Karlsruher Institut für Technologie

Maximilian Katzmann
Hans-Peter Lehmann
Stefan Walzer

Übungsblatt 11

Randomisierte Algorithmik – Wintersemester 2023/2024

Aufgaben zur Vorlesung vom 18.1.2024
Abgabe im ILIAS bis 25.1.2024, 11:30 Uhr
Besprechung am 30.01.2024, 08:00 Uhr

Achte insbesondere bei handschriftlichen Abgaben auf Lesbarkeit. Die Abgabe erfolgt über das Übungsmodul im ILIAS. Gib Deine Ausarbeitungen in *einer* PDF-Datei ab.

Für dieses Blatt ist der Begriff von d-Unabhängigkeit zentral. Weil dieser in der Vorlesung erst kurz vor Ende kam und vielleicht nicht ganz klar wurde, nochmal eine kompakte Definition (etwas anders formuliert als auf den Folien): Eine Familie $\mathcal{X}$ von Zufallsvariablen heißt d-unabhängig, wenn jede Auswahl von bis zu d Zufallsvariablen aus $\mathcal{X}$ unabhängig ist. Eine Auswahl von $d+1$ Zufallsvariablen aus $\mathcal{X}$ könnte dagegen abhängig sein. Aufgabe 3 ist als intuitives Beispiel gedacht.

Eine Familie von Hashfunktionen $\mathcal{H} \subseteq [m]^D$ heißt d-unabhängig, wenn für beliebige d verschiedene Schlüssel $x_1, \dots, x_d$ die Hashwerte $h(x_1), \dots, h(x_d)$ unter einer zufälligen Hashfunktion $h \sim \mathcal{U}(\mathcal{H})$ unabhängige Zufallsvariablen sind, die jeweils gleichverteilt in $[m]$ sind.

***Bemerkung:** Aufgabe 2 ist umfangreich und anspruchsvoll. Es ist keine Schande nicht alle Teilaufgaben zu schaffen.*

Aufgabe 1 – 2−Unabhängigkeit vs. 1-Universalität

Sei $\mathcal{H} \subseteq [m]^D$ eine Familie von Hashfunktionen D nach $[m]$. Zeige oder widerlege, dass folgende Implikationen gelten:

(a) $\mathcal{H}$ ist 2-unabhängig $\Rightarrow$ $\mathcal{H}$ ist 1-universell.

(b) $\mathcal{H}$ ist 1-universell $\Rightarrow$ $\mathcal{H}$ ist 2-unabhängig.

Hinweis: In einem Fall lässt sich die Implikationen leicht zeigen. Im anderen Fall kann man dämliche Gegenbeispiele finden.

Aufgabe 2 – Konzentrationsschranken für Summen d-unabhängiger Zufallsvariablen

Sei $d \in \mathbb{N}$ gerade und $\{X_1, \ldots, X_n\}$ eine d-unabhängige Familie von Zufallsvariablen, die alle Verteilung $\mathrm{Ber}(p)$ mit $p = \Omega(1/n)$ haben. Wir betrachten die Summe $X = \sum_{i=1}^{n} X_i$. Beachte: Weil $X_1, \ldots, X_n$ nicht unabhängig sind ist X nicht unbedingt binomialverteilt!

Ziel dieser Aufgabe ist es, eine Konzentrationsschranke für X zu zeigen, nämlich, dass für beliebige $\delta > 0$ gilt:

$$\Pr[X - \mathbb{E}[X] \geq \delta \mathbb{E}[X]] = O(\delta^{-d}(np)^{-d/2}).$$

Dazu schauen wir die „zentrierten" Zufallsvariablen $Y_i := X_i - p$ an, deren Summe $Y = \sum_{i=1}^{n} Y_i$ und den Erwartungswert $\mathbb{E}[Y^d]$.

(i) Zum Aufwärmen: Sei $d \geq 3$ und $n \geq 3$. Überzeuge dich, dass folgendes gilt und erkläre kurz warum.

(a) $\mathbb{E}[Y_1^5 Y_2^{42}] = \mathbb{E}[Y_1^5]\mathbb{E}[Y_2^{42}]$

(b) $\mathbb{E}[Y_1^5 Y_2^{42} Y_3] = 0$

(c) $\mathbb{E}[Y_1^5] \leq \mathbb{E}[Y_1^2]$

Im weiteren Verlauf darfst du die zugrundeliegenden Einsichten ohne weitere Begründung verallgemeinert verwenden.

(ii) Zeige: $\mathbb{E}[Y_1^2] \leq p$.

(iii) Seien $i_1, \ldots, i_d \in [n]$ (nicht notwendig verschieden) sowie $S = \{i_1, \ldots, i_d\}$. Zeige:

- Falls $|S| > d/2$ dann gilt $\mathbb{E}[Y_{i_1} \cdot \ldots \cdot Y_{i_d}] = 0$.
- Andernfalls gilt $\mathbb{E}[Y_{i_1} \cdot \ldots \cdot Y_{i_d}] \leq p^{|S|}$.

(iv) Zeige $\mathbb{E}[Y^d] = O((np)^{d/2})$. Du darfst annehmen, dass $d = O(1)$ gilt.
Hinweis: Multipliziere $(\sum_{i=1}^{n} Y_i)^d$ aus. Ja, das ergibt n^d Terme.

(v) Beweise das ursprüngliche Ziel dieser Aufgabe indem du die Markov Ungleichung auf Y^d anwendest.

Aufgabe 3 – Bonus: d-Unabhängigkeit ohne Unabhängigkeit

Alice und Bob drehen beide an einem Glücksrad, das m gleich große Segmente mit den Zahlen von 0 bis 9 hat. Seien A und B die Ergebnisse von Alice und Bob. Sei $C = (A + B) \bmod m$.

(a) Zeige: A, B, C sind paarweise unabhängig (d.h. eine 2-unabhängige Familie von Zufallsvariablen).

(b) Zeige: A, B, C sind nicht unabhängig.

(c) Finde für beliebiges $d \in \mathbb{N}$ eine Familie von Zufallsvariablen, die d-unabhängig ist aber nicht unabhängig ist.