Probability Homework Solution

Wang Jindong
201418013229092
✉wangjindong@ict.ac.cn

Jan. 17$^{\text{th}}$ 2016

Contents

Chapter 1 Probability Basics

1.1 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1s_2\cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

1.2 Mutually and Pair-wise independent

Problem
Give an example to show that pairwise independent does not mean mutually independent.

Solution

Proof. Mutually independent does not mean pair-wise independent.For example,A, B, C are 3 pair-wise independent events, which means A and B are independent, B and C are independent and A and C are also independent.This gives the simple fact that

$$Pr(AB) = Pr(A)Pr(B)$$

$$Pr(AC) = Pr(A)Pr(C)$$

$$Pr(BC) = Pr(B)Pr(C)$$

But this has nothing directed to do with

$$Pr(ABC) = Pr(A)Pr(B)Pr(C)$$

which is exactly what mutually independent means.
All in all, mutually independent means pair-wise independent,but not hold vice versa. □

1.3 Monty Hall

Problem
Suppose you are on a game show, and you are given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No.1, and the host, who knows what is behind the doors, opens another door, say No.3, which has a goat. He then says to you, "Do you want to pick door No.2?" Calculate the probability that you wins the car if you switch you choice.

Solution
Consider the events C_1, C_2 and C_3 indicating the car is behind respectively door 1,2 or 3.All these 3 events have probability 1/3.
The player initially choosing door 1 is described by the event X_1.As the first choice of the player is independent of the position of the car,also the conditional probabilities are $Pr(C_i|X_1) = 1/3$.The host opening door 3 is described by H_3.For this event it holds:

$$Pr(H_3|C_1, X_1) = \frac{1}{2}$$

$$Pr(H_3|C_2, X_1) \ = 1$$
$$Pr(H_3|C_3, X_1) = 0$$

Then,if the player initially selects door 1,and the host opens door 3,the conditional probability of winning by switching is:

$$\begin{aligned}
&Pr(C_2|H_3, X_1)\\
&= \frac{Pr(H_3|C_2, X_1)Pr(C_2|X_1)}{Pr(H_3|X_1)}\\
&= \frac{Pr(H_3|C_2, X_1)Pr(C_2|X_1)}{Pr(H_3|C_1, X_1)Pr(C_1|X_1) + Pr(H_3|C_2, X_1)P(C_2|X_1) + Pr(H_3|C_3, X_1)Pr(C_3|X_1)}\\
&= \frac{Pr(H_3|C_2, X_1)}{Pr(H_3|C_1, X_1) + Pr(H_3|C_2, X_1) + Pr(H_3|C_3, X_1)}\\
&= \frac{1}{1/2 + 1 + 0}\\
&= \frac{2}{3}
\end{aligned}$$

1.4 Dice

Problem

Assume that you independently play an unbiased 6-facet dice for n times. Let X be the summation of the results. Calculate the probability that X is divisible by 3.

Solution

Let Y_i denote the result of $i-th$ tossing the dice,while X the sum of n times tossing the dice.It's clear that

$$X = \sum_{i=1}^{n} Y_i$$
$$X_k = \sum_{k} Y_i$$

Therefore according to the Conditional probability theory

$$\begin{aligned}
&Pr(X \text{ is divisible by } 3) = Pr(X \text{ is divisible by } 3|X_{n-1} = x)Pr(X_{n-1} = x)\\
&= \sum_{y} (Y_n + x \text{ is divisible by } 3|X_{n-1} = x)Pr(X_{n-1} = x)\\
&= \frac{1}{3}\sum_{y} Pr(X_{n-1} = x)\\
&= \frac{1}{3}
\end{aligned}$$

1.5 Child Birth

Problem

Assume that on an island, each couple gives birth to babies until a female baby comes out. Suppose that a baby will be male or female with probability 0.5. On average how many male/female babies does a couple have? What if each couple refuses to have more than 5 babies?

Solution

Let M denote the number of male babies and F denote the number of female babies. If each couple gives birth to babies until a female baby comes out, we have

$$E[F] = 1$$

and

$$E[M] = \sum_{i \geq 1} (i-1)\left(\frac{1}{2}\right)^i = 1$$

If each couple refuses to have more than 5 babies, we have

$$E[F] = 1 - \left(\frac{1}{2}\right)^5 = \frac{31}{32}$$

and

$$E[M] = \sum_{i=1}^{5} (i-1)(\frac{1}{2})^i + 5\left(\frac{1}{2}\right)^5 = \frac{31}{32}$$

Chapter 2 Moments and Inequalities

2.1 Quicksort

Problem
Given n,for any $1 \leq i \leq j \leq n$,define Bernoulli random variable Y_{ij} s.t. $Pr(Y_{ij} = 1) = \frac{2}{j-i-1}$.Let $X = \sum_{1 \leq i < j \leq n} Y_{ij}$.Prove that $E(X) \leq 2n \ln n + O(n)$.

Solution

Proof.

$$\begin{aligned}
E(X) &= E\left(\sum_{1 \leq i < j \leq n} Y_{ij}\right) \\
&= \sum_{1 \leq i < j \leq n} E\left(Y_{ij}\right) \\
&= \sum_{1 \leq i < j \leq n} \frac{2}{j-i+1} \\
&= 2\left[\frac{1}{2} + \left(\frac{1}{2} + \frac{1}{3}\right) + \cdots + \left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right)\right] \\
&= 2\left(\frac{n-1}{2} + \frac{n-2}{3} + \cdots + \frac{2}{n-1} + \frac{1}{n}\right) \\
&\leq 2\left(n-1\right)\left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) \\
&= 2\left(n-1\right)\left(\ln n + C\right) \\
&= 2n\left(\ln n + C\right) \\
&= 2n \ln n + O(n)
\end{aligned}$$

Such that the proof ends. $\square$

2.2 Markov's Inequality

Problem
Prove the following extensions of the Chernoff bound. Let $X = \sum_{i=1}^{n} X_i$, where the X_i are independent 0-1 random variables. Let $\mu = E[X]$. Choose any μ_L and μ_H such that $\mu_L \leq \mu \leq \mu_H$. Then, for any $\delta > 0$, $Pr(X \geq (1+\delta)\mu_H) \leq (\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}})\mu_H$. Similarly, for any $0 < \delta < 1$, $Pr(X \leq (1-\delta)\mu_L) \leq (\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}})\mu_L$.

Solution

Proof. At first we will prove

$$\begin{aligned}
E[e^{tX}] &= E\left[e^{t\sum_{i=1}^{n} X_i}\right] \\
&= E\left[t\prod_{i=1}^{n} X_i\right] \\
&= \prod_{i=1}^{n} E\left[e^{tX_i}\right] \\
&= \prod_{i=1}^{n} \left(p_i \cdot e^t + (1-p_i)\cdot 1\right) \\
&= \prod_{i=1}^{n} (1 + p_i(e^t - 1)) \\
&\leq \prod_{i=1}^{n} e^{p_i(e^t-1)} \\
&= e^{\sum_{i=1}^{n} p_i(e^t-1)} \\
&= e^{(e^t-1)\mu} \\
&\leq e^{(e^t-1)\mu_H}
\end{aligned}$$

First inequality proof
Apply Markov's inequality, for any $t > 0$ we have

$$\begin{aligned}
Pr\left(X \geq (1+\delta)\mu_H\right) &= Pr\left(e^{tX} \geq e^{t(1+\delta)\mu_H}\right) \\
&\leq \frac{E[e^{tX}]}{e^{t(1+\delta)\mu_H}} \\
&\leq \frac{e^{(e^t-1)\mu}}{e^{t(1+\delta)\mu_H}} \\
&\leq \frac{e^{(e^t-1)\mu_H}}{e^{t(1+\delta)\mu_H}}
\end{aligned}$$

For any $\delta > 0$, we can set $t = \ln(1+\delta)$ then

$$\begin{aligned}
Pr(X \geq (1+\delta)\mu_H) &\leq \frac{e^{(e^t-1)\mu_H}}{e^{t(1+\delta)\mu_H}} \\
&= \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)\mu_H
\end{aligned}$$

