Solutions to Chapter 2 of Sheldon M. Ross' *Probability Models For Computer Science*

1. Let $q = 1 - p$.

   (a) The vertex set $\{1, ..., k\}$ is a component of the graph if it is connected and there are no edges between nodes from the respective sets $\{1, ..., k\}$ and $\{k + 1, ..., n\}$.

   $$P_k \cdot \left(q^{n-k}\right)^k$$

   (b) The probability that vertex 1 is a member of a component of size k is the sum of the probabilities of all subsets of size k containing vertex 1 is a component.

   $$\binom{n-1}{k-1} \cdot P_k \cdot q^{k(n-k)}$$

   (c) The graph of n nodes is connected if an arbitrary vertex, say vertex 1, belongs to a component of size n; that is, the vertex does not belong to a component of size other than n.

   $$P_n = 1 - \sum_{k=1}^{n-1} \binom{n-1}{k-1} \cdot P_k \cdot q^{k(n-k)}$$

   (d)

   $$P_1 = 1$$
   $$P_2 = p$$
   $$P_3 = 1 - q^2 - 2pq^2$$
   $$P_4 = 1 - 2q^3 - 3pq^4 + 2pq^5$$
   $$P_5 = 1 - 5q^4 - 6q^6 - 4pq^6 + 8q^7 + 6q^8 + 12pq^8 - 8pq^9$$
   $$P_6 = 1 - 6q^5 - 10q^8 - 5pq^8 + 15q^9 + 60q^{11}$$
   $$+ 40pq^{11} - 40q^{12} + 30pq^{12} - 30q^{13} - 80pq^{13} + 40pq^{14}$$

2. Upon choosing a pivot for a sequence of $n > 0$ distinct elements, $n - 1$ comparisons are made before recursively sorting two subsets. If the pivot element is the ith rank, the sizes of the two subsets are $i - 1$ and $n - i$.

   Hence, for $n > 0$, noting also that the pivot is selected uniformly at random among n

elements,

$$M_n = (n-1) + \sum_{i=1}^{n} \frac{1}{n} \left(M_{i-1} + M_{n-i} \right)$$

$$= (n-1) + \frac{1}{n} \left(\sum_{i=1}^{n} M_{i-1} + \sum_{i=1}^{n} M_{n-1} \right)$$

$$= (n-1) + \frac{1}{n} \left(2 \sum_{i=1}^{n} M_{i-1} \right)$$

$$= (n-1) + \frac{2}{n} \sum_{j=0}^{n-1} M_j.$$

Using base case $M_0 = 0$,

$$M_1 = 0$$
$$M_2 = 1$$
$$M_3 = \frac{8}{3}$$
$$M_4 = \frac{29}{6}$$
$$M_5 = \frac{37}{5}$$
$$M_6 = \frac{284}{25}$$
$$M_7 = \frac{2413}{175}$$
$$M_8 = \frac{24167}{1400}.$$

3. Let C denote the number of comparisons, let R denote the number of remaining values in the nonempty pile when the other is first emptied and suppose n is positive.

Every comparison corresponds to a distinct removal of one value whilst both piles are nonempty. Since there are n values in total, the number of comparisons is $n - R$; n less the number of remaining values once one pile is emptied.

Since the remaining values is some maximal suffix of the sorted sequence, R is the number of sequential values, in descending order from the largest value, that were allocated to the same pile. R is certainly at least 1 since the largest value always remains in the nonempty pile and, noting that all values were equally likely to be placed in either pile, the probability of at least $j + 1$ remaining values is half of the probability of at least j remaining values, for $1 \leq j < n$. Hence, for $1 \leq i \leq n$,

$$P(R \geq i) = \left(\frac{1}{2} \right)^{i-1}.$$

Using $E[X] = \sum_{i=1}^{\infty} P(X \geq i)$ for a non-negative random variable X,

$$E[C] = E[n - R]$$
$$= n - \sum_{i=1}^{\infty} P(R \geq i)$$
$$= n - \sum_{i=1}^{n} \left(\frac{1}{2}\right)^{i-1}$$
$$= n - 2 + \left(\frac{1}{2}\right)^{n-1}.$$

4. Let $I_{i,j}$ indicate the event that element e_i precedes e_j at time n. This event can occur if

- neither e_i nor e_j were selected and e_i initially preceded e_j, which happens with probability $\frac{1}{2}(1 - P_i - P_j)^n$ since each element is equally likely to have preceded the other in the initial ordering; or

- e_j is not selected after e_i is selected. Conditioning on the last selection of e_i, this occurs with probability $\sum_{k=1}^{n} P_i (1 - P_i - P_j)^{n-k}$.

