Probabilistische Algorithmen

Alexander May

Fakultät für Mathematik
Ruhr-Universität Bochum

Sommersemester 2016

Organisatorisches

- Vorlesung: **Di 10–12** (2+2 SWS, 6 CP)
- Übung: **tba**
- Assistent: **Robert Kübler**
- Übungsbetrieb: jeweils abwechselnd alle 2 Wochen
  - Präsenzübung, Start 19. April
  - Zentralübung, Start 26. April
- Übungsaufgaben werden korrigiert.
- Gruppenabgaben bis 3 Personen möglich.
- Mündliche Prüfungen: Fr. 29.07.2016 (?)

Ws-Funktion

Definition

Ein *Wsraum* besteht aus

1. *Ergebnismenge* Ω mit Ereignissen $E \subseteq \Omega$,
2. *Ereignismenge* $\mathcal{F} \subseteq 2^{\Omega}$,
3. *Wsfunktion* $\Pr : \mathcal{F} \to \mathbb{R}$.

Ein $e \in \Omega$ heißt *Elementarereignis*.

Definition Wsfunktion

Für eine *Wsfunktion* $\Pr : \mathcal{F} \to \mathbb{R}$ gilt:

1. $0 \leq \Pr(E) \leq 1$ für alle Ereignisse $E \subseteq \Omega$.
2. $\Pr(\Omega) = 1$.
3. Für alle (abzählbar un-)endlichen Sequenzen $E_1, E_2, \ldots$ paarweise disjunkter Ereignisse

$$\Pr(\bigcup_{i \geq 1} E_i) = \sum_{i \geq 1} \Pr(E_i).$$

Eintreten von mindestens einem Ereignis

Lemma

Für beliebige Ereignisse E_1, E_2 gilt:

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2).$$

Beweisidee: Schreibe $E_1 \cup E_2 = E_1 \cup (E_2 \setminus (E_1 \cap E_2))$.

Korollar 1 Union Bound

Für alle (abzählbar un-)endlichen $E_1, E_2, \ldots$ gilt

$$\Pr(\bigcup_{i\geq 1} E_i) \leq \sum_{i \geq 1} \Pr(E_i).$$

Korollar 2 Inklusion-Exklusion

Für alle Ereignisse $E_1, E_2, \ldots, E_n$ gilt

$$\Pr(\bigcup_{i=1}^n E_i) = \sum_{i=1}^n \Pr(E_i) - \sum_{i<j} \Pr(E_i \cup E_j) + \sum_{i<j<k} \Pr(E_i \cap E_j \cap E_k) - \ldots$$

Unabhängigkeit und Bedingte Ws

Definition Unabhängigkeit

Zwei Ereignisse E_1, E_2 sind *unabhängig* gdw

$$\Pr(E_1 \cap E_2) = \Pr(E_1) \cdot \Pr(E_2).$$

Allgemein: $E_1, \ldots, E_k$ sind *unabhängig* gdw für alle $I \subseteq [1, \ldots, k]$ gilt

$$\Pr[\bigcap_{i \in I} E_i] = \prod_{i \in I} \Pr(E_i).$$

Definition Bedingte Ws

Die *bedingte Ws*, dass E_2 eintritt, falls E_1 eintritt, ist

$$\Pr[E_2 \mid E_1] = \frac{\Pr[E_1 \cap E_2]}{\Pr[E_1]} \text{ falls } \Pr[E_1] > 0.$$

Anmerkung: Für unabhängige E_1, E_2 gilt

$$\Pr(E_2 \mid E_1) = \frac{\Pr(E_1 \cap E_2)}{\Pr(E_1)} = \frac{\Pr(E_1)\Pr(E_2)}{\Pr(E_1)} = \Pr(E_2).$$

Polynomvergleich

Problem: Polynomvergleich.

- Überprüfe, ob $(x+1)(x+2)(x+3) \stackrel{?}{=} x^3 + 6x^2 + 10x + 6$.
- Sei d der Grad der zu vergleichenden Polynome.
- Deterministische Lösung: Multipliziere linke Seite in $\mathcal{O}(d^2)$ aus.
- Algorithmus liefert stets die korrekte Lösung.

Notation: $r \in_R A$ bedeutet, wir wählen $r \in A$ uniform gleichverteilt, d.h.

$$\Pr(r = a) = \frac{1}{|A|} \text{ für alle } a \in A.$$

Algorithmus Probabilistischer Polynomvergleich ProbPoly

EINGABE: $F(x), G(x)$

1. FOR $i = 1$ to k
   1. Wähle $r \in_R \{1, \ldots, 100d\}$.
   2. Falls $F(r) \neq G(r)$ Ausgabe “verschieden”, EXIT.

   Ausgabe “gleich”.

Laufzeit: $\mathcal{O}(kd) = \mathcal{O}(d)$ für konstantes k.

Korrektheit

Satz

ProbPoly liefert für $F(x) = G(x)$ stets die korrekte Antwort und für $F(x) \neq G(x)$ die korrekte Antwort mit Ws $\geq 1 - (\frac{1}{100})^k$.

Beweis:

- Falls $F(x) = G(x)$, so gilt auch $F(r) = G(r)$ für alle r.
- Angenommen $F(x) \neq G(x)$. Damit gilt $P(x) = F(x) - G(x) \neq 0$.
- $P(x)$ besitzt $\mathrm{grad}(P(x)) \leq d$ und damit höchstens d Nullstellen.
- Ereignis E_i: ProbPoly liefert nicht "verschieden" in Iteration i.
$$\Pr[E_i] = \Pr[\text{r ist Nullstelle von } P(x)] \leq \frac{d}{100d} = \frac{1}{100} \text{ für alle } i.$$
- Definiere $E = E_1 \cap \ldots \cap E_k$. Aus der Unabhängigkeit der E_i folgt
$$\Pr(E) = \Pr(E_1 \cap \ldots \cap E_k) = \prod_{i=1}^k \Pr(E_i) \leq (\frac{1}{100})^k.$$
- E bedeutet, dass letztlich Ausgabe "gleich" erfolgt. D.h.
$$\Pr[\text{Ausgabe "verschieden"} | F(x) \neq G(x)] = \Pr[\bar{E}] \geq 1 - (\frac{1}{100})^k.$$

Verbesserter Algorithmus

Idee: Hätten gerne paarweise verschiedene r_i in Iteration i.

Algorithmus Probabilistischer Polynomvergleich PROBPOLY2

EINGABE: $F(x), G(x)$

1. FOR $i = 1$ to k
   1. Wähle $r_i \in_R \{1, \ldots, 100d\}$ mit $r_i \neq r_j$ für alle $j = 1, \ldots, k-1$.
   2. Falls $F(r) \neq G(r)$ Ausgabe "verschieden", EXIT.

   Ausgabe "gleich".

Analyse

Analyse der Irrtumsws.

- Angenommen $F(x) \neq G(x)$, d.h. $P(x) = F(x) - G(x) \neq 0$.
- Die Ereignisse E_i, dass $P(r_i) = 0$, sind nun abhängig. D.h.

$$\begin{aligned}\Pr(E_1 \cap \ldots \cap E_k) &= \Pr(E_k \mid E_1 \cap \ldots \cap E_{k-1}) \cdot \Pr(E_1 \cap \ldots \cap E_{k-1}) = \ldots \\ &= \Pr(E_1) \cdot \Pr(E_2 \mid E_1) \cdot \ldots \cdot \Pr(E_k \mid E_1 \cap \ldots \cap E_{k-1}).\end{aligned}$$

- Ziehen der r_i ohne Zurücklegen liefert
$$\Pr(E_j \mid E_1 \cap \ldots \cap E_{j-1}) \leq \tfrac{d-(j-1)}{100d-(j-1)}.$$
- Es folgt für die Fehlerws
$$\Pr(E) = \Pr(E_1 \cap \ldots \cap E_k) \leq \prod_{j=1}^{k} \tfrac{d-(j-1)}{100d-(j-1)}.$$
- Dies ist für $k \ll d$ nur unwesentlich kleiner als $(\frac{1}{100})^k$ zuvor.
- Wir ziehen daher häufig eine vereinfachte Analyse vor.
- D.h. wir verwenden oft unabhängig gleichverteilte r_i.

Matrixmultiplikation

Problem Matrixvergleich.

- Gegeben seien Matrizen $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{F}_2^{n \times n}$. Überprüfe $\mathbf{AB} \stackrel{?}{=} \mathbf{C}$.
- Produkt $\mathbf{A}b_i$ kann für $b_i \in \mathbb{F}_2^n$ in Zeit $\mathcal{O}(n^2)$ berechnet werden.
- Deterministisch: Multipliziere $\mathbf{AB}$ in Zeit $\mathcal{O}(n^3)$ (bzw. $\mathcal{O}(n^{2.37})$) aus.

Algorithmus ProbMatrix

EINGABE: $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{F}_2^{n \times n}$

1. For $i = 1$ to k
   1. Wähle $r \in_R \{0, 1\}^n$.
   2. Falls $\mathbf{A}(\mathbf{B}r) \neq \mathbf{C}r$ Ausgabe “verschieden”, EXIT.

   Ausgabe “gleich”.

Laufzeit: $\mathcal{O}(kn^2) = \mathcal{O}(n^2)$ für $k \ll n$

Alle oder einzelne

Lemma

Wahl von $r = (r_1, \ldots, r_n) \in_R \mathbb{F}_2^n$ ist äquivalent zur Wahl aller $r_i \in_R \mathbb{F}_2$.

Beweis:

- $\Rightarrow$: oBdA sei $i = 1$.
- Es existieren 2^{n-1} Vektoren der Form $0\{0,1\}^{n-1}$ bzw. $1\{0,1\}^{n-1}$.
- D.h. $\Pr(r_1 = 0) = \frac{2^{n-1}}{2^n} = \frac{1}{2} = \Pr(r_1 = 1)$.
- $\Leftarrow$: Wähle alle $r_i \in_R \mathbb{F}_2$ und setze $r = (r_1, \ldots, r_n)$.
- Dann gilt für alle $x \in \mathbb{F}_2^n$

  $$\Pr(r = x) = \Pr(r_1 = x_1 \cap \ldots \cap r_n = x_n) = \prod_{i=1}^n \Pr(r_i = x_i) = \frac{1}{2^n}.$$

Analyse von ProbMatrix

Lemma

Sei $\mathbf{AB} \neq \mathbf{C}$. Dann gilt für alle $r \in_R \mathbb{F}_2^n$

$$\Pr(\mathbf{AB}r = \mathbf{C}r) \leq \tfrac{1}{2}.$$

Beweis:

- Sei $\mathbf{D} = \mathbf{AB} - \mathbf{C} \neq 0$. OBdA $d_{11} \neq 0$.
- Angenommen $\mathbf{AB}r = \mathbf{C}r$, d.h. $\mathbf{D}r = 0$ und insbesondere
$$\textstyle\sum_{j=1}^{n} d_{1j}r_j = 0 \text{ bzw. } r_1 = -\frac{\sum_{j=2}^{n} d_{1j}r_j}{d_{11}}.$$
- Wir wählen in dieser Reihenfolge $r_n, \ldots, r_2, r_1 \in_R \mathbb{F}_2$.
- Die Wahl von $r_n, \ldots, r_2 \in_R \mathbb{F}_2$ determiniert $x := -\frac{\sum_{j=2}^{n} d_{1j}r_j}{d_{11}} \in \mathbb{F}_2$.
- Es folgt $\Pr(r_1 = x) = \frac{1}{2}$ und damit insgesamt
$$\textstyle\Pr(\mathbf{D}r = 0) \leq \Pr(\sum_{j=1}^{n} d_{1j}r_j = 0) = \frac{1}{2}.$$

Korollar

ProbMatrix liefert für $\mathbf{AB} = \mathbf{C}$ stets die korrekte Antwort und für $\mathbf{AB} \neq \mathbf{C}$ die korrekte Antwort mit Ws mindestens $1 - (\frac{1}{2})^k$.

Umdrehen der bedingten Ws

Problem:

- Haben bisher $\Pr(\text{Ausgabe “gleich”}|\mathbf{AB} \neq \mathbf{C})$ analysiert.
- Uns interessiert aber oft $\Pr(\mathbf{AB} \neq \mathbf{C}|\text{Ausgabe “gleich”})$.

Satz von der totalen Ws

Seien $E_1, \ldots, E_n \subset \Omega$ disjunkt mit $\bigcup_{i=1}^n E_i = \Omega$. Dann gilt

$$\Pr(B) = \sum_{i=1}^n \Pr(B \cap E_i) = \sum_{i=1}^n \Pr(B \mid E_i) \Pr(E_i).$$

Beweis: per Bild.

Satz von Bayes

Seien $E_1, \ldots, E_n \subset \Omega$ disjunkt mit $\bigcup_{i=1}^n E_i = \Omega$, $\Pr(B) > 0$. Dann gilt

$$\Pr(E_j \mid B) = \frac{\Pr(E_j \cap B)}{\Pr(B)} = \frac{\Pr(B|E_j) \Pr(E_j))}{\sum_{i=1}^n \Pr(B|E_i) \Pr(E_i)}.$$

Umdrehen der bedingten Ws

Berechnen von $\Pr(\mathbf{AB} \neq \mathbf{C} | \text{Ausgabe “gleich”})$:

- Sei E das Ereignis $\mathbf{AB} = \mathbf{C}$ und B das Ereignis $\mathbf{AB}r = \mathbf{C}r$.
- Starten mit *A priori Modell*, dass $\Pr(E) = \Pr(\bar{E}) = \frac{1}{2}$.
- Es gilt $\Pr(B \mid E) = 1$ und $\Pr(B \mid \bar{E}) \leq \frac{1}{2}$. Mit Satz von Bayes folgt
$$\Pr(E \mid B) = \frac{\Pr(B|E)\Pr(E)}{\Pr(B|E)\Pr(E) + \Pr(B|\bar{E})\Pr(\bar{E})} \geq \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2}} = \frac{2}{3}.$$
- Passen Modell nach 1. Iteration an: $\Pr(E) \geq = 23$ und $\Pr(\bar{E}) = \frac{1}{3}$.
- Bei erneutem Ereignis B liefert der Satz von Bayes
$$\Pr(E \mid B) \geq \frac{\frac{2}{3}}{\frac{2}{3} + \frac{1}{2}\cdot\frac{1}{3}} = \frac{4}{5}.$$
- Allgemein erhalten wir nach dem k-ten Auftreten von B induktiv
$$\Pr(E \mid B) \geq 1 - \frac{1}{2^k + 1}.$$
- D.h. nach z. B. 100 Iterationen erhalten wir
$\Pr(\text{Ausgabe “gleich”} \mid \mathbf{AB} \neq \mathbf{C}) \geq 1 - \frac{1}{2^{100}}$ und
$\Pr(\mathbf{AB} \neq \mathbf{C} \mid \text{Ausgabe “gleich”}) \geq 1 - \frac{1}{2^{100}+1}$.

Randomisierter Min-Cut Algorithmus

Problem Min-Cut

Gegeben: Zusammenhängender ungerichteter Graph $G = (V, E)$
Gesucht: $C \subseteq E$ mit min. $|C|$ und $G = (V, E \setminus C)$ nicht zusammenh.

Algorithmus Kanten-Kontraktion (Karger 1993)

EINGABE: $G = (V, E)$

1. REPEAT UNTIL $|V| = 2$
   1. Wähle $e = \{u, v\} \in_R E$.
   2. Verschmelze u, v zu einem Knoten mit Label u, v.
      Entferne dabei alle Kanten zwischen u und v.

AUSGABE: $C = E$, d.h. alle verbliebenen Kanten

- **Laufzeit:** $\mathcal{O}(|V| + |E|) = \mathcal{O}(n + m)$ für $|V| = n, |E| = m$.
- Bei Terminierung: Zwei Knoten mit Label $S \subset V$ und $V \setminus S$.
- Damit ist C ein Cut, der die Partitionen S und $V \setminus S$ trennt.
- C besitzt aber nicht notwendigerweise minimale Größe.

Analyse KANTEN-KONTRAKTION

Satz

KANTEN-KONTRAKTION berechnet minimalen Cut mit Ws $\geq \frac{2}{n(n-1)}$.

Beweis:

- Sei C_{min} ein minimaler Cut in G mit $|C| = k$.
- Falls nie eine Kante in C_{min} kontrahiert wird, erfolgt Ausgabe C_{min}.
- E_i: Ereignis Kante $\{u, v\} \notin C$ in i-ter Iteration.
- Sei $F_i = \bigcap_{j=1}^{i} E_j$. Wir berechnen zunächst $\Pr(F_1) = \Pr(E_1)$.
- Jedes $v \in V$ besitzt $\deg(v) \geq k$. (Warum?)
- Damit gilt $|V| \geq \frac{kn}{2}$. D.h. $\Pr(\bar{E}_1) \leq \frac{k}{\frac{kn}{2}} = \frac{2}{n}$ bzw. $\Pr(E_1) \geq 1 - \frac{2}{n}$.
- Nach E_1 verbleibt G mit $n-1$ Knoten und minimalem Cut C_{min}.
- D.h. $\Pr(E_2 \mid F_1) \geq 1 - \frac{2}{n-1}$ und allg. $\Pr(E_i \mid F_{i-1}) \geq 1 - \frac{2}{n-(i-1)}$.
- KANTEN-KONTRAKTION liefert nach $n-2$ Kontraktionen C_{min} mit

$$\begin{aligned}\Pr(F_{n-2}) &= \Pr(E_{n-2} \cap F_{n-3}) = \Pr(E_{n-2} \mid F_{n-3}) \cdot \Pr(F_{n-3}) \\ &= \Pr(E_{n-2} \mid F_{n-3}) \cdot \Pr(E_{n-3} \mid F_{n-4}) \cdot \ldots \cdot \Pr(E_2 \mid F_1) \cdot \Pr(F_1)\end{aligned}$$

- Es folgt $\Pr(F_{n-2}) \geq \prod_{i=1}^{n-2} \left(1 - \frac{2}{n-(i-1)}\right) = \frac{2}{n(n-1)}$.

