论题 2-6 作业

姓名：陈劭源　　　　学号：161240004

1 [CS] Problem 5.6-4

Let X denote the amount of money one wins by playing this game once, and X is a random variable. The expectation of X is

$$E(X) = \frac{1}{4}(1+E(X)) + \frac{1}{4}\times 2 + \frac{1}{4}\times 3 + \frac{1}{4}\times 4$$

Solve this equation, we obtain

$$E(X) = \frac{10}{3} \approx 3.33$$

Therefore, the maximum amount of money a rational person would pay to play this game is \$3.33 .

2 [CS] Problem 5.6-8

$$\begin{aligned}
\sum_{i=1}^{n} E(X|F_i)P(F_i) &= \sum_{i=1}^{n} P(F_i) \sum_{s:s\in F_i} X(s)\frac{P(s)}{P(F_i)} \\
&= \sum_{i=1}^{n} \sum_{s:s\in F_i} X(s)P(s) \\
&= \sum_{s:s\in S} X(s)P(s) \\
&= E(X)
\end{aligned}$$

3 [CS] Problem 5.7-1

X follows the binomial distribution with parameters $n = 5$ and $p = 0.6$. The expectation and variance is

$$E(X) = 5\times 0.6 = 3$$

$$V(X) = 5\times 0.6\times 0.4 = 1.2$$

Therefore,

$$E(X-3) = E(X) - E(3) = 3-3 = 0$$

$$E((X-3)^2) = E((X-E(X))^2) = V(X) = 1.2$$

4 [CS] Problem 5.7-2

Every question is one Bernoulli trial with probability 0.6 of success, therefore $E(X_i) = p = 0.6$ $V(X_1) = p(1-p) = 0.24$ The sum of the variances of X_1 through X_5 is 1.2, which equals to the variance of X, because random variables X_1 through X_5 are independent.

5 [CS] Problem 5.7-4

Let random variable X be the number of right answers. X follows the binomial distribution with parameters $n = 100$ and $p = 0.6$. Therefore

$$E(X) = np = 100 \times 0.6 = 60$$

$$V(X) = np(1-p) = 100 \times 0.6 \times 0.4 = 24$$

$$\sigma(X) = \sqrt{V(X)} = 2\sqrt{6} \approx 4.90$$

6 [CS] Problem 5.7-6

Let random variable X_i be the number of right answers in an i-question quiz. X_i follows the binomial distribution with parameters $n = i$ and $p = 0.8$. Therefore

$$V(X_{25}) = 25p(1-p) = 4$$

$$V(X_{100}) = 100p(1-p) = 16$$

$$V(X_{400}) = 400p(1-p) = 64$$

To "correct" these variances, we can use the standard deviation, i.e. the square root of the variance, instead of variance.

7 [CS] Problem 5.7-12

Assume there are n questions on a short-answer test. Let random variable X denote the number of questions a student who knows 80% of the course material answers correctly. X follows the binomial distribution with parameter n and $p = 0.8$.

By the central limit theorem, the distribution of X converges to the normal distribution with expectation np and variance $np(1-p)$, as n grows large. Hence, if we are 95% sure that such student gets a grade between 75% and 85%, about 2 standard deviations should not be greater than 5% of n, i.e.

$$2\sqrt{np(1-p)} \leq 0.05n$$

Solve this inequality and we get

$$n \geq 1600p(1-p) = 256$$

Therefore, about 256 questions are needed.

8 [CS] Problem 5.7-16

a.

$$\begin{aligned} V(X) &= E((X - E(X))^2) \\ &= \sum_{i=1}^{n} (X(x_i) - E(X))^2 P(x_i) \end{aligned}$$

$$
\begin{aligned}
&\geq \sum_{i=1}^{k}(X(x_i)-E(X))^2P(x_i) \\
&> \sum_{i=1}^{k}P(x_i)r^2 \\
&= P(E)r^2
\end{aligned}
$$

b. Dividing by r^2 on both sides yields

$$P(E) < V(X)/r^2$$

That means, the probability of $|X(x)-E(X)| \geq r$ is no more than $V(X)/r^2$.

9 [CS] Problem 5.7-18

a. X follows the binomial distribution with parameter n and p. The expectation of X is np. By Chebyshev's law, we get

$$P(|X(x)-np| \geq sn) = P(|X(x)-E(X)| \geq sn) \leq V(X)/(s^2n^2) = np(1-p)/(s^2n^2) = p(1-p)/(s^2n)$$

b. For every positive ε (arbitrarily small), if we take $n > \frac{p(1-p)}{s^2\varepsilon}$, $P(|X(x)-np| < ns) > 1-\varepsilon$ holds, because

$$
\begin{aligned}
P(|X(x)-np| < ns) &= 1-P(|X(x)-E(X)| \geq sn) \\
&> 1-p(1-p)/(s^2n) \\
&\geq 1-\frac{p(1-p)}{s^2}\frac{s^2\varepsilon}{p(1-p)} \\
&= 1-\varepsilon
\end{aligned}
$$

That means, the probability of $X(x)$ being between $np-sn$ and $np+sn$ can be arbitrarily close to 1 if n is sufficiently large.

10 [TC] Problem 5.2-1

When hiring exactly once, the first candidate is the best candidate among the n candidates. Therefore, the probability of hiring exactly once is $1/n$.

When hiring exactly n times, the candidates are in increasing order of quality. Since there are $n!$ possible permutations of these candidates, the probability of hiring exactly n times is $1/n!$.

