(a) First, we show that the Markov property implies $\mathbb{E}[T \mid X_1 = x_1 \land X_0 = x_0] = \mathbb{E}[T \mid X_1 = x_1]$, for $x_0 \neq 0$:

\[
\mathbb{E}[T \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{t \geq 0} t \cdot \Pr[T = t \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{t \geq 0} t \cdot \sum_{x_{t-1} \ldots x_2} \Pr[X_t = 0 \land X_{t-1} = x_{t-1} \land \ldots \land X_2 = x_2 \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{t \geq 0} t \cdot \sum_{x_{t-1} \ldots x_2} \Pr[X_t = 0 \land X_{t-1} = x_{t-1} \land \ldots \land X_2 = x_2 \mid X_1 = x_1 \land X_0 = x_0] \cdots \Pr[X_2 = x_2 \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{t \geq 0} t \cdot \sum_{x_{t-1} \ldots x_2} \Pr[X_t = 0 \land X_{t-1} = x_{t-1} \land \ldots \land X_2 = x_2 \mid X_1 = x_1] = \\
\sum_{t \geq 0} \Pr[T = t \mid X_1 = x_1] = \\
\mathbb{E}[T \mid X_1 = x_1],
\]

where the sum $\sum_{x_{t-1} \ldots x_2}$ is over the elements of $(S \setminus 0)^{t-2}$. Hence, we can conclude

\[
\sum_{x_1 \in S} \Pr[X_1 = x_1 \mid X_0 = x_0] \cdot \mathbb{E}[T \mid X_1 = x_1] = \\
\sum_{x_1 \in S} \Pr[X_1 = x_1 \mid X_0 = x_0] \cdot \mathbb{E}[T \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{x_1 \in S} \Pr[X_1 = x_1 \mid X_0 = x_0] \cdot \sum_{t \geq 0} t \cdot \Pr[T = t \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{x_1 \in S} \sum_{t \geq 0} t \cdot \Pr[T = t \mid X_1 = x_1 \land X_0 = x_0] = \\
\sum_{x_1 \in S} \sum_{t \geq 0} t \cdot \Pr[T = t \land X_1 = x_1 \mid X_0 = x_0] = \\
\sum_{t \geq 0} \sum_{x_1 \in S} \Pr[T = t \land X_1 = x_1 \mid X_0 = x_0] = \\
\sum_{t \geq 0} t \cdot \Pr[T = t \mid X_1 = x_1 \land X_0 = x_0] = \\
\mathbb{E}[T \mid X_0 = x_0].
\]
(b) We simply calculate
\[
\mathbb{E}[\mathbb{E}[X \mid Y]] = \sum_y \mathbb{E}[X \mid Y = y] \cdot \Pr[Y = y] \\
= \sum_y \sum_x x \cdot \Pr[X = x \mid Y = y] \cdot \Pr[Y = y] \\
= \sum_y \sum_x x \cdot \Pr[X = x \land Y = y] \\
= \sum_x \sum_y x \cdot \Pr[X = x \land Y = y] \\
= \sum_x x \cdot \Pr[X = x] \\
= \mathbb{E}[X].
\]

Solution to Exercise 2

Let $X_t$ be the random variable counting the missing coupons at time $t$ and $T$ be the first point in time $t$ such that $X_t = 0$. First, note that $\mathbb{E}[T \mid X_0 = n] \geq (1 - o(1))n \ln n$ because the process is at most as fast as the Coupon Collector process. Hence, it is left to show an upper bound. Let $\varepsilon > 0$ and $T_{\varepsilon n}$ be the first point in time $t$ such that $X_t = \varepsilon n$. We can write $\mathbb{E}[T \mid X_0 = n] = \mathbb{E}[T_{\varepsilon n} \mid X_0 = n] + \mathbb{E}[T \mid X_0 = \varepsilon n]$.

To compute $\mathbb{E}[T_{\varepsilon n} \mid X_0 = n]$ define the random variables $Y_t$ as
\[
Y_t := \begin{cases} 
X_t, & \text{if } X_t > \varepsilon n, \\
0, & \text{otherwise},
\end{cases}
\]
and denote by $T_Y$ the first point in time $t$ such that $Y_t = 0$. We apply the Multiplicative Drift theorem with $\delta = 1/(2n)$, $s_{\min} = \varepsilon n$, and $s_0 = n$:
\[
\mathbb{E}[T_{\varepsilon n} \mid X_0 = n] = \mathbb{E}[T_Y \mid Y_0 = n] \leq \frac{1 + \ln \frac{1}{2n}}{\varepsilon n} = \mathcal{O}(n).
\]

