Randomized Algorithms
solutions for Tutorial 2

1. Let Y_i be the indicator variable for i being a fixed point. Then, $X = \sum_{i=1}^{n} Y_i$ and $\mathrm{E}[Y_i] = \frac{1}{n}$ for any $i \in [n]$. We have

$$\begin{aligned}
\mathrm{E}[X^2] &= \sum_{i=1}^{n} \mathrm{E}[Y_i^2] + 2 \sum_{1 \le i < j \le n} \mathrm{E}[Y_i Y_j] \\
&= \sum_{i=1}^{n} \mathrm{E}[Y_i] + 2 \sum_{1 \le i < j \le n} \mathrm{E}[Y_i Y_j] \\
&= 1 + 2 \sum_{1 \le i < j \le n} \mathrm{E}[Y_i Y_j].
\end{aligned}$$

Notice that $Y_i Y_j = 1$ if and only if both i and j are fixed points. The probability that happens is $\frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$. Thus,

$$\begin{aligned}
\mathrm{E}[X^2] &= 1 + 2 \sum_{1 \le i < j \le n} \frac{1}{n(n-1)} \\
&= 1 + 2 \cdot \frac{n(n-1)}{2} \cdot \frac{1}{n(n-1)} = 2.
\end{aligned}$$

It implies that $\mathrm{Var}[X] = \mathrm{E}[X^2] - \mathrm{E}[X]^2 = 2 - 1 = 1$.

2. **(a), (b)** Imagine taking our biased coin and flipping it *twice*. After doing this we have the possibility of four outcomes: two "heads", "heads"-then-"tails", "tails"-then-"heads", two "tails".

   Now notice that because flips are independent and identically distributed (with the unknown probability p), that the probabilities of these four outcomes are $p^2, p(1-p), (1-p)p$ and $(1-p)^2$ respectively. In particular, "heads"-then-"tails" and "tails"-then-"heads" have identical probability of being generated. We will use this fact to identify "heads"-then-"tails" with the overall outcome "heads" and "tails"-then-"heads" with the overall outcome "tails", these each having identical probability. If the pair of flips generates two "heads" or two "tails", we re-run the experiment with two new flips of the coin.

Algorithm BIASNOMORE(p)

(a) $\mathsf{flip1} = 0$, $\mathsf{flip2} = 0$;

(b) **while** $\mathsf{flip1} == \mathsf{flip2}$ **do**

(c) $\quad\mathsf{flip1} \leftarrow \mathsf{B}(1, \mathsf{p})$;

(d) $\quad\mathsf{flip2} \leftarrow \mathsf{B}(1, \mathsf{p})$;

(e) **od**

(f) **if** $\mathsf{flip1} == 1$

(g) $\quad$ **return** "heads"

(h) **else**

(i) $\quad$ **return** "tails"

We've already argued that *on any particular two flips*, "heads" and "flips" are equally likely to be returned ($\mathsf{p}(1-\mathsf{p})$ each). These is true regardless of whether we take $2, 4, 6, 8, \ldots, 2i$ flips to return a value - the final pair of flips determines what is returned, and "heads" and "tails" are equally likely at that point. Hence the probability, over all possible sequences of flips that end with a returned value, is equal for "heads" and "tails".

(c) Consider a new geometric random variable Y, where $Y = 1$ if and only if the algorithm succeeds. The algorithm succeeds if and only if we get two distinct coin flips, which has probability $2p(1-p)$. Thus, $E[Y] = \frac{1}{2p(1-p)}$. The expected number of coin flips is $2E[Y] = p^{-1}(1-p)^{-1}$.

3. We start with a bag containing one black ball and one white ball, and repeatedly do the following: choose one ball from the bag uniformly at random, and then put the ball back in the bag with another ball of the same colour. We repeat until there are n balls in the bag.

Claim: by the time that we have n balls (after $n-2$ steps), the number of white balls is equally likely to be any number between 1 and $n-1$.

We will prove this by induction on n.

We should note that no matter what choices are made, we will always have at least one white ball and at least one black ball in the bag.

base case: $n = 2$. In this case we definitely (with probability 1) have exactly 1 white ball in the bag. The range $1, \ldots, n-1$ is just 1, so the hypothesis is trivially correct.