11 [TC] Problem 5.2-2

The first candidate must be hired. Despite the first candidate, exactly one of the remaining candidates is hired. Let p_i denote the probability that the $i-$th candidate is hired, and we have $p_i = 1/i$. Then the probability of hiring twice, denoted as P, is

$$P = \sum_{i=2}^{n}(1-p_2)\cdots p_i \cdots (1-p_n)$$

$$
\begin{aligned}
&= \sum_{i=2}^{n} (1-p_2)\cdots(1-p_n)\frac{p_i}{1-p_i} \\
&= \sum_{i=2}^{n} \left(\prod_{j=2}^{n} (1-p_j) \right) \frac{p_i}{1-p_i} \\
&= \sum_{i=2}^{n} \left(\prod_{j=2}^{n} \frac{j-1}{j} \right) \frac{1/i}{1-1/i} \\
&= \frac{1}{n} \sum_{i=2}^{n} \frac{1}{i-1} \\
&= \frac{\ln n}{n} + O(1/n)
\end{aligned}
$$

12 [TC] Problem 5.2-4

Let X_i be the indicator random variable associated with the event that the $i-$th customer gets his own hat back, the probability of which is $1/n$. We have $X = X_1 + X_2 + \cdots + X_n$, and by the linearity of expectation, we get

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} 1/n = 1$$

13 [TC] Problem 5.3-1

RANDOMIZE-IN-PLACE(A)

1 $n = A.length$
2 swap $A[1]$ with $A[\text{RANDOM}(1, n)]$
3 **for** $i = 2$ **to** n
4   swap $A[i]$ with $A[\text{RANDOM}(i, n)]$

The loop invariant, the **Maintenance** part and the **Termination** part of the modified proof are just the same as the original version. The only difference is the **Initialization** part.

Initialization: Consider the situation just before the first loop iteration, so that $i = 2$. The loop invariant says that for each possible 1-permutation, the subarray contains this 1-permutation with probability $(n-i+1)!/n! = (n-1)!/n! = 1/n$. The subarray $A[1..1]$ is an array contains only one element, which is randomly picked from the original array in line 2. Therefore, $A[1..1]$ contains any 1-permutation with probability $1/n$, and the loop invariant holds prior to the first iteration.

14 [TC] Problem 5.3-2

No. He wants to produce a non-identity permutation randomly, i.e. every possible non-identity permutation can be produced with the same probability. However, because the procedure swaps the first element with one of the rest elements, it cannot produce permutations such that the $A[1]$ is the original element, but $A[2..n]$ contains a non-identity permutation of the rest elements.

15 [TC] Problem 5.3-3

When $n \leq 2$, this code obviously produces a uniform random permutation. However, when $n > 2$, it doesn't.

In each iteration the procedure swaps the current element with a random element in the whole array. Therefore, there are n^n ways to swap the elements, each having the probability $1/n^n$. However, there are $n!$ possible permutations in total. We define a equivalence relation on these n^n ways, that two ways of swapping are equivalent if and only if the two ways produces the same permutation. If this procedure produces a uniform random permutation, the equivalence relation divides the n^n ways into $n!$ equivalence classes of the same size. Hence, the size of each equivalent class is $n^n/n!$. However, when $n > 2$, $n^n/n!$ is not an integer, because $(n-1)$ is a factor of the denominator but not a factor of the numerator when $n > 2$.

16 [TC] Problem 5.3-4

$A[i]$ winds up in position j in B when $\mathit{offset} \mod n = (j-i) \mod n \in \{0, 1, \cdots, n-1\}$, where *offset* is uniformly distributed from 1 to n, i.e $\mathit{offset} \mod n$ is uniformly distributed from 0 to $n-1$. Therefore, each element $A[i]$ has a $1/n$ probability of winding up in any particular position in B.

This procedure performs a right circular shift by $\mathit{offset} \mod n$ positions on the original array, where $\mathit{offset} \mod n$ is a random integer uniformly distributed from 0 to $n-1$. Therefore, the procedure can only produce n of the $n!$ possible permutations of the original array, i.e. the resulting permutation is not uniformly random.

17 [TC] Problem 5.2

a. Pseudocode:

```
RANDOMIZE-SEARCH(A, x)
n = A.length
r = n
let T be a new array of Boolean
for i = 1 to n
    T[i] = false
while r ≠ 0
    t = RANDOM(1, n)
    if T[t] = x
        return t
    else
        if T[t] = false
            T[t] = true
            r = r - 1
// x is not in the array
return -1
```

b. Assume X is the number of indices into A that we pick before we find x. Then we have

$$E(X) = \frac{1}{n} \times 1 + \frac{n-1}{n}(E(X)+1)$$

Solve this equation and we get

$$E(X) = n$$

c. Assume X is the number of indices into A that we pick before we find x. Then we have

$$E(X) = \frac{k}{n} \times 1 + \frac{n-k}{n}(E(X)+1)$$

Solve this equation and we get

$$E(X) = n/k$$

d. Let X_i denote the number of indices into A that we pick before we find an element that has not checked before, when i elements in A has been checked. We have

$$E(X_i) = \frac{n-i}{n} \times 1 + \frac{i}{n}(E(X_i)+1)$$

Solve this equation, we get

$$E(X_i) = \frac{n}{n-i}$$

Let X denote the number of indices into A that we must pick before we have checked all elements of A and RANDOM-SEARCH terminates. We have $X = \sum_{i=0}^{n-1} X_i$, and by the linearity of expectation, we get

$$\begin{aligned} E[X] &= E\left[\sum_{i=0}^{n-1} X_i\right] \\ &= \sum_{i=0}^{n-1} E[X_i] \\ &= \sum_{i=0}^{n-1} \frac{n}{n-i} \\ &= n\sum_{i=1}^{n} \frac{1}{i} \\ &= n \ln n + \Theta(n) \end{aligned}$$