Let R_i denote the position of e_i at time n.

$$R_i = 1 + \sum_{j \neq i} I_{j,i}$$

Let R denote the position of the element requested at time n. Conditioning on the element requested at time n, its expected value is

$$E[R] = \sum_{i} P_i E[R_i]$$
$$= \sum_{i} P_i + \sum_{i} P_i \sum_{j \neq i} E[I_{j,i}]$$
$$= 1 + \sum_{i} P_i \sum_{j \neq i} \left(\frac{1}{2}(1 - P_i - P_j)^n + P_j \sum_{k=1}^{n} (1 - P_i - P_j)^{n-k}\right)$$
$$= 1 + \sum_{i} \frac{P_i}{2} \sum_{j \neq i} (1 - P_i - P_j)^n + \sum_{i} \sum_{j \neq i} P_i P_j \sum_{k=1}^{n} (1 - P_i - P_j)^{n-k}.$$

5. Let $p_{i,j}$ denote the probability that e_i precedes e_j, and R the position of the requested element. Given the requested element, its position is 1 more the number of elements that precede it. Hence, conditioning on the requested element, the expected position of the requested element is

$$E[R] = 1 + \sum_{i} P_i \sum_{j \neq i} p_{j,i}$$
$$= 1 + \sum_{i} \sum_{j < i} (p_{j,i} P_i + p_{i,j} P_j)$$
$$= 1 + \sum_{i} \sum_{j < i} ((1 - p_{i,j}) P_i + p_{i,j} P_j)$$
$$= 1 + \sum_{i} \sum_{j < i} (P_i + p_{i,j}(P_j - P_i)).$$

3

The expected position can be minimised by $p_{i,j} = 1$ where $P_i \geq P_j$; that is, ordering the elements in decreasing order of their probabilities of being selected.

6. Let F denote the number of fixed points and, for $1 \leq i \leq n$, let I_i indicate that position i is a fixed point of the random permutation. Then,

$$F = I_1 + ... + I_n.$$

The expected value of the Bernoulli random variable I_i is $\frac{1}{n}$, so the expected number of fixed points is

$$E[F] = \sum_{i=1}^{n} E[I_i] = n \cdot \frac{1}{n} = 1.$$

Let $p_{i,j}$ denote the joint probability mass function of I_i and I_j, for $i \neq j$. It is clear that the function is symmetric; $p_{i,j}(a,b) = p_{i,j}(b,a)$.

The covariance of I_i and I_j, for $i \neq j$, is

$$\begin{aligned}
\mathrm{Cov}(I_i, I_j) &= E[(I_i - E[I_i])(I_j - E[I_j])] \\
&= \sum_{x=0}^{1} \sum_{y=0}^{1} p_{i,j}(x,y) \cdot \left(x - \frac{1}{n}\right)\left(y - \frac{1}{n}\right) \\
&= p_{i,j}(0,0) \cdot \frac{1}{n^2} + 2p_{i,j}(1,0) \cdot \frac{1-n}{n^2} + p_{i,j}(1,1) \cdot \frac{(n-1)^2}{n^2} \\
&= \left(\frac{1}{n} + \frac{n-2}{n}\frac{n-2}{n-1}\right) \cdot \frac{1}{n^2} + 2\left(\frac{1}{n}\frac{n-2}{n-1}\right) \cdot \frac{1-n}{n^2} + \left(\frac{1}{n}\frac{1}{n-1}\right) \cdot \frac{(n-1)^2}{n^2} \\
&= \frac{n^2 - 3n + 3}{n^3(n-1)} + \frac{-2n^2 + 6n - 4}{n^3(n-1)} + \frac{n^2 - 2n + 1}{n^3(n-1)} \\
&= \frac{1}{n^2(n-1)}.
\end{aligned}$$

