Analysis and Design of Algorithms

NUS
National University of Singapore

Algorithms
CS3230

Tutorial

Week 5

Question 1

The TA would like to encourage everyone in class to prepare before coming to class. To do that, the TA wants to select students who have not prepared to answer questions in class.

The TA selects k students in each tutorial to answer the questions and wants to come up with an algorithm for selecting the students. The worst case for the TA would be the case where only the k selected students have prepared, but the other n-k students in the class did not prepare. The TA thinks that an uncooperative class can make the worst case happen every time if the algorithm for selecting the students is made known. Hence there is no choice but to hide the selection algorithm **(source code or pseudocode)** from the class. Is it true that if the TA makes the selection algorithm known, the students will be able to force the worst case to happen?

- Yes
- No

Bernoulli Trial

An instance of a ***Bernoulli trial*** has probability p of success and probability $q = 1 - p$ of failure.

Geometric Distribution

Suppose we have a sequence of independent Bernoulli trials, each with prob p of success. Let X be the number of trials needed to obtain success for the first time.

Then, X follows the ***geometric distribution***

$$\Pr[X = k] = q^{k-1}p$$

$$E[X] = \frac{1}{p}$$

Question 2

The TA would like to encourage everyone in class to prepare before coming to class. To do that, the TA wants to select students who have not prepared to answer questions in class.

The TA decides to select students at random (with replacement) to answer the questions. Assume that there are n students and k of them have not prepared. What is the expected number of questions required for finding a student who has not prepared?

---

1. n
2. n/k
3. $n/2k$
4. k/n

Coupon Collector Problem

- There are *n* types of coupon that a collector would like to collect. The coupons are put into a box and randomly drawn with replacement.

- What is the expected number of times that the collector needs to draw from the box before he has at least one of each type of coupon?

Coupon Collector Problem

T_i is the time to collect the i-th coupon after i-1 coupon has been collected.

The prob of collecting a new coupon after i-1 coupon has been collected is $p_i = (n-(i-1))/n$.

T_i has a geometric distribution with expectation $1/p_i = n/(n-(i-1))$.

Let T_i be the number of draws used to collect the i-th coupon and T be the total number of draws.

$$T = \sum_{i=1}^{n} T_i$$

Total number of draws

$$E[T] = E\left[\sum_{i=1}^{n} T_i\right]$$

Expected value

$$= \sum_{i=1}^{n} E[T_i]$$

Linearity of expectation

Linearity of Expectations

- For any two events X, Y (does not matter whether dependent or independent) and a constant a.

$$E[X+Y] = E[X] + E[Y]$$

$$E[aX] = aE[X]$$

$$
\begin{aligned}
E[T] &= \sum_{i=1}^{n} E[T_i] \\
&= \sum_{i=1}^{n} \frac{n}{n-(i-1)} \\
&= \frac{n}{n} + \frac{n}{n-1} + \cdots + \frac{n}{1} \\
&= n \cdot \left(\frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \right) \\
&= n \cdot H_n \\
&= \Theta(n \lg n)
\end{aligned}
$$

where H_n is the n-th harmonic number and is $\Theta(\lg n)$

Question 3

The TA would like to encourage everyone in class to prepare before coming to class. To do that, the TA wants to select students who have not prepared to answer questions in class.

The TA decides to select students at random to answer questions. However, the TA is worried that not everyone in class would get to participate even though many questions will be asked throughout the semester. Assume that there are n students in the class. What is the expected number of questions that need to be asked before every student has been asked a question?

- $\Theta(\lg n)$
- $\Theta(n)$
- $\Theta(n \lg n)$
- $\Theta(n^2)$

Question 4

- Let A[1..n] be an array of n distinct names. Suppose m of them are male names. We hope to select q male names from A[1..n]. We propose the following algorithm to obtain q male names.

- Since personal data is sensitive, we hope to estimate the expected number of accesses to the array *A*.

- Please compute the expected number of access of Query(A, q).

Query(*A, q*)
Let S=Φ;
for *j* = 1 **to** *q*
  Repeat
    Randomly select *k* from
      {1, 2, …, n};
    Set *B*=*A*[*k*];
  Until *B* is a male and *k*∉*S*;
  S = {*k*}∪*S*;
Report S;