Amplification

Lemma Amplification

$n(n-1)\ln n$-maliges Wiederholen von KANTEN-KONTRAKTION und Ausgabe des kleinsten Cuts liefert minimalen Cut mit Ws $\geq 1 - \frac{1}{n^2}$.

Beweis:

- Ereignis E_i: kein minimaler Cut in i-ter Wiederholung
- Damit liefert KANTEN-KONTRAKTION keinen minimalen Cut mit $\Pr(E_1 \cap \ldots \cap E_{n(n-1)\ln n}) = \prod_{i=1}^{n(n-1)\ln n} \Pr(E_i) \leq \left(1 - \frac{2}{n(n-1)}\right)^{n(n-1)\ln n}$.
- Mittels $1 - x \leq e^{-x}$ erhalten wir
$$\Pr(E_1 \cap \ldots \cap E_{n(n-1)\ln n}) \leq e^{-2\ln n} = \frac{1}{n^2}.$$

Diskrete Zufallsvariablen

Definition Zufallsvariable

Eine *Zufallsvariable (ZV)* X ist eine Abbildung $X : \Omega \to \mathbb{R}$. Eine *diskrete Zufallsvariable* nimmt nur (abzählbar un-)endlich viele Werte an.

Bsp:

- Sei $\Omega = \{(1,1),(1,2),\ldots\}$ ein 2-maliger Münzwürf.
- ZV X: Summe der beiden Würfe. X nimmt Werte in $\{2, \ldots, 12\}$ an.
- $\Pr(X = 4) = \Pr((1,3)) + \Pr((2,2)) + \Pr((3,1)) = \frac{3}{36} = \frac{1}{12}$

Definition Unabhängigkeit von Zufallsvariablen

Zwei ZV X, Y sind *unabhängig* gdw

$$\Pr((X = x) \cap (Y = y)) = \Pr(X = x) \cdot \Pr(Y = y) \text{ für alle } x, y.$$

Allgemein: $X_1, \ldots, X_k$ sind *unabhängig* gdw für alle $I \subseteq \{1, \ldots, k\}$ gilt

$$\Pr\left(\bigcap_{i \in I} X_i = x_i\right) = \prod_{i \in I} \Pr(X_i = x_i) \text{ für alle } x_i \text{ mit } i \in I.$$

Erwartungswert

Definition Erwartungswert

Der *Erwartungswert* einer diskreten ZV ist definiert als

$$\mathbb{E}[X] = \sum_i i \cdot \Pr(X = i).$$

$\mathbb{E}[X]$ ist endlich, falls $\sum_i |i| \cdot \Pr(X = i)$ konvergiert, sonst unendlich.

Bsp:

- Sei X die Summe zweier Würfe eines Würfels. Dann gilt
$$\mathbb{E}[X] = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + \ldots + 12 \cdot \frac{1}{36} = 7.$$
- Sei X eine ZV mit $\Pr(X = 2^i) = \frac{1}{2^i}$ für $i \geq 1$. Dann gilt
$$\mathbb{E}[X] = \sum_{i \geq 1} 2^i \cdot \frac{1}{2^i} = \infty.$$

Linearität des Erwartungswerts

Satz Linearität des Erwartungswerts

Seien $X_1, \ldots, X_n$ diskrete ZV mit endlichem Erwartungswert. Dann gilt

$$\mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i].$$

Beweis: Nur für $n = 2$, für allgemeine n per Induktion.

$$\begin{aligned}
\mathbb{E}[X+Y] &= \sum_i \sum_j (i+j) \Pr((X=i) \cap (Y=j)) \\
&= \sum_i i \sum_j \Pr((X=i) \cap (Y=j)) + \sum_j j \sum_i \Pr((X=i) \cap (Y=j))
\end{aligned}$$

- Der Satz von der totalen Ws liefert damit
$$\mathbb{E}[X+Y] = \sum_i i \Pr(X=i) + \sum_j j \Pr(Y=j) = \mathbb{E}[X] + \mathbb{E}[Y].$$

Beispiel zuvor:

- Sei X_1, X_2 ZV für 1. bzw 2. Wurf und $X = X_1 + X_2$.
- Dann gilt $\mathbb{E}[X] = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{7}{2}$ und $\mathbb{E}[X_i] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = 7$.

Linearität des Erwartungswerts

Lemma

Für alle $c \in \mathbb{R}$ und alle diskrete ZV X gilt

$$\mathbb{E}[cX] = c\mathbb{E}[X].$$

Beweis:

- Für $c = 0$ sind beide Seiten 0. Für $c \neq 0$ gilt

$$\begin{aligned}
\mathbb{E}[cX] &= \sum_j j \Pr(cX = j) \\
&= c \sum_j \frac{j}{c} \Pr\left(X = \frac{j}{c}\right) \\
&= c \sum_i i \Pr(X = i) \\
&= c\mathbb{E}[X].
\end{aligned}$$

Linearität des Erwartungswerts

Lemma

Es gilt $\mathbb{E}[X^2] \geq (\mathbb{E}[X])^2$.

Beweis:

- Definiere $Y = (X - \mathbb{E}[X])^2 > 0$. Es folgt

$$\begin{aligned} 0 \leq \mathbb{E}[Y] &= \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X])^2] \\ &= \mathbb{E}[X^2] - \mathbb{E}[2X\mathbb{E}[X]] + (\mathbb{E}[X])^2 = \mathbb{E}[X^2] - (\mathbb{E}[X])^2. \end{aligned}$$

Eine Verallgemeinerung liefert die folgende Jensen Ungleichung.

Jensen Ungleichung

Satz Jensen Ungleichung

Sei $f : \mathbb{R} \to \mathbb{R}$ konvex. Dann gilt $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$.

Beweis:

- Wir nehmen an, dass f eine Taylor-Entwicklung besitzt.
- Sei $\mu = \mathbb{E}[X]$. Dann gilt $f(X) = f(\mu) + f'(\mu)(X - \mu) + \frac{f''(c)(X-\mu)^2}{2}$.
- Konvexität von f is äquivalent zu $f''(c) \geq 0$. Wir erhalten
$$f(X) \geq f(\mu) + f'(\mu)(X - \mu).$$
- Anwenden des Erwartungswerts auf beiden Seiten liefert
$$\mathbb{E}[f(x)] \geq \mathbb{E}[f(\mu)] + f'(\mu)(\mathbb{E}[X] - \mu) = f(\mu) = f(\mathbb{E}[X]).$$

Bernoulli und Binomial ZV

- Betrachten ein Zufallsexperiment E, das mit Ws p erfolgreich ist.
- Definiere für $i = 1, \ldots, n$ die *Bernoulli*- bzw. *Indikator*-ZV (IV)

$$Y_i = \begin{cases} 1 & \text{falls } E \text{ erfolgreich} \\ 0 & \text{sonst} \end{cases}.$$

- Für alle IV Y_i gilt $\mathbb{E}[Y_i] = 0 \cdot \Pr[Y_i = 0] + 1 \cdot \Pr[Y_i = 1] = \Pr[Y_i = 1]$.
- Definiere $X = Y_1 + \ldots + Y_n$ als ZV für die Anzahl der Erfolge.

Definition Binomialverteilung

Eine ZV X ist *binomial verteilt* gemäß $B(n, p)$ falls

$$\Pr(X = j) = \binom{n}{j} p^j (1-p)^{n-j} \text{ für } j = 0, \ldots, n.$$

Anmerkungen

- $X = j$, falls wir genau j Erfolge und $n - j$ Misserfolge erhalten.
- Ws-Verteilung: $\sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} = (p + (1-p))^n = 1$.
- Wegen $\mathbb{E}[Y_i] = p$ gilt $\mathbb{E}[X] = \mathbb{E}[Y_1] + \ldots + \mathbb{E}[Y_n] = np$.

Bedingter Erwartungswert

Definition Bedingter Erwartungswert

$\mathbb{E}[X \mid Y = y] = \sum_x x \cdot \Pr(X = x \mid Y = y)$.

Lemma

Für alle ZV X, Y gilt $\mathbb{E}[X] = \sum_y \Pr(Y = y)\mathbb{E}[X \mid Y = y]$.

Beweis:

$$\begin{aligned}
\sum_y \Pr(Y = y)\mathbb{E}[X \mid Y = y] &= \sum_x \sum_y x \Pr(X = x \mid Y = y) \Pr(Y = y) \\
&= \sum_x \sum_y x \Pr(X = x \cap Y = y) \\
&= \sum_x x \Pr(X = x) = \mathbb{E}[X].
\end{aligned}$$

Bedingter Erwartungswert

Definition

$\mathbb{E}[X \mid Y]$ ist eine ZV in Y, die den Wert $\mathbb{E}[X|Y=y]$ für $Y=y$ annimmt.

Beispiel:

- 2-facher Münzwurf: X_1, X_2 ZV für 1. bzw. 2. Wurf und $X = X_1 + X_2$.
$$\mathbb{E}[X \mid X_1] = \sum_x \Pr(X = x \mid X_1) = \sum_{x=X_1+1}^{X_1+6} x \cdot \tfrac{1}{6} = X_1 + \tfrac{7}{2}.$$
- Es folgt $\mathbb{E}[\mathbb{E}[X \mid X_1]] = \mathbb{E}[X_1 + \frac{7}{2}] = \mathbb{E}[X_1] + \frac{7}{2} = 7 = \mathbb{E}[X]$.

Satz

$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]]$

Beweis:

- Da $\mathbb{E}[X \mid Y]$ eine ZV in Y ist, folgt aus obiger Definition
$$\mathbb{E}[\mathbb{E}[X \mid Y]] = \sum_y \Pr(Y = y)\mathbb{E}[X \mid Y = y].$$
- Mit dem Lemma auf voriger Folie ist die rechte Seite gleich $\mathbb{E}[X]$.

Geometrische Verteilung

Definition Geometrische Verteilung

Eine ZV X ist *geometrisch* verteilt mit Parameter $0 < p < 1$, falls

$$\Pr(X = n) = (1-p)^{n-1}p \text{ für } n \geq 1.$$

Anmerkung:

- D.h. $X = n$ beschreibt, dass der 1. Erfolg im n-ten Versuch eintritt.

$$\sum_{i\geq 1}(1-p)^{i-1}p = p \cdot \sum_{i \geq 0}(1-p)^i = p \cdot \frac{1}{1-(1-p)} = 1$$

Lemma

Sei X eine diskrete ZV, die nur nicht-negative Werte annimmt. Es gilt

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr(X \geq i).$$

Beweis:

$$\sum_{i=1}^{\infty} \Pr(X \geq i) = \sum_{i=1}^{\infty}\sum_{j=i}^{\infty} \Pr(X = j) = \sum_{i=1}^{\infty} i \Pr(X = i) = \mathbb{E}[X].$$

Geometrische Verteilung

Für geometrisch verteilte X gilt $\Pr(X \geq i) = (1-p)^{i-1}$ und daher

$$\mathbb{E}[X] = \sum_{i=1}^{\infty}(1-p)^{i-1} = \sum_{i=0}^{\infty}(1-p)^{i} = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Alternative Rechnung mittels bedingter Erwartungswerte:

- Sei Y eine IV mit $Y = 1$ für Erfolg im 1. Versuch.
- Mit dem Lemma auf Folie 25 gilt

$$\begin{aligned} \mathbb{E}[X] &= \Pr(Y=0)\mathbb{E}[X \mid Y=0] + \Pr(Y=1)\mathbb{E}[X \mid Y=1] \\ &= (1-p)\mathbb{E}[X+1] + p = (1-p)\mathbb{E}[X] + 1. \end{aligned}$$

- Auflösen nach $\mathbb{E}[X]$ liefert $\mathbb{E}[X] = \frac{1}{p}$.

Coupon Collector Problem

Coupon Collector Problem: Wie oft muss man mit Zurücklegen aus $\{1, \ldots, n\}$ ziehen, bis alle Zahlen gezogen wurden?

Analyse Coupon Collector Problem:

- Sei X die Anzahl von Zügen, bis alle Zahlen gezogen wurden.
- X_i: Anzahl Züge, für die man exakt $i-1$ verschiedene Zahlen hat.
- Offenbar ist $X = 1 + \sum_{i=1}^{n} X_i$. Jedes X_i ist geometrisch verteilt mit
$$p_i = \frac{n-(i-1)}{n}.$$
- Es folgt $\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}$ und

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

Quicksort

Algorithmus QUICKSORT

EINGABE: Menge S verschiedener $x_1, \ldots, x_n \subset \mathbb{Z}$

1. IF $|S| \leq 1$, Ausgabe S.
2. Wähle Pivotelement $x \in_R S$.
3. Partitioniere in $L = \{x_i \in S \mid x_i < x\}$ und $R = \{x_i \in S \mid x_i > x\}$.
4. Ausgabe QUICKSORT(L), x, QUICKSORT(R).

AUSGABE: Aufsteigende Sortierung von S.

Anmerkungen:

- Jede Iteration kostet für die Partitionierung $\mathcal{O}(n)$ Vergleiche.
- Im worst case benötigt man $\Omega(n)$ viele Rekursionen.
- Im best case benötigt man $\mathcal{O}(\log n)$ viele Rekursionen.
- Wir zeigen, dass man auch im average case $\mathcal{O}(\log n)$ benötigt.

Analyse Quicksort

Theorem Analyse Quicksort

QUICKSORT benötigt erwartet $2n \ln + \Theta(n)$ Vergleiche.

Beweis:

- Sei $y_1, \ldots, y_n$ die sortierte Reihenfolge von $x_1, \ldots, x_n$.
- IV X_{ij}= 1, falls y_i und y_j von QUICKSORT verglichen werden.
- Wir erhalten für die ZV X für die Gesamtzahl von Vergleichen
$$X = \sum_{1 \le i < j \le n} X_{ij} \text{ und } \mathbb{E}[X] = \sum_{1 \le i < j \le n} \mathbb{E}[X_{ij}].$$
- Es gilt $\mathbb{E}[X_{ij}] = \Pr(X_{ij} = 1)$.
- y_i, y_j werden verglichen gdw eines von beiden als erstes Pivot aus der Menge $\{y_i, \ldots, y_j\}$ ausgewählt werden. (Warum?)
- Dies geschieht mit $\Pr(X_{ij} = 1) = \frac{2}{j-i+1}$. Wir erhalten

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{1 \le i < j \le n} \frac{2}{j-i+1} = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{2}{j-i+1} = \sum_{i=1}^{n} \sum_{j=2}^{n-i+1} \frac{2}{j} = \sum_{j=2}^{n} \sum_{i=2}^{n+1-j} \frac{2}{j} = \\
&= \sum_{j=2}^{n} (n-1+j)\frac{2}{j} = ((n+1)\sum_{j=2}^{n} \frac{2}{j}) - 2(n-1) = (2n+2)\sum_{j=1}^{n} \frac{1}{j} - 4n = 2n \ln n + \Theta(n).
\end{aligned}$$

Markov-Ungleichung

Satz Markov-Ungleichung

Sei $X \geq 0$ eine ZV. Dann gilt

$$\Pr(X \geq a) \leq \frac{\mathbb{E}[X]}{a} \text{ für alle } a > 0.$$

Beweis:

- Definiere IV mit $I = 1$ gdw $X \geq a$. Wegen $X \geq 0$ gilt
$$I \leq \frac{X}{a} \text{ und } \mathbb{E}[I] \leq \mathbb{E}[\tfrac{X}{a}] = \frac{\mathbb{E}[X]}{a}.$$
- Da I eine ZV ist, folgt
$$\Pr(X \geq a) = \mathbb{E}[I] \leq \frac{\mathbb{E}[X]}{a}.$$

Bsp:

- n-facher Münzwürf: Obere Ws-Schranke für mehr als $\frac{3}{4}n$-mal Kopf.
- Sei X_i IV für Kopf im i-ten Wurf. Definiere $X = X_1 + \ldots + X_n$.
- Es gilt $\mathbb{E}[X_i] = \frac{1}{2}$ und $\mathbb{E}[X] = \frac{n}{2}$. Markov-Ungleichung liefert damit
$$\Pr(X \geq \tfrac{3}{4}n) \leq \frac{n/2}{3n/4} = \frac{2}{3}.$$

Momente einer ZV

Definition k-tes Moment

Das *k-te Moment* einer ZV X ist $\mathbb{E}[X^k]$.

