formula

21 October 2021 20:33

Lemma 3: $P(\bigcup_{i=1}^{n} A_i) = \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|-1} \cdot P(\bigcap_{i \in S} A_i)$

Pf.

Derangement: #(none of elements being correct pos) $= n! \cdot [\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \cdots (-1)^n \frac{1}{n!}]$

P(K classes correctly) $= 1 - \frac{1}{1}$ (correct $-$ 2 corr $\cdots$ k corr)

(only) $\approx 1 - \frac{1}{1} + \frac{1}{2!} - \frac{1}{3!} \cdots \frac{1}{k!} \approx e^{-1}$

▷ $P(A|B) = \frac{P(A \cap B)}{P(A \cap B) + P(\bar{A} \cap B)}$

▷ $P(A \cap B) = P(B) \cdot P(A|B)$

▷ $= P(A) \cdot P(B|A)$.

$\Omega = \bigcup_{i=1}^{n} B_i$, where B_i's are mutually disjoint & cover Ω.

$P(A) = \sum_{i=1}^{n} P(B_i) \cdot P(A|B_i)$ $= \sum_i P(A \cap B_i)$

A,B independent $\iff P(B|A) = P(B) \iff P(A \cap B) = P(A) \cdot P(B)$.

$A \cap B = \phi \Rightarrow A, B$ dependent

: · Collection Σ of subsets of Ω is called Sigma-algebra if:

(i) $\Omega \in \Sigma$

(ii) Σ closed under complement $[A \in \Sigma \Rightarrow A^c \in \Sigma]$

(iii) " " " countable union.

Theorem (Bayes): $P(B|A) = P(B) \cdot P(A|B) / P(A)$

$= \frac{P(B) \cdot P(A|B)}{P(B) \cdot P(A|B) + P(B^c) \cdot P(A|B^c)}$

$P(A) = P(A \cap B) + P(A \cap B^c) = P(B) \cdot P(A|B) + P(B^c) P(A|B^c)$

Expected value of X is $E[X] := \sum_{x \in range(X)} P(X=x) \cdot x$

Usually range(X) is a finite

Conditional expectation $E[X|B] := \sum_{w \in B} P(w|B) \cdot X(w)$

$= \frac{1}{P(B)} \cdot \sum_{w \in B} P(w) \cdot X(w)$. $P(w|B) = \frac{P(w)}{P(B)}$ here

$E[X] = \sum_{i \in [k]} E[X|B_i] \cdot P(B_i)$

$E[\alpha \cdot X] = \alpha \cdot E[X]$; $\forall \alpha \in \mathbb{R}$.

rem: $E[X+Y] = E[X] + E[Y]$.

$E[\sum \alpha_i \cdot X_i] = \sum_i \alpha_i \cdot E[X_i]$

linear combination $(\alpha_i \in \mathbb{R}, \forall i)$

Bernoulli random variable.

· Toss a coin with $P(H) =: p$

Equality Checking Protocol

Protocol: 1) Turn A into number $N_A := \sum_i a_i \cdot 2^i$.

2) Pick a random prime $p \in [t]$. a_i: i-th bit

at most $\log t$ bits

3) Compute residue $R_A := N_A \mod p$ $[\Rightarrow \lg t$ -bits$]$

4) Send (R_A, p) to Bob. $[2 \lg t$ -bits$]$

5) Bob checks $R_A \stackrel{?}{=} R_B$. [Output YES iff $R_A = R_B$]

$P(R_A = R_B \mid A \neq B) < 1/n$

Poisson random variable with parameter α,

X takes values $\{0,1,2,\ldots\} =: \mathbb{W}$, and

$P(X=k) = e^{-\alpha} \cdot \alpha^k / k!$; $\forall k \in \mathbb{W}$.

$E(X) = \alpha$

Markov inequality $P(X \geq a) \leq E[X]/a$

Variance of X is $var(X) := E[(X - E[X])^2]$

Standard-deviation of X is $\sigma(X) := \sqrt{var(X)}$.

$\forall a \in \mathbb{R}$, $var(a \cdot X) = a^2 \cdot var(X)$

$var(X)$ $E[X^2] - E[X]^2$.