Second inequality proof
Apply Markov's inequality, for any $t < 0$ we have

$$\begin{aligned}
Pr(X \leq (1-\delta)\mu_L) &= Pr\left(e^{tX} \geq e^{t(1-\delta)\mu_L}\right) \\
&\leq \frac{E[e^{tX}]}{e^{t(1-\delta)\mu_L}} \\
&\leq \frac{e^{(e^t-1)\mu}}{e^{t(1-\delta)\mu_L}} \\
&\leq \frac{e^{(e^t-1)\mu_L}}{e^{t(1-\delta)\mu_L}}
\end{aligned}$$

For any $0 < \delta < 1$, we can set $t = \ln(1-\delta) < 0$ then

$$\begin{aligned}
Pr(X \leq (1-\delta)\mu_L) &\leq \frac{e^{(e^t-1)\mu_L}}{e^{t(1-\delta)\mu_L}} \\
&= \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)\mu_L
\end{aligned}$$

□

2.3 Moment and Chernoff Bound

Problem

Let $X_1, \cdots X_n$ be independent Poisson trials such that $Pr(X_i = 1) = p_i$ and let $a_1, \cdots a_n$ be real numbers in $[0, 1]$. let $X = \sum_{i=1}^{n} a_i X_i$ and $\mu = E[X]$. Then the following Chernoff bound holds: for any $\delta > 0$, $Pr(X \geq (1+\delta)\mu) \leq (\frac{e^\delta}{(1+\delta)^{(1+\delta)}})^\mu$. Also prove a similar bound for the probability $Pr(X \leq (1-\delta)\mu)$ for any $0 < \delta < 1$.

Solution

Proof. First we can see that $\mu = E[X] = \sum_{i=1}^{n} a_i p_i$, and use this to calculate $M_X(t)$. We can see that $M_{X_i}(t) = 1 + p_i(e^t - 1)$, so

$$
\begin{aligned}
E(e^{tX}) = M_X(t) &= \prod_{i=1}^{n}(a_i + a_i p_i(e^t - 1)) \\
&\leq \prod_{i=1}^{n}(1 + a_i p_i(e^t - 1)) \\
&\leq \prod_{i=1}^{n} e^{a_i p_i(e^t-1)} \\
&= e^{\sum_{i=1}^{n} a_i p_i(e^t-1)} \\
&= e^{(e^t-1)\mu}
\end{aligned}
$$

Then we get the same the product of the n generating functions as the independent Poisson trials has. So we get the same Chernoff bound. That is to say $Pr(X \geq (1+\delta)\mu) \leq (\frac{e^\delta}{(1+\delta)^{(1+\delta)}})^\mu$ for any $\delta > 0$ and $Pr(X \leq (1-\delta)\mu) \leq (\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}})^\mu$ for any $0 < \delta < 1$. □

2.4 Convex Function and Chernoff Bound

Problem

Recall that a function f is said to be convex if, for any x_1, x_2 and for $0 \leq \lambda \leq 1$, $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$.

- Let Z be a random variable that takes on a finite set of values in $[0, 1]$, and let $p = E[Z]$. Define the Bernoulli random variable X by $Pr(X = 1) = p$ and $Pr(X = 0) = 1 - p$. Show that $E[f(Z)] \leq E[f(X)]$ for any convex function f.
- Use the fact that $f(x) = e^{tx}$ is convex for any fixed $t \geq 0$ to obtain a Chernoff-like bound for Z.

Solution

Proof.

- According to convex definition,we have

$$
\begin{aligned}
E[f(Z)] &= \sum_{i=1}^{n} f(z_i)p_i \\
&= \sum_{i=1}^{n}((1-z_i)\cdot 0 + z_i \cdot 1)p_i) \\
&\leq \sum_{i=1}^{n}((1-z_i)f(0) + z_i f(1)p_i) \\
&= \left(1 - \sum_{i=1}^{n} p_i z_i\right) f(0) + \sum_{i=1}^{n} p_i z_i f(1) \\
&= (1 - E[Z])f(0) + E[Z]f(1) \\
&= E[f(X)].
\end{aligned}
$$

Such that the proof ends.

- Simply we have $E[e^{tX}] = (1-p) + pe^t$. According to *Markov's* inequality, for any $t > 0$ we have

$$\begin{aligned} Pr(Z \geq a) &= Pr(e^{tZ} \geq e^{ta}) \\ &\leq \frac{E(e^{tZ})}{e^{ta}} \\ &\leq \frac{E(e^{tX})}{e^{ta}} \\ &= \frac{pe^t - p + 1}{e^{ta}} \end{aligned}$$

So we have $Pr(Z \geq a) \leq \frac{pe^t - p+1}{e^{ta}}$ for any $t > 0$ as a *Chernoff*-like bound.

□

2.5 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1 s_2 \cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

Chapter 3 Bins and balls: handling dependence

3.1 Maxload

Problem
Suppose that balls are thrown randomly into n bins. Show, for some constant c_1, that if there are $c_1\sqrt{n}$ balls then the probability that no two land in the same bin is at most $1/e$. Similarly, show for some constant c_2 (and sufficiently large n) that, if there are $c_2\sqrt{n}$ balls, then the probability that no two land in the same bin is at least $1/2$. Make these constants as close to optimum as possible. Hint: you may need the fact that $e^{-x} \geq 1 - x$ and $e^{-x-x^2} \leq 1 - x$ for $x \leq 1/2$.

Solution

- Find c_1
  The probability that max load is 1 is

$$\begin{aligned}
Pr(MaxLoad = 1) &= (1 - \frac{1}{n})(1 - \frac{2}{n}) \cdots (1 - \frac{c_1\sqrt{n} - 1}{n}) \\
&\leq \prod_{i=1}^{c_1\sqrt{n}-1} e^{-\frac{i}{n}} \\
&= e^{\frac{c_1\sqrt{n} - c_1^2 n}{2n}} \\
&\leq e^{\frac{c_1 - c_1^2}{2}}
\end{aligned}$$

  Then we set $\frac{c_1 - c_1^2}{2} \leq -1$, and get $c_1 \geq 2$. □

- The probability that max load is 1 is

$$\begin{aligned}
Pr(MaxLoad = 1) &= (1 - \frac{1}{n})(1 - \frac{2}{n}) \cdots (1 - \frac{c_2\sqrt{n} - 1}{n}) \\
&\geq \prod_{i=1}^{c_2\sqrt{n}-1} e^{-\frac{i}{n} - \frac{i^2}{n^2}} \\
&= e^{-\frac{c_2\sqrt{n}(c_2\sqrt{n})}{2n} - \frac{(c_2\sqrt{n}-1)c_2\sqrt{n}(2c_2\sqrt{n}-1)}{6n^2}} \\
&\geq e^{-\frac{c_2^2}{2} - \frac{2c_2^3 n\sqrt{n}}{6n^2}} \\
&\geq e^{-\frac{c_2^2}{2} - \frac{c_2^3}{3}} \\
&\geq 1 - \frac{c_2^2}{2} - \frac{c_2^3}{3}
\end{aligned}$$

  Then we set $1 - \frac{c_2^2}{2} - \frac{c_2^3}{3} \geq \frac{1}{2}$, and get $0 \leq c_2 \leq 0.8$. □

3.2 Poisson Variables

Problem
Let X be a Poisson random variable with mean μ, representing the number of errors on a page of this book. Each error is independently a grammatical error with probability p and a spelling error with probability $1 - p$. If Y and Z are random variables representing the numbers of grammatical and spelling

errors (respectively) on a page of this book, Prove that Y and Z are Poisson random variables with means $p\mu$ and $(1-p)\mu$, respectively. Also, prove that Y and Z are independent.