Definition Varianz

Die Varianz einer ZV X ist definiert als

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

Die Standardabweichung von X ist $\sigma[X] = \sqrt{\mathbf{Var}[X]}$.

Bsp:

- Sei $X = k$ konstant. Dann ist $\mathbf{Var}[X] = \mathbb{E}[(k - \mathbb{E}[k])^2] = 0$.
- Sei k eine Konstante und
$$X = \begin{cases} k\mathbb{E}[X] & \text{mit Ws } \frac{1}{k} \\ 0 & \text{mit Ws } 1 - \frac{1}{k} \end{cases}, \text{ d.h. } X^2 = \begin{cases} k^2(\mathbb{E}[X])^2 & \text{mit Ws } \frac{1}{k} \\ 0 & \text{mit Ws } 1 - \frac{1}{k} \end{cases}.$$
- Es folgt $\mathbf{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (k-1)(\mathbb{E}[X])^2$.
- D.h. $\mathbf{Var}[X]$ wird beliebig groß, falls X stark von $\mathbb{E}[X]$ abweicht.

Linearität + Kovarianz

Definition Kovarianz

Die *Kovarianz* von zwei ZV X, Y ist

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

Satz Linearität + Kovarianz

Für zwei ZV X, Y gilt

$$\mathbf{Var}[X + Y] = \mathbf{Var}[X] + \mathbf{Var}[Y] + 2\mathbf{Cov}(X, Y).$$

Beweis:

$$\begin{aligned}
\mathbf{Var}[X + Y] &= \mathbb{E}[(X + Y - \mathbb{E}[X + Y])^2] = \mathbb{E}[(X - \mathbb{E}[X] + Y - \mathbb{E}[Y])^2] \\
&= \mathbb{E}[(X - \mathbb{E}[X])^2 + (Y - \mathbb{E}[Y])^2 + 2(X - \mathbb{E}[X])(2Y - \mathbb{E}[Y])] \\
&= \mathbb{E}[(X - \mathbb{E}[X])^2] + E[(Y - \mathbb{E}[Y])^2] + 2\mathbb{E}[(X - \mathbb{E}[X])(2Y - \mathbb{E}[Y])] \\
&= \mathbf{Var}[X] + \mathbf{Var}[Y] + 2\mathbf{Cov}(X, Y).
\end{aligned}$$

Kovarianz

Satz Linearität + Kovarianz allgemein

Für ZV $X_1, \dots, X_n$ gilt allgemein

$$\mathbf{Var}[\sum_{i=1}^n X_i] = \sum_{i=1}^n \mathbf{Var}[X_i] + \sum_{1\le i<j\le n} 2\mathbf{Cov}(X_i, X_j).$$

Beweis: analog

Satz Multiplikativität von $\mathbb{E}$ für unabhängige ZV

Für unabhängige ZV X, Y gilt $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

Beweis:

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= \sum_i \sum_j (i \cdot j) \cdot \Pr((X = i) \cap (Y = j)) \\
&= \sum_i \sum_j (i \cdot j) \cdot \Pr(X = i) \cdot \Pr(Y = j) \\
&= (\sum_i i \cdot \Pr(X = i)) \cdot (\sum_j j \cdot \Pr(Y = j)) = \mathbb{E}[X] \cdot \mathbb{E}[Y]
\end{aligned}$$

Kovarianz

Anmerkung: Für abhängige X, Y gilt i. Allg. $\mathbb{E}[X \cdot Y] \neq \mathbb{E}[X] \cdot \mathbb{E}[Y]$.

- Wir betrachten einen 2-fachen Münzwurf.
- IV X für Kopf im 1. Wurf, ZV Y für Anzahl Kopf in beiden Würfen.
- Es gilt $\mathbb{E}[X] \cdot \mathbb{E}[Y] = \frac{1}{2} \cdot 1 = \frac{1}{2}$, aber $\mathbb{E}[X \cdot Y] = \frac{3}{4}$.

Satz Linearität von **Var** für unabhängige ZV

Für unabhängige ZV X, Y gilt

$$\mathbf{Cov}(X, Y) = 0 \text{ und damit } \mathbf{Var}[X + Y] = \mathbf{Var}[X] + \mathbf{Var}[Y].$$

Beweis:

$$\begin{aligned} \mathbf{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[X - \mathbb{E}[X]] \cdot \mathbb{E}[Y - \mathbb{E}[Y]] \\ &= (\mathbb{E}[X] - \mathbb{E}[X])(\mathbb{E}[Y] - \mathbb{E}[Y]) = 0 \end{aligned}$$

Korollar Linearität von **Var** für unabhängige ZV

Für paarweise unabhängige ZV $X_1, \ldots X_n$ gilt

$$\mathbf{Var}[\textstyle\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbf{Var}[X_i].$$

Chebyshev-Ungleichung

Bsp: Varianz der Binomialverteilung $B(n,p)$

- Seien $X_1, \dots, X_n$ ZV aus $B(1,p)$ und $X = X_1 + \ldots + X_n$. Es gilt $\mathbf{Var}[X_i] = \mathbb{E}[(X_i - \mathbb{E}[X_i])^2] = p(1-p)^2 + (1-p)(-p)^2 = p(1-p)$.
- Damit ist $\mathbf{Var}[X] = np(1-p)$.

Satz Chebyshev-Ungleichung

Für eine ZV X und alle $a > 0$ gilt

$$\Pr(|X - \mathbb{E}[X]| \geq a) \leq \frac{\mathbf{Var}[X]}{a^2}.$$

Beweis:

- Es gilt $\Pr(|X - \mathbb{E}[X]| \geq a) = \Pr((X - \mathbb{E}[X])^2 \geq a^2)$.
- Da $(X - \mathbb{E}[X])^2 > 0$ liefert die Markov-Ungleichung

$$\Pr((X - \mathbb{E}[X])^2 \geq a^2) \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2} = \frac{\mathbf{Var}[X]}{a^2}.$$

Chebyshev-Ungleichung

Korollar

Für eine ZV X und alle $t > 1$ gilt

1. $\Pr(|X - \mathbb{E}[X]| \geq t \cdot \sigma) \leq \frac{1}{t^2}$,
2. $\Pr(|X - \mathbb{E}[X]| \geq t \cdot \mathbb{E}[X]) \leq \frac{\mathbf{Var}[X]}{t^2(\mathbb{E}[X])^2}$.

Bsp: *n*-facher Münzwurf

- Schranke für die Ws, dass $\geq \frac{3}{4}n$-mal Kopf auftritt.
- Sei X_i IV für Kopf. Es gilt $\mathbb{E}[(X_i^2)] = \mathbb{E}[X_i] = \frac{1}{2}$ und damit
$$\mathbf{Var}[X_i] = \mathbb{E}[(X_i)^2] - (\mathbb{E}[X_i])^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}.$$
- Für $X = X_1 + \ldots + X_n$ folgt $\mathbf{Var}[X] = \frac{n}{4}$. Chebyshev liefert
$$\Pr(X \geq \tfrac{3}{4}n) \leq \Pr(|X - \mathbb{E}[X]| \geq \tfrac{n}{4}) \leq \frac{\mathbf{Var}[X]}{(n/4)^2} = \frac{4}{n}.$$

Varianz einer geometrischen ZV

Lemma Varianz einer geometrischen ZV

Sei X eine geometrische ZV mit Parameter p. Dann gilt $\mathbf{Var}[X] = \frac{1-p}{p^2}$.

Beweis:

- Sei X ein ZV für die Anzahl Würfe, bis zum 1. Mal Kopf aufritt.
- Sei Y IV für Kopf im 1. Wurf. Dann gilt

$$\begin{aligned} \mathbb{E}[X^2] &= \Pr(Y=0)\mathbb{E}[X^2 \mid Y=0] + \Pr(Y=1)\mathbb{E}[X^2 \mid Y=1] \\ &= (1-p)\mathbb{E}[(X+1)^2] + p = (1-p)\mathbb{E}[X^2] + 2(1-p)\mathbb{E}[X] + 1. \end{aligned}$$

- Einsetzen von $\mathbb{E}[X] = \frac{1}{p}$ und Auflösen liefert $\mathbb{E}[X^2] = \frac{2-p}{p^2}$. Es folgt
$$\mathbf{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Analyse von Coupon Collector

Frage: Ws bei Coupon Collector $\geq 2n\sum_{i=1}^{n}\frac{1}{i} = 2nH_n$-mal zu ziehen?

Vergleich von Markov, Chebyshev und Union Bound:

- Markov liefert
$$\Pr(X \geq 2nH_n) \leq \frac{\mathbb{E}[X]}{2nH_n} = \frac{nH_n}{2nHn} = \frac{1}{2}.$$
- Wir verwenden die (vereinfachte) Ungleichung $\mathbf{Var}[X_i] \leq \frac{1}{p^2}$:
$$\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i] \leq \sum_{i=1}^{n} (\tfrac{n}{n-i+1})^2 = n^2 \sum_{i=1}^{n} \tfrac{1}{i^2} \leq n^2 \tfrac{\pi^2}{6}.$$
- Chebyshev liefert damit
$$\Pr(|X - \mathbb{E}[X]| \geq nH_n) \leq \frac{\mathbf{Var}[X]}{n^2 H_n^2} \leq \frac{\pi^2}{6H_n^2} = \mathcal{O}(\frac{1}{\ln^2 n}).$$
- Ereignis E_i: Element i ist nach $2\ln n$ Schritten nicht gezogen.
$$\Pr(E_i) = (1 - \tfrac{1}{n})^{2n\ln n} < e^{-2\ln n} = \tfrac{1}{n^2}.$$
- Ereignis E: Irgendein Element nach $2\ln n$ Schritten nicht gezogen.
$$\Pr(E) \leq \sum_{i=1}^{n} \Pr(E_i) < n\tfrac{1}{n^2} = \tfrac{1}{n}.$$
- D.h. Union Bound ist hier besser als Markov und Chebyshev.

Randomisierter Median

Problem Median-Berechung

Gegeben: $S \subset \mathbb{Z}$ mit $|S| = n$ ungerade
Gesucht: Median m von S, d.h. das $\frac{n+1}{2}$-te Element in Sortierung

Anmerkungen:

- 1. Lösung: Sortiere in $\mathcal{O}(n \log n)$ und gebe $\frac{n+1}{2}$-tes Element aus.
- $\exists$ komplexer deterministischer Algorithmus mit Laufzeit $\mathcal{O}(n)$.
- Idee eines probabilistischen Ansatzes:
  - Sample eine (kleine) Menge R, so dass $\ell, u \in R$ mit $\ell \leq m \leq u$.
  - Betrachte die Menge $C = \{s \in S \mid \ell \leq s \leq u\}$ mit $|C| = o(\frac{n}{\log n})$.
  - Bestimme die Anzahl x_ℓ der Element in S, die kleiner als ℓ sind.
  - Sortiere C und bestimme das $(\frac{n+1}{2} - x_\ell)$-kleinste Element in C.
- Führt zu einfachem $\mathcal{O}(n)$-Algorithmus mit besserer $\mathcal{O}$-Konstante. (im Vergleich zum oben erwähnten deterministischen Algorithmus)

Randomisierter Median

Algorithmus MEDIAN

EINGABE: S mit $|S| = n$ ungerade

1. Wähle Multimenge $R \subset_R S$, $|R| = n^{\frac{3}{4}}$ (Ziehen mit Zurücklegen).
2. Sortiere R und setze ℓ, u auf das $(\frac{1}{2}n^{\frac{3}{4}} \pm \sqrt{n})$-kleinste Element.
3. Berechne $C = \{s \in S \mid \ell \leq s \leq u\}$.
4. Berechne $x_\ell = |\{s \in S \mid s < \ell\}|$ und $x_u = |\{s \in S \mid s > u\}|$.
5. Falls $|C| > 4n^{\frac{3}{4}}$, $x_\ell > \frac{n}{2}$ oder $x_u > \frac{n}{2}$, Ausgabe FAIL.
6. Sortiere C und gib das $(\frac{n+1}{2} - x_\ell)$-kleinste Element in C aus.

AUSGABE: Median m von S

Korrektheit von MEDIAN

Satz Korrektheit von MEDIAN

MEDIAN gibt in $\mathcal{O}(n)$ entweder den Median oder FAIL aus.

Beweis:

- Der Median ist in C gdw $x_\ell \leq \frac{n}{2}$ und $x_u \leq \frac{n}{2}$.
- Die Laufzeit in Schritt 3 und 4 ist jeweils $\mathcal{O}(n)$.
- $|C| \leq 4n^{\frac{3}{4}}$ sichert, dass Schritt 6 nur Zeit $o(n)$ benötigt. $\square$

Wir erhalten FAIL für eines der folgenden Ereignisse:

1. $E_1 : Y_1 = |\{r \in R \mid r \leq m\}| < \frac{1}{2}n^{\frac{3}{4}} - \sqrt{n}$
   Für dieses Ereignis gilt $\ell > m$ und damit $x_\ell > \frac{n}{2}$.
2. $E_2 : Y_2 = |\{r \in R \mid r \geq m\}| < \frac{1}{2}n^{\frac{3}{4}} - \sqrt{n}$
   Für dieses Ereignis gilt $u < m$ und damit $x_u > \frac{n}{2}$.
3. $E_3 : |C| > 4n^{\frac{3}{4}}$

Ws von FAIL

Lemma Ws von E_1 (E_2 analog)

$\Pr(E_1) \leq \frac{1}{4}n^{-\frac{1}{4}}$

Beweis:

- IV $X_i = 1$, falls das i-te Element in R kleiner oder gleich m ist.

$$\Pr(X_i = 1) = \frac{\frac{n+1}{2}}{n} = \frac{1}{2} + \frac{1}{2n}$$

- ZV $Y_1 = \sum_{i=1}^{n^{\frac{3}{4}}} X_i$ ist verteilt gemäß $B(n', p) = B(n^{\frac{3}{4}}, \frac{1}{2} + \frac{1}{2n})$. D.h.

$$\mathbf{Var}[Y_1] = n'p(1-p) = n^{\frac{3}{4}}(\tfrac{1}{2} + \tfrac{1}{2n})(\tfrac{1}{2} - \tfrac{1}{2n}) < \tfrac{1}{4}n^{\frac{3}{4}}.$$

- Chebyshev liefert

$$\begin{aligned} \Pr(E_1) &= \Pr(Y_1 < \frac{1}{2}n^{\frac{3}{4}} - \sqrt{n}) \\ &\leq \Pr(|Y_1 - \mathbb{E}[Y_1]| > \sqrt{n}) \leq \frac{\mathbf{Var}[Y_1]}{n} < \frac{1}{4}n^{-\frac{1}{4}}. \square \end{aligned}$$

Ws von FAIL

Lemma Ws von E_3

$\Pr(E_3) \leq \frac{1}{2}n^{-\frac{1}{4}}$

Beweis:

- Falls $|C| > 4n^{\frac{3}{4}}$, sind $> 2n^{\frac{3}{4}}$ Elemente in C größer/kleiner als m.
- Ereignis $E_>$: Mehr als $2n^{\frac{3}{4}}$ Elemente in C sind größer als m.
- D.h. die Ordnung von u in S's Sortierung ist mindestens $\frac{1}{2}n + 2n^{\frac{3}{4}}$, bzw. $\geq \frac{1}{2}n^{\frac{3}{4}} - \sqrt{n}$ Elem. in R sind unter den $\frac{1}{2}n - 2n^{\frac{3}{4}}$ größten in S.
- Sei X ZV für die $(\frac{1}{2}n - 2n^{\frac{3}{4}})$-größten Elemente von S in R.
- X ist gemäß $B(n^{\frac{3}{4}}, \frac{1}{2} - 2n^{-\frac{1}{4}})$ verteilt mit $\mathbb{E}[X] = \frac{1}{2}n^{\frac{3}{4}} - 2\sqrt{n}$ und
$$\mathbf{Var}[X] = n^{\frac{3}{4}}(\tfrac{1}{2} - 2n^{-\frac{1}{4}})(\tfrac{1}{2} + 2n^{-\frac{1}{4}}) < \tfrac{1}{4}n^{\frac{3}{4}}.$$
- Chebyshev liefert $\Pr(E_>) =$
$$\Pr(X \geq \tfrac{1}{2}n^{\frac{3}{4}} - \sqrt{n}) \leq \Pr(|X - \mathbb{E}[X]| \geq \sqrt{n}) \leq \frac{\mathbf{Var}[X]}{n} < \tfrac{1}{4}n^{-\frac{1}{4}}.$$
- Analog $\Pr(E_<) < \frac{1}{4}n^{-\frac{1}{4}}$ und $\Pr(E_3) \leq \Pr(E_>) + \Pr(E_<) < \frac{1}{2}n^{-\frac{1}{4}}$.

Monte Carlo und Las Vegas

Korollar Ws von FAIL

MEDIAN gibt FAIL mit Ws $\geq n^{-\frac{1}{4}}$ aus.

Definition Monte Carlo Algorithmus

Ein *Monte Carlo Algorithmus* ist ein probabilistischer Algorithmus, der FAIL oder eine inkorrekte Ausgabe liefern kann.

Definition Las-Vegas Algorithmus

Ein *Las Vegas Algorithmus* ist ein probabilistischer Algorithmus, der stets die korrekte Antwort liefert.

