Probabilistic Method and Random Graphs

Lecture 2. Moments and Inequalities [1]

Xingwu Liu

Institute of Computing Technology
Chinese Academy of Sciences, Beijing, China

[1] The slides are partially based on Chapters 3 and 4 of Probability and Computing.

Preface

Questions, comments, or suggestions?

Monty Hall Problem?

Review

1. Probability axioms
2. Union Bound
3. Independence
4. Conditional probability and chain rule
    - $\Pr(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} \Pr(A_i | \bigcap_{j=1}^{i-1} A_j)$
5. Random variables: expectation, linearity

Bernoulli distribution

Bernoulli random variable

- $\Pr(X = 1) = p, \Pr(X = 0) = 1 - p$
- Modeling coin flip
- $\mathbb{E}[X] = p * 1 + (1 - p) * 0 = p$
- $X^k = X$

Bernoulli distribution

Bernoulli random variable

- $\Pr(X = 1) = p, \Pr(X = 0) = 1 - p$
- Modeling coin flip
- $\mathbb{E}[X] = p * 1 + (1 - p) * 0 = p$
- $X^k = X$

An example

How many triangles among 4 nodes when the links appear independently randomly?

Binomial distribution

Binomial random variable

- The number of successes in n independent trials of the Bernoulli experiment with success probability p
- For any $0 \le i \le n$, $\Pr(X = i) = C_n^i p^i (1-p)^{n-i}$
- $X = \sum_{i=1}^{n} X_i$
- $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$

Binomial distribution

Binomial random variable

- The number of successes in n independent trials of the Bernoulli experiment with success probability p
- For any $0 \le i \le n$, $\Pr(X = i) = C_n^i p^i (1-p)^{n-i}$
- $X = \sum_{i=1}^{n} X_i$
- $\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$

Application

Router packets sampling

The story of Farmer&Rabbit

Geometric distribution

The story of Farmer&Rabbit

Geometric random variable

- The number of independent trials until success, where each trial has success probability p
- $\Pr(X = i) = (1-p)^{i-1}p$ for $i \geq 1$
- $\mathbb{E}[X] = \sum_{i \geq 1} i(1-p)^{i-1}p = 1/p$

The story of Farmer&Rabbit

Geometric random variable

- The number of independent trials until success, where each trial has success probability p
- $\Pr(X = i) = (1-p)^{i-1}p$ for $i \geq 1$
- $\mathbb{E}[X] = \sum_{i\geq 1} i(1-p)^{i-1}p = 1/p$

Suppose the daily probability that God throws a rabbit at the trunk is 10^{-4}. How many years does the farmer has to wait?

Geometric distribution

The story of Farmer&Rabbit

Geometric random variable

- The number of independent trials until success, where each trial has success probability p
- $\Pr(X = i) = (1-p)^{i-1}p$ for $i \geq 1$
- $\mathbb{E}[X] = \sum_{i \geq 1} i(1-p)^{i-1}p = 1/p$

Suppose the daily probability that God throws a rabbit at the trunk is 10^{-4}. How many years does the farmer has to wait?

Memoryless: particular to geometric distribution

For geometric random variable X, if $n > 0$,
$\Pr(X = n + k | X > k) = \Pr(X = n)$

Application: coupon collector's problem

Problem statement

The # of purchases for collecting n coupon types?

Application: coupon collector's problem

Problem statement

The # of purchases for collecting n coupon types?

- X_i: the number of purchases after you get $i-1$ types of coupons until getting the ith one
- $X = \sum_{i=1}^{n} X_i$
- X_i: geometric random variable with parameter $p_i = 1 - \frac{i-1}{n}$
- $\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}$

Application: coupon collector's problem

Problem statement

The # of purchases for collecting n coupon types?

- X_i: the number of purchases after you get $i-1$ types of coupons until getting the ith one
- $X = \sum_{i=1}^{n} X_i$
- X_i: geometric random variable with parameter $p_i = 1 - \frac{i-1}{n}$
- $\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}$

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \sum_{i=1}^{n} \frac{1}{i} = n \ln n + \Theta(n).$$

Fight the salesman

When $n = 12$, $\mathbb{E}[X] \approx 30 <<<<<<<< 200$. Impossible! Cheater!

Coupon collector's problem: fight the salesman

Expectation is too weak

The probability of exceeding expectation can be arbitrarily big, Guy!

Example

- Random variables Y_α with $\alpha \geq 1$
- Let $\Pr(Y_\alpha = \alpha) = \frac{1}{\alpha}$ and $\Pr(Y_\alpha = 0) = 1 - \frac{1}{\alpha}$
- $\mathbb{E}[Y_\alpha] = 1$
- $\Pr(Y_\alpha \geq 1) = \frac{1}{\alpha}$ can be arbitrarily close to 1

But, mh...
Is it possible to exceed so much with high probability?

An inequality for tail probability

Markov's inequality

If $X \geq 0$ and $a > 0$, $\Pr(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$.

Proof:

$$\mathbb{E}[X] = \sum_{i \geq 0} i * \Pr(X = i) \geq \sum_{i \geq a} i * \Pr(X = i)$$
$$\geq \sum_{i \geq a} a * \Pr(X = i) = a * \Pr(X \geq a).$$

Observations

- Intuitive meaning (level of your income)
- With 12 coupons, $\mathbb{E}[X] \approx 30, \Pr(X \geq 200) < 1/6$
- Loose? Tight when only expectation is known!