Standard deviation of Bernoulli (with $P(H) =: p$)

$Var(X) = E[X^2] - E[X]^2 = p - p^2$.

$\sigma(X) = \sqrt{p(1-p)} \leq 1/2$

Chebyshev inequality

random variable X & $a > 0$,

$P(|X - E[X]| \geq a) \leq Var(X)/a^2 \leq (\sigma(X)/a)^2$.

$P(X \geq E[X] + 2\sigma$ OR $X \leq E[X] - 2\sigma) \leq 1/4$.

Weak Linearity of Variance $var(\sum_i X_i) = \sum_i var(X_i)$

$E[X_1 X_2] = E[X_1] \cdot E[X_2]$

Weak Law of large numbers $\bar{X} := (\sum_{i=1}^{n} X_i)/n$

$P(|\bar{X} - E[X]| \geq a) \leq var(X)/n a^2$

$E[\bar{X}] = \sum E[X_i]/n = E[X]$.

$var(\bar{X})$ $\frac{n \cdot var(X)}{n^2}$

Chernoff inequality $P(X=1) =: p$.

sum $S := \sum_{i=1}^{n} X_i$ & $\delta \in (0,1)$.

decay is exponential in n:= #repetitions of X.

$P(S < (1-\delta) \cdot E[S]) < (e^{-E[X] \cdot \delta^2/2})^n$

Setting:

$P(S < (1-\delta)\mu) < (e^{-\delta}/(1-\delta)^{1-\delta})^{\mu} \leq (e^{-\delta^2/2})^{\mu}$

(i,j)-th entry is $T_{ij} := P(X_1 = j \mid X_0 = i)$.

T is the transition matrix of a (homog.) Markov chain.

initial probability distribution

$\mu \in [0,1]^{|S|}$, with $\mu_i := P(X_0 = i)$

$|\mu| = \sum \mu_i = 1$

Each row (~~or column~~) of T sums to 1.

Such matrices are called stochastic. (not doubly-stochastic)

Many examples from other fields use Markov chain modeling (i.e. memorylessness!):

Let the process be given by column vector $\vec{\mu}$ & matrix M. S is its state-space.

(Evolution): $\forall n \geq 1$, $\vec{p}_n^T = \vec{p}_0^T \cdot M^n$

Theorem (Perron-Frobenius 1907): If M is the transition matrix of a regular Markov chain, then

$\lim_{n \to \infty} M^n =: \mathbf{1} \cdot w^T$, where $\mathbf{1}$:= column vector with all 1's & w is some prob. distribution.

↳ w is called the stationary distribution.

↳ For initial distribution μ: $\lim (\mu^T M^n) = \mu^T \cdot \mathbf{1} \cdot w^T = w^T$ is independent of μ! (Memorylessness?)

$(M_1 - m_1) = (M_0 - m_0) \cdot (1 - 2\delta) < M_0 - m_0$ $[\because \delta > 0]$

$\lim_{n \to \infty} v_n = \lim_{n \to \infty} M^n \cdot v_0 = c_{v_0} \cdot \mathbf{1}$

[Recall: $\mu^T \cdot \mathbf{1} \cdot w^T = w^T$.

Exercise 2: M is doubly-stochastic $\Rightarrow$ stationary distribution is uniform distribution!

▷ $M^n \xrightarrow{n \to \infty} \frac{1}{|S|} \cdot J$ [J is all 1 matrix.]

↳ Random walk in undirected graphs gives a doubly-stochastic process.

↳ visiting all vertices in the end!

(ii)' The web-surfer is allowed to "stray" to a random page with prob $=: p > 0$, or "follow" a link in the current-page.

Strategy 3: Thus, $M'_{ji} := \begin{cases} p \cdot \frac{1}{n} + (1-p) \cdot \frac{1}{m_j}, & \text{if } (j,i) \in E \\ p \cdot \frac{1}{n}, & \text{if } (j,i) \notin E. \end{cases}$

(follow) (stray)

Row sum of M' = 1

Defn: Hashing $\varphi_R: S \to T$ is called pairwise independent (p.i.) if:

(i) Rnd. variables $\{\varphi_R(s) \mid s \in S\}$ are p.i.