Anmerkung: Beim Monte Carlo Alg. ist die Laufzeit im Gegensatz zum Las Vegas Alg. typischerweise keine Zufallsgröße.

Übung: Wandeln Sie MEDIAN von einem Monte-Carlo Alg. mit Laufzeit $\mathcal{O}(n)$ in einen Las-Vegas Alg. mit erwarteter Laufzeit $\mathcal{O}(n)$.

Moment Erzeugendenfunktion

Definition Moment Erzeugendenfunktion

Die *Moment Erzeugendenfunktion* einer ZV X ist $M_X(t) = \mathbb{E}[e^{tX}]$.

Anmerkungen:

- Uns interessiert $M_X(t)$ für t in der Nähe von Null.
- Annahme im Folgenden: Wir können die Operatoren für Erwartungswert und Ableitung vertauschen.
  (korrekt für alle von uns betrachteten Verteilungen)
- Sei $M_X^{(n)}(0)$ die n-te Ableitung von $M_X(t)$ an der Stelle $t = 0$.

Satz $M_X(t)$ beschreibt alle Momente von X

Für alle $n \geq 1$ gilt $\mathbb{E}[X^n] = M_X^{(n)}(0)$.

Beweis:

- Vertauschen von Ableitung und $\mathbb{E}$ liefert $M_X^{(n)}(t) = \mathbb{E}[X^n e^{tX}]$.
- Damit folgt $M_X^{(n)}(0) = \mathbb{E}[X^n]$.

Moment Erzeugendenfunktion

Bsp: Geometrische ZV X mit Parameter p.

- Es gilt
  $M_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=1}^{\infty}(1-p)^{k-1}pe^{tk} = \frac{p}{1-p}\sum_{k=1}^{\infty}((1-p)e^t)^k$.
- Für $t < \ln(\frac{1}{1-p})$ folgt $M_X(t) = \frac{p}{1-p}\left(\frac{1}{1-(1-p)e^t} - 1\right)$.
- Ableiten nach t liefert $M_X^{(1)}(t) = \frac{pe^t}{(1-(1-p)e^t)^2}$.
- Auswerten an der Stelle $t = 0$ ergibt $\mathbb{E}[X] = \frac{p}{p^2} = \frac{1}{p}$.
- Analog folgt, dass $\mathbb{E}[X^2] = \frac{2-p}{p^2}$. (Übung)

Satz Momente einer ZV determinieren Verteilung

Seien X, Y ZV. Falls für ein $\delta > 0$

$$M_X(t) = M_Y(t) \text{ für alle } t \in (-\delta, \delta),$$

dann besitzen X und Y dieselbe Verteilung.

(ohne Beweis)

Multiplikativität der Erzeugendenfunktion

Satz Multiplikativität der Erzeugendenfunktion

Seien X, Y unabhängige ZV. Dann gilt $M_{X+Y}(t) = M_X(t)M_Y(t)$.

Beweis:

$$M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}e^{tY}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = M_X(t)M_Y(t) \; \square$$

Anwendung:

- Sei $M_X(t)M_Y(t)$ die Moment Erzeugendenfkt einer Verteilung D.
- Dann muss D die Verteilung $X+Y$ sein.

Herleitung von Chernoff Schranken:

- Aus der Markov Ungleichung folgt für alle $t > 0$
$$\Pr(X \geq a) = \Pr(e^{tX} \geq e^{ta}) \leq \frac{\mathbb{E}[e^{tX}]}{e^{ta}}.$$
- Daraus folgt insbesondere $\Pr(X \geq a) \leq \min_{t>0} \frac{\mathbb{E}[e^{tX}]}{e^{ta}}$.
- Man wählt nun für die gewünschte Verteilung ein geeignetes t.
- Oft ist man an einer gut handhabbaren Schranke interessiert.

Poisson Proben

$M_X(t)$ für Poisson Proben:

- Seien $X_1, \ldots, X_n$ unabhängige 0-1 ZV mit $\Pr(X_i = 1) = p_i$. (sogenannte *Poisson Proben*)
- Sei $X = \sum_{i=1}^n X_i$ mit $\mu = \mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n p_i$.
- Für die Moment Erzeugendenfunktion von X_i gilt
$$M_{X_i}(t) = \mathbb{E}[e^{tX_i}] = p_i e^t + (1 - p_i) = 1 + p_i(e^t - 1) \leq e^{p_i(e^t-1)}.$$
- Mit Hilfe der Multiplikativität von $M_X(t)$ folgt

$$\begin{aligned}
\mathbb{E}[e^{tX}] = M_X(t) &= \prod_{i=1}^n M_{X_i}(t) \\
&\leq \prod_{i=1}^n e^{p_i(e^t-1)} = e^{\sum_{i=1}^n p_i(e^t-1)} = e^{(e^t-1)\mu}.
\end{aligned}$$

Chernoff Schranken

Satz Chernoff Schranken

Seien $X_1, \ldots, X_n$ unabhängige Poisson Proben mit $\Pr(X_i = 1) = p_i$. Sei $X = \sum_{i=1}^n X_i$ und $\mu = \mathbb{E}[X]$. Dann gilt

1. für alle $\delta > 0$: $\Pr(X \geq (1+\delta)\mu) < \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$.
2. für $0 \leq \delta \leq 1$: $\Pr(X \geq (1+\delta)\mu) \leq e^{-\frac{\mu\delta^2}{3}}$.
3. für $R \geq 6\mu$: $\Pr(X \geq R) \leq 2^{-R}$.

Beweis:

- Aus der Markov Ungleichung folgt

$$\Pr(X \geq (1+\delta)\mu) = \Pr(e^{tX} \geq e^{t(1+\delta)\mu}) \leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu)}} \leq \left(\frac{e^{(e^t-1)}}{e^{t(1+\delta)}}\right)^\mu .$$

- Für $\delta > 0$ können wir $t = \ln(1+\delta) > 0$ wählen. Aussage 1 folgt.
- Für Aussage 2 kann man die Ungleichung $\frac{e^\delta}{(1+\delta)^{1+\delta}} \leq e^{-\frac{\delta^2}{3}}$ zeigen.
- Für 3. sei $R = (1+\delta)\mu$. Für $R \geq 6\mu$ folgt $1+\delta \geq 6$ und

$$\Pr(X \geq (1+\delta)\mu) < \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu \leq \left(\frac{e}{(1+\delta)}\right)^{(1+\delta)\mu} \leq \left(\frac{e}{6}\right)^R \leq 2^{-R}.$$

Chernoff Schranken

Satz Chernoff Schranken (Abweichung nach unten)

Seien $X_1, \ldots, X_n$ unabhängige Poisson Proben mit $\Pr(X_i = 1) = p_i$.
Sei $X = \sum_{i=1}^n X_i$ und $\mu = \mathbb{E}[X]$. Dann gilt

1. für alle $\delta > 0$: $\Pr(X \leq (1-\delta)\mu) < \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu}$.
2. für $0 \leq \delta \leq 1$: $\Pr(X \leq (1-\delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$.

Beweis: analog zum Beweis zuvor.

Korollar Chernoff Schranke

Seien $X_1, \ldots, X_n$ unabhängige Poisson Proben mit $\Pr(X_i) = p_i$.
Sei $X = \sum_{i=1}^n X_i$ und $\mu = \mathbb{E}[X]$. Dann gilt für $0 \leq \delta \leq 1$:

$$\Pr(|X - \mu| \geq \delta\mu) \leq 2e^{-\frac{\mu\delta^2}{3}}.$$

Anwendung Chernoff

n-facher Münzwurf:

- Sei X die Anzahl von Köpfen bei n Münzwürfen. Mit Chernoff folgt
$$\Pr(|X - \tfrac{n}{2}| \geq \tfrac{1}{2}\sqrt{6n\ln n}) \leq 2e^{-\frac{1}{3}\frac{n}{2}\frac{6\ln n}{n}} = \frac{2}{n}.$$
- D.h. die Abweichung vom Mittelwert ist meist $\mathcal{O}(\sqrt{n\ln n})$.
- Mit Chebychev hatten wir $\Pr(|X - \frac{n}{2}| \geq \frac{n}{4}) \leq \frac{4}{n}$.
- Chernoff liefert die deutlich bessere exponentielle Schranke
$$\Pr(|X - \tfrac{n}{2}| \geq \tfrac{n}{4}) \leq 2e^{-\frac{1}{3}\frac{n}{2}\frac{1}{4}} = 2e^{-\frac{n}{24}}.$$

Chernoff: Spezialfälle

Satz Chernoff für ± 1-ZV

Seien $X_1, \ldots, X_n$ unabhängige ZV mit $\Pr(X_i = 1) = \Pr(X_i = (-1)) = \frac{1}{2}$.
Sei $X = \sum_{i=1}^{n} X_i$. Für alle $a > 0$ gilt $\Pr(X \geq a) \leq e^{-\frac{a^2}{2n}}$.

Beweis:

- Mit Taylor-Entwicklung $e^{-t} = 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \ldots$ gilt für alle $t > 0$
  $$\mathbb{E}[e^{tX_i}] = \frac{1}{2}e^{t} + \frac{1}{2}e^{-t} = \sum_{i \geq 0} \frac{t^{2i}}{(2i)!} \leq \sum_{i \geq 0} \frac{t^{2i}}{2^i i!} = \sum_{i \geq 0} \frac{(t^2/2)^i}{i!} = e^{\frac{t^2}{2}}.$$
- Es folgt aus der Unabhängigkeit $\mathbb{E}[e^{tX}] = \prod_{i=1}^{n} \mathbb{E}[e^{tX_i}] \leq e^{\frac{t^2 n}{2}}$ und
  $$\Pr(X \geq a) \leq \frac{\mathbb{E}[e^{tX}]}{e^{ta}} \leq e^{\frac{t^2 n}{2} - ta}.$$
- Setzung von $t = \frac{a}{n}$ liefert $\Pr(X \geq a) \leq e^{-\frac{a^2}{2n}}$. $\square$

Korollar Chernoff für ± 1-ZV

Seien $X_1, \ldots, X_n$ unabhängige ZV mit $\Pr(X_i = 1) = \Pr(X_i = (-1)) = \frac{1}{2}$.
Sei $X = \sum_{i=1}^{n} X_i$. Für alle $a > 0$ gilt $\Pr(|X| \geq a) \leq 2e^{-\frac{a^2}{2n}}$.

Chernoff: Spezialfälle

Satz Chernoff für 0, 1-ZV

Seien $Y_1, \ldots, Y_n$ unabhängige ZV mit $\Pr(Y_i = 1) = \Pr(Y_i = 0) = \frac{1}{2}$. Sei $Y = \sum_{i=1}^n Y_i$ und $\mu = \mathbb{E}[Y] = \frac{n}{2}$. Es gilt

1. für alle $a > 0$: $\Pr(Y \geq \mu + a) \leq e^{-\frac{2a^2}{n}}$.
2. für alle $\delta > 0$: $\Pr(Y \geq (1+\delta)\mu) \leq e^{-\delta^2 \mu}$.

Beweis:

- Mit der Ersetzung $Y_i = \frac{X_i+1}{2}$ erhalten wir
$$Y = \sum_{i=1}^n Y_i = \sum_{i=1}^n \frac{X_i+1}{2} = (\frac{1}{2}\sum_{i=1}^n X_i) + \frac{n}{2} = \frac{1}{2}X + \mu.$$
- Aus voriger Folie folgt $\Pr(Y \geq \mu + a) = \Pr(X \geq 2a) \leq e^{-\frac{2a^2}{n}}$.
- Mit der Setzung $a = \delta\mu$ folgt analog
$$\Pr(Y \geq (1+\delta)\mu) = \Pr(X \geq 2\delta\mu) \leq e^{-\frac{2\delta^2\mu^2}{n}} = e^{-\delta^2\mu}. \quad \square$$

Korollar

Wir gelten selbige Schranken für $\Pr(Y \leq \mu - a)$ und $\Pr(Y \leq (1-\delta)\mu)$.

Mengen Balancierung

Problem Mengen Balancierung

Gegeben: $A \in \mathbb{F}_2^{n\times m}$
Gesucht: $\mathbf{b} \in \{-1,1\}^m$, so dass $||A\mathbf{b}||_\infty$ minimal ist.

Satz

Für zufällige $b \in_R \{-1,1\}^m$ gilt $\Pr(||A\mathbf{b}||_\infty \geq \sqrt{4m\ln n}) \leq \frac{2}{n}$.

Beweis:

- Sei $A\mathbf{b} = \mathbf{c}$ und k die Anzahl Einsen in der i-ten Zeile $\mathbf{a}_i$ von A.
- Falls $k \leq \sqrt{4m\ln n}$ dann gilt $c_i \leq \sqrt{4m\ln n}$. Sei also $k > \sqrt{4m\ln n}$.
- Die k Nicht-Null Terme in $\langle \mathbf{a}_i, \mathbf{b}\rangle$ sind unabhängige (± 1)-ZV X_i.
- Es gilt $\Pr(X_i = 1) = \Pr(X_i = (-1)) = \frac{1}{2}$. Mittels Chernoff folgt

$$\Pr(|\langle \mathbf{a}_i, \mathbf{b}\rangle| \geq \sqrt{4m\ln n}) \leq 2e^{-\frac{4m\ln n}{2k}} \leq \frac{2}{n^2}, \text{ wegen } k \leq m. \quad \square$$

Übung: Es existieren A, für die $||A\mathbf{b}|| = \Omega(\sqrt{n})$ für alle $\mathbf{b}$.

Geburtstags-Paradoxon (informal)

Problem Geburtstags-Paradoxon

Gegeben: n mögliche Geburtstage

Gesucht: m Personen, so dass 2 Personen am selben Tag Geburtstag mit Ws $\geq \frac{1}{2}$ haben.

Analyse:

- m Personen haben verschiedene Geburtstage mit Ws
$$(1-\tfrac{1}{n})\cdot(1-\tfrac{2}{n})\cdot\ldots\cdot(1-\tfrac{m-1}{n}) = \prod_{j=1}^{m-1}(1-\tfrac{j}{n}).$$
- Aus der Taylorentwicklung von e^x folgt $1-\frac{k}{n} \approx e^{-\frac{k}{n}}$ für $k \ll n$. D.h.
$$\prod_{j=1}^{m-1}(1-\tfrac{j}{n}) \approx \prod_{j=1}^{m-1} e^{-\frac{j}{n}} = e^{-\sum_{j=1}^{m-1}\frac{j}{n}} \approx e^{-\frac{m^2}{2n}}.$$
- Wir erhalten $e^{-\frac{m^2}{2n}} = \frac{1}{2}$ für $m = \sqrt{2n\ln 2}$.
- Approximation liefert für $n = 365$ den Wert $m \approx 22.49$.

Das Bälle-Urnen Modell

Definition Bälle-Urnen Modell

Im *Bälle-Urnen Modell* werfen wir m Bälle in n Urnen.

Interessante Fragestellungen:

- Wieviele Urnen bleiben leer?
- Wieviele Bälle sind in der vollsten Urne?
- Wann enthält eine Urne mehr als einen Ball? (Geburtstags-Paradoxon)
- Wann enthalten alle Urnen mindestens einen Ball? (Coupon Collector)

Maximale Beladung

Satz Maximale Beladung einer Urne

Für n Bälle in n Urnen und hinreichend große n enthält kein Urne mehr als $3\frac{\ln n}{\ln\ln n}$ Bälle mit Ws höchstens $\frac{1}{n}$.

Beweis:

- Ereignis E_i: Urne i enthält M Bälle.
- Es existieren $\binom{n}{M}$ Mengen mit M Bällen.
- Jede dieser Mengen ist mit Ws $(\frac{1}{n})^M$ komplett in Urne i, d.h.
$$\Pr(E_i) \le \binom{n}{M}(\tfrac{1}{n})^M \le \tfrac{1}{M!}.$$
- Es gilt $e^k = \sum_{i=0}^{\infty} \frac{k^i}{i!} > \frac{k^k}{k!}$, d.h. $k! > (\frac{k}{e})^k$. Damit $\Pr(E_i) \le (\frac{e}{M})^M$.
- Für $M \le 3\frac{\ln n}{\ln\ln n}$ gilt für hinreichend große n

$$\begin{aligned}\Pr(E) &\le \Pr(E_1) + \ldots + \Pr(E_n) \le n\left(\frac{e}{M}\right)^M \le n\left(\frac{e\ln\ln n}{3\ln n}\right)^{3\frac{\ln n}{\ln\ln n}} \\ &\le n\left(\frac{\ln\ln n}{\ln n}\right)^{3\frac{\ln n}{\ln\ln n}} = e^{\ln n}(e^{\ln\ln\ln n - \ln\ln n})^{3\frac{\ln n}{\ln\ln n}} = \frac{1}{n^2}\cdot n^{o(1)} \le \frac{1}{n}.\end{aligned}$$

Anwendung BUCKET SORT

Algorithmus BUCKET SORT

EINGABE: $n = 2^m$ Zahlen $x_1, \ldots, x_n \in_R [0, 2^k)$ mit $k \geq m$.

1. For $i = 1$ to n: Sortiere x_i in Bucket $MSB_m(x_i)$. (m oberste Bits)
2. For $i = 0$ to $n - 1$: Sortiere Bucket i aufsteigend mit INSERTION SORT.