Solution

- Y and Z are Poisson random variables

  Proof. Since $X \sim Poisson(\mu)$, So

  $$Pr(X = k) = e^{-\mu}\frac{\mu^k}{k!}$$

  And

  $$\begin{aligned} Pr(Y = k) &= \sum_{i=k}^{\infty} Pr(X = i)\binom{i}{k}p^k(1-p)^{i-k} \\ &= \frac{e^{-\mu}p^k\mu^k}{k!}\sum_{i=k}^{\infty}\frac{\mu^{i-k}(1-p)^{i-k}}{(i-k)!} \\ &= \frac{e^{-\mu}p^k\mu^k e^{\mu(1-p)}}{k!} \\ &= e^{-\mu p}\frac{(\mu p)^k}{k!} \end{aligned}$$

  So $Y \sim Poisson(p\mu)$.The same goes for Z. □

- Y and Z are independent

  Proof.

  $$\begin{aligned} Pr(Y = k_1, Z = k_2) &= Pr(X = k_1 + k_2)\binom{k_1+k_2}{k_1}p^{k_1}(1-p)^{k_2} \\ &= e^{-\mu}\frac{\mu^{k_1+k_2}}{(k_1+k_2)!}\frac{(k_1+k_2)!}{k_1!k_2!}p^{k_1}(1-p)^{k_2} \\ &= e^{-p\mu}\frac{(p\mu)^(k_1)}{k_1!} \cdot e^{-(1-p)\mu}\frac{((1-p)\mu)^{k_2}}{k_2!} \\ &= Pr(Y = k_1) \cdot Pr(Z = k_2) \end{aligned}$$

  So Y and Z are independent. □

3.3 Birthday

Problem
There are n students in a classroom. Assume that their birthdays are uniformly randomly distributed and that every year has 365 days. Calculate the probability that there are two students having the same birthday and the probability that randomly choosing a student, there exists another student having the same birthday with him/her.

Solution

- The first question "two people share the same birthday" is a bit **ambiguous** because we really don't know the purpose is to calculate there are **EXACTLY** two people sharing the same birthday or there are **AT LEAST** two people sharing the same birthday.So we calculate them both:
  AT LEAST two people:
  Consider the event of no people share the same birthday as A,it's easy to know that

  $$Pr(A) = \frac{A(365, n)}{365^n}$$

So the probability that at least two people share the same birthday is

$$Pr\left(\overline{A}\right) = 1 - Pr(A) = 1 - \frac{A(365, n)}{365^n}.$$

EXACTLY two people:

$$Pr(EXACTLY) = \frac{\binom{n}{2} A(364, n-2)}{365^n}.$$

- Let B denote that randomly choosing a student b,there is no other students having the same birthday as him/her,so

$$Pr(B) = \binom{n}{1} \frac{1}{365} \left(\frac{364}{365}\right)^{n-1}$$

So the probability that there is at least another student sharing the same birthday as b is

$$Pr\left(\overline{B}\right) = 1 - Pr(B) = 1 - \binom{n}{1} \frac{1}{365} \left(\frac{364}{365}\right)^{n-1}.$$

3.4 Poisson and Chernoff Bound

Problem
Prove Chernoff-like boounds for Poisson random variable X_μ with expectation μ:
(1)If $x > \mu$,then $Pr(X_\mu \leq x) \leq \frac{e^{-\mu}(e\mu)^x}{x^x}$.
(2)If $x < \mu$,then $Pr(X_\mu \leq x) \leq \frac{e^{-\mu}(e\mu)^x}{x^x}$.

Solution

Proof. For any $t > 0$ and $x > \mu$,we have

$$Pr(X \geq x) = Pr(e^{tX_\mu} \geq e^{tx}) \leq \frac{E[e^{tX_\mu}]}{e^{tx}}$$

Considering the generating function of Poisson distribution, we have

$$Pr(X_\mu \geq x) \leq e^{\mu(e^t - 1) - xt}.$$

Choosing $t = \ln(x/\mu) > 0$ gives

$$\begin{aligned} Pr(X_\mu \geq x) &\leq e^{x - \mu - x\ln(x/\mu)} \\ &= \frac{e^{-\mu}(e\mu)^x}{x^x}. \end{aligned}$$

For any $t < 0$ and $x < \mu$,

$$Pr(X_\mu \leq x) = Pr(e^{tX_\mu} \geq e^{tx}) \leq \frac{E[e^{tX_\mu}]}{e^{tx}}$$

Hence

$$Pr(X_\mu \leq x) \leq e^{\mu(e^t - 1) - xt}.$$

Choosing $t = \ln(x/\mu) < 0$,it follows that

$$\begin{aligned} Pr(X_\mu \leq x) &\leq e^{x - \mu - x\ln(x/\mu)} \\ &= \frac{e^{-\mu}(e\mu)^x}{x^x}. \end{aligned}$$

Such that the proof ends. □

3.5 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1 s_2 \cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

Chapter 4 Bins and Balls - Poisson Approximation

4.1 Poisson Distribution

Problem

Assume that for each n,there are Bernoulli random variables $B_1^n, \cdots, B_n^n$ with indicators X_i^n.Let $Y_n = \sum_{i=1}^n$,suppose that
(a) $\lim\limits_{n\to\infty} E[Y_n] = \lambda$,and
(b)For any k, $\lim\limits_{1\le i_1<\cdots<i_k\le n} Pr(\bigcap_{r=1}^k B_{i_r}^n) = \frac{\lambda^k}{k!}$.
Prove that $\lim\limits_{n\to\infty} Pois(\lambda)$,i.e. $\lim\limits_{n\to\infty} Pr(Y_n = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ for any k.

Solution

Proof. Since $Y_n = \sum_{i=1}^n X_i^n$,so Y_n satisfies Binomial distribution.Suppose $Y_n \sim Bin(n,p)$.
Since $E[Y_n] = \sum_{i=1}^n Y_i \cdot p = \lambda$,so $p = \frac{\lambda}{n}$. Therefore

$$\begin{aligned}
\lim_{n\to\infty} Pr(Y_n = k) &= \lim_{n\to\infty} \binom{n}{k} p^k (1-p)^{n-k} \\
&= \lim_{n\to\infty} \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \lim_{n\to\infty} \underbrace{\frac{n!}{n^k(n-k)!}}_{\to 1} \frac{\lambda^k}{k!} \underbrace{\left(1-\frac{\lambda}{n}\right)^{n-k}}_{\to e^{-\lambda}} \\
&= \frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}$$

Such that the proof ends.
I also noticed that this problem is the case of weak independence,the following proof goes when $Y_n > 0$,but as for $Y_n = k$,I basically have no idea how to figure it out.I tried Inclusion-exclusion principle and Taylor's expansion,but it turned out not working.
Since $\lim\limits_{n\to\infty} E[Y_n] = \lim\limits_{n\to\infty} E\left[\sum_{i=1}^n X_i^n\right] = \lim\limits_{n\to\infty} \sum_{i=1}^n E[X_i^n] = \lambda$,and X_i^n satisfies Bernoulli distribution,so $\lim\limits_{n\to\infty} Pr(X_i^n) = \lambda$.
According to $inclusion - exclusion$ principle

$$\begin{aligned}
Pr(Y_n = k) &= Pr\left(\sum_{i=1}^n X_i^n = k\right) \\
&= \sum_{k=1}^n (-1)^{k-1} \sum_{1\le i_1\le i_2\le\cdots\le i_k\le n} Pr\left(\bigcap_{r=1}^k X_{i_r}^n\right) \\
&= Pr\left(\sum_{i=1}^n X_i^n\right) - \sum_{1\le i_j\le i_k\le n} Pr(X_i^n \cap X_j^n) + \cdots + (-1)^{n-1} \sum_{1\le i_1\le i_2\le\cdots\le i_k\le n} Pr\left(\bigcap_{r=1}^n X_{i_r}^n\right) \\
&= \lambda - \sum_{k=1}^n (-1)^{k-1}\frac{\lambda^k}{k!}
\end{aligned}$$

Besides,according to $Taylor's$ expansion,we have

$$\lim_{n\to\infty}\left(\lambda - \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} - \cdots + (-1)^{n-1}\frac{\lambda^{n-1}}{(n-1)!}\right) = e^{-\lambda}$$

Such that we can simply get $\lim_{n\to\infty} Pr(Y_n = k) = \frac{e^{-\lambda}\lambda^k}{k!}$. □

4.2 Poisson Distribution 2

Problem
Dene the Poisson experiment: n bins, each is filled with independent $Poisson(m/n)$ balls. Let E be the event that no bins are empty, and X be the total number of balls filled with. Prove that $Pr(E|X = k)$ increases with k.

