4. Übungsblatt

Abgabe: Montag, 6. Dezember 2004, *zu Beginn* der Vorlesung

Besprechung: Donnerstag, 9. Dezember 2004, Raum -101, 11:30 Uhr

Aufgabe 1
Sie werfen wieder Bälle zufällig und gleichverteilt in k Körbe.

(a) Wie oft müssen Sie werfen (Erwartungswert), um jedem Korb mit mindestens einem Ball zu füllen?

Tipp: Sie können z.B. folgendermaßen vorgehen: Überlegen Sie sich die Wahrscheinlichkeit, den i-ten Korb zu füllen, wenn schon $i-1$ Körbe gefüllt sind. Welche Verteilung liegt hier zu Grunde? Sie kennen den zugehörigen Erwartungswert bereits – verwenden Sie ihn.

(b) Wie groß ist die Varianz?

Tipp: Sie könnten zunächst die Varianz der obigen Verteilung bestimmen $(1-p)/p^2 \ldots$

Aufgabe 2
Seien $X_i : \Omega \to \mathbb{R}_+$ unabhängige und beschränkte ($X_i \leq C$) Zufallsvariablen. Höffdings-Schranke (eine der vielen Variationen der Chernoff-Schranken) lautet dann:

$$\mathbf{Pr}\left(\left|\frac{X_1 + \ldots + X_k}{k} - \mathbf{E}\left[\frac{X_1 + \ldots + X_k}{k}\right]\right| \geq \delta\right) \leq e^{-2k\delta^2/C^2}$$

Gegeben sei ein ungerichteter und einfacher Graph G in Adjazenzmatrix Repräsentation. Eine *l-Clique* ist ein vollständiger Subgraph von G mit l Knoten. Die Anzahl aller l-Cliquen eines Graphen wird mit $K_l(G)$ angegeben. Die *l-Cliquen-Dichte* von G sei definiert durch:

$$D_l(G) = \frac{K_l(G)}{\binom{n}{l}}$$

(a) Überlegen Sie sich einen sehr einfachen Algorithmus, um die l-Cliquen-Dichte in $\Theta\left(\binom{n}{l}\binom{l}{2}\right)$ Zeit zu berechnen.

(b) Schätzen sie nun die l-Cliquen-Dichte mit Stichprobenerhebung. Ihr Algorithmus sollte einen Wert im Intervall $[D_l - \delta, D_l + \delta]$ mit Wahrscheinlichkeit p berechnen. Geben Sie die Laufzeit in Abhängigkeit von l, δ und p an.

(c) Können Sie die Anzahl der l-Cliquen abschätzen?

Aufgabe 3
In der Vorlesung haben Sie ein inkrementell-randomisierten Algorithmus gesehen, der die Trapezzerlegung $H(N)$ von einer Menge N von Segmenten in allgemeiner Lage bestimmt.

(a) Basierend auf diesem Algorithmus, beschreiben Sie einen Algorithmus, der auch für degenerierte Eingaben $H(N)$ liefert. Zeigen Sie, dass die erwartete Laufzeit dadurch nicht geändert wird.

Sei $\bar{N}$ eine Menge von $\bar{n}$ monotonen Ketten. Eine *Kette* ist eine Folge $(s)_{1 \leq i \leq k}$ von Segmenten, in der jede Segment s_i einen Endpunkt gemeinsam mit s_{i-1} und den anderen mit s_{i+1} hat, für $2 \leq i \leq k-1$. Eine Kette heißt *monoton*, falls jede Gerade, die parallel zur y-Achse ist, diese Kette höchstens in einem Punkt schneidet. Sei jetzt N die Menge aller Segmente, die zu diesen Ketten gehören, n die Anzahl der Segmente in N und k die Anzahl der Kreuzungen zwischen Segmenten in N.

(b) Beschreiben Sie einen randomiserten Algorithmus, der die Trapezzerlegung $H(N)$ in erwartet $O(n + k + \bar{n} \log \bar{n})$ Zeit berechnet. Beweisen Sie diese Laufzeit.

Tipp: Ändern Sie den inkrementell-randomisierten Algorithmus, den Sie in der Vorlesung gesehen haben, indem Sie in jedem Schritt zufällig eine neue *Kette* einfügen.