AUSGABE: Zahlen in den Buckets $0, \ldots, n - 1$

Korrektheit:

- Schritt 1: Elemente in Bucket i sind kleiner als die in Bucket $i + 1$.
- Schritt 2: Zusätzliche Sortierung der Elemente pro Bucket.

Analyse BUCKET SORT

Satz Laufzeit BUCKET SORT

BUCKET SORT läuft in erwarteter Zeit $\mathcal{O}(n)$.

Beweis:

- Die Zahlen entsprechen Bällen, die Buckets entsprechen Urnen.
- Schritt 1 läuft in deterministischer Zeit $\mathcal{O}(n)$.
- ZV X_i für die Anzahl Zahlen in Bucket i.
- Die Laufzeit für Bucket i ist höchstens cX_i^2 für eine Konstante c.
- Damit ist die erwartete Laufzeit von Schritt 2 höchstens
$$\mathbb{E}[\sum_{i=0}^{n-1} c(X_i)^2] = c\sum_{i=0}^{n-1} \mathbb{E}[X_i^2] = cn\mathbb{E}[X_0^2].$$
- Da $X_0 \sim B(n, \frac{1}{n})$ wissen wir bereits
$$\mathbb{E}[X_0^2] = n(n-1)p^2 + np = \frac{n(n-1)}{n^2} + 1 < 2.$$
- Damit läuft Schritt in erwarteter Zeit $\mathcal{O}(n)$ $\quad\square$.

Die Poisson Verteilung

Motivation:

- Wir betrachten den Besetzungsgrad von Urnen.
- ZV $X_i = 1$ gdw die j-te Urne leer ist. D.h. $\mathbb{E}[X_i] = (1 - \frac{1}{n})^m \approx e^{-\frac{m}{n}}$.
- Es folgt für die Anzahl X leerer Urnen
$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n(1 - \frac{1}{n})^m \approx ne^{-\frac{m}{n}}.$$
- D.h. der relative Anteil leerer Urnen ist approximativ $e^{-\frac{m}{n}}$.
- Generell: Ws, dass eine feste Urne genau j Bälle enthält, ist
$$p_j = \binom{m}{j}(\frac{1}{n})^j(1 - \frac{1}{n})^{m-j} = \frac{1}{j!}\frac{m(m-1)\ldots(m-j+1)}{n^j}(1 - \frac{1}{n})^{m-j} \approx \frac{e^{-\frac{m}{n}}(\frac{m}{n})^j}{j!}.$$

Die Poisson Verteilung

Definition Poisson Verteilung

Eine ZV X ist *Poisson* verteilt mit Parameter μ, falls für alle $j \geq 0$

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

Anmerkungen:

- Ws-Verteilung: $\sum_{j\geq 0} \Pr(X = j) = e^{-\mu} \sum_{j\geq 0} \frac{\mu^j}{j!} = e^{-\mu} e^{\mu} = 1.$
- Für den Erwartungswert einer Poisson verteilten ZV X gilt
$$\mathbb{E}[X] = \sum_{j\geq 1} j \frac{e^{-\mu}\mu^j}{j!} = \mu \sum_{j\geq 1} \frac{e^{-\mu}\mu^{j-1}}{(j-1)!} = \mu \sum_{j\geq 0} \frac{e^{-\mu}\mu^j}{j!} = \mu.$$
- Bälle-Urnen: Die Verteilung ist approximativ Poisson mit $\mu = \frac{m}{n}$.
- $\mu = \frac{m}{n}$ entspricht der durchschnittlichen Belegung der Urnen.

Summe unabhängiger Poisson ZV

Satz Moment Erzeugendenfunktion einer Poisson ZV

Sei eine ZV X Poisson verteilt mit μ. Dann gilt $M_X(t) = e^{\mu(e^t-1)}$.

Beweis:

$$\mathbb{E}[e^{tX}] = \sum_{j\geq 0} \frac{e^{-\mu}\mu^j}{j!} e^{tj} = e^{-\mu} \sum_{j\geq 0} \frac{(\mu e^t)^j}{j!} = e^{-\mu} e^{\mu e^t} = e^{\mu(e^t-1)} \quad \square$$

Satz Summe unabhängiger Poisson ZV

Die endliche Summe unabhängiger Poisson ZV ist eine Poisson ZV.

Beweis:

- Wir betrachten nur die Summer zweier ZV. Der Satz folgt induktiv.
- Seien X, Y ZV mit Erwartungswerten μ_1, μ_2.
- Wir erhalten $M_{X+Y}(t) = M_X(t) \cdot M_Y(t) = e^{(\mu_1+\mu_2)(e^t-1)}$.
- Dies ist Erzeugendenfkt einer Poisson ZV mit Parameter $\mu_1 + \mu_2$.
- Mit Folie 47 ist $X + Y$ damit eine ZV mit Erwartungswert $\mu_1 + \mu_2$. $\square$

Chernoff Schranke für Poisson ZV

Satz Chernoff Schranke für Poisson ZV

Sei X eine Poisson ZV mit Parameter μ. Dann gilt

1. für $x > \mu$: $\Pr(X \geq x) \leq \frac{e^{-\mu}(e\mu)^x}{x^x}$.
2. für $x < \mu$: $\Pr(X \leq x) \leq \frac{e^{-\mu}(e\mu)^x}{x^x}$.

Beweis:

- Wir zeigen nur die 1. Ungleichung, die 2. folgt analog. Es gilt

$$\Pr(X \geq x) = \Pr(e^{tX} \geq e^{tx}) \leq \frac{\mathbb{E}[e^{tX}]}{e^{tx}} \leq e^{\mu(e^t - 1) - xt}.$$

- Für die Wahl $t = \ln(\frac{x}{\mu}) > 0$ folgt

$$\Pr(X \geq x) \leq e^{x - \mu - x\ln(\frac{x}{\mu})} = e^{x - \mu - x\ln x + x\ln\mu} = \frac{e^{-\mu}(e\mu)^x}{x^x}. \qquad \square$$

Poisson als Grenzwert der Binomial-Verteilung

Satz Poisson ist Grenzwert der Binomial-Verteilung für kleine p

Sei $X_n \sim B(n,p)$, wobei $\lim_{n\to\infty} np = \lambda$ konstant ist. Dann gilt

$$\lim_{n\to\infty} \Pr(X_n = k) = \frac{e^{-\lambda}\lambda^k}{k!} \text{ für alle festen } k.$$

Beweisskizze:

- Es gilt unter Verwendung von $1 + x \leq e^x$ für $|x| \leq 1$
$$\Pr(X_n = k) = \binom{n}{k} p^k (1-p)^{n-k} \leq \frac{n^k}{k!} p^k \frac{(1-p)^n}{(1-p)^k} \leq \frac{(np)^k}{k!} \frac{e^{-pn}}{1-pk}.$$
- Wegen $\lim_{n\to\infty} np = \lambda$ folgt $\lim_{n\to\infty} p = \frac{\lambda}{n} = 0$. Damit gilt
$$\lim_{n\to\infty} \Pr(X_n = k) \leq \frac{e^{-pn}(np)^k}{k!} = \frac{e^{-\lambda}\lambda^k}{k!}.$$
- Ähnlich kann man $\lim_{n\to\infty} \Pr(X_n = k) \geq \frac{e^{-\lambda}\lambda^k}{k!}$ zeigen. $\square$

Poisson Approximation für Balle und Urnen

Bälle-Urnen Modell:

- ZV $X_i^{(m)}$, $i = 1, \ldots, n$, der Bälle pro Urne sind nicht unabhängig.
- Z.B. gilt offenbar $X_n^{(m)} = m - \sum_{i=1}^{n-1} X_i^{(m)}$.
- Wir würden gerne die $X_i^{(m)}$ als **unabhängige** Poisson-ZV $Y_i^{(m)}$, $i = 1, \ldots, n$, mit $\mu = \frac{m}{n}$ behandeln (*Poisson-Fall*).

Lemma Poisson versus exakt

Die Verteilung $(Y_1^{(m)}, \ldots, Y_n^{(m)})$ eingeschränkt auf $\sum_i Y_i^{(m)} = k$ ist identisch zur Verteilung $(X_1^{(k)}, \ldots, X_n^{(k)})$, unabhängig von m.

Beweis:

- Für $\sum_i k_i = k$: $p_1 = \Pr((X_1^{(k)}, \ldots, X_n^{(k)}) = (k_1, \ldots, k_n)) = \frac{\binom{k}{k_1, \ldots, k_n}}{n^k}$.
- Für $p_2 = \Pr((Y_1^{(m)}, \ldots, Y_n^{(m)}) = (k_1, \ldots, k_n) \mid \sum_i Y_i^{(m)} = k)$ gilt

$$p_2 = \frac{\Pr((Y_1^{(m)}=k_1) \cap \ldots \cap (Y_1^{(m)}=k_n) \cap (\sum_i k_i = k))}{\sum_{i=1}^n Y_i^{(m)}=k} = \frac{\prod_{i=1}^n e^{-\frac{m}{n}} (\frac{m}{n})^{k_i}/(k_i!)}{e^{-m} m^k / k!} = p_1. \square$$

Poisson versus exakt

Satz Poisson versus exakt

Sei $f(x_1, \ldots, x_n)$ eine nicht-negative Funktion. Dann gilt

$$\mathbb{E}[f(X_1^{(m)}, \ldots, X_n^{(m)})] \leq e\sqrt{m} \cdot \mathbb{E}[f(Y_1^{(m)}, \ldots, Y_n^{(m)})].$$

Beweis: Unter Verwendung der Abschätzung $m! \leq e\sqrt{m}(\frac{m}{e})^m$ gilt

$$\begin{aligned}
\mathbb{E}[f(Y_1^{(m)}, \ldots, Y_n^{(m)})] &= \sum_{k=0}^{\infty} \mathbb{E}[f(Y_1^{(m)}, \ldots, Y_n^{(m)}) \mid \sum_{i=1}^{n} Y_i^{(m)} = k] \cdot \Pr(\sum_{i=1}^{n} Y_i^{(m)} = k) \\
&\geq \mathbb{E}[f(Y_1^{(m)}, \ldots, Y_n^{(m)}) \mid \sum_{i=1}^{n} Y_i^{(m)} = m] \cdot \Pr(\sum_{i=1}^{n} Y_i^{(m)} = m) \\
&= \mathbb{E}[f(X_1^{(m)}, \ldots, X_n^{(m)})] \cdot \frac{e^{-m} m^m}{m!} \\
&\geq \frac{1}{e\sqrt{m}} \cdot \mathbb{E}[f(X_1^{(m)}, \ldots, X_n^{(m)})]. \quad \square
\end{aligned}$$

Poisson versus exakt

Korollar Poisson versus exakt

Jedes Ereignis E, dass Ws p im Poisson-Fall besitzt,
besitzt Ws $\leq e\sqrt{m}p$ im exakten Bälle/Urnen-Fall.

Beweis: Sei f die Indikatorfunktion von E. Dann ist $\mathbb{E}[f] = \Pr(E)$. $\square$

Übung: Für spezielle f sind noch bessere Schranken möglich.

Untere Schranke für maximale Beladung

Satz Untere Schranke für maximale Beladung

Wir werfen n Bälle in n Urnen. Dann besitzt für hinreichend große n eine Urne mindestens $\frac{\ln n}{\ln \ln n}$ Bälle mit Ws $\geq 1 - \frac{1}{n}$.

Beweis:

- Modelliere Anzahl X_i der Bälle in Urne i als Poisson-ZV mit $\mu = 1$.
- Es gilt $\Pr(X_i \geq M) \geq \frac{e^{-\mu}\mu^M}{M!} = \frac{1}{eM!}$.
- D.h. alle Urnen haben weniger als M Bälle mit Ws höchstens
$$\left(1 - \frac{1}{eM!}\right)^n \leq e^{-\frac{n}{eM!}}.$$
- Falls $e^{-\frac{n}{eM!}} \leq \frac{1}{n^2}$, ist diese Gegenws im exakten Fall $\leq \frac{e\sqrt{n}}{n^2} < \frac{1}{n}$.
- Für die Wahl $M = \frac{\ln n}{\ln \ln n}$ folgt $e^{-\frac{n}{eM!}} \leq \frac{1}{n^2}$ (nach etwas Rechnen). $\square$

Anwendung Hash-Ketten

Szenario: Wörterbuchsuche

- Besitzen sortierte Black-List mit n nicht erlaubten Passworten.
- Überprüfen eines einzelnen Passworts benötigt $\Theta(\log n)$ Schritte.
- Frage: Effizientere Datenstruktur für die Wörterbuchsuche?

Datenstruktur Hash-Kette:

- Hashfunktion $f : U \to [1, n]$ mit $\Pr(f(x) = j) = \frac{1}{n}$ für alle $x \in U$.
- Hashkollisionen werden mittels verketteter Listen behandelt.
- D.h. die n Passwörter entsprechen Bällen, die n Hashwerte Urnen.
- Suche eines Wortes: erwartet $\Theta(1)$ und $\mathcal{O}(\frac{\ln n}{\ln \ln n})$ mit hoher Ws.

Scheduling und Leader Election

Szenario: Scheduling

- n Nutzer $u_i \in U$ wollen gleichzeitig einen Rechner nutzen.
- Müssen Reihenfolge festlegen, d.h. eine Permutation wählen.

Lösung mittels Hashing:

- Wähle eine Hashfunktion $f : U \to [1, n^3]$ mit $\Pr(x = j) = \frac{1}{n^3}$.
- Nutzer u_i mit kleinstem Hashwert $f(u_i)$ kommt zuerst, usw.
- Benötigen: Keine zwei Nutzer erhalten denselben Hashwert.
- Für ein festes u_i gilt $f(u_i) = f(u_j)$ für ein $j \neq i$ mit Ws
$$1 - (1 - \frac{1}{n^3})^{n-1} \leq \frac{n-1}{n^3} < \frac{1}{n^2}.$$
- Union Bound: 2 Nutzer besitzen denselben Hashwert mit Ws $< \frac{1}{n}$.

Leader Election: Wähle Leader u_i mit kleinstem Hashwert $f(u_i)$.

Die Probabilistische Methode

Beobachtung: Besitzt ein Ereignis Ws > 0, so muss es existieren!

Notation: Sei K_n der komplette Graph mit n Knoten und $\binom{n}{2}$ Kanten.

Satz

Falls $2^{\binom{k}{2}-1} > \binom{n}{k}$, existiert eine 2-Kantenfärbung des K_n, so dass kein K_k Subgraph monochromatisch ist.

Beweis:

- Färbe jede Kante zufällig und unabhängig mit Ws $\frac{1}{2}$.
- Ereignis A_i: i-te Clique $K_k^{(i)}$, $i = 1, \ldots, \binom{n}{k}$, ist monochromatisch.
- $K_k^{(i)}$ besitzt $\binom{k}{2}$ Kanten. D.h. $\Pr(A_i) = \frac{1}{2^{\binom{k}{2}-1}}$.
- Mit Union Bound existiert ein monochromatischer $K_k^{(i)}$ mit
$$\Pr(\bigcup_{i=1}^{\binom{n}{k}} A_i) \leq \sum_{i=1}^{\binom{n}{k}} \Pr(A_i) = \binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1.$$
- D.h. kein monochromatischer $K_k^{(i)}$ existiert mit
$$\Pr(\bigcap_{i=1}^{\binom{n}{k}} \bar{A}_i) = 1 - \Pr(\bigcup_{i=1}^{\binom{n}{k}} A_i) > 0. \quad \Box$$

Nicht alle kleiner als der Erwartungswert.

Lemma

Sei X eine ZV mit $\mathbb{E}[X] = \mu$. Dann gilt

$$\Pr(X \geq \mu) > 0 \text{ und } \Pr(X \leq \mu) > 0.$$

Beweis:

- Wir zeigen nur $\Pr(X \geq \mu) > 0$, $\Pr(X \leq \mu) > 0$ folgt analog.
- Angenommen $\Pr(X \geq \mu) > 0$. Wir erhalten den Widerspruch

$$\mu = \sum_x x \Pr(X = x) = \sum_{x<\mu} x \Pr(X = x) < \mu \sum_{x<\mu} \Pr(X = x) = \mu.$$

Problem Max-Cut

Gegeben: Graph $G = (V, E)$ mit $|V| = n$ und $|E| = m$
Gesucht: Maximaler Cut $V = A \,\dot{\cup}\, B$

Max-Cut ist ein NP-hartes Problem.

Mindestgröße eines Max-Cuts

Satz Mindestgröße eines Max-Cuts

Jeder Graph $G = (V, E)$ besitzt einen Cut der Größe mindestens $\frac{m}{2}$.