Conditional expectation

Definition

$$\mathbb{E}_Y[Y|Z] \triangleq \mathbb{E}[Y|Z=z] \triangleq \sum_y y * \Pr(Y=y|Z=z)$$

Theorem

$$\mathbb{E}_Z[\mathbb{E}_Y[Y|Z]] \triangleq \sum_z \mathbb{E}[Y|Z=z]\Pr(Z=z) = \mathbb{E}[Y]$$

Proof.

$$\begin{aligned}\sum_z \Pr(Z=z)\mathbb{E}[Y|Z=z] &= \sum_z \Pr(Z=z)\sum_y y\frac{\Pr(Y=y,Z=z)}{\Pr(Z=z)}\\ &= \sum_y y\sum_z \Pr(Y=y,Z=z)\\ &= \sum_y y\Pr(Y=y) = \mathbb{E}[Y]\end{aligned}$$

□

Application: expected run-time of Quicksort

Via conditional expectation

- X_n: the runtime of sorting an n-sequence.
- K: the rank of the pivot.
- If $K = k$, the pivot divides the sequence into a $(k-1)$-sequence and an $(n-k)$-sequence.
- Given $K = k$, $X_n = X_{k-1} + X_{n-k} + n - 1$.
- $\mathbb{E}[X_n | K = k] = \mathbb{E}[X_{k-1}] + \mathbb{E}[X_{n-k}] + n - 1$.
- $\mathbb{E}[X_n] = \sum_{k=1}^{n} \Pr(K = k)(\mathbb{E}[X_{k-1}] + \mathbb{E}[X_{n-k}] + n - 1)$
  $= \sum_{k=1}^{n} \frac{\mathbb{E}[X_{k-1}] + \mathbb{E}[X_{n-k}]}{n} + n - 1$.
- Please verify that $\mathbb{E}[X_n] = 2n \ln n + O(n)$.

Application: expected run-time of Quicksort

Via linearity + indicators

- y_i: the i-th biggest element
- Y_{ij}: indicator for the event that y_i, y_j are compared
- $Y_{ij} = 1$ iff the first pivot in $\{y_i, y_{i+1}, ...y_j\}$ is y_i or y_j
- $\mathbb{E}[Y_{ij}] = \Pr(Y_{ij} = 1) = \frac{2}{j-i+1}$
- $X_n = \sum_{i=1}^{n} \sum_{j=i+1}^{n} Y_{ij}$
- $\mathbb{E}[X_n] = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{E}[Y_{ij}]$
- It is easy to see that $\mathbb{E}[X_n] = (2n+2) \sum_{i=1}^{n} \frac{1}{i} + O(n)$

Moments of random variables

Why moments?

- *Global* features of a random variable.
- Expectation is too weak: can't distinguish Y_α

Definition

- kth moment: $\mathbb{E}[X^k]$.
- Variance: $Var[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$, whose square root indicates how far the variable deviates from the expectation.
- Example: $Var[Y_\alpha] = \alpha - 1$

- Covariance: $Cov(X, Y) \triangleq \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$.
- It's zero in case of independence.

Properties of the variance

$Var[X+Y] = Var[X] + Var[Y] + 2Cov(X,Y)$

$Var[X+Y] = Var[X] + Var[Y]$ if X and Y are independent.

$Var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

$Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$

Variances of some random variables

Binomial random variable with parameters n and p

- $X = \sum_{k=1}^{n} X_i$ with the X_i's independent.
- $Var[X_i] = p - p^2 = p(1-p)$.
- $Var[X] = \sum_{k=1}^{n} Var[X_i] = np(1-p)$

Geometric random variable with parameter p

Straightforward computing shows that $Var[X] = \frac{1-p}{p^2}$

Coupon collector's problem

- We know that $Var[X_i] = \frac{1-p_i}{p_i^2}$.
- $Var[X] = \sum_{k=1}^{n} Var[X_i] \leq \sum_{k=1}^{n} \frac{n^2}{(n-k+1)^2} \leq \frac{\pi^2 n^2}{6}$

A new argument against the salesman

Chebyshev's inequality

- $\Pr(|X - \mathbb{E}[X]| \geq a) \leq \frac{Var[X]}{a^2}$.
- An immediate corollary from Markov's inequality.

Coupon collector's problem

$\Pr(X \geq 200) = \Pr(|X - \mathbb{E}[X]| \geq 170) \leq \frac{255}{170^2} < 0.01$

A new argument against the salesman

Chebyshev's inequality

- $\Pr(|X - \mathbb{E}[X]| \geq a) \leq \frac{Var[X]}{a^2}$.
- An immediate corollary from Markov's inequality.

Coupon collector's problem

$\Pr(X \geq 200) = \Pr(|X - \mathbb{E}[X]| \geq 170) \leq \frac{255}{170^2} < 0.01$

Trump card

- By union bound, $\Pr(|X - nH_n| \geq 5nH_n) \leq \frac{1}{n^5}$.
- Hint: Consider the probability of not containing the ith coupon after $(c+1)n \ln n$ steps.

Union bound beats the others. What a surprise!

Brief introduction to Chebyshev

- May 16, 1821 – December 8, 1894
- A founding father of Russian mathematics

- Probability, statistics, mechanics, geometry, number theory
- Chebyshev inequality, Bertrand-Chebyshev theorem, Chebyshev polynomials, Chebyshev bias
- Aleksandr Lyapunov, Markov brothers