(ii) $\forall s \in S$, $\varphi_R(s)$ is uniformly distributed in T.

Exercise: (i) $\iff \forall s \neq s' \in S, \forall t, t' \in T$, $P(\varphi_R(s) = t \wedge \varphi_R(s') = t') = (1/|T|)^2$.

(ii) $\iff \forall s \in S, t \in T$, $P(\varphi_R(s) = t) = 1/|T|$.

Analyse: · Suppose E steals (X,Y) & instead sends (X', Y') on the channel to B.

· Let Σ be the event: $Y' = \varphi_R(X')$; in which case, B wrongly accepts msg $X' \neq X$. $= \frac{1}{|T|}$

K_n := complete, undirected graph on n vertices.

$\forall S \in \binom{V}{k}$, $P(S$ is monochromatic$) = 2/2^{\binom{k}{2}}$

(by union bound) $P(\exists S$ monochromatic$) \leq \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$

[# colorings $= 2^{|E|} = 2^{\binom{n}{2}} \approx 2^{n^2/2}$.

ercise: $R(k,\ell) \leq \binom{k+\ell-2}{k-1}$

2) Large Cut in a graph

- Let $G=(V,E)$ be an undirected graph. For $A \subseteq V$, define $cut(A) := \{(u,v) \in E \mid (u \in A, v \in \bar{A}) \vee (u \in \bar{A}, v \in A)\}$

undirected subgraph of G. $\bar{A} := V \setminus A$

orem 2: Given $G=(V,E)$, a cut of size $\geq |E|/2$ can be found efficiently (randomized algo.).

↳ it proves existence of A: $|cut(A)| \geq |E|/2$.

) Sum-Free Subset

For subset $S \subseteq \mathbb{Z}$, define $S+S := \{s_1 + s_2 \mid s_1, s_2 \in S\}$.

Defn: S is sum-free if $S \cap (S+S) = \phi$.

rem 3: For any set S of n nonzero integers, there is a subset $S' \subseteq S$: (i) S' is sum-free, & (ii) $|S'| > n/3$.

4) Discrepancy

Theorem 4: Given n unit vectors $v_i \in \mathbb{R}^n$, $i \in [n]$. Then, $\exists$ "binary"-vector $b \in \{-1,1\}^n$ s.t. $\|\sum_{i \in [n]} b_i \cdot v_i\| \leq \sqrt{n}$.

$\exists \bar{b} \in \{-1,1\}^n$, $X^2 \leq n$ (resp. $\geq n$)

" , $X \leq \sqrt{n}$ (resp $\geq \sqrt{n}$).

Bernoulli random variable.

- Toss a coin with $P(H) =: p$

- Prob. mass fn. $P(X=1) = P(H) = \boxed{p}$

$E[X] = P(X=1)\cdot 1 + P(X=0)\cdot 0 = \boxed{p}$.

Binomial random variable.

many times 'H appears? Say X.

Prob. Mass fn.: $P(X=i) = \binom{n}{i}\cdot p^i \cdot (1-p)^{n-i}$

$E[X]$ $\boxed{np}$.

Geometric Random Variable

tosses to get H.

Prob. mass fn.: $P(X_1=k) = (1-p)^{k-1}\cdot p$

$E[X]$ $\boxed{\frac{1}{p}}$.

Negative binomial random variable

$P(H) =: p$, till you get n H's.

Prob. mass fn.: $P(X_n=k) = \binom{k-1}{n-1}\cdot p^n \cdot (1-p)^{k-n}$

$E[X_n]$ $\boxed{n/p}$

Continuous random variable X

$E[X] := \int_{-\infty}^{\infty} x\cdot(f_X(x)\cdot dx)$

Exponential Random Variable.

For parameter $\lambda>0$, define $f_X(x) := \begin{cases}\lambda\cdot e^{-\lambda x}, & x\geq 0\\ 0 & x<0\end{cases}$

$E[X] = \int_0^\infty x\cdot\lambda e^{-\lambda x}\cdot dx = \boxed{1/\lambda}$

Normal / Gaussian random variable.

