Assignment 2

2023-12675 박지호

EXERCISE 2.9

a. (a)

- $$\begin{aligned}\mathbf{E}[\max(X_1, X_2)] &= \sum_{n=1}^{k} n \times \Pr(\max(X_1, X_2) = n) \\ &= \sum_{n=1}^{k} \sum_{m=1}^{n} \Pr(\max(X_1, X_2) = n) \\ &= \sum_{m=1}^{k} \sum_{n=m}^{k} \Pr(\max(X_1, X_2) = n) \\ &= \sum_{m=1}^{k} \Pr(\max(X_1, X_2) \geq m) \\ &= \sum_{m=1}^{k} (1 - \Pr(\max(X_1, X_2) < m)) \\ &= \sum_{m=1}^{k} \left(1 - \frac{(m-1)^2}{k^2}\right) \\ &= k - \frac{1}{k^2} \times \frac{k(k-1)(2k-1)}{6} \\ &= \frac{4k^2 + 3k - 1}{6k}\end{aligned}$$

- $$\begin{aligned}\mathbf{E}[\min(X_1, X_2)] &= \sum_{n=1}^{k} n \times \Pr(\min(X_1, X_2) = n) \\ &= \sum_{n=1}^{k} \sum_{m=1}^{n} \Pr(\min(X_1, X_2) = n) \\ &= \sum_{m=1}^{k} \sum_{n=m}^{k} \Pr(\min(X_1, X_2) = n) \\ &= \sum_{m=1}^{k} \Pr(\min(X_1, X_2) \geq m) \\ &= \sum_{m=1}^{k} \frac{(k+1-m)^2}{k^2} \\ &= \frac{1}{k^2} \times \frac{k(k+1)(2k+1)}{6} \\ &= \frac{2k^2 + 3k + 1}{6k}\end{aligned}$$

b. (b)

- From (a), $\mathbf{E}[\max(X_1, X_2)] + \mathbf{E}[\min(X_1, X_2)] = k + 1$

- $\mathbf{E}[X_1] = \mathbf{E}[X_2] = \sum_{n=1}^{k} n \times \Pr(X_1 = n)$
  $= \sum_{n=1}^{k} n \times \frac{1}{k}$
  $= \frac{1}{k} \times \frac{k(k+1)}{2}$
  $= \frac{k+1}{2}$
  - $\mathbf{E}[X_1] + \mathbf{E}[X_2] = k + 1$
- $\therefore \mathbf{E}[\max(X_1, X_2)] + \mathbf{E}[\min(X_1, X_2)] = \mathbf{E}[X_1] + \mathbf{E}[X_2]$

c. (c)

- $\max(X_1, X_2) + \min(X_1, X_2) = X_1 + X_2$
- $\mathbf{E}[\max(X_1, X_2) + \min(X_1, X_2)] = \mathbf{E}[X_1 + X_2]$
  $= \mathbf{E}[\max(X_1, X_2)] + \mathbf{E}[\min(X_1, X_2)] = \mathbf{E}[X_1] + \mathbf{E}[X_2]$

2.15

- Let X be the number of total coin flips until the k-th head
- Let X_i be the number of coin flips after the i-th head until the $i+1$-th head. (Excluding the coin flip that results in i-th head, and including the coin flip that results in $i+1$-th head)
  - X_i all follows geometric distribution with parameter p, $\mathbf{E}[X_i] = \frac{1}{p}$
  - $X = \sum_{i=0}^{k-1} X_i$
- $\mathbf{E}[X] = \sum_{i=0}^{k-1} \mathbf{E}[X_i] = \frac{k}{p}$
- The expected number of coin flips until the k-th head is $\frac{k}{p}$

2.18

- After the n-th item,
- In order for the k-th item to be stored in memory
  - It should have replaced the item when it appeared $\left(\frac{1}{k}\right)$
  - And it should not have been replaced afterwards $\left(\prod_{i=k+1}^{n} \frac{i-1}{i}\right)$
- Therefore, the overall probability of k-th item being stored in memory is

$$\frac{1}{k} \times \prod_{i=k+1}^{n} \frac{i-1}{i} = \frac{1}{k} \times \frac{k}{k+1} \times \ldots \times \frac{n-1}{n} = \frac{1}{n}$$

  - There is an equal chance for each of n items encountered to be the item stored in memory
- $\therefore$ The desired property is achieved