Beweis:

- Weise jedes $v \in V$ mit jeweils Ws $\frac{1}{2}$ zu A oder B.
- Sei $E = \{e_1, \ldots, e_m\}$ und IV $X_i = 1$, falls e_i zum Cut beiträgt.
- Es gilt $\mathbb{E}[X_i] = \frac{1}{2}$. Sei C ZV für die Größe des Cuts. Dann gilt
$$\mathbb{E}[C] = \sum_{i=1}^{m} \mathbb{E}[X_i] = \frac{m}{2}.$$
- D.h. es existiert eine Partition $V = A \,\dot{\cup}\, B$ mit Cutgröße $|C| \geq \frac{m}{2}$. $\square$

Algorithmus(Las Vegas) BIG-CUT

EINGABE: $G = (V, E)$

1. REPEAT
   1. Weise jedes $v \in V$ jeweils mit Ws $\frac{1}{2}$ zu A oder B.

   UNTIL $C(A, B) \geq \frac{m}{2}$

AUSGABE: $V = A \,\dot{\cup}\, B$ mit Cutgröße $C \geq \frac{m}{2}$

Konstruktion eines großen Cuts

Satz Laufzeit von BIG-CUT

BIG-CUT läuft in erwarteter Zeit $\mathcal{O}(nm^2)$.

Beweis:

- Jede Zuweisung in 1.1 läuft in $\mathcal{O}(n)$.
- Die Überprüfung von $C(A, B) \geq \frac{m}{2}$ benötigt $\mathcal{O}(m)$.
- Definiere $p = \Pr(C \geq \frac{m}{2})$. Wegen $C \leq m$ gilt

$$\frac{m}{2} = \mathbb{E}[C] = \sum_{i<\frac{m}{2}} i \Pr(C = i) + \sum_{i \geq \frac{m}{2}} i \Pr(C = i) \leq (1-p)\frac{m-1}{2} + pm.$$

- Auflösen nach p liefert $p \geq \frac{1}{m+1}$.
- Die Anzahl Iterationen von Schritt 1 ist geometrisch verteilt mit p.
- D.h. wir benötigen erwartet $\frac{1}{p} = \mathcal{O}(m)$ viele Iterationen. □

Derandomisierung von BIG-CUT

Algorithmus DETERMINISTIC BIG-CUT

EINGABE: $G = (V, E)$

1. Setze $A = \{v_1\}$ und $B = \emptyset$.
2. FOR $i = 2$ to n
   1. Falls v_i weniger Nachbarn in A als in B besitzt, setzte $A = A \cup \{v_i\}$.
   2. Sonst setze $B = B \cup \{v_i\}$.

AUSGABE: $V = A \dot\cup B$ mit Cutgröße $C \geq \frac{m}{2}$

Laufzeit: $\mathcal{O}(n + m)$, d.h. linear in der Eingabe.

Korrektheit DETERMINISTIC BIG-CUT

Satz Korrektheit von DETERMINISTIC BIG-CUT

DETERMINISTIC BIG-CUT berechnet einen Cut der Größe $C \geq \frac{m}{2}$.

Beweis:

- Sei $x_i \in \{A, B\}$ die Menge, in die v_i platziert wird.
- Sei $\mathbb{E}[C|x_1, \ldots, x_k]$ die erwartete Cut-Größe nach Platzierung von $x_1, \ldots, x_k$, wobei die verbliebenen Knoten zufällig platziert werden.
- Dann gilt $\mathbb{E}[C] \geq \frac{m}{2}$ (BIG-CUT) und $\mathbb{E}[C] = \mathbb{E}[C \mid x_1]$.
- Ferner gilt
  $$\mathbb{E}[C \mid x_1, \ldots, x_k] = \frac{1}{2}\mathbb{E}[C \mid x_1, \ldots, x_k, A] + \frac{1}{2}\mathbb{E}[C \mid x_1, \ldots, x_k, B].$$
- Alle Kanten, die nicht inzident zu v_{k+1} sind, tragen dasselbe zu den beiden Erwartungswerten auf der rechten Seite bei.
- DETERM. BIG-CUT wählt das Maximum beider Erwartungswerte.
- Es folgt $\mathbb{E}[C \mid x_1, \ldots, x_k] \leq \mathbb{E}[C \mid x_1, \ldots, x_{k+1}]$ für alle $0 < k < n$.
- D.h. bei Terminierung von DETERMINISTIC BIG-CUT gilt $\mathbb{E}[C \mid x_1, \ldots, x_n] \geq \mathbb{E}[C \mid x_1, \ldots, x_{n-1}] \geq \ldots \geq \mathbb{E}[C] \geq \frac{m}{2}$. $\square$

Sample and Modify

Sample and Modify Strategie

1. Sample: Erzeuge zufällige Struktur.
2. Modify: Modifiziere, bis die gewünschte Eigenschaft erreicht ist.

Algorithmus INDEPENDENT SET

EINGABE: $G = (V, E)$ mit $|V| = n, |E| = m$

1. Setze $d = \frac{2m}{n}$ (durchschnittlicher Grad eines Knoten).
2. Lösche jedes $v \in V$ und dessen inzidente Kanten mit Ws $1 - \frac{1}{d}$.
3. Lösche alle verbliebenen $e \in E$ und einen der adjazenten Knoten.

AUSGABE: Knotenmenge $V' \subset V$ mit $\{u, v\} \notin E$ für alle $u, v \in V'$

Korrektheit:

- Nach dem Samplen in Schritt 2 erhalten wir kein Independent Set.
- Die Modifikation in Schritt 3 stellt die gewünschte Eigenschaft her.

Größe des Independent Sets

Satz Größe des Independent Sets

INDEPENDENT SET berechnet ein V' erwarteter Größe $\frac{n^2}{4m}$.

Beweis:

- Sei X ZV für die Anzahl Knoten nach Schritt 1. Dann gilt $\mathbb{E}[X] = \frac{n}{d}$.
- Sei Y ZV für die Anzahl Kanten nach Schritt 2.
- Kante überlebt, falls beide adjazenten Knoten überleben. D.h.
$$\mathbb{E}[Y] = m\frac{1}{d^2} = \frac{dn}{2d^2} = \frac{n}{2d}.$$
- Schritt 3 löscht höchstens Y Knoten. D.h. $|V'| \geq X - Y$ mit
$$\mathbb{E}[X - Y] = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d} = \frac{n^2}{4m}. \quad \square$$

Korollar

Jedes $G = (V, E)$ enthält eine unabhängige Menge der Größe $\geq \frac{n^2}{4m}$.

Zufällige Graphen

Definition Zufälliger Graph $G_{n,p}$

Für $G_{n,p}$ wähle n Knoten und setze jede der $\binom{n}{2}$ Kanten mit Ws p.

Satz

Für alle $\epsilon > 0$ und hinreichend große n gilt:
$G_{n,p}$ mit $p = o(n^{-\frac{2}{3}})$ enthält eine Clique der Größe 4 mit Ws kleiner ϵ.

Beweis:

- Seien $C_1, \ldots, C_{\binom{n}{4}}$ alle Mengen mit 4 Knoten.
- Sei IV $X_i = 1$ falls C_i eine Clique ist und $X = \sum_{i=1}^{\binom{n}{4}} X_i$. Dann gilt
$$\mathbb{E}[X] = \sum_{i=1}^{\binom{n}{4}} \Pr[X_i = 1] = \binom{n}{4} p^6 \le n^4 \cdot o(n^{-4}) = o(1).$$
- D.h. $\mathbb{E}[X] < \epsilon$ für hinreichend große n. Da $X \ge 0$ folgt
$$\Pr(X \ge 1) \le \mathbb{E}[X] < \epsilon. \quad \square$$

Frage: Enthält $G_{n,p}$ für $p = \omega(n^{-\frac{2}{3}})$ eine 4er-Clique mit großer Ws?

2. Moment Methode

Satz 2. Moment Methode

Sei X eine nicht-negative ganzzahlige ZV. Dann gilt

$$\Pr(X = 0) \leq \frac{\mathbf{Var}[X]}{(\mathbb{E}[X])^2}.$$

Beweis: Mit Chebyshevs Ungleichung gilt

$$\Pr(X = 0) \leq \Pr(|X - \mathbb{E}[X]| \geq \mathbb{E}[X]) \leq \frac{\mathbf{Var}[X]}{(\mathbb{E}[X])^2} \quad \square.$$

Nachteil: Die Berechnung von **Var**$[X]$ ist oft aufwändig.

2. Moment Methode

Satz

Sei $X = \sum_{i=1}^{n} X_i$ mit 0,1-ZV X_i. Dann gilt

$$\Pr(X > 0) \geq \sum_{i=1}^{n} \frac{\Pr(X_i{=}1)}{\mathbb{E}[X \mid X_i{=}1]}.$$

Beweis: Sei $Y = \frac{1}{X}$ für $X > 0$ und $Y = 0$ sonst. Dann gilt

$$\begin{aligned}
\Pr(X > 0) &= \mathbb{E}[XY] = \mathbb{E}\left[\sum_{i=1}^{n} X_i Y\right] = \sum_{i=1}^{n} \mathbb{E}[X_i Y] \\
&= \sum_{i=1}^{n} \mathbb{E}[X_i Y \mid X_i = 1] \Pr(X_i = 1) + \mathbb{E}[X_i Y \mid X_i = 0] \Pr(X_i = 0) \\
&= \sum_{i=1}^{n} \mathbb{E}[Y \mid X_i = 1] \Pr(X_i = 1) = \sum_{i=1}^{n} \mathbb{E}\left[\frac{1}{X} \mid X_i = 1\right] \Pr(X_i = 1) \\
&\geq \sum_{i=1}^{n} \frac{\Pr(X_i = 1)}{\mathbb{E}[X \mid X_i = 1]}. \qquad \text{(mit Jensens Ungleichung } \mathbb{E}[f(X)] \geq f(\mathbb{E}[X])\text{)} \qquad \square
\end{aligned}$$

Zufällige Graphen

Satz

Für alle $\epsilon > 0$ und hinreichend große n gilt:
$G_{n,p}$ mit $p = \omega(n^{-\frac{2}{3}})$ enthält keine Clique der Größe 4 mit Ws kleiner ϵ.

Beweis:

- Sei wie zuvor IV $X_i = 1$ falls C_i eine Clique ist und $X = \sum_{i=1}^{\binom{n}{4}} X_i$.
- Zeigen, dass $\Pr(X > 0) \xrightarrow[n\to\infty]{} 1$. D.h. $\Pr(X = 0) < \epsilon$.
- Es gilt für ein festes X_j

$$\mathbb{E}[X \mid X_j = 1] = \sum_{i=1}^{\binom{n}{4}} \mathbb{E}[X_i \mid X_j = 1] = \sum_{i=1}^{\binom{n}{4}} \Pr(X_i = 1 \mid X_j = 1).$$

- Für $\binom{n-4}{4}$ Knotenmengen C_i mit $|C_i \cap C_j| = 0$: $X_i = 1$ mit Ws p^6.
- Für $4\binom{n-4}{3}$ Knotenmengen C_i mit $|C_i \cap C_j| = 1$: $X_i = 1$ mit Ws p^6.
- Für $6\binom{n-4}{2}$ Knotenmengen C_i mit $|C_i \cap C_j| = 2$: $X_i = 1$ mit Ws p^5.
- Für $4\binom{n-4}{1}$ Knotenmengen C_i mit $|C_i \cap C_j| = 3$: $X_i = 1$ mit Ws p^3.
- Es folgt $\Pr(X > 0) \geq \frac{\binom{n}{4}p^6}{1+\binom{n-4}{4}p^6+4\binom{n-4}{3}p^6+6\binom{n-4}{2}p^5+4\binom{n-4}{1}p^3} \xrightarrow[n\to\infty]{} 1.$ $\square$

Lovasz Local Lemma

Motivation: Probabilistische Methode

- Seien $E_1, \ldots, E_n$ schlechte Ereignisse.
- Weiter seien $E_1, \ldots, E_n$ unabhängig. D.h. für alle $I \subseteq \{1, \ldots, n\}$ gilt
$$\Pr(\bigcap_{i \in I} E_i) = \prod_{i \in I} \Pr(E_i).$$
- Mit $E_1, \ldots, E_n$ sind auch die Ereignisse $\overline{E_1}, \ldots, \overline{E_n}$ unabhängig.
- Falls $\Pr(E_i) < 1$ für alle i, dann folgt
$$\Pr(\bigcap_{i \in I} \overline{E_i}) = \prod_{i \in I} \Pr(\overline{E_i}) > 0.$$
- D.h. es existiert ein Element des Wsraums, das in keinem der schlechten Ereignisse auftaucht.
- **Frage:** Was passiert für limitierte Formen von Unabhängigkeit?

Definition Abhängigkeitsgraph

E heißt *unabhängig von* $E_1, \ldots, E_n$ falls für alle $I \subseteq \{1, \ldots, n\}$ gilt
$$\Pr(E \mid \bigcap_{i \in I} E_i) = \Pr(E).$$

Der *Abhängigkeitsgraph* $G = (V, E)$ für $E_1, \ldots, E_n$ ist definiert als
$$V = \{1, \ldots, n\} \text{ und } E = \{(i, j) \mid E_i \text{ ist abhängig von } E_j\}.$$

Lovasz Local Lemma

Lemma Lovasz Local Lemma (1975)

Seien $E_1, \ldots, E_n$ Ereignisse mit

1. $\Pr(E_i) \leq p$ für alle $i = 1, \ldots, n$,
2. Abhängigkeitsgraph $G = (V, E)$ von $E_1, \ldots, E_n$ besitzt Grad $\leq d$,
3. $4dp \leq 1$.

Dann gilt $\Pr(\bigcap_{i=1}^{n} \overline{E_i}) > 0$.

Beweis:

- Sei $S \subseteq \{1, \ldots, n\}$. Wir zeigen für alle $k \notin S$
  $\Pr(E_k \mid \bigcap_{j \in S} \overline{E_j}) \leq 2p$ per Induktion über $|S| = s$, $s = 0, \ldots, n$.
- Die Aussage des Theorems folgt aus

$$\Pr(\bigcap_{i=1}^{n} \overline{E_i}) \;=\; \prod_{i=1}^{n} \Pr(\overline{E_i} \mid \bigcap_{j=1}^{i-1} \overline{E_j}) = \prod_{i=1}^{n} (1 - \Pr(E_i \mid \bigcap_{j=1}^{i-1} \overline{E_j})) \geq \prod_{i=1}^{n} (1 - 2p) > 0.$$

- Benötigen $\Pr(\bigcap_{j \in S} \overline{E_j}) > 0$. Fall $s = 1$ folgt aus $\Pr(\overline{E_j}) \geq 1 - p > 0$.

Lovasz Local Lemma

Beweis: (Fortsetzung)

- Für $s > 1$ sei OBdA $S = \{1, \ldots, s\}$. Es gilt analog wie zuvor
$$\Pr(\bigcap_{i=1}^{s} \overline{E_i}) = \prod_{i=1}^{s} \left(1 - \Pr(E_i \mid \bigcap_{j=1}^{i-1} \overline{E_j})\right) \overset{\text{IV}}{\geq} \prod_{i=1}^{s}(1-2p) > 0.$$
- Sei $S_1 = \{j \in S \mid (k,j) \in E\}$ und $S_2 = S \setminus S_1$. Für $S_2 = S$ gilt
$$\Pr(E_k \mid \bigcap_{j \in S} \overline{E_j}) = \Pr(E_k) \leq p.$$
- Sei also $|S_2| < s$. Sei $F_S = \bigcap_{j \in S} \overline{E_j}$. Es gilt $F_S = F_{S_1} \cap F_{S_2}$ und
$$\Pr(E_k \mid F_S) = \frac{\Pr(E_k \cap F_S)}{\Pr(F_S)} = \frac{\Pr(E_k \cap F_{S_1} \cap F_{S_2})}{\Pr(F_{S_1} \cap F_{S_2})} = \frac{\Pr(E_k \cap F_{S_1} | F_{S_2})}{\Pr(F_{S_1} | F_{S_2})}.$$
- Es gilt $\Pr(E_k \cap F_{S_1} \mid F_{S_2}) \leq \Pr(E_k \mid F_{S_2}) = \Pr(E_k) \leq p$.
- Es genügt nun, $\Pr(F_{S_1} \mid F_{S_2}) \geq \frac{1}{2}$ zu zeigen. Es gilt

$$\begin{aligned} \Pr(F_{S_1} \mid F_{S_2}) &= \Pr(\bigcap_{i \in S_1} \overline{E_i} \mid \bigcap_{j \in S_2} \overline{E_j}) \geq 1 - \sum_{i \in S_1} \Pr(E_i \mid \bigcap_{j \in S_2} \overline{E_j}) \\ &\overset{\text{IV}}{\leq} 1 - \sum_{i \in S_1} 2p \geq 1 - 2pd \geq \frac{1}{2}. \quad \square \end{aligned}$$

Anwendung: Erfüllbarkeit

Problem Erfüllbarkeit von k-SAT

Gegeben: k-SAT Formel (keine Klausel enthält Variable doppelt)
Gesucht: erfüllende Belegung

Satz Existenz erfüllender Belegung

Eine k-SAT Formel ist erfüllbar, falls keine Variable in mehr als $T = \frac{2^k}{4k}$ Klauseln vorkommt.