Define $f_X(x) := \frac{1}{\sqrt{2\pi}}\cdot e^{-x^2/2}$, for $x\in\mathbb{R}$.

Then, $(\bar{X}_n - \mu)$, for large n, behaves like the standard-normal distribution!

substituting:

$P(S<(1-\delta)u) < \left(e^{-\delta}/(1-\delta)^{1-\delta}\right)^u \leq \left(e^{-\delta^2/2}\right)^u$

sign change

$P(S>(1+\delta)\cdot u) < \left(e^{\delta}/(1+\delta)^{1+\delta}\right)^u$

ne: $P(S<(1-\delta).E[S]$ OR $S>(1+\delta).E[S])$ $<$ (exponentially small in n).

$\{X, Y, X+Y, X-Y\}$ is a fa

- pairwise (or 2-wise)
- not 3-wise indep.
- not mutually indep.

$\Rightarrow P(\ell) < n^2/\ell! < 1/n^3$. $\Rightarrow$ Load on each server is almost always $\leq 6\lg n/\lg\lg n$ $\leftarrow$ (exp. smaller than n)

[Stirling's Estimate] $\ell! \approx (\ell/e)^\ell\cdot\sqrt{2\pi\ell}$

$P(X_1=x_1\wedge\cdots\wedge X_k=x_k) = P(X_1=x_1)\cdot P(X_2=x_2\mid X_1=x_1)\cdots P(X_k=x_k\mid X_1=x_1\wedge\cdots\wedge X_{k-1}=x_{k-1})$

Process is independent iff $P(X_1=x_1\wedge\cdots\wedge X_k=x_k) = \prod_{1\leq i\leq k} P(X_i=x_i)$.

Markov Chain $P(X_k=x_k\mid X_{k-1}=x_{k-1}\wedge\cdots\wedge X_1=x_1) = P(X_k=x_k\mid X_{k-1}=x_{k-1})$,

homogeneous Markov Chain if
obability does not depend on time
$P(X_k=i\mid X_{k-1}=j) = P(X_1=i\mid X_0=j)$ "difference" is same

(stay) $p\cdot\frac{1}{n}$, if $(j,i)\notin E$.

Row sum of M' = 1

$M' = p\cdot J_n + (1-p)\cdot M$; where $J_n := \frac{1}{n}\cdot J$ all-1 matrix

from Strategy-2

M' is a regular, homog. Markov chain.

emma: Let $M' := \frac{1}{2}\cdot I + \frac{1}{2}\cdot M$. If M is ergodic then M' is regular.

Martingale.

A parent in E_j is expected to give birth to a child in the same stat. $= j$

j-th-entry in $M^n\cdot\begin{pmatrix}0\\1\\\vdots\\N\end{pmatrix}$

$p_{00} = p_{NN} = 1$, while other $p_{jk}<1$

Uniform sampling from $[0\ldots m-1]$

- Given an integer m (& unbiased coin), design an algorithm to pick random $X\in[0\ldots m-1]$.

$\forall t\in[0\ldots m-1]$: $P(\text{Output is } t) = \frac{1}{m}$

$E[\#\text{times (1) is executed}]$ $\mu = 2^k/m < 2$.

Sampling k numbers from $[0\ldots m-1]$

Given k, m; you want to pick a random subset $S\subseteq[0\ldots m-1]$ of size $|S|=k$.

Let $t_1\neq t_2\in[0\ldots m-1]$. $P(S=\{t_1,t_2\})$ $1/\binom{m}{2}$

$P(S=\{t_1,\ldots,t_k\}) = 1/\binom{m}{k}$ Extremal set families jaise?

$E[\#\text{iterations for an } i] = m/(m-i+1)$

$E[\#\text{steps in the algo.}] = \sum_{i\in[k]} m/(m-i+1)$

Uniform Sampling a permutation of $[n]$

- Given n, you want to pick a permutation, say as a string $\underline{S}$. Eg. 132 of [3].