Induction step: Suppose we have shown the claim for $n = k$ *(Induction Hypothesis (IH))*. We now show it must also hold for $n = k+1$.

For all k, j, where $2 \le k \le n$ and $1 \le j \le k-1$, let us use $A_{k,j}$ to denote the event that, at the stage where there are k balls in the bag, the number of white balls in the bag is j.

If the claim holds for $n = k$, then we know that $\Pr[A_{k,j}] = \frac{1}{k-1}$ for every $j = 1 \ldots, k-1$. This is our IH.

Now we want to inductively compute the probability of the event $A_{k+1,j}$ for all $j = 1, \ldots, k$, and show that $\Pr[A_{k+1,j}] = \frac{1}{k}$. Doing so would complete the proof by induction.

Suppose we already have k balls in the bag, and are about to choose a random ball and add a $(k+1)$'st ball to the bag.

Let $B_{k,\text{white}}$ be the event that the random ball chosen from the bag with k balls in it is white. let $B_{k,\text{black}}$ be the event that the random ball chosen from the bag with k balls in it is white.

Crucial observation: note that for any $k \geq 2$, and any $2 \leq j \leq k-1$, we have the following equality between events:

$$A_{k+1,j} = (A_{k,j} \cap B_{k,\text{black}}) \cup (A_{k,j-1} \cap B_{k,\text{white}})$$

This is because if we started with j white balls and we draw a black ball then there remain j white balls after this extra step; on the other hand, if there were $j-1$ white balls and we draw a white ball, then there are j white balls after this extra step. Moreover, those are the only two possibilities: the only way we could have ended up with j white balls in the bag at step $k+1$ is by either one or the other of these two scenarios.

What about the cases when $j = 1$ and $j = k$. In those cases, following the same reasoning, we have:

$$A_{k+1,1} = (A_{k,1} \cap B_{k,\text{black}})$$

and

$$A_{k+1,k} = (A_{k,k-1} \cap B_{k,\text{white}})$$

Note, firstly, that $(A_{k,j} \cap B_{k,\text{black}}) \cap (A_{k,j-1} \cap B_{k,\text{white}}) = \emptyset$, meaning the two events $(A_{k,j} \cap B_{k,\text{black}})$ and $(A_{k,j-1} \cap B_{k,\text{white}})$ are mutually exclusive.

Hence, for $j \in \{1, \ldots, k-1\}$, using the induction hypothesis, and the definition of conditional probability, we see that:

$$\begin{aligned} \Pr[A_{k,j} \cap B_{k,\text{black}}] &= \Pr[A_{k,j}] \cdot \Pr[B_{k,\text{black}} \mid A_{k,j}] \\ &= \frac{1}{k-1} \cdot \left(\frac{k-j}{k}\right) \end{aligned}$$

Likewise, for $j \in \{2, \ldots, k\}$, using the induction hypothesis, and the definition of conditional probability, we see that:

$$\begin{aligned} \Pr[A_{k,j-1} \cap B_{k,\text{white}}] &= \Pr[A_{k,j-1}] \cdot \Pr[B_{k,\text{white}} \mid A_{k,j-1}] \\ &= \frac{1}{k-1} \cdot \left(\frac{j-1}{k}\right) \end{aligned}$$

Hence, for $j \in \{2, \ldots, k-1\}$, we have

$$\begin{aligned}
\Pr[A_{k+1,j}] &= \Pr[(A_{k,j} \cap B_{k,\text{black}}) \cup (A_{k,j-1} \cap B_{k,\text{white}})] \\
&= \Pr[(A_{k,j} \cap B_{k,\text{black}})] + \Pr[(A_{k,j-1} \cap B_{k,\text{white}})] \\
&= \frac{1}{k-1} \cdot (\frac{k-j}{k}) + \frac{1}{k-1} \cdot (\frac{j-1}{k}) \\
&= \frac{1}{k-1} \cdot (\frac{k-j}{k} + \frac{j-1}{k}) \\
&= \frac{1}{k-1} \cdot \frac{k-1}{k} \\
&= \frac{1}{k}
\end{aligned}$$

