CSC 284/484 - homework 1 (basic probability)

http://www.cs.rochester.edu/~stefanko/Teaching/16CS484

Students that take the course as 484 are required to do **both** 284/484 and 484 parts of the homework. Students that take the course as 284 are only required to do 284/484 part of the homework (of course you are welcome to solve/turn-in the 484 part as well).

---

1 284/484 homework - solve and turn in

1.1 Theoretical part

Exercise 1.1 (**due 1/26/2016**) We throw m balls into n bins (each ball is thrown into a uniformly random bin independently of the other balls). What is the expected value of the number of empty bins?

Exercise 1.2 (**due 1/26/2016**) What is the expected running time of the following algorithm:

$X = n$
while $X > 0$ do $X =$ uniformly random integer in $\{0, \ldots, X-1\}$.

Asymptotic (Θ) answer is sufficient.

1.2 Applied part

Exercise 1.3 (**due 1/26/2016**) We are given n random variables $X_1, \ldots, X_n$ and we want to find the largest k such that the following statement is true: "the variables $X_1, \ldots, X_n$ are k-wise independent".

Your implementation should read the input from `stdin`. The first line of the input contains an integer $m = |\Omega|$ (the size of the sample space; w.l.o.g, $\Omega = \{1, \ldots, m\}$; each atomic event has the same probability $1/m$) and an integer n (the number of random variables). Each of the next n lines contain m integers—the i-th line contains $X_i(1), \ldots, X_i(m)$.

Example input:

```
4 3
0 0 1 1
0 1 0 1
0 1 1 0
```

Example output:

```
2
```

2 484 homework - solve and turn in

2.1 Theoretical part

Exercise 2.1 (**due 1/26/2016**) Let $A, B \subseteq U$ be disjoint and such that $|A| = \Theta(n)$ and $|B| = \Theta(n)$. Let R be a random subset of U where each element is picked independently with probability $\Theta(1/n)$. Show

$$P(A \cap R = \emptyset \text{ and } B \cap R \neq \emptyset) > c,$$

where c is a constant (depending on the constants in the Θ's).

Exercise 2.2 (**due 1/26/2016**) Let $0 < p < 1$. Let $X_1, X_2, \ldots$ be independent identically distributed random variables $P(X_i = -1) = 1 - p$, $P(X_i = 1) = p$. Let T be the smallest t such that $\sum_{i=1}^{t} X_i < 0$ (note that $P(T = 1) = 1 - p$). What is $E[T]$? (Your answer should be a function of p.)

Exercise 2.3 (**due 1/26/2016**) Let p be a prime. Let n be a number $n \leq p$.

- pick random $a_0, \ldots, a_{k-1} \in \{0, \ldots, p-1\}$,
- let $X_g = a_0 + a_1 g + a_2 g^2 + \cdots + a_{k-1} g^{k-1} \mod p$ for $g = 0, \ldots, n-1$.

Prove that the random variables $X_0, \ldots, X_{n-1}$ are k-wise independent.

2.2 Applied part

The objective of the next exercise is to make you familiar with the most useful construction of k-wise independent random variables (and also k-wise independent hash functions). It is the same idea as **??** except it uses finite fields of size 2^t. Don't be scared— You only need to understand very little about finite fields in order to implement the method:
(`https://en.wikipedia.org/wiki/Finite_field_arithmetic`).

Do implement your own finite field arithmetic (do not use other's implementations). It is fine if you use brute force to find an irreducible polynomial and generator of the multiplicative group. It is fine to have quadratic algorithm for polynomial multiplication and division. Do not optimize.

Exercise 2.4 (**due 1/26/2016**) Implement a program that generates samples from k-wise independent random variables $X_0, \ldots, X_{n-1}$ with values in $GF(2^t)$. For simplicity assume $n < 2^t$. Use the following method:

- pick random $a_0, \ldots, a_{k-1} \in GF(2^t)$,
- let g be a generator of the multiplicative group of $GF(2^t)$,
- let $X_i = a_0 + a_1 g^i + a_2 g^{2i} + \cdots + a_{k-1} g^{(k-1)i}$ (all computations done in $GF(2^t)$).

The input to your program is: n, k, t (where $n < 2^t$); the output is a sample from $X_0, \ldots, X_{n-1}$.