Let $t_1\neq t_2\in[n]$. $P(S=t_1t_2) = \frac{1}{n}\cdot\frac{1}{n-1}$

$P(S=t_1\cdots t_n) = 1/n(n-1)(n-2)\cdots 1 = 1/n!$

$E[\#\text{steps}] = \approx n\cdot\log n$

$\exists t\in[-1,1]$, $X\leq n$ (resp. $=n$)

" , $X\leq\sqrt{n}$ (resp. $\geq\sqrt{n}$).

Extremal Set Families

Defn: Let $\mathcal{F} =: \{(A_i,B_i)\mid i\in[h]\}$ be a family of set-pairs. It is called (k,ℓ)-system if $|A_i|=k$, $|B_j|=\ell$ & $\begin{cases}A_i\cap B_i = \phi, & \text{(no intersec' i-i)}\\ A_i\cap B_j\neq\phi, & \forall i\neq j\in[h].\end{cases}$

5: (Bollobás, 1965) $\mathcal{F}$ is (k,ℓ)-system $\Rightarrow |\mathcal{F}|\leq\binom{k+\ell}{\ell}$.

$E_i :=$ event that elements of A_i precede B_i.

$P(E_i) = 1/\binom{k+\ell}{k}$ Girls wala example in assembly

laim: $\forall i\neq j\in[h]$, E_i, E_j are disjoint.

6) Super-concentrator

- Defn: Super-concentrator is dag $G=(V,E)$ with n special input nodes $I\subset V$ & n output nodes $O\subset V$: $\forall k$, $\forall S\in\binom{I}{k}$, $\forall T\in\binom{O}{k}$, vertices in S connects to T with k disjoint paths. (physically wired)

Exercise: A superconcentrator exists with $|V|=2n$ & $|E|=n^2$.

First, we show: Lemma: An efficient randomized algo. constructs $(6j, 4j, 3j)$-concentrator; with $|E|=O(j)$.

thm 6: A randomized poly-time algo. designs a superconcentrator $G=(V,E)$ with $|V|=20j$ & $|E|=O(j)$.

$3j<k\leq 6j \Rightarrow$ # vertices in (S,T) that are matched by M are $\geq(k-3j)$.

$\Rightarrow$ #unmatched vertices in S is $\leq 3j$.

Streaming Algorithms

m, n are very large

Defn: In data stream model, there's an input stream $\sigma =: \langle a_1, \ldots, a_m \rangle$ whose elements are tokens a_i from the universe $[n] = \{1, \ldots, n\}$.

or each token $j \in [n]$ & $r \geq 0$, define

$X_{r,j} := \begin{cases} 1, & \text{if } \nu_2(h(j)) \geq r \\ 0, & \text{else} \end{cases}$

$Y_r := \sum \{X_{r,j} \mid j \text{ s.t. } j \in \sigma\}$.

$T :=$ terminal value of z (when algo. st

$E[Y_r] = d/2^r$.

$\mathrm{var}(Y_r) = d/2^r$. $P(\frac{d}{3} \leq \hat{d} \leq 3d) \geq 1 - \frac{2\sqrt{2}}{3}$

Thm 1: $P(d/3 \leq \text{output} \leq 3d) \geq 1 - 2^{-\Omega(k)}$ & $s \leq k \cdot \log n$.

Let $f_j :=$ frequency of j in σ.

$\sum_{j \in [n]} f_j = m =: F_1$ (first mom

second moment $F_2 := \sum_j f_j^2$

$Z = \sum_{j \in [n]} f_j \cdot h(j)$

$E[Z] = \sum_j f_j \cdot E[h(j)] = 0$ $E[Y] = F_2$

$\Rightarrow P(|Y - F_2| > \alpha \cdot F_2) \leq \frac{2F_2^2}{(\alpha F_2)^2} = 2/\alpha^2$.

Y approx. F_2 well!

OUTPUT $Y' := \frac{Y_1 + \ldots + Y_k}{k}$

$E[Y'] = E[Y] = F_2$ &

$\mathrm{var}(Y') = \mathrm{var}(Y)/k$.