X]$ of $\mathbf{E}[X]$
- For the entire process, we need $\mathrm{O}\left(\frac{r^2}{\varepsilon^2}\right)$ samples $\mathrm{O}\left(\ln\frac{1}{\delta}\right)$ times.
  - Therefore, we need a total of $\mathrm{O}\left(\frac{r^2}{\varepsilon^2}\ln\frac{1}{\delta}\right)$ samples

4.10

- Let X_i be the payout of the i-th game in dollars
  - $\forall i, j$, X_i and X_j are independent
  - $$\mathbf{E}[X_i] = \frac{4}{25} \times 3 + \frac{1}{200} \times 100 = \frac{49}{50}$$
  - $$M_{X(t)} = \mathbf{E}[e^{X_i t}] = \frac{4}{25} \times e^{3t} + \frac{1}{200} \times e^{100t}$$
- Let X be the total payout of the machine over the first million games
  - Let $N = 1,000,000$
  - $$X = \sum_{i=1}^{N} X_i$$
  - $$\mathbf{E}[X] = \sum_{i=1}^{N} \mathbf{E}[X_i] = \frac{49}{50} N$$
  - $$M_X(t) = \prod_{i=1}^{N} M_{X_i}(t) = \left(\frac{4}{25} e^{3t} + \frac{1}{200} e^{100t}\right)^N = e^{3Nt} \left(\frac{1}{200} e^{97t} + \frac{4}{25}\right)^N$$
- The net profit of the machine after N games is $N - X$
  - Let $M = 10,000$
  - $\Pr[N - X \leq -M] = \Pr[X \geq N + M]$
- The Chernoff bound (Using the formula for the sum of Poisson trials):
  - $$\Pr[X \geq N + M] \leq \Pr\left[X \geq \left(1 + \left(\frac{50(N+M)}{49N} - 1\right)\right) \times \frac{49}{50} N\right]$$
  - Let $\delta = \left(\frac{50(N+M)}{49N} - 1\right) = \frac{3}{98}, \mu = 980000$
  - The Chernoff bound (Using the formula for the sum of Poisson trials):

    $$\Pr[X \geq N + M] \leq \left(\frac{e^{\delta}}{(\delta + 1)^{\delta + 1}}\right)^{\mu} = 3.83165 \times 10^{-198}$$
- The Chernoff bound (General):
  - $\Pr[X \geq N + M] \leq \min_{t>0} \frac{M_X(t)}{e^{(N+M)t}} = \min_{t>0} \left(e^{(2N-M)t} \left(\frac{1}{200} e^{97t} + \frac{4}{25}\right)^N\right)$

- $$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N}\right)$$
$$= (2N-M)e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N}+\frac{97}{200}Ne^{97t}e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N-1}$$
$$= \left((2N-M)\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)+\frac{97}{200}Ne^{97t}\right)\times e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N-1}$$
$$= \left(\frac{99N-M}{200}e^{97t}+\frac{4}{25}(2N-M)\right)\times e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N-1}$$
$$> 0$$
  - $$\therefore \min_{t>0}\left(e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N}\right)$$
$$= \lim_{t\to 0+}\left(e^{(2N-M)t}\left(\frac{1}{200}e^{97t}+\frac{4}{25}\right)^{N}\right)$$
$$= \left(\frac{33}{200}\right)^{N} \approx 9.6085\times 10^{-795881}$$
- The Chernoff bound (General): $\Pr[X \geq N+M] \leq 8.7946\times 10^{-782517}$

4.13

(a)

- If $x = 1$
  - $\Pr[X \geq nx] = \Pr[X \geq n] = p^n = e^{-nF(x,p)}$
- If $p = 0$
  - The given function cannot be defined, so we ignore this case
- Oherwise $(0 < p < x < 1)$
  - The Chernoff bound for $\Pr[X \geq xn]$
    - $M_{X_i}(t) = pe^t + (1-p)$
    - $$M_X(t) = \prod_{i}^{n} M_{X_i}(t) = \left(pe^t + 1 - p\right)^n$$
    - $$\Pr(X \geq xn) \leq \min_{t>0} \frac{\left(pe^t+1-p\right)^n}{e^{nxt}}$$

- $\dfrac{\partial}{\partial t}\left(\dfrac{(pe^t+1-p)^n}{e^{nxt}}\right)=0$

  $\Leftrightarrow npe^t(pe^t+1-p)^{n-1}e^{-nxt}-nx(pe^t+1-p)^n e^{-nxt}=0$

  $\Leftrightarrow pe^t-x(pe^t+1-p)=0$

  $\Leftrightarrow p(1-x)e^t-x(1-p)=0$

  $\Leftrightarrow t=\ln\dfrac{x(1-p)}{p(1-x)}$

  - $\frac{x}{1-x}$ is an increasing function for $0<x<1$
    - $1>x>p>0\Rightarrow\dfrac{x}{1-x}>\dfrac{p}{1-p}>0\Rightarrow\dfrac{x(1-p)}{p(1-x)}>1\Rightarrow t>0$
  - For this t, the Chernoff bound is