Solution

Proof. If we set variable Y as the number of empty bins, so we can get that

$$Pr(E|X = k) = Pr(Y = 0|X = k) = 1 - \sum_{i=1}^{n-1} Pr(Y = i|X = k)$$

Because $\{Y = i|X = k + 1\}$ can be derived by two situations:

- By $\{Y = i|X = k\}$ and the next ball fall into the $n - k$ bins who are not empty.
- By $\{Y = i + 1|X = k\}$ and the next ball fall into any one bin of the $i + 1$ bins who are empty.

So we can get that

$$Pr(Y = i|X = k + 1) = Pr(Y = i|X = k)\frac{n - i}{n} + Pr(Y = i|X = k)\frac{i + 1}{n}.$$

Therefore

$$\begin{aligned}
Pr(E|X = k + 1) &= Pr(Y = 0|X = k + 1) \\
&= 1 - \sum_{i=1}^{n-1} Pr(Y = i|X = k + 1) \\
&= 1 - \sum_{i=1}^{n-1} Pr(Y = i|X = k)\frac{n - i}{n} - \sum_{i=1}^{n-i} Pr(Y = i + 1|X = k)\frac{i + 1}{n} \\
&= 1 - \sum_{i=1}^{n-1} Pr(Y = i|X = k)\frac{n - i}{n} - \sum_{i=2}^{n} Pr(Y = i|X = k)\frac{i}{n} \\
&= 1 - \frac{n - 1}{n}Pr(Y = 1|X = k) - \sum_{i=2}^{n-1} Pr(Y = i|X = k)\frac{n - i}{n} - \sum_{i=2}^{n} Pr(Y = i|X = k)\frac{i}{n} \\
&= 1 - \frac{n - 1}{n}Pr(Y = 1|X = k) - \sum_{i=2}^{n-1} \left(\frac{n - i}{n} + \frac{i}{n}\right) Pr(Y = i|X = k) \\
&= 1 - \sum_{i=2}^{n-1} Pr(Y = i|X = k) - Pr(Y = 1|X = k) + \frac{1}{n}Pr(Y = 1|X = k) \\
&= 1 - \sum_{i=1}^{n-1} Pr(Y = i|X = k) + \frac{1}{n}Pr(Y = 1|X = k) \\
&= Pr(E|X = k) + \frac{1}{n}Pr(Y = 1|X = k) \\
&> Pr(E|X = k).
\end{aligned}$$

Such that the proof ends. □

Proof using Poisson approximation

Proof. The number of balls in the bins satisfies Poisson distribution,so

$$Pr(k = 0) = \frac{e^{-\lambda}\lambda^k}{k!} = e^{-\frac{m}{n}}$$

So the probability of no bins are empty is

$$Pr(\epsilon) = \left(1 - e^{-\frac{m}{n}}\right)^n = e^{-ne^{-\frac{m}{n}}}$$

According to Poisson Approximation

$$Pr(\epsilon|X = k) = e^{-ne^{-\frac{k}{n}}}$$

It's clear this function goes monotonically with k. $\square$

4.3 Poisson Distribution 3

Problem
Follow the Poisson experiment and the notation in the last question. Assume that $m = n \ln n + cn$. Rigorously prove that $Pr(\epsilon|X = m + \sqrt{2m \ln m}) - Pr(\epsilon|X = m - \sqrt{2m \ln m}) \to 0$ as $n \to \infty$.

Solution

Proof. From the book we know that

$$Pr(\epsilon|X = m + \sqrt{2m \ln m}) - Pr(\epsilon|X = m - \sqrt{2m \ln m}) \geq |Pr(\epsilon||X - m| \leq \sqrt{2m \ln m}) - Pr(\epsilon|X = m)|$$

And because

$$|Pr(\epsilon||X - m| \leq \sqrt{2m \ln m}) - Pr(\epsilon|X = m)| = o(1)$$

So the limit is $o(1)$. $\square$

4.4 Agent and Resource

Problem
The following problem models a simple distributed system wherein agents contend for resources but back off in the face of contention. Balls represent agents, and bins represent resources.
The system evolves over rounds. Every round, balls are thrown independently and uniformly at random into n bins. Any ball that lands in a bin by itself is served and removed from consideration. The remaining balls are thrown again in the next round. We begin with n balls in the first round, and we will finish when every ball is served.

- If there are b balls at the start of a round, what is the expected number of balls at the start of the next round?
- Suppose that every round the number of balls served was exactly the expected number of balls to be served. Show that all the balls would be served in $O(lnlnn)$ rounds. (Hint: If x_j is the expected number of balls left after j rounds, show and use that $x_{j+1} \leq x_j^2/n$.)

Solution

- Consider the probability that a bin has 1 ball,denote it as X_i,then the probability satisfies *binomial* distribution,so

  $$Pr(X_i) = \binom{b}{1} \frac{1}{n} \left(1 - \frac{1}{n}\right)^{b-1}$$

  Such that for n bins,the expected number of balls of n bins after 1 round is

  $$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i) = b\left(1 - \frac{1}{n}\right)^{b-1}$$

  Therefore the expected number of balls left is $b - b\left(1 - \frac{1}{n}\right)^{b-1}$.

- *Proof.* From above we have

$$\begin{aligned} x_{j+1} &= x_j - x_j\left(1-\frac{1}{n}\right)^{x_j-1} \\ &\le x_j - x_j\left(1-\frac{x_j-1}{n}\right) \\ &= \frac{x_j^2}{n} \end{aligned}$$

So

$$\begin{aligned} x_1 &\le \frac{x_0^2}{n} \\ x_2 \le \frac{x_1^2}{n} &\le \frac{x_0^4}{n^3} \\ &\cdots \\ x_j &\le \frac{x_0^{2^j}}{n^{2^j-1}} \end{aligned}$$

Note that $x_0 = n$,we simply let $x_{j-1} = 1$,which means after j rounds,there will be no balls left. Since $x_j \le \frac{x_0^{2^j}}{n^{2^j-1}}$,so

$$\begin{aligned} 2^j &\le 2j\ln x_0 - \ln x_j + 1 \\ j &\le \ln(2j\ln x_0 - \ln x_j + 1) \\ j &\le \ln\ln x_0 + \ln 2j \end{aligned}$$

Since $x_{j-1} = 1, x_0 = n$,we can simply get $j \le \ln\ln n$,which means $j = O(\ln\ln n)$. □

4.5 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1s_2\cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

Chapter 5 Bins and Balls - Applications

5.1 Balls Expectation

Problem
Let X_i be the number of balls in bin i when m balls are independently and uniformly thrown at random into n bins,and $Y_i^{(m)}$,$1 \le i \le n$,are independent Poisson random variables each having expectation m/n.Assume that f is a nonnegative function.
(1)Prove that if $E[f(X_1^{(m)},\cdots,X_n^{(m)})]$ is monotonically increasing in m,then $E[f(X_1^{(m)},\cdots,X_n^{(m)})] \le 2E[f(Y_1^{(m)},\cdots,Y_n^{(m)})]$.
(2)if $E[f(X_1^{(m)},\cdots,X_n^{(m)})]$ is monotonically decreasing in m,then $E[f(X_1^{(m)},\cdots,X_n^{(m)})] \le 2E[f(Y_1^{(m)},\cdots,Y_n^{(m)})]$.

Solution

Proof. Firstly we'll prove 2 Lemmas:
Lemma 1:If Z is a Poisson variable of mean μ, where $\mu \ge 1$ is an integer,then $Pr(Z = \mu+h) \ge Pr(Z = \mu - h - 1)$.

Proof.

$$\begin{aligned}
\frac{Pr(Z=\mu+h)}{Pr(Z=\mu-h-1)} &= \frac{e^{-\mu}\mu^{\mu+h}}{(\mu+h)!} \Big/ \frac{e^{-\mu}\mu^{\mu-h-1}}{(\mu-h-1)!} \\
&= \frac{\mu^{2h+1}}{(\mu-h)(\mu-h-1)\cdots(\mu+h)} \\
&= \frac{\mu^2}{\mu^2-h^2}\cdot\frac{\mu^2}{\mu^2-(h-1)^2}\cdots\frac{\mu^2}{\mu^2-1^2}\cdot\frac{\mu}{\mu} \\
&\ge 1.
\end{aligned}$$

Such that the proof ends. □

Lemma 2: $Pr(Z \ge \mu) \ge 1/2$.