Hence, given also that the variance of the Bernoulli random variable I_i is $\frac{1}{n}\left(1 - \frac{1}{n}\right) = \frac{n-1}{n^2}$, the variance of the number of fixed points is

$$\begin{aligned}
\mathrm{Var}(F) &= \sum_{i=1}^{n} \mathrm{Var}(I_i) + \sum_{i} \sum_{j \neq i} \mathrm{Cov}(I_i, I_j) \\
&= n \cdot \frac{n-1}{n^2} + n(n-1) \cdot \frac{1}{n^2(n-1)} \\
&= 1 - \frac{1}{n} + \frac{1}{n} \\
&= 1.
\end{aligned}$$

An alternative method of calculating the variance computes the expected value of F^2.

$$E[F^2] = E\left[\left(\sum_{i=1}^{n} I_i\right)^2\right]$$

$$= E\left[\sum_{i=1}^{n} I_i^2\right] + E\left[\sum_{i=1}^{n}\sum_{j\neq i}^{n} I_i I_j\right]$$

$$= \sum_{i=1}^{n} E[I_i^2] + \sum_{i=1}^{n}\sum_{j\neq i}^{n} E[I_i I_j]$$

$$= \sum_{i=1}^{n} E[I_i] + \sum_{i=1}^{n}\sum_{j\neq i}^{n} p_{i,j}(1,1)$$

$$= n\cdot\frac{1}{n} + n(n-1)\cdot\left(\frac{1}{n}\frac{1}{n-1}\right)$$

$$= 2$$

Then the variance is $\mathrm{Var}(F) = E[F^2] - E[F]^2 = 2 - 1^2 = 1$.

7. Let $j_1, ..., j_n$ be an arbitrary permutation of $1, ..., n$ and suppose X is a random permutation. Since each j_i could represent a distinct index of the vector X, the probability of

$$X(j_1) = 1, ..., X(j_n) = n \tag{1}$$

is equal to the probability of

$$(X(1), ..., X(n)) = (k_1, ..., k_n)$$

where $k_1, ..., k_n$ is the permutation of $1, ..., n$ such that $k_i = X(i)$. Since X is a random permutation, this probability is $\frac{1}{n!}$. Equation (1) may be restated

$$j_1 = X^{-1}(1), ..., j_n = X^{-1}(n)$$

so

$$P\left\{(X^{-1}(1), ..., X^{-1}(n)) = (j_1, ..., j_n)\right\} = \frac{1}{n!}.$$

8. Let P denote the number of matched pairs, and let I_i for $0 < i \leq n$ indicate that i belongs to a matched pair.

$$P = \frac{I_1 + ... + I_n}{2}$$

(a) If there is a single person, $n = 1$, there are no matched pairs, $P = 0$.

Otherwise, supposing $n > 1$, the expected value of I_i is the probability that i belongs to a matched pair,

$$E[I_i] = \frac{n-1}{n}\frac{1}{n-1} = \frac{1}{n},$$

and the expected value of P is

$$\frac{1}{2}\sum_{i=1}^{n} I_i = \frac{1}{2}\left(n\cdot\frac{1}{n}\right) = \frac{1}{2}.$$

Hence,

$$E[P] = \begin{cases} 0 & \text{if } n = 1 \\ \frac{1}{2} & \text{otherwise} \end{cases}.$$

(b) The probability that both i and j belong to a matched pair, for $i \neq j$, can be calculated by conditioning on whether they belong to the same matched pair (i.e., i chooses j's hat and j chooses i's hat).