Beweis:

- Wähle eine zufällige Belegung der Variablen.
- Ereignis E_i, $i = 1, \ldots, m$: i-te Klausel ist nicht erfüllt.
- Da jede Klausel k Literale enthält, gilt $p = \Pr(E_i) = \frac{1}{2^k}$.
- E_i, E_j abhängig, falls Klauseln i, j gemeinsame Variable besitzen.
- Jeder der k Variablen in Klausel i kommt in $\leq T = \frac{2^k}{4k}$ Klauseln vor.
- D.h. wir erhalten $d \leq kT \leq \frac{2^k}{4}$. Es folgt $4dp \leq 4\frac{2^k}{4}2^{-k} = 1$.
- Mit Lovasz Local Lemma folgt $\Pr(\bigcap_{i=1}^m \overline{E_i}) > 0$. $\square$

Anwendung: Kanten-disjunkte Pfade

Problem Wahl kanten-disjunkter Pfade

Gegeben: n Paare Nutzer mit Mengen F_i von m Pfaden pro Nutzer
Gesucht: Auswahl n kanten-disjunkter Pfade

Satz Existenz kanten-disjunkter Auswahl

Es existiert eine kanten-disjunkte Auswahl, falls für alle $F_i, F_j, i \neq j$ die Anzahl nicht kanten-disjunkter Pfade höchstens $k \leq \frac{m}{8n}$ ist.

Beweis:

- Wähle jeden Pfad aus F_i unabhängig gleichverteilt mit Ws $\frac{1}{m}$.
- Ereignis $E_{i,j}$: Pfade aus F_i, F_j besitzen gemeinsame Kante.
- Es gilt $p = \Pr(E_{i,j}) \leq \frac{k}{m}$ für alle $i \neq j$.
- Sei d der Grad des Abhängigkeitsgraphen der $E_{i,j}$.
- $E_{i,j}$ ist unabhängig von $E_{i',j'}$ für $\{i,j\} \cap \{i',j'\} = \emptyset$. D.h. $d \leq 2n$.
- Wir erhalten $4dp \leq \frac{8nk}{m} \leq 1$.
- Mit Lovasz Local Lemma folgt $\Pr(\bigcap_{i \neq j} \overline{E_{i,j}}) > 0$. $\square$

Markov Kette

Definition Markov Kette

Ein *stochastischer Prozess* $\mathbf{X} = \{X_t \mid t \in \mathbb{N}_0\}$ ist eine Menge von ZV.
Ein stochastischer Prozess $\mathbf{X}$ heißt *Markov Kette*, falls

$$\Pr(X_t = a_t \mid X_{t-1} = a_{t-1}, \ldots, X_0 = a_0) = \Pr(X_t = a_t \mid X_{t-1} = a_{t-1}).$$

Anmerkungen:

- D.h Zustand X_t hängt nur von X_{t-1} ab (unabhängig von Historie).
- Sei $\{0, 1, \ldots, n\}$ bzw. $\{0, 1, \ldots\}$ der Zustandsraum der X_t.
- Wir gehen von Zustand i nach Zustand j mit Übergangsws
$$p_{i,j} = \Pr(X_t = j \mid X_{t-1} = i).$$
- Wir definieren die Übergangsmatrix $\mathbf{P} = (p_{i,j})_{0 \le i,j \le n \text{ (bzw. } \infty)}$.
- Es gilt $\sum_{j \ge 0} p_{i,j} = 1$ für alle i.
- Sei $p_i(t)$ die Ws: Prozess besitzt zum Zeitpunkt t Zustand i.
- $p(t) = (p_0(t), p_1(t), \ldots, p_n(t))$ ist Zustandsverteilung. Es gilt
$$p_i(t) = \sum_{j \ge 0} p_j(t-1)p_{j,i} \text{ bzw. } p(t) = p(t-1)\mathbf{P}.$$

Eigenschaften von Markov Ketten

Anmerkungen:

- Wir bezeichnen die Ws in m Schritten von i nach j zu wechseln
$$p_{i,j}^m = \Pr(X_{t+m} = j \mid X_t = i).$$
- Offenbar gilt $p_{i,j}^m = \sum_{k \geq 0} p_{i,k} p_{k,j}^{m-1}$.
- Sei $\mathbf{P^{(m)}} = (p_{i,j}^m)_{0 \leq i,j \leq n}$. Dann gilt $\mathbf{P^{(m)}} = \mathbf{P} \cdot \mathbf{P^{(m-1)}} = \mathbf{P}^m$ bzw.
$$p(t+m) = p(t)\mathbf{P^{(m)}}.$$
- **P** wird oft als gerichteter Graph $G(V, E)$ veranschaulicht.

Anwendung: Random Walk 2-SAT Algorithmus

Problem 2-SAT

Gegeben: 2-SAT Formel $\phi(x_1, \ldots, x_n)$
Gesucht: erfüllende Belegung oder Ausgabe “nicht erfüllbar”

Algorithmus 2-SAT

EINGABE: $\phi(x_1, \ldots, x_n)$, m (Parameter für Erfolgsws)

1. Starte mit einer zufälligen Belegung.
2. FOR $i = 1$ to $2mn^2$
   1. Falls Belegung erfüllend, Ausgabe der Belegung, EXIT.
   2. Wähle eine beliebige nicht-erfüllte Klausel k.
   3. Ändere für ein zufälliges Literal in k die Belegung der Variable.
3. Ausgabe “nicht erfüllbar”.

Analyse 2-SAT

Satz 2-SAT

2-SAT findet für erfüllbare ϕ eine erfüllende Belegung nach erwartet $\mathcal{O}(n^2)$ Iterationen. Die Ausgabe "nicht-erfüllbar" erfolgt mit Ws $\leq \frac{1}{2^m}$.

Beweis:

- Sei S eine erfüllende Belegung und A_i die Belegung in Iteration i.
- Sei X_i eine ZV für die Anzahl der Übereinstimmungen in S und A_i.
- Falls $X_i = n$, so gibt 2-SAT die Belegung S aus.
- Es gilt $\Pr(X_{i+1} = 1 \mid X_i = 0) = 1$. In Schritt 2.2 ist k nicht erfüllt.
- Daher stimmen A_i, S an ≥ 1 Stelle nicht überein. D.h.

  $$\Pr(X_{i+1} = j+1 \mid X_i = j) \geq \tfrac{1}{2} \text{ bzw. } \Pr(X_{i+1} = j-1 \mid X_i = j) \leq \tfrac{1}{2}.$$
- Wir betrachten die Markov Kette $\mathbf{Y} = \{Y_t \mid t \geq 0\}$ mit

  $$Y_0 = X_0,\ \Pr(Y_{i+1} = 1 \mid Y_i = 0) = 1 \text{ und für } 1 \leq j < n:$$
  $$\Pr(Y_{i+1} = j+1 \mid Y_i = j) = \Pr(Y_{i+1} = j-1 \mid Y_i = j) = \tfrac{1}{2}.$$

Analyse 2-SAT

Beweis: (Fortsetzung)

- **Y** benötigt zum Erreichen von n mindestens solange wie $X_0, X_1, \ldots$
- **Y** modelliert einen Random Walk auf dem Intervall $0, 1, \ldots, n$.
- ZV Z_i: Anzahl Schritte bei Startwert i bis zum Erreichen von $i + 1$.
- Sei $h_i = \mathbb{E}[Z_i]$. Es gilt $h_0 = 1$. Für $1 \leq i < n$ folgt
$$h_i = \mathbb{E}[Z_i] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \mathbb{E}[1 + Z_{i-1} + Z_i] = 1 + \frac{1}{2} h_{i-1} + \frac{1}{2} h_i.$$
- Wir erhalten $h_i = 2 + h_{i-1} = 2 + 2 + h_{i-2} = \ldots = 2i + h_0 = 2i + 1$.
- D.h. unabhängig vom Startwert Y_0 benötigen wir Schrittzahl max.
$$\mathbb{E}[\sum_{i=0}^{n-1} Z_i] = \sum_{i=0}^{n-1} \mathbb{E}[Z_i] = \sum_{i=0}^{n-1} 2i + 1 = n^2.$$
- Teile 2-SAT in Segmente der Größe von $2n^2$ Iterationen.
- Pro Segment benötigen wir $\leq n^2$ Iterationen zum Erreichen von n.
- Die Markov- Ungleichung liefert $\Pr(\sum_{i=0}^{n-1} Z_i \geq 2n^2) \leq \frac{n^2}{2n^2} = \frac{1}{2}$.
- D.h. wir finden S nicht in m Segmenten mit Ws höchstens $\frac{1}{2^m}$. □

Anwendung: randomisierter 3-SAT Algorithmus

Modifikation für 3-SAT Algorithmus:
Setze im Algorithmus 2-SAT für die FOR-Schleife $i = 1$ to ∞.

Satz Komplexität für 3-SAT

3-SAT benötigt auf einer erfüllbaren 3-SAT Formel erwartet Zeit $\mathcal{O}(2^n)$.

Beweis:

- Seien S, A_i, X_i und Z_i analog zum vorigen Beweis. Es gilt nun
  $$\Pr(X_{i+1} = j+1 \mid X_i = j) \geq \tfrac{1}{3} \text{ bzw. } \Pr(X_{i+1} = j-1 \mid X_i = j) \leq \tfrac{2}{3}.$$
- Wir betrachten die Markov Kette $\mathbf{Y} = \{Y_t \mid t \geq 0\}$ mit
  $$Y_0 = X_0,\ \Pr(Y_{i+1} = 1 \mid Y_i = 0) = 1 \text{ und für } 1 \leq j < n\text{:}$$
  $$\Pr(X_{i+1} = j+1 \mid X_i = j) = \tfrac{1}{3} \text{ und } \Pr(X_{i+1} = j-1 \mid X_i = j) = \tfrac{1}{2}.$$
- Nun folgt $\mathbb{E}[Z_i] = h_i = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \mathbb{E}[1 + Z_{i-1} + Z_i] = 1 + \frac{2}{3} h_{i-1} + \frac{2}{3} h_i$.
  $\Rightarrow h_i = 3 + 2h_{i-1} = 3(1+2) + 2^2 h_{i-2} = \ldots = 3(2^0 + \ldots + 2^{i-1}) + 2^i h_0 = 2^{i+2} - 3$
- D.h. $h_{n-1} = \Theta(2^n)$. Aus $\sum_{i=0}^{n-1} h_i = \mathcal{O}(2^n)$ folgt die Laufzeit. $\square$

3-SAT Algorithmus mit Reset

Idee: Falls X_i nicht nahe bei n ist, starte neu.

Algorithmus 3-SAT$_{Reset}$

EINGABE: $\phi(x_1, \ldots, x_n)$

1. REPEAT
2. Wähle eine zufällige Belegung.
   1. FOR $i = 1$ to $3n$
      1. Wähle beliebige nicht-erfüllte Klausel k. Falls nicht vorhanden, EXIT.
      2. Ändere für ein zufälliges Literal in k die Belegung der Variable.
3. UNTIL (erfüllende Belegung gefunden)

AUSGABE: erfüllende Belegung von $\phi(x_1, \ldots, x_n)$

3-SAT Algorithmus mit Reset

Satz 3-SAT Algorithmus mit Reset

Für erfüllbare ϕ besitzt 3-SAT$_{Reset}$ erwartete Laufzeit $\mathcal{O}(n^{\frac{3}{2}}(4/3)^n)$.

Beweis:

- Sei q_j die Ws, dass S in $3n$ Schritten erreicht wird, wenn anfangs in Schritt 2 genau j Variablen nicht mit S übereinstimmen.
- Man muss in Summe j-mal in die "richtige" Richtung gehen. D.h.

$$q_j \geq \max_{k=0,\ldots,j}\{(\tfrac{1}{3})^{j+k}(\tfrac{2}{3})^k\binom{j+2k}{k}\} \stackrel{k=j}{\geq} (\tfrac{1}{3})^{2j}(\tfrac{2}{3})^j\binom{3j}{j} \text{ für } 0 < j \leq n.$$

- Aus der Stirling-Formel folgt $\binom{3j}{3} \geq \frac{c}{\sqrt{j}}(\frac{27}{4})^j$, $c < 1$ konstant. D.h.

$$q_j \geq \frac{c}{\sqrt{j}}(\tfrac{1}{3})^{2j}(\tfrac{2}{3})^j(\tfrac{27}{4})^j \geq \frac{c}{\sqrt{j}}(\tfrac{1}{2})^j.$$

- Sei q die Ws, dass wir in Schleife 2.1 erfolgreich sind. Wir erhalten

$$\begin{aligned} q &= \tfrac{1}{2^n} + \textstyle\sum_{j=1}^n \binom{n}{j}(\tfrac{1}{2})^n q_j \geq \tfrac{c}{\sqrt{n}}(\tfrac{1}{2})^n \textstyle\sum_{j=0}^n \binom{n}{j}(\tfrac{1}{2})^j(1)^{n-j} \\ &= \Omega(n^{-\frac{1}{2}})(\tfrac{1}{2})^n(\tfrac{3}{2})^n = \Omega(n^{-\frac{1}{2}})(\tfrac{3}{4})^n. \end{aligned}$$

- Insgesamt benötigt man erwartet Laufzeit $\frac{1}{q}3n = \mathcal{O}(n^{\frac{3}{2}}(4/3)^n)$. $\square$

Gambler's Ruin

Szenario: Spielen bis zum Ruin

- Spieler 1 besitze ℓ_1 Euro, Spieler 2 besitze ℓ_2 Euro.
- Pro Runde gewinne jeder Spieler 1 Euro vom anderen mit Ws $\frac{1}{2}$.
- Ein Spieler gewinnt, falls er das Geld des anderen besitzt.

Frage: Mit welcher Ws q_0 gewinnt Spieler 1?

Modellierung als Random Walk:

- Wir betrachten den Gewinn von Spieler 1.
- D.h. der Random Walk beginnt in 0 und endet in $-\ell_1$ bzw. ℓ_2.

Gambler's Ruin

Satz Spielen bis zum Ruin

Spieler 1 gewinnt mit Ws $q_0 = \frac{\ell_1}{\ell_1+\ell_2}$.

Beweis:

- Sei q_j die Ws des Gewinns bei Zwischengewinn von j Euro.
- Uns interessiert q_0. Es gilt $q_{-\ell_1} = 0, q_{\ell_2} = 1$ und

$$q_j = \frac{q_{j-1}}{2} + \frac{q_{j+1}}{2} \text{ für } -\ell_1 < j < \ell_2.$$

$$\overset{(1)}{\Rightarrow} \quad 1 = q_{\ell_2} = 2q_{\ell_2-1} - q_{\ell_2-2} = 3q_{\ell_2-2} - 2_{\ell_2-3} = (\ell_2+1)q_0 - \ell_2 q_{-1}$$

$$\overset{(2)}{\Rightarrow} \quad 0 = q_{-\ell_1} = 2q_{-\ell_1+1} - q_{-\ell_1+2} = 3q_{-\ell_1+2} - 2q_{-\ell_1+3} = \ell_1 q_{-1} - (\ell_1 - 1)q_0$$

- Addiere ℓ_1-mal Gleichung (1) zu ℓ_2-mal Gleichung (2).
- Es folgt $\ell_1 = \ell_1(\ell_2+1)q_0 - \ell_2(\ell_1-1)q_0 = (\ell_1+\ell_2)q_0$. $\square$

Übung: Zeigen Sie, dass $q_j = \frac{\ell_1+j}{\ell_1+\ell_2}$ für $-\ell_1 < j < \ell_2$.

Random Walks auf Graphen

Definition Stationäre Zustandsverteilung

Ein Ws-Verteilung $\bar{\pi}$ für eine Markov Kette heißt *stationär*, falls $\bar{\pi} = \bar{\pi}\mathbf{P}$.

Anmerkung:

- Wir berechnen stationäres $\bar{\pi}$ durch Lösen des linearen Systems
$$\bar{\pi} = \bar{\pi}\mathbf{P}.$$
- Sei $h_{i,i}$ die erwartete Anzahl von Schritten von Zustand i nach i.
- Man kann zeigen, dass $h_{i,i} = \frac{1}{\bar{\pi}_i}$.

Random Walk auf $G = (V, E)$:

- Wir betrachten endliche, ungerichtete, zusammenhängende G.
- Zusätzlich soll G nicht bipartit sein. D.h. für jeden Knoten $v \in V$ existiert ein Pfad von v nach v ungerader Länge.
- Für jedes $v \in V$ bezeichne $d(v)$ den Grad von v.
- Angenommen wir sind im Zeitpunkt t in Knoten v. Dann gehen wir zum Zeitpunkt $t + 1$ mit Ws jeweils $\frac{1}{d(v)}$ zu einem der Nachbarn.

Random Walk besitzt stationäre Verteilung.

Satz Random Walk besitzt stationäre Verteilung.

Random Walks auf G konvergieren zu einem stationären π mit

$$\pi_v = \frac{d(v)}{2|E|}.$$

Beweis:

- Es gilt $\sum_{v \in V} \pi_v = \sum_{v \in V} \frac{d(v)}{2|E|} = 1$. D.h. π_v ist eine Ws-Verteilung.
- Sei $\mathbf{P}$ die Übergangsmatrix und $N(v)$ die Nachbarn von v.
- Das lineare Gleichungssystem $\bar{\pi} = \bar{\pi}\mathbf{P}$ ist äquivalent zu
$$\pi_v = \sum_{u \in N(v)} \pi_u \frac{1}{d(u)}.$$
- Die Setzung $\pi_v = \frac{d(v)}{2|E|}$ löst das System, denn
$$\sum_{u \in N(v)} \pi_u \frac{1}{d(u)} = \sum_{u \in N(v)} \frac{d(u)}{2|E|} \frac{1}{d(u)} = \frac{d(v)}{2|E|} = \pi_v. \qquad \square$$

Korollar

Es gilt für $h_{v,v} = \frac{2|E|}{d(v)}$ für alle $v \in V$.