2.22

a. BubbleSort mutation count

- Proposition: The number of mutations in the bubble sort is equal to the total number of inverted pair in the initial permutation
  - Proposition: A bubble sort mutation decreases the number of inverted pair by 1
    - Let
      - The permutation before the mutation be $a_1, ..., a_k, a_{k+1}, ..., a_n$
      - And the permutation after the mutation be $b_1, ..., b_k, b_{k+1}, ..., b_n$
    - Let's say the mutation swapped a_k and a_{k+1}
      - $\forall i \neq k, k+1, b_i = a_i, b_k = a_{k+1}, b_{k+1} = a_k$
    - After the permutation:
      - (b_k, b_{k+1}) is not inverted, while (a_k, a_{k+1}) was inverted
      - $\forall i, j$ such that $i, j \neq k, k+1 \wedge i < j$, the invertedness of the pair (b_i, b_j) is equal to (a_i, a_j)
      - $\forall i < k$, the invertedness of the pair (b_i, b_k) is equal to (a_i, a_{k+1}), and the invertedness of the pair (b_i, b_{k+1}) is equal to (b_i, b_k), leaving the total number of inverted pairs the same.
      - Likewise, $\forall j > k+1$, the invertedness of the pair (b_k, b_j) is equal to (a_{k+1}, a_j), and the invertedness of the pair (b_{k+1}, b_j) is equal to (a_k, a_j), leaving the total number of inverted pairs the same.
    - Therefore, the number of inverted pairs is descreased by exactly one after the mutation.
  - Since the result of the bubble sort is a sorted array with no inverted pairs,
    - The number of inverted pairs in the initial permutation is the number of mutations that happened in the process.

b. Expected Number of Inverted Pairs in Any Permutation

- We can compute the expected number of bubble sort mutation for any permutation by computing the expected number of inverted pair in any permutation
- For any permutation $a_1, ..., a_n$, there exists a permutation where the order of elements are completely reversed $b_1 = a_n, ..., b_n = a_1$
  - Exactly one of (a_i, a_j) and (b_i, b_j) is a inverted pair
  - Therefore, among the two permutations, there is exactly $\binom{n}{2} = \frac{n(n-1)}{2}$ inverted pairs
  - The average number of inverted pair among the two permutations is $\frac{n(n-1)}{4}$
- Since every permutation can be coupled like that, the average number, or the expected number of inverted pairs for every permutation is $\frac{n(n-1)}{4}$
- The expected number of bubble sort mutation for any permutation is also $\frac{n(n-1)}{4}$

EXERCISE 2.27

- $$\begin{aligned} \mathbf{E}[X] &= \sum_{k=1}^{\infty} k \Pr(X = k) \\ &= \sum_{k=1}^{\infty} k \frac{6}{\pi^2} \frac{1}{k^2} \\ &= \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k} \end{aligned}$$
- The expected value of X diverges

EXERCISE 2.32

a. (a)

- For $i \leq m$, trivially $\Pr(E_i) = 0$
- For $i > m$, In order for E_i to happened
  - The i-th candidate must be the best $\left(\frac{1}{n}\right)$
  - We must not hire $m + 1$-th to $i - 1$-th candidate
    - The best candidate among the first $i - 1$ candidate should have been in the first m candidate $\left(\frac{m}{i-1}\right)$
- $$\Pr(E_i) = \begin{cases} 0 \text{ if } i \leq m \\ \frac{1}{n} \times \frac{m}{i-1} \text{ otherwise} \end{cases}$$
  $$\Pr(E) = \sum_{i=1}^{n} \Pr(E_i) = \frac{m}{n} \sum_{i=m+1}^{n} \frac{1}{i-1}$$

b. (b)

- $$\begin{aligned} \forall j \text{ such that } i - 1 < j < i, \frac{1}{j} < \frac{1}{i-1} < \frac{1}{j-1} \\ \sum_{i=m+1}^{n} \int_{i-1}^{i} \frac{1}{j} \mathrm{d}j < \sum_{i=m+1}^{n} \int_{i-1}^{i} \frac{1}{i-1} \mathrm{d}j < \sum_{i=m+1}^{n} \int_{i-1}^{i} \frac{1}{j-1} \mathrm{d}j \\ \int_{m}^{n} \frac{1}{j} \mathrm{d}j < \sum_{i=m+1}^{n} \frac{1}{i-1} < \int_{m}^{n} \frac{1}{j-1} \mathrm{d}j \\ [\![\ln |j|]\!]_m^n < \sum_{i=m+1}^{n} \frac{1}{i-1} < [\![\ln |j-1|]\!]_m^n \\ \ln n - \ln m < \sum_{i=m+1}^{n} \frac{1}{i-1} < \ln(n-1) - \ln(m-1) \\ \frac{m}{n}(\ln n - \ln m) < \Pr(E) < \frac{m}{n}(\ln(n-1) - \ln(m-1)) \end{aligned}$$

c. (c)

- $$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}m}\left(\frac{m}{n}(\ln n - \ln m)\right) &= \frac{1}{n}(\ln n - \ln m) - \frac{m}{n}\frac{1}{m} \\ &= \frac{1}{n}(\ln n - \ln m - 1) = 0 \\ &\Leftrightarrow m = \frac{n}{e}\end{aligned}$$
- $\frac{m}{n}(\ln n - \ln m)$ is maximized when $m = \frac{n}{e}$
- For this m,
- $\Pr(E) > \frac{m}{n}(\ln n - \ln m) = \frac{1}{e}$