Proof.

$$\begin{aligned}
1 &= \sum_{k=0}^{\infty} Pr(Z=k) \\
&= \sum_{k=0}^{\mu-1} Pr(Z=k) + Pr(Z \ge \mu) \\
&\le \sum_{k=\mu}^{2\mu-1} Pr(Z=k) + Pr(Z \ge \mu) \\
&\le \sum_{k=\mu}^{\infty} Pr(Z=k) + Pr(Z \ge \mu) \\
&= Pr(Z \ge \mu) + Pr(Z \ge \mu) \\
&= 2Pr(Z \ge \mu)
\end{aligned}$$

Such that $Pr(Z \ge \mu) \ge 1/2$.

□

Then comes the problem proof.

$$\begin{aligned}
E\left[f({Y_1}^{(m)},\cdots,{Y_n}^{(m)})\right] &= \sum_{k\geq 0} E\left[f({Y_1}^{(m)},\cdots,{Y_n}^{(m)})|\sum {Y_i}^{(m)}=k\right] Pr\left(\sum {Y_1}^{(m)}=k\right)\\
&\geq \sum_{k\geq m} E\left[f({Y_1}^{(m)},\cdots,{Y_n}^{(m)})|\sum {Y_i}^{(m)}=k\right] Pr\left(\sum {Y_1}^{(m)}=k\right)\\
&= E\left[f({X_1}^{(k)},\cdots,{X_n}^{(k)})\right] Pr\left(\sum {Y_i}^{(m)}=k\right)\\
&\geq E\left[f({X_1}^{(m)},\cdots,{X_n}^{(m)})\right] Pr\left(\sum {Y_i}^{(m)}=k\right)\\
&= E\left[f({X_1}^{(k)},\cdots,{X_n}^{(k)})\right] Pr\left(\sum {Y_i}^{(m)}\geq m\right)
\end{aligned}$$

Let $Z=\sum {Y_i}^{(m)}$.Since each ${Y_i}^{(m)}$ is a Poisson random variable, their sum Z is also a Poisson random variable.Further,the mean value of Z is m.Thus,by results from above,$Pr(Z\geq m)\geq 1/2$.Combining it with above,we have

$$E\left[f({X_1}^{(m)},\cdots,{X_n}^{(m)})\right]\leq 2E\left[f({Y_1}^{(m)},\cdots,{Y_n}^{(m)})\right]$$

□

5.2 Bloom Filter

Problem

Bloom filters can be used to estimate set differences. Suppose Alice has a set X and Bob has a set Y , both with n elements. For example, the sets might represent their 100 favorite songs. Alice and Bob create Bloom filters of their sets respectively, using the same number of bits m and the same k hash functions. Determine the expected number of bits where our Bloom filters differ as a function of m,n,k and $|X\bigcap Y|$. Explain how this could be used as a tool to find people with the same taste in music more easily than comparing lists of songs directly.

Solution

Proof. Let Z be a random variable denoting the number of bits where the Bloom filters differ.Let Z_i be an indicator such that

$$\begin{aligned}
Z_i &= 1 \quad \text{if the } i\text{th bit of the Bloom filters differ}\\
Z_i &= 0 \quad \text{otherwise}
\end{aligned}$$

Thus,$Z=Z_1+Z_2+\cdots+Z_m$.
When $|X\cap Y|=r, Z_i=1$ only happens when each of the r common elements are not mapped to the ith bit,together with exactly one of the following cases(that causes the ith bit different):
(a) Some elements of $X-(X\cap Y)$ is mapped to the ith bit,but all elements of $Y-(X\cap Y)$ are not;
(b) Some elements of $Y-(X\cap Y)$ is mapped to the ith bit,but all elements of $X-(X\cap Y)$ are not.
Let Q_i denote the event that the r common elements are not mapped to the ith bit.By assuming that the hash functions we choose will map elements independently and uniformly at random to one of the m bits,we have

$$\begin{aligned}
Pr(Z_i=1) &= Pr(Q_i\cap(Case(a)orCase(b))\\
&= Pr(Q_i)Pr(Case(a)orCase(b))\\
&= \left(1-\frac{1}{m}\right)^{rk}\cdot 2\cdot\left(1-\left(1-\frac{1}{m}\right)^{(n-r)k}\right)\cdot\left(\left(1-\frac{1}{m}\right)^{(n-r)k}\right)\\
&= 2\left(1-\frac{1}{m}\right)^{nk}\left(1-\left(1-\frac{1}{m}\right)^{(n-r)k}\right)
\end{aligned}$$

As Z_i is an indicator,$E[Z_i]=Pr(Z_i=1)$.Thus,

$$E[Z]=\sum_{i=0}^{m-1}E[Z_i]=m\cdot E[Z_i]=2m\left(1-\frac{1}{m}\right)^{nk}\left(1-\left(1-\frac{1}{m}\right)^{(n-r)k}\right)$$

□

5.3 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1 s_2 \cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

Chapter 6 Set 6-NULL

Chapter 7 The Method of Counting and Expectation

7.1 Turan Theorem

Problem

We mentioned a probabilistic proof of Turan theorem in the lecture notes. Recall the random process generating an independent set S. Let p be the probability that the independent set S has size at least $\frac{|V|}{D+1}$.Show that $p > \frac{1}{2D|V|^2}$.

Solution

Proof. For a vertex i,the probability it is picked is $Pr(X_i = 1) = \frac{1}{d_i+1}$.
The expectation of the size of the independent set S is

$$\begin{aligned} E(|S|) &= E\left(\sum_{i=1}^{|V|} X_i\right) \\ &= \sum_{i=1}^{|V|} E(X_i) \\ &= \sum_{i=1}^{|V|} \frac{1}{d_i+1} \end{aligned}$$

According to the *Harmonic mean* property,we know that

$$\frac{|V|}{\sum_{i=1}^{|V|} \frac{1}{d_i+1}} \le \frac{\sum_{i=1}^{|V|} d_i + 1}{|V|} \le D + \frac{1}{|V|} \le D+1$$

which means $E(|S|) \ge \frac{|V|}{D+1}$. Therefore

$$\begin{aligned} \frac{|V|}{D+1} \le E(|S|) &= \sum_{i=1}^{|V|} Pr(|S| = i)i \\ &= \sum_{i=1}^{\frac{|V|}{D+1}-1} Pr(|S| = i)i + \sum_{i=\frac{|V|}{D+1}}^{|V|} Pr(|S| = i)i \\ &\le \left(\frac{|V|}{D+1} - 1\right)(1-p) + |V| \cdot p \end{aligned}$$

The above equation derives

$$p \ge \frac{1}{1 + \frac{D|V|}{D+1}}$$

For a complete graph,$D = |V| - 1$;otherwise $D < |V| - 1$.So $D \le |V| - 1$.And $D = \frac{2E}{|V|}$,sufficing

$$p \ge \frac{1}{1 + \frac{D|V|}{D+1}} \ge \frac{1}{|V|} \ge \frac{1}{4E|V|} \ge \frac{1}{2D|V|^2}.$$

Such that the proof ends. □

7.2 Independent Set

Problem
Given an n-vertex undirected graph $G = (V, E)$, consider the following method of generating an independent set. Given a permutation σ of the vertices, define a subset $S(\sigma)$ of the vertices as follows: for each vertex $i, i \in S(\sigma)$ if and only if no neighbor j of i precedes i in the permutation σ.

- Show that each $S(\sigma)$ is an independent set in G.
- Suggest a natural randomized algorithm to produce σ for which you can show that the expected cardinality of $S(\sigma)$ is $\sum_{i=1}^{n} \frac{1}{d_i+1}$, where d_i denotes the degree of vertex i.

