1. Ans. For each $i$, we have
\[ E[X_i] = \Pr(X_i = 1) = 1/n \quad \text{and} \quad E[X_i^2] = \Pr(X_i^2 = 1) = 1/n. \]

For any $i, j$ with $i \neq j$, we have
\[ E[X_iX_j] = \Pr(X_iX_j = 1) = \Pr(X_i = 1 \cap X_j = 1) = \frac{1}{n(n-1)}. \]

Let $X$ be the number of fixed points. So, $X = \sum_{i=1}^n X_i$, and $E[X] = \sum_{i=1}^n E[X_i] = 1$.

Then, we have
\[ \text{Var}[X] = E[X^2] - (E[X])^2 = E \left( \left( \sum_{i=1}^n X_i \right)^2 \right) - 1 = \sum_{i=1}^n E[X_i^2] + \sum_{i \neq j} E[X_iX_j] - 1 = n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} - 1 = 1. \]

2. Ans. Firstly,
\[ \Pr \left( \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| > \varepsilon \right) = \Pr \left( |(X_1 + X_2 + \cdots + X_n) - n\mu| > n\varepsilon \right). \]

By Chebyshev’s inequality and the independence of $X_i$’s, we have:
\[ \Pr \left( |(X_1 + X_2 + \cdots + X_n) - n\mu| > n\varepsilon \right) \leq \frac{\text{Var}[X_1 + X_2 + \cdots + X_n]}{(n\varepsilon)^2} = \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}. \]

Combining, we have
\[ 0 \leq \lim_{n \to \infty} \Pr \left( \left| \frac{X_1 + X_2 + \cdots + X_n}{n} - \mu \right| > \varepsilon \right) \leq \lim_{n \to \infty} \frac{\sigma^2}{n\varepsilon^2} = 0, \]

which completes the proof of the weak law of large numbers.

3. (a) Ans. Let $X = n\tilde{p}$ denote the number of heads that came up. So, $X = \text{Bin}(n, p)$ and $E[X] = np$. Then we have:
\[ \Pr(|p - \tilde{p}| > \varepsilon p) = \Pr(|np - n\tilde{p}| > n\varepsilon p) = \Pr(np - n\tilde{p} > n\varepsilon p) + \Pr(n\tilde{p} - np > n\varepsilon p) = \Pr(X < (1 - \varepsilon)E[X]) + \Pr(X > (1 + \varepsilon)E[X]) \leq \exp \left( \frac{-n\varepsilon^2}{2} \right) + \exp \left( \frac{-n\varepsilon^2}{3} \right). \]
(b) **Ans.** When

\[ n > \frac{3 \ln(2/\delta)}{a\varepsilon^2}, \]

we have:

\[ \frac{na\varepsilon^2}{3} > \ln(2/\delta) \quad \text{so that} \quad \delta > 2 \exp\left(\frac{-na\varepsilon^2}{3}\right). \]

Combining this with the result of part (a), we have:

\[ \Pr(|p - \tilde{p}| > \varepsilon p) \leq \exp\left(\frac{-na\varepsilon^2}{2}\right) + \exp\left(\frac{-na\varepsilon^2}{3}\right) < 2 \exp\left(\frac{-na\varepsilon^2}{3}\right) < \delta. \]

4. (a) **Ans.** We shall make use of the following claim:

**Claim 1.** For any \( r \in [0, 1], e^{tr} - 1 \leq r(e^t - 1). \)

**Proof.** Let \( f(r) = r(e^t - 1) - e^{tr} + 1. \) Then we have \( f'(r) = (e^t - 1) - te^{tr}, \) and \( f''(x) = -t^2e^{tr} \leq 0. \) This implies that \( f \) is a concave function.

In other words, for \( r \in [0, 1], f \) achieves minimum value either at the boundaries \( f(0) \) or \( f(1). \) Thus, \( f(r) \geq \min\{f(0), f(1)\} = 0 \) for all \( r \in [0, 1], \) and the claim follows. \( \Box \)

Back to the answer. Since \( W = \sum_{i=1}^{n} a_i X_i, \) we have

\[ \nu = E[W] = \sum_{i=1}^{n} a_i E[X_i] = \sum_{i=1}^{n} a_i p_i. \]

For any \( i, \)

\[ E[e^{ta_i X_i}] = p_i e^{ta_i} + (1 - p_i) = 1 + p_i(e^{ta_i} - 1) \leq 1 + p_i(e^t - 1), \]

where the last inequality is from Claim 1.

Hence,

\[ E[e^{ta_i X_i}] \leq e^{p_i a_i (e^t - 1)}, \]

and by the independence of \( X_i \)'s and property of MGF,

\[ E[e^{tW}] = \prod_{i=1}^{n} E[e^{ta_i X_i}] \leq \prod_{i=1}^{n} e^{a_i p_i (e^t - 1)} = e^{\nu(e^t - 1)}. \]

For any \( t > 0, \) we have

\[ \Pr(W \geq (1 + \delta)\nu) = \Pr(e^{tW} \geq e^{t(1+\delta)\nu}) \leq \frac{E[e^{tW}]}{e^{t(1+\delta)\nu}} \leq \frac{e^{\nu(e^t - 1)}}{e^{t(1+\delta)\nu}}. \]