Laufzeit für Pfade

Lemma Laufzeit für Pfade

Falls $(u, v) \in E$, so gilt $h_{v,u} < 2|E|$.

Beweis:

- Seien $N(u)$ die Nachbarn von u. Es gilt
$$\frac{2|E|}{d(u)} = h_{u,u} = \sum_{w \in N(u)} \frac{1}{d(u)}(1 + h_{w,u}).$$
- Es folgt $2|E| = \sum_{w \in N(u)}(1 + h_{w,u})$.
- Wegen $v \in N(u)$ folgt sicherlich $h_{v,u} < 2|E|$. $\square$

Definition Überdeckungszeit

Für $G = (V, E)$ sei T_v die erwartete Zeit bis ein Random Walk gestartet in $v \in V$ alle Knoten von V besucht.
Wir bezeichnen $T_G = \max_{v \in V}\{T_v\}$ als *Überdeckungszeit* von G.

Überdeckungszeit

Satz Überdeckungszeit

Für alle $G = (V, E)$ gilt $T_G < 4|V| \cdot |E|$.

Beweis:

- Wähle einen Spannbaum S von G. S besitzt $|V| - 1$ Kanten.
- Wir starten Tiefensuche auf S in einem beliebigem Startknoten.
- Dies liefert eine Traversierung, die jede Kante genau einmal in beide Richtungen durchläuft.
- Ferner entspricht der Startknoten dem Endknoten.
- Sei $v_0, v_1, \ldots, v_{2(|V|-1)} = v_0$ die Tour dieser Traversierung.
- Die erwartete Zeit für diese Tour ist eine obere Schranke für T_G.
- D.h. $T_G \leq \sum_{i=0}^{2|V|-3} h_{v_i, v_{i+1}} < (2|V| - 2) \cdot 2|E| < 4|V| \cdot |E|$. $\square$

Probabilistische Pfadsuche

Frage: Existiert ein Pfad in G von s nach t?

- Deterministisch mit Breitensuche lösbar in Zeit $\mathcal{O}(|V| + |E|)$.
- Erfordert allerdings auch Speicher $\Omega(|V|)$.

Algorithmus PATH

EINGABE: $G = (V, E), s, t \in V$

1. Starte einen Random Walk in s.
2. Falls t in $4n^3$ Schritten erreicht wird, Ausgabe "Pfad".
   Sonst Ausgabe "kein Pfad".

Probabilistische Pfadsuche

Satz

Falls ein Pfad von s nach t existiert, so gibt PATH mit Ws $\geq \frac{1}{2}$ die korrekte Antwort. PATH benötigt $\mathcal{O}(\log(|V|))$ Speicher.

Beweis:

- Sei X ZV für die erwartete Zeit von s nach t per Random Walk.
- Es gilt offenbar $\mathbb{E}[X] \leq T_G < 4|V| \cdot |E| < 2n^3$. Mit Markov folgt
$$\Pr(X > 4n^3) \leq \frac{\mathbb{E}[X]}{4n^3} < \frac{1}{2}.$$
- PATH muss die jetzige Position und die Anzahl Schritte speichern.
- Dies benötigt $\mathcal{O}(\log(|V|))$ Speicher. $\square$

Motivation Monte Carlo Methode

Algorithmus APPROX-π

EINGABE: m (Anzahl der Samples)

1. Setze $Z = 0$.
2. FOR $i = 1$ to m
    1. Wähle zufälligen Punkt $P = (X, Y)$ mit $X, Y \in_R [-1, 1]$.
    2. Falls $\sqrt{X^2 + Y^2} \leq 1$, setze $Z = Z + 1$. (P ist im Einheitskreis)

AUSGABE: $Z \cdot \frac{4}{m}$ als Approximation für π

Anmerkungen:

- ZV $Z_i = 1$ gdw $\sqrt{X^2 + Y^2} \leq 1$ in der i-ten Iteration.
- Es gilt $\Pr(Z_i = 1) = \frac{\pi}{4}$ und daher $\mathbb{E}[Z] = \sum_{i=1}^{m} \mathbb{E}[Z_i] = \frac{m\pi}{4}$.
- D.h. $Z' = \frac{4Z}{m}$ ist eine gute Approximation für π.
- Die Chernoff Schranke auf Folie 52 liefert für $0 \leq \epsilon \leq 1$

$$\Pr(|Z' - \pi| \geq \epsilon\pi) = \Pr(|Z - \frac{m\pi}{4}| \geq \frac{\epsilon m\pi}{4}) = \Pr(|Z - \mathbb{E}[Z]| \geq \epsilon\mathbb{E}[Z]) \leq 2e^{-\frac{m\pi\epsilon^2}{12}}.$$

- D.h. für hinreichend großes m wird die Approximation beliebig gut.

(ϵ, δ)-Approximation

Definition (ϵ, δ)-Approximation

Die Ausgabe X eines Alg. ist eine (ϵ, δ)-Approximation für V, falls

$$\Pr(|X - V| \leq \epsilon V) \geq 1 - \delta.$$

Anmerkungen:

- APPROX-π liefert eine (ϵ, δ)-Approximation für $\epsilon \leq 1$, falls

$$2e^{-\frac{m\pi\epsilon^2}{12}} < \delta, \text{ d.h. } m \geq \frac{12\ln(\frac{2}{\delta})}{\pi\epsilon^2}.$$

(ϵ, δ)-Approximation mittels Chernoff

Satz (ϵ, δ)-Approximation mittels Chernoff

Seien $X_1, \ldots, X_m$ unabhängige IV mit $\mu = \mathbb{E}[X_i]$. Es gilt

$$\Pr\left(\left|\frac{1}{m}\sum_{i=1}^{m} X_i - \mu\right| \geq \epsilon\mu\right) \leq \delta \text{ für } m \geq \frac{3\ln(\frac{2}{\delta})}{\epsilon^2\mu}.$$

D.h. m Samples liefern eine (ϵ, δ)-Approximation für μ.

Beweis:

- Sei $X = X_1 + \ldots + X_m$. Sei $\mu' = \mathbb{E}[X] = m\mu$.
- Wir verwenden die Chernoff Schranke von Folie 52

$$\Pr(|X - \mu'| \geq \delta\mu') \leq 2e^{-\frac{\mu'\delta^2}{3}}.$$

- Es folgt

$$\Pr\left(\left|\frac{1}{m}\sum_{i=1}^{m} X_i - \mu\right| \geq \epsilon\mu\right) = \Pr(|X - \mu'| \geq \epsilon\mu') \leq 2e^{-m\cdot\frac{\mu\epsilon^2}{3}} \leq \delta. \; \square$$

DNF Counting

Szenario:

- Betrachten Probleme, die Eingaben x auf Werte $V(x)$ abbilden.

Problem DNF Counting

Gegeben:	Formel ϕ in disjunktiver Normalform (DNF)
Gesucht:	Anzahl der erfüllenden Belegungen $V(\phi)$ von ϕ

Anmerkungen:

- Beispiel für DNF-Formel: $\phi = (x_1 \wedge \overline{x_2}) \vee (x_1 \wedge x_3)$.
- Es ist einfach, die Erfüllbarkeit von DNF-Formeln zu entscheiden.

SAT $\leq_p$ DNF Counting, d.h. DNF Counting ist NP-schwer.

- Sei ϕ eine SAT-Formel. Wir betrachten $\bar{\phi}$.
- Schreibe $\bar{\phi}$ mit de Morgans Regel als DNF-Formel.
- ϕ erfüllbar gdw. es existiert eine Belegung, die $\bar{\phi}$ nicht erfüllt.
- Zähle die Anzahl der erfüllenden Belegungen von $\bar{\phi}$.
- Ist diese weniger als 2^n, so ist ϕ erfüllbar.

FPRAS

Sei $|x|$ die Eingabegröße von x.

Definition FPRAS

Ein Algorithmus A ist ein *FPRAS (fully polynomial randomized approximation scheme)*, falls A bei Eingabe x, ϵ, δ mit $0 < \epsilon, \delta < 1$ eine (ϵ, δ)-Approximation von $V(x)$ in Zeit polynomiell in $\frac{1}{\epsilon}, \ln(\frac{1}{\delta}), |x|$ liefert.

Algorithmus NAIVE-DNF COUNTING

EINGABE: DNF-Formel $\phi(x_1, \ldots, x_n)$, m

1. Setze $X = 0$.
2. FOR $i = 1$ to m
   1. Wähle uniform eine Belegung B von $x_1, \ldots, x_n$.
   2. Falls B erfüllend, setze $X := X + 1$.

AUSGABE: $Y = X \cdot \frac{2^n}{m}$ als Approximation für $V(\phi)$

Analyse NAIVE-DNF COUNTING

Satz Analyse NAIVE-DNF COUNTING

Für $V(\phi) \geq \frac{2^n}{\mathrm{poly}(n)}$ ist NAIVE-DNF COUNTING ein FPRAS.

Beweis:

- Sei die IV $X_i = 1$ gdw. B in Iteration i erfüllend. Sei $X = \sum_{i=1}^m X_i$.
- Es gilt $\mu = \Pr(X_i = 1) = \frac{V(\phi)}{2^n}$, d.h. $\mathbb{E}[Y] = \mathbb{E}[X] \cdot \frac{2^n}{m} = V(\phi)$.
- D.h. $\frac{X}{m}$ liefert eine (ϵ, δ)- Approximation für $\mu = \frac{V(\phi)}{2^n}$, falls

$$m \geq \frac{3\ln(\frac{2}{\delta})}{\epsilon^2 \mu} = \frac{3\ln(\frac{2}{\delta}) \cdot 2^n}{\epsilon^2 V(\phi)}.$$

- Damit ist $Y = \frac{X}{m} \cdot 2^n$ eine (ϵ, δ)- Approximation für $V(\phi)$.
- Für $V(\phi) \geq \frac{2^n}{\mathrm{poly}(n)}$ ist m polynomiell in $\frac{1}{\epsilon}, \ln(\frac{1}{\delta}), n$.
- D.h. NAIVE-DNF COUNTING ist ein FPRAS für DNF Counting. $\square$

Problem: Für $V(\phi) = \mathrm{poly}(n)$ benötigen wir exp. viele Samples m.

Samplen von erfüllenden Belegungen

Verbessertes Samplen:

- Sei $\phi = C_1 \vee \ldots \vee C_t$.
- OBdA enthalte keine Klausel C_i eine Variable und deren Negation.
- Sei B_i die Menge der erfüllenden Belegungen von C_i.
- Sei ℓ_i die Anzahl der Literale in C_i. Es gilt $|B_i| = 2^{n-\ell_i}$.
- Definiere $U = \{(i, b) \mid 1 \leq i \leq t \text{ und } b \in B_i\}$ mit $|U| = \sum_{i=1}^{t} 2^{n-\ell_i}$.
- Belegungen können mehrmals in U auftauchen.
- Die Anzahl erfüllender Belegungen von ϕ ist $d(\phi) = |\bigcup_{i=1}^{t} B_i|$.
- Wir zählen nur das erste Auftreten einer Belegung durch

$$S = \{(i, b) \mid 1 \leq i \leq t, b \in B_i, b \notin B_j \text{ für } j < i\} \text{ mit } |S| = d(\phi).$$

Idee: Sample uniform aus U, zähle wie oft man dabei in S landet.

DFN-COUNTING

Uniformes Samplen aus U:

- Wähle Klausel i mit Ws $\frac{|B_i|}{|U|}$.
- Wähle zufällig eine erfüllende Belegung $b \in B_i$.
- Die Literale aus C_i müssen dafür auf wahr gesetzt werden.
- Die in C_i nicht auftretenden Variablen werden uniform gesetzt.
- Damit wird jedes $(i, b) \in U$ ausgewählt mit Ws $\frac{|B_i|}{|U|} \cdot \frac{1}{|B_i|} = \frac{1}{|U|}$.

Algorithmus DNF-COUNTING

EINGABE: $\phi(x_1, \ldots, x_n) = C_1 \vee \ldots \vee C_t$ (C_i enthalte ℓ_i Literale), m

1. Setze $X = 0$.
2. Berechne $|B_i| = 2^{n-\ell_i}$ für $i = 1, \ldots, t$. Berechne $|U| = \sum_{i=1}^{t} |B_i|$.
3. FOR $k = 1$ to m
    1. Wähle Klausel i mit Ws $\frac{|B_i|}{|U|}$ aus.
    2. Wähle uniform eine erfüllende Belegung $b \in B_i$.
    3. Falls $b \notin B_j$ für $1 \leq j < i$, setze $X = X + 1$. (effizient testbar)

AUSGABE: $Y = \frac{X}{m} \cdot |U|$

Analyse von DFN-Counting

Satz DFN-Counting ist FPRAS

DFN-Counting ist ein FPRAS für $m = \lceil \frac{3t}{\epsilon^2} \ln(\frac{2}{\delta}) \rceil$.

Beweis:

- Wir wählen in 3.2 ein uniformes $b \in U$. Es gilt $\Pr(b \in S) \geq \frac{1}{t}$.
- Sei IV $X_k = 1$ gdw in Iteration k der Wert von X erhöht wird.
- D.h. $\mu = \mathbb{E}[X_k] \geq \frac{1}{t}$ bzw. $t \geq \frac{1}{\mu}$. Für die Wahl $m = \lceil \frac{3t \ln(\frac{2}{\delta})}{\epsilon^2} \rceil$ folgt
$$m \geq \frac{3 \ln(\frac{2}{\delta})}{\epsilon^2 \mu}.$$
- Mit Chernoff-Schranke: $\frac{X}{m}$ liefert eine (ϵ, δ)-Approximation von $\frac{|S|}{|U|}$.
- Damit liefert $\frac{X}{m} \cdot |U|$ eine (ϵ, δ)-Approximation von $|S|$. $\square$

Beobachtung: Geeignetes Samplen erlaubt approximatives Zählen.

Markov Ketten Monte Carlo Methode (MCMC)

Bsp: Sample in $G = (V, E)$ uniform eine unabhängige Menge.

Idee: Konstruiere eine Markov Kette mit folgenden Eigenschaften

- Die Zustände bestehen aus den unabhängigen Mengen in G.
- Der stationäre Zustand π ist die Gleichverteilung.

- Sei X_0 ein Startzustand und $X_0, X_1, \ldots$ ein Lauf der Kette.
- Nach einer hinreichend großen Zahl Schritte r erreichen wir π.
- Verwende $X_r, X_{2r}, X_{3r}, \ldots$ als Approximation uniformer Samples.
- Man kann r und die Qualität der Samples explizit bestimmen.

MCMC mit uniformer Verteilung

Frage: Wann ist π uniform?

- In Graphen ist die stationäre Verteilung abhängig vom Grad.
- Idee: Erzeuge gleichen Grad M durch Selbstkanten.
- Graphen mit Selbstkanten sind nicht bipartit.

Satz Uniforme stationäre Verteilung

Sei Ω ein endlicher Zustandsraum mit Nachbarn $\{N(x) \mid x \in \Omega\}$. Sei $N = \max_{x \in \Omega} |N(x)|$ und $M \in \mathbb{N}$ mit $M \geq N$. Eine Markov Kette mit

$$P_{x,y} = \begin{cases} \frac{1}{M} & \text{für } x \neq y \text{ und } y \in N(x) \\ 0 & \text{für } x \neq y \text{ und } y \notin N(x) \\ 1 - \frac{N(x)}{M} & \text{für } x = y \end{cases}$$

besitzt uniforme stationäre Verteilung.

Beweis: (ohne Beweis)

Uniformes Samplen unabhängiger Mengen

Algorithmus Markov Kette ISET

EINGABE: $G = (V, E)$

1. Wähle Startzustand $X_0 = \emptyset$.
2. Berechnung von X_{i+1} aus X_i für alle $i \geq 0$:
   1. Wähle $v \in_R V$.
   2. Falls $v \in X_i$ setze $X_{i+1} = X_i \setminus \{v\}$.
   3. Falls $v \notin X_i$ und $X_i \cup \{v\}$ unabhängig ist, setze $X_{i+1} = X_i \cup \{v\}$.
   4. Sonst setze $X_{i+1} = X_i$.

Anmerkungen:

- Jedes X_i ist nach Konstruktion eine unabhängige Menge.
- Benachbarte X_i unterscheiden sich in höchstens einem Knoten.
- Jede unabhängige Menge von G kann von ISET erreicht werden.

Uniformes Samplen unabhängiger Mengen

Satz Unifomes Samplen unabhängiger Mengen

ISET besitzt uniforme stationäre Verteilung π, falls $|E| \geq 1$.

Beweis:

- Sei $\{u, v\} \in E$. Angenommen wir sind im Zustand $X_i = \{u\}$.
- Dann gilt $P_{u,u} > 0$, d.h. wir haben eine Selbstkante.
- Für $x \neq y$ gilt entweder $P_{x,y} = \frac{1}{|V|}$ oder $P_{x,y} = 0$.
- Ferner gilt $\sum_y P_{x,y} = 1$.
- Damit ist voriger Satz (Folie 116) anwendbar, und π ist uniform. $\square$