Solution

- According to the algorithm describe above, we denote the permutation σ as $\sigma = v_1, v_2, \ldots, v_i, \ldots, v_n$ where n is the number of vertex. And $S(\sigma) = d_1, d_2, \ldots, d_k$, where k is the size of $S(\sigma)$. Obviously, $d_1 = v_1 \in S(\sigma)$ as the first element. Let's look at the d_i which belongs to $S(\sigma)$. There's no connection with the d_j where $j < i$, because no neighbor of it precedes before. And any d_j where $j > i$ has no connection with d_i for the same reason. So the vertices in $S(\sigma)$ is an independent set in G.
- With the natural randomized algorithm to produce σ. The expectation of vertex v_i in the subset $S(\sigma)$ is $E(v_i) = \frac{1}{d_i+1}$, for the reason of that v_i need occur first within its d_i neighbors and itself, so the probability of that is $\frac{1}{d_i+1}$.
So consider all of n vertices $E(S(\sigma)) = \sum_{i=1}^{n} \frac{1}{d_i+1}$.

7.3 2-Coloring Edge

Problem
Prove that, for every integer n, there exists a way to 2-color the edges of K_x so that there is no monochromatic clique of size k when $x = n - \binom{n}{k} 2^{1-\binom{k}{2}}$. Note that K_x stands for the x-vertex complete graph. (Hint, start by 2-coloring the edges of K_n and fix things up.)

Solution

Proof. In K_n graph chose a k-cliques from $\binom{n}{k}$ k-cliques in K_n. Let X_i be an random variable, such that $X_i = 1$ if clique i is monochromatic and 0 otherwise. Further let $X = \sum X_i$.

$$E[X] = \sum E[X_i] = \binom{n}{k} 2^{-\binom{k}{2}+1}$$

Then we remove a vertex from each monochromatic k-clique. Then the residual graph is a complete graph without monochromatic k-cliques. The expectation of number of vertices in the residual graph is

$$n - E[X] = n - \binom{n}{k} 2^{-\binom{k}{2}+1} = x$$

So there is a 2-coloring of K_n where removing a vertex from each monochromatic k-clique results in a K_y, $y \geq x$ that has no monochromatic k-clique. If K_y does not have a monochromatic K_k, because $y \geq x$ K_x does not have too. $\square$

7.4 2-Coloring Edge Proof

Problem
Prove the following claims.

- For every integer n, there exists a coloring of the edges of the complete graph K_n by two colors so that the total number of monochromatic copies of K_4 is at most $\binom{n}{4} 2^{-5}$.
- Give a randomized algorithm for finding a coloring with at most $\binom{n}{4} 2^{-5}$ monochromatic copies of K_4 that runs in expected time polynomial in n.

Solution

- First we consider random 4 vertices in n-vertices graph. Once one of edges is colored, then the remain $\binom{4}{2} - 1 = 5$ edges have the probability $Pr(A_i) = 2^{-5}$ to color to the same color. Where A_i denote the event that clique i is monochromatic in $\binom{n}{4}$ cliques. Also we define that if clique i is monochromatic then random variable $A_i = 1$, otherwise $A_i = 0$. So $E(A_i) = 2^{-5}$.
  In order to calculate $E(\sum A_i)$ we yields:
  $$E(\sum A_i) = \binom{n}{4}2^{-5}$$
  Using the Lemma 6.2 we have $Pr(\sum A_i \leq \binom{n}{4}2^{-5}) > 0$ So there exist one 2-coloring that has at most $\binom{n}{4}2^{-5}$ K_4 are monochromatic.

- Color the edge independently and uniformly. Denote $X = \sum A_i$. Let $p = Pr(X \leq \binom{n}{4}2^{-5}$. Then we have
  $$\begin{aligned}\binom{n}{4}2^{-5} &= E[X] \\ &= \sum_{i \leq \binom{n}{4}2^{-5}} iPr(X = i) + \sum_{i \geq \binom{n}{4}2^{-5}+1} iPr(X = i) \\ &\geq p + (1-p)\binom{n}{4}2^{-5} + 1\end{aligned}$$
  So we have
  $$\frac{1}{p} \leq \binom{n}{4}2^{-5}$$
  Thus, the expected number of samples is at most $\binom{n}{4}2^{-5}$. Testing to see if $X \leq \binom{n}{4}2^{-5}$ can be done in $O(n^4)$ time. So the algorithm can be done in polynomial time.

7.5 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1 s_2 \cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

Chapter 8 Derandomization, Second Moment Method, Lovasz Local Lemma

8.1 Find Edge Coloring

Problem
For every integer n, there exists a coloring of the edges of the complete graph K_n by two colors so that the total number of monochromatic copies of K_4 is at most $\binom{n}{4}2^{-5}$. Design a deterministic, efficient algorithm to find such a coloring.

Solution
Assign values to variables deterministically, one at a time, in an arbitrary order $x_1, x_2, \ldots, x_n$. Suppose that we have assigned the first k edge. Let $y_1, y_2, \ldots, y_k$ be the corresponding assigned values. We compute the two quantities,

$$E[X|x_1 = y_1, x_2 = y_2, ..., x_k = y_k, x_{k+1} = color_1]$$
$$E[X|x_1 = y_1, x_2 = y_2, ..., x_k = y_k, x_{k+1} = color_2]$$

and then choose the setting with larger expectation.

8.2 Second Moment Method

Problem
Consider a graph in $G_{n,p}$ with $p = c\frac{\ln n}{n}$. Use the second moment method to prove that if $c < 1$ then, for any constant $\epsilon > 0$ and for n sufficiently large, the graph has isolated vertices with probability at least $1-\epsilon$.

Solution

Proof. We consider the event X_i denotes that the i^{th} vertex is isolated. So

$$X_i = \begin{cases} 1 & if\ v_i\ is\ isolated \\ 0 & otherwise \end{cases}$$

Let

$$X = \sum_{i=1}^{n}(1-p)^{n-1}.$$

so that

$$E[X] = n(1-p)^{n-1}$$

In order to prove that if $c < 1$ then, for any constant $\epsilon > 0$ and for n sufficiently large, the graph has no isolated vertex with probability at most ϵ. That means $Pr(X = 0) = o(1)$.
We wish to compute

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

Applying Lemma 6.9, we see that we need to consider the covariance of the X_i.

$$\begin{aligned} Cov[X_iX_j] &= E[X_iX_j] - E[X_i]E[X_j] \\ &= (1-p)^{2n-3} - (1-p)^{n-1} * (1-p)^{n-1} \\ &= p(1-p)^{2n-3} \end{aligned}$$

So

$$Var[X] \leq E[X] + \sum Cov[X_i X_j] = E[X] + o(pn^2(1-p)^{2n-3})$$

Then

$$\begin{aligned} Pr(X=0) &\leq \frac{Var[X]}{E[X]^2} \\ &= \frac{1}{n(1-p)^{n-1}} + \frac{p}{1-p} \end{aligned}$$

for $p = c\frac{\ln n}{n}$ and $c < 1$ with $n \to \infty$, $Pr(X = 0) \to o(1)$. So the graph has isolated vertices with probability at least $1 - \epsilon$. □

8.3 Asymmetric Lovasz Lemma

Problem

Prove the Asymmetric Lovasz Local Lemma: Let $\mathbb{A} = \{A_1, \ldots, A_n\}$ be a set of finite events over a probability space, and for each $1 \leq i \leq n$, $\tau(A_i) \in \mathbb{A}$ is such that A_i is mutually independent of all events not in $\tau(A_i)$. If $\sum_{A_j \in \tau(A_i)} Pr(A_j) \leq 1/4$ for all i, then $Pr(\bigwedge_{i=1}^n \bar{A}_i) \geq \prod_{i=1}^n (1 - 2Pr(A_i)) > 0$. [Hint: let $x(A_i) = 2Pr(A_i)$ and use the general Lovasz Local Lemma.]

Solution

Proof. First we need to prove a lemma that if $0 \leq a_i \leq 1/2$ for all $i = 1, 2, \ldots, n$, then $\prod_{i=1}^n (1 - 2a_i) \geq 1 - 2\sum_{i=1}^n a_i$.

Induction for n. When $n = 1$, the inequality holds obviously. Assume that when $n = k$, the inequality holds. Consider the case when $n = k + 1$,

$$\begin{aligned} \prod_{i=1}^{k+1}(1-2a_i) &= \prod_{i=1}^{k}(1-2a_i)(1-2a_{k+1}) \\ &\geq (1 - 2\sum_{i=1}^{k} a_i)(1 - 2a_{k+1}) \\ &= 1 - 2\sum_{i=1}^{k+1} a_i + 4\sum_{i=1}^{k} a_i a_{k+1} \\ &\geq 1 - 2\sum_{i=1}^{k+1} a_i \end{aligned}$$

So the inequality holds.