Then, for any \( \delta > 0, \) we can set \( t = \ln(1 + \delta) > 0 \) and obtain:

\[ \Pr(W \geq (1 + \delta)\nu) < \left(\frac{e^{\delta}}{(1 + \delta)^{1+\delta}}\right)^\nu. \]
(b) **Ans.** For any $t < 0$, we have

$$
\Pr(W \leq (1 - \delta)\nu) = \Pr(e^{tW} \geq e^{(1-\delta)\nu} \leq \frac{E[e^{tW}]}{e^{t(1-\delta)\nu}} \leq \frac{e^{\nu(e^t-1)}}{e^{(1-\delta)\nu}}.
$$

Then, for any $0 < \delta < 1$, we can set $t = \ln(1 - \delta) < 0$ and obtain:

$$
\Pr(W \leq (1 - \delta)\nu) < \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}}\right)^\nu.
$$

5. (a) **Ans.** The physical meaning of $X$ is the number of times we flipped a fair coin to get the $n$th head. Thus, the event $X > (1 + \delta)2n$ is saying that the $n$th head does not appear among the first $(1 + \delta)2n$ flips.

Let $Y = \text{Bin}((1 + \delta)2n, 0.5)$ be a binomial random variable that counts the number of heads appearing in a sequence of $(1 + \delta)2n$ fair coin flips. Then we have:

$$
\Pr(X > (1 + \delta)2n) = \Pr(Y < n) = \Pr(Y < E[Y]/(1 + \delta))
$$

$$
= \Pr \left( Y < \left(1 - \frac{\delta}{1 + \delta} \right) E[Y] \right) \leq \exp \left( - \frac{E[Y](\delta/(1 + \delta))^2}{2} \right) = \exp \left( \frac{n\delta^2}{2(1 + \delta)} \right).
$$

(b) i. **Ans.** Assume $e^t < 2$. Then,

$$
E[e^{tX}] = \frac{1}{2}e^t + \frac{1}{4}e^{2t} + \frac{1}{8}e^{3t} + \frac{1}{16}e^{4t} + \cdots = \frac{e^t/2}{1 - (e^t/2)} = \frac{e^t}{2 - e^t}.
$$

ii. **Ans.** Consider $t \in (0, \ln 2)$. Let

$$
f(t) = (2 - e^t)e^{t(1+2\delta)}
$$

so that $f(t) > 0$ for all $t \in (0, \ln 2)$. Then,

$$
f'(t) = 2(1 + 2\delta)e^{t(1+2\delta)} - (2 + 2\delta)e^{t(2+2\delta)},
$$

which is 0 only if $t = t^* = \ln(1 + \delta/(1 + \delta))$. Then,

$$
f''(t) = 2(1 + 2\delta)^2e^{t(1+2\delta)} - (2 + 2\delta)^2e^{t(2+2\delta)},
$$

so that

$$
f''(t^*) = 2(1 + 2\delta)^2e^{t(1+2\delta)} - (2 + 2\delta)^2 \left( \frac{1 + 2\delta}{1 + \delta} \right) e^{t(1+2\delta)}
$$

$$
= 2(1 + 2\delta)^2e^{t(1+2\delta)} - 4(1 + \delta)(1 + 2\delta)e^{t(1+2\delta)}
$$

$$
= 2(1 + 2\delta)e^{t(1+2\delta)}((1 + 2\delta) - 2(1 + \delta)) < 0.
$$

This shows that $f$ attains maximum when $t = t^*$, which implies $1/f$ attains minimum at $t = t^*$ as desired.

---

1. The original problem has a bug by not including this assumption.
2. The original problem has a bug by not restricting $t$ to this range.
iii. **Ans.** For any \( t \in (0, \ln 2) \) and \( t^* = \ln(1 + \delta/(1 + \delta)) \),

\[
\Pr(X > (1 + \delta)2n) = \Pr(tX > t(1 + \delta)2n) = \Pr(e^{tX} > e^{t(1+\delta)2n}) \\
\leq E[e^{tX}] / e^{t(1+\delta)2n} = \prod_i E[e^{tX_i}] / e^{t(1+\delta)2n} \\
= \left( \frac{e^t}{(2 - e^t)e^{t(1+\delta)^2}} \right)^n \\
= ((2 - e^t)e^{t(1+\delta)} e^{t^*(1+2\delta)})^{-n} \quad \text{from b(ii)} \\
= \left( \left( 1 - \frac{\delta}{1 + \delta} \right) \left( 1 + \frac{\delta}{1 + \delta} \right)^{1+2\delta} \right)^{-n}.
\]

(c) **Ans.** By substituting the result of part (b) using the three formulas, we have:

\[
\Pr(X > (1 + \delta)2n) \leq \left( \left( 1 - \frac{\delta}{1 + \delta} \right) \left( 1 + \frac{\delta}{1 + \delta} \right)^{1+2\delta} \right)^{-n} \\
< \left( e^{-\varepsilon} (e^{1-\varepsilon})^{(1+2\delta)/(1+\delta)} \right)^{-n} \\
< \left( e^{-\varepsilon} (e^{1-\varepsilon}) \delta^2 \right)^{-n} \\
= \exp\left( -n(1 - \varepsilon)(\delta^2 - \varepsilon) \right).
\]

In the limit case where \( \varepsilon \) tends to 0, the bound obtained in part (b) will be slightly tighter than the one in part (a).