The probability i and j belong to the same matched pair is

$$P(I_i = 1, I_j = 1, i \text{ and } j \text{ matched pair}) = \frac{1}{n}\frac{1}{n-1}.$$

The probability that both i and j belong to matched pairs, but not the same, depends on n. If $n < 4$, it is not possible for i and j to belong to distinct pairs. Otherwise, if $n \geq 4$, it is the probability that i matches with some other than j, with probability $\frac{n-2}{n}\frac{1}{n-1}$, and j belongs to a matched pair amongst the remaining, with probability $\frac{1}{n-2}$. That is,

$$P(I_i = 1, I_j = 1, i \text{ and } j \text{ not matched pair}) = \begin{cases} 0 & \text{if } 1 < n < 4 \\ \frac{n-2}{n}\frac{1}{n-1}\frac{1}{n-2} & \text{otherwise} \end{cases}.$$

Hence, for $i \neq j$, summing the above expressions,

$$P(I_i = 1, I_j = 1) = \begin{cases} \frac{1}{n(n-1)} & \text{if } 1 < n < 4 \\ \frac{2}{n(n-1)} & \text{otherwise} \end{cases}.$$

If there is a single person, $n = 1$, there is P has zero variance since $P = 0$.
Otherwise, supposing $n > 1$, the variance of P is

$$\begin{aligned}
E[P^2] - E[P]^2 &= E\left[\left(\frac{I_1 + \ldots + I_n}{2}\right)^2\right] - E[P]^2 \\
&= \frac{1}{4}\left(\sum_{i=1}^{n} E[I_i^2] + \sum_i \sum_{j \neq i} E[I_i I_j]\right) - \left(\frac{1}{2}\right)^2 \\
&= \frac{1}{4}\left(n \cdot \frac{1}{n} + \sum_i \sum_{j \neq i} P(I_i = 1, I_j = 1)\right) - \frac{1}{4} \\
&= \frac{n(n-1)}{4} \cdot P(I_i = 1, I_j = 1).
\end{aligned}$$

Hence,

$$\text{Var}(P) = \begin{cases} 0 & \text{if } n = 1 \\ \frac{1}{4} & \text{if } 1 < n < 4 \\ \frac{1}{2} & \text{otherwise} \end{cases}.$$

(c) Let C denote the length of the cycle containing an arbitrarily chosen value, say the first. The predicate there are no pairs is equivalent to the predicate there are no cycles

6

of length 2. Then, defining $S_n = \{x \in \mathbb{Z} : x \neq 2, 1 \leq x \leq n\}$, for $n > 0$

$$Q_n = \sum_{i \in S_n} P(C = i) \cdot Q_{n-i}$$

$$= \frac{1}{n} \sum_{i \in S_n} Q_{n-i}.$$

(d) Using base case $Q_0 = 1$,

$$Q_1 = 1$$
$$Q_2 = \frac{1}{2}$$
$$Q_3 = \frac{1}{2}$$
$$Q_4 = \frac{5}{8}$$
$$Q_5 = \frac{5}{8}$$
$$Q_6 = \frac{29}{48}$$
$$Q_7 = \frac{29}{48}$$
$$Q_8 = \frac{233}{384}.$$

9. Let C_1 denote the size of the cycle that contains 1.

   (a) With base case $M_0 = 0$,

$$M_n = E[N]$$
$$= \sum_{i=1}^{n} E[N|C_1 = i] \cdot P(C_1 = i)$$
$$= \sum_{i=1}^{n} (1 + M_{n-i}) \cdot \frac{1}{n}$$
$$= 1 + \frac{1}{n} \sum_{i=1}^{n} M_{n-i}.$$

   (b) There are i values in a cycle of size i, and $i \cdot \frac{1}{i} = 1$, so the term in the summation corresponding to a given value contributes that value's part in its cycle.

$$E[N] = \sum_{j=1}^{n} E\left[\frac{1}{C_j}\right]$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} P(C_j = i) \cdot \frac{1}{i}$$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{1}{i}$$

$$= \sum_{i=1}^{n} \frac{1}{i}$$

This result can be confirmed by induction to be the solution to the recurrence relation (9a).

When $n = 0$,

$$M_0 = \sum_{i=1}^{0} \frac{1}{i} = 0$$

as required.

