Homework3

2.18

When $Y = 4$, there are 2 cases. $p(X = AAAA) = \frac{1}{16}, p(X = BBBB) = \frac{1}{16}$

When Y = 5, there are $2 \times 4 = 8$ cases, with $p = \frac{1}{32}$

When $Y = 6$, there are $2 \times C_5^2 = 20$ cases, with $p = \frac{1}{64}$

When $Y = 7$, there are $2 \times C_6^3 = 40$ cases, with $p = \frac{1}{128}$

$$
\begin{aligned}
H(X) &= -(\sum_{i=1}^{2} \frac{1}{16}\log\frac{1}{16} + \sum_{i=1}^{8} \frac{1}{32}\log\frac{1}{32} + \sum_{i=1}^{20} \frac{1}{64}\log\frac{1}{64} + \sum_{i=1}^{40} \frac{1}{128}\log\frac{1}{128}) \\
&= \frac{1}{2} + \frac{5}{4} + \frac{15}{8} + \frac{35}{16} \\
&= 5.8125
\end{aligned}
$$

$$
\begin{aligned}
H(Y) &= -(\frac{1}{8}\log\frac{1}{8} + \frac{1}{4}\log\frac{1}{4} + \frac{5}{16}\log\frac{5}{16} + \frac{5}{16}\log\frac{5}{16}) \\
&= 1.92379
\end{aligned}
$$

Since Y is a function of X, then $H(Y|X) = 0$.

$\because H(X|Y) + H(Y) = H(Y|X) + H(X)$

$\therefore H(X|Y) = H(X) - H(Y) = 3.88871$

2.20

Denote $H(X_1, X_2, \ldots, X_n) = H(\mathcal{X})$

Since $\mathbf{R}$ is the function of $\mathcal{X}$, and $H(\mathbf{R}) + H(\mathcal{X}|\mathbf{R})$thus $H(\mathbf{R}) \le H(\mathcal{X})$

Besides, if we know $H(\mathbf{R}, X_i), i = 1, 2, 3, \cdots n$, then we could determine $\mathcal{X}$. Therefore, $H(\mathbf{R}, X_i) \ge H(\mathcal{X})$

$$H(\mathcal{X}) \le H(\mathbf{R}, X_i) = H(\mathbf{R}) + H(X_i|\mathbf{R}) \le H(\mathbf{R}) + H(X_i)$$

Since $H(X_i) = -p\log p - (1-p)\log(1-p) \le 1$. Therefore,

$$H(\mathcal{X}) \le H(\mathbf{R}, X_i) = H(\mathbf{R}) + H(X_i|\mathbf{R}) \le H(\mathbf{R}) + H(X_i) \le H(\mathbf{R}) + 1$$

2.21

$H(X) = -\sum_p p(x)\log p(x) \ge -\sum_{p(x)\le d} p(x)\log p(x) = \sum_{p(x)\le d} p(x)\log\frac{1}{p(x)} \ge \log\frac{1}{d} \times Pr\{p(x) \le d$

2.27

$$H(\mathbf{p}) = -\sum_{i=1}^{m} p_i \log p_i$$

$$H(\mathbf{q}) = -\sum_{i=1}^{m-1} q_i \log q_i = -[\sum_{i=1}^{m-2} p_i \log p_i + (p_{m-1} + p_m)\log(p_{m-1} + p_m)]$$

$$
\begin{aligned}
H(\mathbf{p})-H(\mathbf{q}) &= -(p_{m-1}\log p_{m-1}+p_m\log p_m)+(p_{m-1}+p_m)\log(p_{m-1}+p_m) \\
&= (p_{m-1}+p_m)\times(-1)\times(\frac{p_{m-1}}{p_{m-1}+p_m}\log p_{m-1}+\frac{p_m}{p_{m-1}+p_m}\log p_m-\log(p_{m-1}+p_m)) \\
&= (p_{m-1}+p_m)\times(-1)\times(\frac{p_{m-1}}{p_{m-1}+p_m}\log\frac{p_{m-1}}{p_{m-1}+p_m}+\frac{p_m}{p_{m-1}+p_m}\log\frac{p_m}{p_{m-1}+p_m}) \\
&= (p_{m-1}+p_m)H(\frac{p_{m-1}}{p_{m-1}+p_m},\frac{p_m}{p_{m-1}+p_m})
\end{aligned}
$$

2.29

$H(X,Y|Z)=H(X|Z)+H(Y|X,Z)\geq H(X|Z)$

With equality when $H(Y|X,Z)=0$, meaning that the Y is the function of X,Z

$I(X,Y;Z)=I(X;Z)+I(Y;Z|X)\geq I(X;Z)$

With equality when $I(Y;Z|X)=0$, meaning that Y and Z are conditionally independent given X

$H(X,Y,Z)-H(X,Y)=H(Z|X,Y)=H(Z|X)-I(Y;Z|X)\leq H(Z|X)=H(X,Z)-H(X)$

With equality when $I(Y;Z|X)=0$, meaning that Y and Z are conditionally independent given X

$I(X;Z|Y)+I(Z;Y)=I(X,Y;Z)=I(Z;X)+I(Z;Y|X)$

Therefore, this is actually an equality.