For the remaining cases of $j = 1$ and $j = k$, we have

$$\begin{aligned}
\Pr[A_{k+1,1}] &= \Pr[A_{k,1} \cap B_{k,\text{black}}] \\
&= \frac{1}{k-1} \cdot (\frac{k-1}{k}) \\
&= \frac{1}{k}
\end{aligned}$$

and

$$\begin{aligned}
\Pr[A_{k+1,k}] &= \Pr[A_{k,k-1} \cap B_{k,\text{white}}] \\
&= \frac{1}{k-1} \cdot (\frac{k-1}{k}) \\
&= \frac{1}{k}
\end{aligned}$$

This completes the proof, by induction, that for all $k \geq 2$, and for all $j \in \{1, \ldots, k-1\}$, $\Pr[A_{k,j}] = \frac{1}{k-1}$.

4. Let Y be a nonnegative integer-valued random variable with (strictly) positive expectation. Prove that

$$\frac{(\mathrm{E}[Y])^2}{\mathrm{E}[Y^2]} \;\leq\; \Pr[Y \neq 0] \;\leq \mathrm{E}[Y].$$

Proof: First let's establish the inequality on the right, namely $\Pr[Y \neq 0] \leq \mathrm{E}[Y]$. For this, notice that since Y's range is non-negative and integer, we know

$$\begin{aligned}
E[Y] &= \sum_{j=0}^{\infty} j \cdot \Pr[Y = j] \\
&= \sum_{j=1}^{\infty} j \cdot \Pr[Y = j] \\
&\geq 1 \cdot \sum_{j=1}^{\infty} \Pr[Y = j], \\
&= \Pr[Y \geq 1] \;=\; \Pr[Y \neq 0],
\end{aligned}$$

where the first step (expansion of $E[Y]$) and final step (equality of $\Pr[Y \geq 1]$ and $\Pr[Y \neq 0]$) both follow from the fact that Y only takes on non-negative integer values.

Next let's establish the inequality on the left, namely , $\frac{(E[Y])^2}{E[Y^2]} \leq \Pr[Y \neq 0]$.

Here is one way to prove this inequality, using Jensen's inequality:

Consider the conditional expectations $E[Y \mid Y \neq 0]$ and $E[Y^2 \mid Y \neq 0]$. Note that the function $f(x) = x^2$ is convex. Therefore, by Jensen's inequality, we know that

$$(E[Y \mid Y \neq 0])^2 \leq E[Y^2 \mid Y \neq 0]$$

We have that

$$\begin{aligned}
E[Y \mid Y \neq 0] &= \sum_{j=0}^{\infty} j \cdot \frac{\Pr[Y = j, Y \neq 0]}{\Pr[Y \neq 0]} \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=0}^{\infty} j \cdot \Pr[Y = j, Y \neq 0] \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=1}^{\infty} j \cdot \Pr[Y = j] \\
&= \frac{1}{\Pr[Y \neq 0]} E[Y]
\end{aligned}$$

Similarly, we have

$$
\begin{aligned}
E[Y^2 \mid Y \neq 0] &= \sum_{j=0}^{\infty} j^2 \cdot \frac{\Pr[Y = j, Y \neq 0]}{\Pr[Y \neq 0]} \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=0}^{\infty} j^2 \cdot \Pr[Y = j, Y \neq 0] \\
&= \frac{1}{\Pr[Y \neq 0]} \sum_{j=1}^{\infty} j^2 \cdot \Pr[Y = j] \\
&= \frac{1}{\Pr[Y \neq 0]} E[Y^2]
\end{aligned}
$$

Therefore, we know that

$$
\left(\frac{1}{\Pr[Y \neq 0]} E[Y] \right)^2 \leq \frac{1}{\Pr[Y \neq 0]} E[Y^2]
$$

Multiplying both sides by $\frac{(\Pr[Y \neq 0])^2}{E[Y^2]}$ we get

$$
\frac{(E[Y])^2}{E[Y^2]} \leq \Pr[Y \neq 0]
$$

as claimed.

It is also possible (but a bit more involved and tedious) to prove $\frac{(E[Y])^2}{E[Y^2]} \leq \Pr[Y \neq 0]$ by just expanding out the terms $\Pr[Y \neq 0], E[Y^2]$ and $E[Y]^2$, using their definitions.