$$
\begin{aligned}
\Pr(X\geq xn)&\leq\frac{(pe^t+1-p)^n}{e^{nxt}}\\
&=\frac{\left(p\times\frac{x(1-p)}{p(1-x)}+1-p\right)^n}{\left(\frac{x(1-p)}{p(1-x)}\right)^{nx}}=\frac{\left(\frac{1-p}{1-x}\right)^n}{\left(\frac{x(1-p)}{p(1-x)}\right)^{nx}}\\
&=\left(\left(\frac{1-x}{1-p}\right)^{1-x}\left(\frac{x}{p}\right)^x\right)^{-n}\\
&=e^{-n\left(x\ln\frac{x}{p}+(1-x)\ln\frac{1-x}{1-p}\right)}=e^{-nF(x,p)}
\end{aligned}
$$

- $\therefore\Pr(X\geq xn)\leq e^{-nF(x,p)}$

(b)

- Let $G(x,p)=F(x,p)-2(x-p)^2$
  - $$
    \begin{aligned}
    \frac{\partial G}{\partial x}&=\frac{\partial}{\partial x}\left(x\ln\frac{x}{p}+(1-x)\ln\frac{1-x}{1-p}-2(x-p)^2\right)\\
    &=\ln\frac{x}{p}+1-\ln\frac{1-x}{1-p}-1-4(x-p)\\
    &=\ln\frac{x}{p}-\ln\frac{1-x}{1-p}-4(x-p)
    \end{aligned}
    $$
  - $$
    \begin{aligned}
    \frac{\partial^2 G}{\partial x^2}&=\frac{\partial}{\partial x}\left(\ln\frac{x}{p}-\ln\frac{1-x}{1-p}-4(x-p)\right)\\
    &=\frac{1}{x}+\frac{1}{1-x}-4\\
    &=\frac{1-x+x-4x(1-x)}{x(1-x)}\\
    &=\frac{(1-2x)^2}{x(1-x)}\geq 0
    \end{aligned}
    $$
- Conjecture: $\forall x,p>0, G(x,p)\geq 0$
  - $x=p: G(p,p)=0\geq 0$, the conjecture holds

- $\frac{\partial G}{\partial x}(p,p) = 0$
- $x > p$:
  - $\frac{\partial G}{\partial x}(x,p) = \frac{\partial G}{\partial x}(p,p) + \int_p^x \frac{\partial^2 G}{\partial x^2}(t,p)\,\mathrm{d}t \geq 0$
  - $G(x,p) = G(p,p) + \int_p^x \frac{\partial G}{\partial x}(t,p)\,\mathrm{d}t \geq 0$
- $x < p$:
  - $\frac{\partial G}{\partial x}(x,p) = \frac{\partial G}{\partial x}(p,p) + \int_p^x \frac{\partial^2 G}{\partial x^2}(t,p)\,\mathrm{d}t \leq 0$
  - $G(x,p) = G(p,p) + \int_p^x \frac{\partial G}{\partial x}(t,p)\,\mathrm{d}t \geq 0$

- $\therefore \forall x, p, G(x,p) = F(x,p) - 2(x-p)^2 \geq 0$

(c)

- $\Pr[X \geq (p+\varepsilon)n] \leq e^{-nF(p+\varepsilon,p)} \leq e^{-n\times 2(\varepsilon+p-p)^2} = e^{-2n\varepsilon^2}$

(d)

- Conjecture: $\forall x < p, \Pr[X \leq xn] \leq e^{-nF(x,p)}$
  - If $x = 0$
    - $\Pr[X \leq xn] = \Pr[X \leq 0] = (1-p)^n = e^{-nF(x,p)}$
  - If $p = 1$
    - The given function cannot be defined, so we ignore this case
  - Otherwise (0 < x < p < 1)
    - The Chernoff bound for $\Pr[X \leq xn]$
      - $\Pr[X \leq xn] \leq \min_{t<0} \frac{(pe^t+1-p)^n}{e^{nxt}}$
    - The value of t that minimizes the right hand side is the same as (a), $t = \ln \frac{x(1-p)}{p(1-x)}$
      - This time, $t < 0$ since $0 < x < p < 1$
      - The minimum value of the right hand side is also the same as (a)
    - Hence, following the same step as (a), we get $\Pr(X \geq xn) \leq e^{-nF(x,p)}$
  - $\therefore \Pr(X \geq xn) \leq e^{-nF(x,p)}$
- $\Pr[X \leq (p-\varepsilon)n] \leq e^{-nF(p-\varepsilon,p)} \leq e^{-2n\varepsilon^2}$
- $\therefore \Pr[|X - pn| \geq \varepsilon n] = \Pr[X \leq (p-\varepsilon)n] + \Pr[X \geq (p+\varepsilon)n] \leq 2e^{-2n\varepsilon^2}$