Using the general Lovasz Local Lemma, we set $x(A_i) = 2Pr(A_i)$. Then

$$\begin{aligned} x(A_i) \prod_{A_j \in \Gamma(A_i)} (1 - x(A_j)) &= 2Pr(A_i) \prod_{A_j \in \Gamma(A_i)} (1 - 2Pr(A_j)) \\ &\geq 2Pr(A_i)(1 - 2\sum_{A_j \in \Gamma(A_i)} Pr(A_j) \\ &\geq 2Pr(A_i)(1 - 2 * 1/4) \\ &= Pr(A_i) \end{aligned}$$

So the general Lovasz Local Lemma condition holds. Then we have the result

$$\begin{aligned} Pr(\bigwedge_{i=1}^{n} \bar{A}_i) &\geq \prod_{i=1}^{n}(1 - x(A_i)) \\ &= \prod_{i=1}^{n}(1 - 2Pr(A_i)) \\ &> 0. \end{aligned}$$

□

8.4 Vertex Coloring

Problem
Given $\beta > 0$, a vertex-coloring of a graph G is said to be β-frugal if (i) each pair of adjacent vertices has different colors, and (ii) no vertex has β neighbors that have the same color.
Prove that if G has maximum degree $\Delta \geq \beta^\beta$ with $\beta \geq 2$, then G has a β-frugal coloring with $16\Delta^{1+1/\beta}$ colors. [Hint: you may want to define two types of events corresponding to the two conditions of being β-frugal. Then the result in question 2 can be used.]

Solution

Proof. By the following equation
$$\binom{\Delta+1}{\beta} = \binom{\Delta}{\beta} + \binom{\Delta}{\beta-1}$$
we can prove that $\binom{\Delta}{\beta}$ is monotonically increasing for Δ when β is given.
Let the number of colors used to β-frugal coloring be $N = 16\Delta^{1+1/\beta}$, and the algorithm assigns each vertex a uniformly random color.
Now we define two types of events with total number of $m+n$, when n is the number of vertices, and m is the number of edges:

- The pair vertices of e_i has the same color;
- The vertex v_i has β neighbors that have the same color.

Define d_i is the degree of vertex i.
For each event A_i in type I,
$$Pr(A_i) = \frac{1}{N}$$
For each event A_i in type II,
$$\begin{aligned} Pr(A_i) &= \binom{d_i}{\beta}\left(\frac{1}{N}\right)^{\beta-1} \\ &\leq \binom{\Delta}{\beta}\left(\frac{1}{N}\right)^{\beta-1} \end{aligned}$$

Consider the number of dependent events of each event in type I. First, each edge connected to the two vertices in the given edge has an event in type I, whose total number is at most $2(\Delta?1)$. Second, each vertex of the edge has an event in type II, whose total number is exactly 2. Thus, for each event A_i in type I,
$$\begin{aligned} \sum_{A_j \in \Gamma(A_i)} Pr(A_j) &\leq 2(\Delta-1)\frac{1}{N} + 2\binom{\Delta}{\beta}\left(\frac{1}{N}\right)^{\beta-1} \\ &= 2\left[(\Delta-1)\frac{1}{16\Delta^{1+1/\beta}}\right] + \binom{\Delta}{\beta}\left(\frac{1}{16\Delta^{1+1/\beta}}\right)^{\beta-1} \\ &\leq 2\left[\frac{1}{16} + \Delta\dot(\Delta-1)\cdots(\Delta-\beta+1)\right] \end{aligned}$$

This definition for events is hard to prove. Another proof from Alistair Sinclair is in the last section.
□

8.5 Read Paper

Problem
Read the paper "A constructive proof of the general Lovasz Local Lemma".

Chapter 9 A Brief Introduction to Markov Chains

9.1 Graph and Matrix

Problem
Consider an arbitrary Markov chain with transition matrix P and transition diagram G. Recall that state i leads to state j if and only if multi-step transition probability $p(n) > 0$ for some $n > 0$. On the other hand, j can be reached from i in G if and only if there is a directed path from i to j. Please show that i leads to state j in the Markov chain if and only if j can be reached from i in G.

Solution

Proof. Since i can lead to j,there always exists a path from i to j,say $i \to k_1 \to k_2 \cdots k_{n-1} \to k_n \to j$,sufficing

$$P_{ik_1} > 0, P_{k_1 k_2} > 0, \cdots, P_{k_{n-1}k_{n-2}} > 0, P_{k_n j} > 0$$

For the path with length 2,there exists a state k such that $i \to k \to j$,which means $P_{ik} > 0, P_{kj} > 0$,sufficing $P_{ij}^{(2)} > P_{ik}P_{kj} > 0$.
For the path with length n,According to C-K equation

$$\begin{aligned} P_{ij}^{n} &= \sum_{k=0}^{\infty} P_{ik}^{m} P_{kj}^{n-m} \\ &\geq P_{ik_1} P_{k_1 k_2} \cdots P_{k_{n-1}k_n} P_{k_n j} \\ &> 0 \end{aligned}$$

This proof holds for all the paths with length n.Similarly for all the other paths longer than n we can use **induction** to make the conclusion hold. Such that the proof ends. □

9.2 Aperiodic

Problem
We say a Markov chain is aperiodic if and only if all states in the chain are aperiodic. Given a finite-state aperiodic Markov chain, assume that each pair of states communicates. Then prove that if n is large enough, $p_{ij}^{(n)} > 0$ for all states i, j.

Solution

Proof. For an aperiodic state i, $\exists r$, sufficing $P_{ii}^{(r)} > 0$. On the other hand,since each pair of states communicates,$\exists t$,sufficing $P_{ij}^{(t)} > 0$.
Thus if n is large enough,say $n \geq r + t$,it will suffice

$$P_{ij}^{(n)} > P_{ii}^{(r)} P_{ij}^{(t)} > 0$$

Such that the proof ends. □

9.3 Hitting Time

Problem
Given a Markov chain,let i be a state.Define $f_{ii}^{(k)}$ to be the probability that starting with state i at time

0,the chain first returns to state i at time k,and $p_{ii}^{(k)}$ be the k-step probability of returning to i from i.Prove that $p_{ii}^{(n)} = \sum_{i=1}^{n} f_{ii}^{(k)} p_{ii}^{(n-k)}$,where $p_{ii}^{(0)} = 1$.

Solution

Proof. We now proof the normal situation:

$$P_{ij}^{(n)} = \sum_{m=1}^{n} f_{ij}^{(n)} p_{jj}^{(n-m)}, n \geq 1.$$

From Elementary Probability,

$$\begin{aligned} p_{ij}^{(n)} = P_i(X_n = j) &= \sum_{m=1}^{n} P_i(T_j = m, X_n = j) \\ &= \sum_{m=1}^{n} P_i(T_j = m) P_i(X_n = j | T_j = m) \end{aligned}$$

The first term equals $f_{ij}^{(m)}$,by definition.For the second term,we use the Strong Markov Property:

$$P_i(X_n = j | T_j = m) = P_i(X_n = j | X_m = j) = P_{jj}^{(n-m)}.$$

Let $j = i$ such that the proof ends. □

9.4 Finite and Recurrent

Problem
Given a finite markov chain, where finiteness means that there are a finite number of states, prove that

1. At least one state is recurrent.
2. All recurrent states are positive recurrent.

Solution

1. Since there are a finite number of communicating classes, and since once the chain leaves a communicating class it cannot return, it must eventually settle into one communicating class. Thus, at least one state in this class is visited an unbounded number of times after the chain enters it.

2. Let $C(i)$ denote the communicating class containing state i. Let p be the largest transition probability less than 1 from any states in $C(i)$.

$$h_{i,i} = \sum_t t r_{i,i}^t \leq \sum_t t p^t$$

Since there are finite state so $r_{i,i}^t = 0, t > n$. Therefore

$$h_{i,i} \leq \sum_{t=1}^{n} t p^t \leq \frac{n}{1-p} < \infty$$

Then i is positive recurrent.