Suppose $M_i = \sum_{j=1}^{i} \frac{1}{j}$ for $0 \leq i \leq k$ and arbitrary integer k. The below shows that $M_{k+1} = \sum_{i=1}^{k+1} \frac{1}{i}$.

$$M_{k+1} = 1 + \frac{1}{k+1} \sum_{i=1}^{k+1} M_{k+1-i}$$

$$= 1 + \frac{1}{k+1} \sum_{i=1}^{k} M_i$$

$$= 1 + \frac{1}{k+1} \sum_{i=1}^{k} \sum_{j=1}^{i} \frac{1}{j}$$

$$= 1 + \frac{1}{k+1} \sum_{1 \leq j \leq i \leq k} \frac{1}{j}$$

$$= 1 + \frac{1}{k+1} \sum_{j=1}^{k} \sum_{i=j}^{k} \frac{1}{j}$$

$$= 1 + \frac{1}{k+1} \sum_{j=1}^{k} \frac{k-j+1}{j}$$

$$= 1 + \frac{1}{k+1} \sum_{j=1}^{k} \left(\frac{k+1}{j} - 1\right)$$

$$= 1 - \frac{k}{k+1} + \sum_{j=1}^{k} \frac{1}{j}$$

$$= \sum_{j=1}^{k+1} \frac{1}{j}$$

(c) Let D denote the event that $1, 2, ..., k$ are all in the same cycle, supposing $1 \le k \le n$. If D occurs, the size of the cycle containing 1 must be at least k. Hence, conditioning on C_1 and using the relationship

$$\binom{i}{j-1} + \binom{i}{j} = \binom{i+1}{j},$$

$$
\begin{aligned}
P(D) &= \sum_{x=k}^{n} P(C_1 = x) P(D|C_1 = x) \\
&= \sum_{x=k}^{n} \frac{1}{n} \left(\frac{x-1}{n-1} \cdot ... \cdot \frac{x-(k-1)}{n-(k-1)} \right) \\
&= \frac{(n-k)!}{n!} \sum_{x=k}^{n} \frac{(x-1)!}{(x-k)!} \\
&= \frac{(n-k)!}{n!} \sum_{x=k}^{n} \binom{x-1}{x-k} (k-1)! \\
&= \frac{k!}{k} \frac{(n-k)!}{n!} \sum_{y=0}^{n-k} \binom{y+k-1}{y} \\
&= \frac{1}{k} \binom{n}{k}^{-1} \left[\sum_{y=0}^{n-k} \binom{y+k-1}{y} \right] \\
&= \frac{1}{k} \binom{n}{k}^{-1} \left[\binom{0+k-1}{0} + \sum_{y=1}^{n-k} \binom{y+k-1}{y} \right] \\
&= \frac{1}{k} \binom{n}{k}^{-1} \left[1 + \sum_{y=1}^{n-k} \left(\binom{y+k}{y} - \binom{y-1+k}{y-1} \right) \right] \\
&= \frac{1}{k} \binom{n}{k}^{-1} \left[1 + \sum_{y=1}^{n-k} \binom{y+k}{y} - \sum_{y=0}^{n-k-1} \binom{y+k}{y} \right] \\
&= \frac{1}{k} \binom{n}{k}^{-1} \left[1 + \binom{n-k+k}{n-k} - \binom{0+k}{0} \right] \\
&= \frac{1}{k} \binom{n}{k}^{-1} \left[1 + \binom{n}{k} - 1 \right] \\
&= \frac{1}{k}.
\end{aligned}
$$

(d) Let E denote the event that $1, 2, ..., k$ is a cycle, supposing $1 \le k \le n$. The probability of this event is the probability 1 is in a cycle of size k, $P(C_1 = k) = \frac{1}{n}$; and the other values in that cycle are $2, ..., k$, which denotes one of the $\binom{n-1}{k-1}$ possible cycles of size k.

Putting this together,

$$
\begin{aligned}
P(E) &= \frac{1}{n}\binom{n-1}{k-1}^{-1} \\
&= \frac{(k-1)!(n-k)!}{n(n-1)!} \\
&= \frac{k!}{k}\frac{(n-k)!}{n!} \\
&= \frac{1}{k}\binom{n}{k}^{-1}.
\end{aligned}
$$