Day 6
Probabilistic method
確率的手法

- Tools in probability
  - Linearity of expetation 期待値の線形性
    - Max cut problem 最大カット問題
  - Birthday trick 誕生日のトリック
  - Coupon collector problem クーポン収集問題

Max Cut　最大カット

- Given a graph G = (V,E), a cut is a partition of V into two sets X and V-X

  カット:グラフの頂点の分割
- The size of cut is the number of edges between X and V-X　カットの大きさ
  - Max cut: the maximum size cut
  - Min cut: the minimum size cut
- Problem: Show the size of max cut is at least |E|/2.　最大カットのサイズは辺数の過半数

Birthday trick　誕生日トリック

- We have k students in a class.
  - What is the probability that there is a pair of students with the same birthday.
    - How large k to attain probability 1?
    - How about k＝30？
- Ｋ人の学生で同じ誕生日が居る確率は？
  - 確率１になるのは？

Applications of birthday trick 1

- We have n nuts and n bolts
- Each nut has exactly one fitting bolt
- We can test a bolt and a nut, and find which is of larger size.
- We can compare two bolts and see which is larger.
- How to find a fitting pair of nuts and bolts with a small number of comparisons?
- n 個のボルトとナット（n対の合致がある）から、合致する対を見つける問題

Discrete Logarithm 離散対数問題

- We have a large prime **p** and a natural number **g**
- Given x, its discrete logarithm with base g and modulo p is a number y such that

  $g^{y} = x \bmod p$
- We want to compute y efficiently
- Find an O($p^{½}$) time method
  - You can use the fact that $g^{p} = g$

Coupon collector problem
クーポン収集問題

- There are n types of coupons and at each trial a coupon is chosen at random. Each trial is independent from others. Let X be the random variable showing the number of trials required to have all kinds of coupons. n 種類のクーポンを揃えるために、何枚のクーポンを集めるだろうか。
  - Find the expected value E(X) of X
  - Find the variance E(X) of X

Some beautiful formulas

$$1+1/2+1/3+...+1/n \le n \ln n+\gamma$$

$$\varsigma(s)=1+1/2^s+1/3^s+1/4^s+....=\sum_n 1/n^s$$

$$\varsigma(2)=\pi^2/6$$

Chebyshev's inequality: For a random variable X with standard deviation σ and expectation μ, Prob(|X – μ | > t σ) < 1/t^2

Stronger tools: Chernoff's ineuality, Azuma's inequality, which I omit here

Coupon collector

- x(i): the time to find a new coupon after having i different coupons.
  - p_i= (n-i）/n : probability to find a new coupon
- $E(x(i))= \quad 1/p_i = n/(n-i)$
- $V(x(i)) = (1-p_i)/p_i^2 = ni/(n-i)^2$
- Expectation and variance are linear!, so add x(i) for i=1,2,..,n

http://www-stat.stanford.edu/~susan/surprise/Collector.html

Probability is not linear

- The probability of A or B happens is not always Pr(A) + Pr(B)
- The probability of A and B happne is not always Pr(A)Pr(B)
  - When it happens??
  - You will see that you should always aware of above, even in real life….

Puzzle of 100 prisoners

- Story: The names of 100 prisoners are placed in 100 wooden boxes, one name to a box, and the boxes are lined up on a table in a room. One by one, the prisoners are led into the room, each may look in at most 50 boxes, but must leave the room then, and is permitted no further communication with the others. The prisoners have a chance to plot their strategy in advance, and they are going to need it, because unless every single prisoner finds his own name all will subsequently be executed.
- Problem: Find a strategy for them which has probability of success exceeding 30%

100人の囚人の問題

- 100 人の囚人の氏名が、テーブルに一列に並んだ100 個の木箱に一つづつ入っている。囚人は一人づつ呼ばれて、高々50個の箱の中身を見ることができる。終了後は退室し、他の囚人との情報交換はできない。囚人たちは前もって作戦を立てることができる。囚人全員が自分の名前を見つけない限り、全員の囚人が一斉に処刑される運命にある
- 問題: 成功確率が30%以上になる作戦を立案せよ

Intuition

- The success probability of each prisoner is exactly 0.5.
- Thus, the probability that all prisoners find their names is $(0.5)^{100}$, which is less than 0.000,000,000,000,000,000,000,000,000,001.
- How to increase it to 0.3?
- Expectation number of names found is 50.
- Concentrate the success!
  - If we make "at least 75 prisoners fail" or "all prisoners success", we are done. (why??)