2.32

a)

$$
\hat{X}(Y)=\begin{cases}1, Y=a\\2, Y=b\\3, Y=c\end{cases}
$$

$P_e=\frac{1}{2}$

b)

$H(X|Y)=H(X,Y)-H(Y)$

$H(X,Y)=-3\times\frac{1}{6}\times\log\frac{1}{6}-6\times\frac{1}{12}\log\frac{1}{12}=\log 3+1.5$

$H(Y)=-3\times\frac{1}{3}\log\frac{1}{3}=\log 3$

$\therefore H(X|Y)=1.5$

$$
Fano\ inequality: P_e\geq\frac{H(X\mid Y)-1}{\log(|X|-1)}
$$

$P_e=\frac{1}{2}$ satisfies the Fano's inequality, and is also the lower bound.

2.40

a)

$H(X)=-\frac{1}{8}\sum_{i=1}^{8}\log\frac{1}{8}=3$

b)

$H(Y) = -\sum_{k=1}^{\infty} 2^{-k} \cdot (-k) = \lim_{k\to\infty}(2 - \frac{1}{2^{k-1}} - \frac{k}{2^k}) = 2$

c)

$H(X+Y, X-Y) = H(X,Y) = H(X) + H(Y) - I(X;Y) = 5$

3.1

a)

Since $X \geq 0, t > 0$

$Pr\{X \geq t\} \cdot t \leq \sum_{x \geq t} p(x) \cdot x \leq EX$

$$X = \begin{cases} 0, p = 1/2 \\ 1, p = 1/2 \end{cases}$$

let $t = 1$, then the equality is reached.

b)

Let $X = |Y - \mu|^2$

$$\Pr\{|Y - \mu| > \epsilon\} \leq \frac{\sigma^2}{\epsilon^2} \Leftrightarrow \Pr\{X > \epsilon^2\} \leq \frac{EX}{\epsilon^2}$$

By the Markov's inequality in (a), Chebyshev's inequality holds

c)

Let $Y = \overline{Z_n}, \ \ \sigma' = \frac{\sigma}{\sqrt{n}}$

$$\Pr\left\{\left|\overline{Z}_n - \mu\right| > \epsilon\right\} \leq \frac{\sigma^2}{n\epsilon^2} \Leftrightarrow \Pr\{|Y - \mu| > \epsilon\} \leq \frac{\sigma'^2}{\epsilon^2}$$

By Chebyshev's inequality, it holds.

3.3

Let R_n be the cake remained after n^{th} cut. Let C_i denote the bigger fraction of cake in i^{th} cut.

Then we have:

$$\lim_{n\to\infty} \frac{1}{n} \log R_n = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log C_i = E(\log C_1) = \frac{3}{4} \cdot \log \frac{2}{3} + \frac{1}{4} \cdot \log \frac{3}{5}$$

3.13

a)

$H(X) = -0.6 \log 0.6 - 0.4 \log 0.4 = 0.970951$

b)

$$| -\frac{1}{25} \sum \log p(x^{25}) - H(X)| < \epsilon$$

Which means $-\frac{1}{25} \sum \log p(x^{25}) \in (H(X) - \epsilon, H(X) + \epsilon) = (0.870951, 1.070951)$

Check the last column of the table, we could see $k = 11, 12, 13, \cdots, 19$ satisfy the requirement.

```python
from math import comb as C
from math import pow
n = 25
p = 0.6
F = [0]
c = [0]
for k in range(0, n + 1):
    tmp = C(n, k) * pow(p,k) * pow((1-p),n - k)
    F.append(F[k] + tmp)

for k in range(0, n + 1):
    tmp = C(n, k)
    c.append(c[k] + tmp)

print(F[20] - F[11]) # 0.9362462771170672
print(c[20]- c[11]) # 26366510
```

$$\left|A_\epsilon^{(n)}\right| = \sum_{k=11}^{19}\binom{n}{k} = \sum_{k=0}^{19}\binom{n}{k} - \sum_{k=0}^{10}\binom{n}{k} = 33486026 - 7119516 = 26366510$$

c)

Since $p = 0.6 > 0.5$, obviously more one in a sequence(i.e. bigger k) , larger the probability of this sequence. Thus, if we want to make the 0.9 set as small as possible, we should start filling it in this order: $k = 25, 24, \cdots$. Once the probability of the set is over 0.9, then we terminate the procedure.

```python
from math import comb as C
from math import pow
n = 25
p = 0.6
F = [0]
c = [0]
for k in range(0, n + 1):
    tmp = C(n, k) * pow(p,k) * pow((1-p),n - k)
    F.append(F[k] + tmp)

print(F[26] - F[13]) # [13,26] 0.8462322310242368
print(F[26] - F[12]) # [12,26] 0.9221989361329268
```

From the output of the code we know that the smallest set includes k from 13 to 25 and some sequences with k = 12

For $k = 12$, we still need some sequence to gap the distance from $d = 0.9 - 0.8462322310242368 \approx 0.053768$

Then we need at least $d/[p^{12} \cdot (1-p)^{13}] = 3680687.875 \approx 3680688$

Then the smallest set with probability 0.9 has 16777216 + 3680688 = 20457904 sequences.

Note that this number might no be the exactest, since we have applied some approximation. However, the size of the smallest set with probability 0.9 objectively exists.

d)

In (b), k is from 11 to 19, while in (c), k is from 12 to 25, and some sequences with k = 12 are not included

So the intersection set includes sequences with k from 13 to 19, and some sequences with k = 12.

It's size is 33486026 - 16777216 + 3680688 = 20389498

Let s denote the probability of some sequences with k = 12

In (c), we have known that s = 0.053768

$F(19) - F(12)$ = 0.8168700262319829

The probability is $F(19) - F(12) + s = 0.870638$