9.5 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1 s_2 \cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

Chapter 10 Excursions and Stationary Distributions of Markov Chains

10.1 Strong Markov Property

Problem
In the proof of strong Markov property, we claim that there is an event $\mathbb{A}$ such that
(1)$\mathbb{A}$ is determined by $X_1, X_2, \cdots, X_{n-1}$
(2)$(X_{n-s} = i_s, 0 \leq s \leq n, \tau = n) = (X_n = i_0, \mathbb{A}(X_{n-1}, \cdots, X_0))$.
Prove this claim. (hint: use the definition of stopping time)

Solution

Proof. Let T be a stopping time and $\mathbb{A}$ is determined by the past $(X_0, X_1, ..., X_n)$ if $T < \infty$. Since $(X_0, ..., X_T)\mathbf{I}(T < \infty) = \sum_{n \in \mathbb{Z}+}(X_0, ..., X_n)\mathbf{I}(T = n)$,any such $\mathbb{A}$ must have the property that,for all $n \in \mathbb{Z}+, \mathbb{A} \cap T = n$ is determined by $(X_0, ..., X_n)$.We are going to show that for any such $\mathbb{A}$,

$P(X_{T+1} = j_1, X_{T+2} = j_2, ...; \mathbb{A}|X_T, T < \infty) = P(X_{T+1} = i_1, X_{T+2} = i_2, ...|X_T = i, T < \infty)P(\mathbb{A}|X_T, T < \infty)$

and that

$$P(X_{T+1} = j_1, X_{T+2} = j_2, ...|X_T = i, T < \infty) = p_{i,j_1} p_{j_1,j_2} \cdots$$

We have:

$$P(X_{T+1} = j_1, X_{T+2} = j_2, ...; \mathbb{A}, X_T = i, T = n) = P(X_{n+1} = j_1, X_{n+2} = j_2, ...; \mathbb{A}, X_n = i, T = n).$$

Such that the proof ends. $\square$

10.2 Expectation Markov

Problem
Consider an irreducible Markov chain $X_n, n \geq 0$ with state space S and a positive recurrent state a. It has a unique stationary distribution π over S. For any function $f : S \to R$,prove that $\frac{E[\sum_{n=0}^{T_{aa}-1} f(X_n)]}{E[T_{aa}]}$ equals $E_\pi[f]$,the expectation of f with respect to π.

Solution

Proof. Since a is positive recurrent, $\lim_{t\to\infty} N_t = \infty$.By law of large number,$\lim_{t\to\infty} \frac{1}{N_t} \sum_{r=1}^{mN_{t-1}} G_r = E\left[\sum_{n=0}^{T_{aa}-1} f(X_n)\right]$.
So:

$$\frac{T_a^{N_t}}{N_t} \leq \frac{t}{N_t} \leq \frac{T_a^{N_t+1}}{N_t}$$

$$T_a^r = T_a^1 + \sum_{k=2}^{r}(T_a^k - T_a^{k-1}) = T_a^1 + \sum_{k=1}^{r-1} T_{aa}^k$$

$$P(T_a < \infty) = 1, so \lim_{t\to\infty} \frac{T_a^1}{n_t} = 0$$

By law of large number

$$\lim_{t\to\infty} \frac{T_{N_t}^a}{N_t} = \lim_{t\to\infty} \frac{T^{N_t+1}}{N_t} = ET_{aa}.$$

So

$$\lim_{t\to\infty} \frac{t}{N_t} = E[T_{aa}]$$

Such that $\lim_{t\to\infty} \frac{f(X_0+...+f(X_t))}{t} = \frac{E\left[\sum_{n=0}^{T_{aa}-1} f(X_n)\right]}{E[T_{aa}]}$. □

10.3 Excursion and Recurrent

Problem

Use the concept of an excursion to prove that for a Markov chain, a state j is recurrent if i leads to j and i is recurrent.

Solution

Proof. **Proof without excursion**

Suppose that state i is recurrent, so $\sum_n p_{ii}^{(n)} = +\infty$. As we know state j and i belong to the same communicating class. So there exist k_1 and k_2 that $p_{ij}^{(k_1)} > 0$ and $p_{ji}^{(k_2)} > 0$.

For all $n > 0$ that

$$p_{jj}^{(k_1+k_2+n)} \geq p_{ji}^{(k_1)} p_{ii}^{(n)} p_{ij}^{(k_2)}$$

so we have

$$\begin{aligned}
\sum_n p_{jj}^{(n)} &\geq \sum_n p_{jj}^{(k_1+k_2+n)} \\
&\geq \sum_n p_{ji}^{(k_1)} p_{ii}^{(n)} p_{ij}^{(k_2)} \\
&= p_{ji}^{(k_1)} p_{ij}^{(k_2)} \sum_n p_{ii}^{(n)} \\
&\geq +\infty
\end{aligned}$$

So j is recurrent.

Proof using excursion

Start with $X_0 = i$.If i is recurrent and $i \rightsquigarrow j$ then there is a positive probability ($= f_{ij}$) that j will appear in one of the i.i.d. excursions $\mathscr{X}_i^{(0)}, \mathscr{X}_i^{(1)}, ...$, and so the probability that j will appear in a specific excursion is positive. So the random variables

$$\delta_{j,r} := \mathbf{1}(j \text{ appears in excursion } \mathscr{X}_i^{(r)}), r = 0, 1, ...$$

are i.i.d. and since they take the value 1 with positive probability,infinitely many of them will be 1(with probability 1),showing that j will appear in infinitely many of the excursions for sure. Hence,not only $f_{ij} > 0$,but also $f_{ij} = 1$. Hence,$j \rightsquigarrow i$. The last statement is simply an expression in symbols of what we said above. Indeed,the probability that j will appear in a specific excursion equals $g_{ij} = P_i(T_j < T_i)$. □

10.4 Two Chains Stability

Problem

In the proof of the stability theorem, assume that the two independent Markov chains $X_n, n \geq 0$ and $Y_n, n \geq 0$ encounter at time T. Prove that $Pr(T < \infty) = 1$.

Solution

Proof. Consider the process

$$W_n := (X_n, Y_n).$$

Then $(W_n, n \geq 0)$ is a Markov chain with state process $S \times S$.Its initial state $W_0 = (X_0, Y_0)$ has distribution $P(W_0 = (x, y)) = \mu(x)\mu(y)$.Its (1-step) transition probabilities are

$$q_{(x,x'),(y,y')} := P(W_{n+1} = (x', y')|W_n = (x, y)) = P(X_{n+1} = x'|X_n = x)P(Y_{n+1} = y'|Y_n = y) = p_{x,x'}p_{y,y'}$$

Its n-step transition probabilities are

$$q^{(n)}_{(x,x'),(y,y')} = p^{(n)}_{x,x'}, p^{(n)}_{y,y'}.$$

From aperiodicity assumption we have that $p^{(n)}_{x,x'} > 0$ and $p^{(n)}_{y,y'}$ for all large n,implying that $q^{(n)}_{(x,x'),(y,y')}$ for all large n,and so $(W_n, n \geq 0)$ is an irreducible chain.Notice that

$$\sigma(x,y) := \pi(x)\pi(y), (x,y) \in S \times S,$$

is a stationary distribution for $W_n, n \geq 0$.By positive recurrence,$\pi(x) > 0$ for all $x \in S$.Therefore $\sigma(x,y) > 0$ for all $(x,y) \in S \times S$.Hence $W_n, n \geq 0$ is positive recurrent.In particular,

$$P(T < \infty) = 1.$$

Such that the proof ends. □

10.5 Coin

Problem
Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin.Write down the results in a string $s_1 s_2 \cdots s_{20}$, where s_i is 1 if the i-th trial gets Head,and otherwise is 0.
Solution
The result of tossing a coin 20 times is:
11101000111010101000.

TUPLE $=< T_s, T_e, Loc >$, define set **TIME** $= \{t | T_s \leq t \leq T_e\}, \forall T1, T2 \in$ **TUPLE**, $T1 \cap T2 = \emptyset$ and $T1.$**TIME** $\cap\, T2.$**TIME** $= \emptyset$