Now we determine $\mathbb{E}[T \mid X_0 = \varepsilon n]$. Let $S$ be the event that Alice loses all coupons. We have
\[
\mathbb{E}[T \mid X_0 = \varepsilon n] = (1 - \frac{1}{2n})^{\varepsilon n} \cdot \mathbb{E}[T \mid X_0 = \varepsilon n \land \overline{S}] + (1 - (1 - \frac{1}{2n})^{\varepsilon n}) \cdot \mathbb{E}[T \mid X_0 = \varepsilon n \land \overline{S}]
\leq \mathbb{E}[T \mid X_0 = \varepsilon n \land \overline{S}] + \frac{\varepsilon}{2} \cdot \mathbb{E}[T \mid X_0 = \varepsilon n \land \overline{S}],
\]
where we used Bernoulli’s inequality in the last line. Note that $\mathbb{E}[T \mid X_0 = \varepsilon n \land \overline{S}] = (1 + o(1))n \ln n$, since conditioned on the event $\overline{S}$ the process behaves like the Coupon Collector process. Furthermore, we can bound $\mathbb{E}[T \mid X_0 = \varepsilon n \land \overline{S}] \leq 2 \mathbb{E}[T \mid X_0 = n]$. Solving yields $\mathbb{E}[T \mid X_0 = \varepsilon n] \leq (1 + o(1))n \ln n$.

Solution to Exercise 3

We consider the case $a \leq 1$. Define the random variables $Y_t$ as
\[
Y_t := \begin{cases} 
\ln X_t, & \text{if } X_t > 0, \\
0, & \text{otherwise}.
\end{cases}
\]

Note that $Y_t$ is 0 if $X_t$ is 0 or 1. Hence, to show the lower bound for the Random Decline process, it suffices to show $\mathbb{E}[T \mid Y_0 = \ln n] = \Omega(\ln n)$. The drift of $(Y_t)_{t \geq 0}$ is
\[
\mathbb{E}[Y_t - Y_{t+1} \mid Y_t = \ln x] = \ln x - \sum_{i=2}^{[ax]} \frac{\ln i}{1 + [ax]} \\
\leq \ln x - \frac{1}{1 + [ax]} \cdot ([ax] \ln [ax] - [ax]) \\
= \mathcal{O}(1).
\]
Thus, we can apply the Additive Drift theorem to obtain that $\mathbb{E}[T \mid Y_0 = \ln n] = \Omega(\ln n)$. For $a > 1$ note that we can couple the processes: Let $(X_t)_{t \geq 0}$ be the process with $a > 1$ and $(X'_t)_{t \geq 0}$ be the process with $a = 1$. The coupling is as follows: if $X_{t+1} \leq X'_t$ we set $X'_{t+1} = X_{t+1}$ and otherwise $(X'_t)_{t \geq 0}$ does a regular step. Hence, $X_t \geq X'_t$ holds for all $t \geq 0$ and $(X_t)_{t \geq 0}$ can not be faster than $(X'_t)_{t \geq 0}$.

Solution to Exercise 4

(i) If there is a weight $w_i < 0$, then we can just apply the automorphism $\varphi$ to the hypercube that flips 0 and 1 in the $i$-th bit, and replace the weight $w_i < 0$ by $w'_i := -w_i > 0$.

Note that then we have $f_w(x) - f_w(y) = f_w'(\varphi(x)) - f_w'(\varphi(y))$ for all $x, y \in H_n$. Therefore, by setting $y = 0$, we get $f_w(x) = f_w'(\varphi(x)) - f_w'(\varphi(0))$ for all $x \in H_n$, so both problems are equivalent. Not only do they take the miminum in the same places, have the same optimum, but the optimization algorithm makes the same choices on both instances, modulo $\varphi$.

Similarly, changing the order of the bits does not change the problem.

(ii) The upper bound is trivial. For the lower bound, observe that for fixed $i$, the probability that we flip exactly the $i$-th bit is $1/n \cdot (1 - 1/n)^{n-1}$. We use the well-known formula $(1 - x/n)^n \leq e^{-x} \leq (1 - x/n)^{n-1}$ for $x = 1$. (It holds for all $x \geq 0$.) Now we sum over all $i$ and obtain

$$p_0 = \left(1 - \frac{1}{n}\right)^{n-1} \geq \frac{1}{e}.$$

(iii) For $a = 0$ the statement is trivial. For $a > 0$, we compute directly

$$\mathbb{E}[X_t - X_{t+1} \mid X_t = a] \geq \mathbb{E}[X_t - X_{t+1} \mid X_t = a \text{ and we flip exactly one bit.}] \cdot \frac{1}{e} = \frac{1}{en} \cdot \sum_{\text{wrong i-th bit}} Pr[\text{flip exactly i-th bit}] \cdot w_i \geq \frac{1}{en} \cdot \sum_{\text{wrong i-th bit}} w_i = \frac{a}{en}.$$

(iv) We apply the Variable Drift theorem, where $s_{\min} = w_{\min}$ and $h(x) = x/(en)$. We obtain

$$\mathbb{E}[T \mid X_0 = a] \leq \frac{w_{\min}}{h(w_{\min})} + \int_{w_{\min}}^{a} \frac{1}{h(x)} dx = en \left(1 + \int_{w_{\min}}^{a} \frac{1}{x} dx\right) = en \left(1 + \log \left(\frac{a}{w_{\min}}\right)\right).$$

Observing that the starting value is $a = \sum_{i=1}^{n} w_i$ yields the result.