Aufgaben zu Kapitel 4 der Vorlesung „Randomisierte Algorithmen“

Aufgabe 4.1
Zeigen Sie: Sind X und Y unabhängige Zufallsvariablen (die numerische Werte haben), dann sind auch e^X und e^Y unabhängige Zufallsvariablen.

Lösung 4.1
Im vorliegenden Fall reicht eine ganz simple Argumentation.

Zwei ZV heißen unabhängig, wenn für alle x, y gilt:

$$\mathbf{Pr}\left[X = x \wedge Y = y\right] = \mathbf{Pr}\left[X = x\right] \cdot \mathbf{Pr}\left[Y = y\right]$$

Daher:

$$\begin{aligned}
\mathbf{Pr}\left[e^X = x \wedge e^Y = y\right] &= \mathbf{Pr}\left[X = \ln x \wedge Y = \ln y\right] && \text{Injektivität von } \exp() \\
&= \mathbf{Pr}\left[X = \ln x\right] \cdot \mathbf{Pr}\left[Y = \ln y\right] && \text{Unabhängigkeit von } X \text{ und } Y \\
&= \mathbf{Pr}\left[e^X = x\right] \cdot \mathbf{Pr}\left[e^Y = y\right]
\end{aligned}$$

Aufgabe 4.2
Eine Zufallsvariable X_i heißt *geometrisch verteilt mit Parameter p*, wenn sie Werte $t \in \mathbb{N} = \{1, 2, 3, \ldots\}$ annimmt und gilt:

$$\mathbf{Pr}\left[X_i = t\right] = p(1-p)^{t-1} .$$

1. Berechnen Sie den Erwartungswert einer geometrisch verteilten Zufallsvariablen.
2. Es seien nun $X_1, \ldots, X_n$ unabhängige, identisch und geometrisch (mit gleichem Parameter p) verteilte Zufallsvariablen und $X = X_1 + \cdots + X_n$. Berechnen Sie den Erwartungswert μ von X.

Lösung 4.2

1.

$$\begin{aligned}
\mathbf{E}[X_i] &= \sum_{t=1}^{\infty} tp(1-p)^{t-1} \\
&= p\sum_{t=1}^{\infty} t(1-p)^{t-1} \\
&= p \cdot \frac{1}{(1-(1-p))^2} && \text{siehe Aufgabenblatt 1} \\
&= 1/p
\end{aligned}$$

2. $\mathbf{E}[X] = \mathbf{E}[\sum X_i] = \sum \mathbf{E}[X_i] = n/p$

Aufgabe 4.3
Eine Münze, die bei jedem Wurf mit Wahrscheinlichkeit $1/2$ Zahl zeigt, wird n Mal (unabhängig) geworfen. Die binäre ZV X_i sei 1, falls beim i-ten Wurf Zahl kommt und 0 sonst. Es sei $X = \sum X_i$.

- Welche Schranke liefert die Chebyshev-Ungleichung für die Wahrscheinlichkeit
$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right] \text{ ?}$$
- Welches Ergebnis liefert die Anwendung der Chernoff-Schranken für
$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right] \text{ ?}$$
- Welches Ergebnis liefert die Anwendung der Chernoff-Schranken für
$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{1}{2}\sqrt{6n\ln n}\right] \text{ ?}$$

Lösung 4.3
Wir bentuzen folgende Formulierung der Chebyshev-Ungleichung:

$$\mathbf{Pr}\left[|X - \mu_X| \geq a\right] \leq \frac{\mathbf{var}[X]}{a^2} .$$

Außerdem nutzen wir die Tatsache, dass für unabhängige ZV X_i gilt:

$$\mathbf{var}\left[\sum_i X_i\right] = \sum_i \mathbf{var}[X_i]$$

Im Falle der Münzwürfe ist $X_i^2 = X_i$, daher $\mathbf{E}\left[X_i^2\right] = 1/2$, $\mathbf{var}\left[X_i\right] = 1/2 - 1/4 = 1/4$ und $\mathbf{var}\left[X\right] = n/4$.

Damit ergibt sich:

- Chebyshev-Ungleichung:

$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right] \leq \frac{n/4}{(n/4)^2} = \frac{4}{n}$$

- Chernoff-Schranke:

$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right]$$

  ist wegen $\mu = n/2$ und $0 < \delta = 1/2 < 1$ eine Abschätzung nach oben durch $2\exp(-\mu\delta^2/5)$ möglich:

$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right] \leq 2\exp(-\frac{1}{5}\frac{n}{2}\frac{1}{2^2}) = 2e^{-n/40}$$

- Chernoff-Schranke im zweiten Fall:

$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{1}{2}\sqrt{6n\ln n}\right]$$

  ist $\delta = \frac{1}{2}\sqrt{6n\ln n}/\mu = \frac{1}{2}\sqrt{6n\ln n}/(n/2)$ und daher $\delta^2 = \frac{6\ln n}{n}$ und $-\mu\delta^2/5 = 6\ln n/10$. Also

$$\mathbf{Pr}\left[\left|X - \frac{n}{2}\right| \geq \frac{1}{2}\sqrt{6n\ln n}\right] \leq 2e^{-0.6\ln n} = 2\left(\frac{1}{n}\right)^{0.6}$$

Aufgabe 4.4

Betrachten Sie die folgende Variante randBitFixing des Bit-Fixing-Algorithmus:

- Solange aktuelle Adresse x und Zieladresse y verschieden sind, wird aus den Bitpositionen, an denen sich x und y unterscheiden, zufällig gleichverteilt ein i gewählt und der Pfad von x nach $x \oplus e_i$ fortgesetzt.

Es soll bewiesen werden, dass wie beim deterministischen Bit-Fixing auch bei dieser Vorgehensweise beim Routen der Permutation „Matrix-Transposition" mit großer Wahrscheinlichkeit noch „große" Staus entstehen.

Hier ein paar Hinweise zu einer möglichen Vorgehensweise:

- Es sei $c = d/2$.

- Betrachten Sie die Pakete, die in den $\sqrt{N}$ Knoten mit den Adressen

$$x = (\underbrace{x_1, x_2, \ldots, x_c}_{c \text{ Bits}}, \underbrace{0, 0, \cdots, 0}_{c \text{ Bits}})$$

  starten.
- Betrachten Sie eine beliebige aber feste Zahl k mit $1 \le k \le c$ (die Sie später geeignet wählen) und die Menge

$$S_k = \{(x_1, x_2, \ldots, x_c, 0, 0, \cdots, 0) \mid \text{genau } k \text{ der ersten } c \text{ Bits sind } 1\}$$

Aufgaben:

1. Beweisen Sie für alle $1 \le k \le n$:
   (a) $\left(\frac{n}{k}\right)^k \le \binom{n}{k}$ (b) $\binom{n}{k} \le \left(\frac{en}{k}\right)^k$.
   Hinweis zu (b): Stirlings Formel.

2. Wie groß ist S_k? Geben Sie eine Abschätzung ohne Binomialkoeffizienten an.

3. Es sei $x \in S_k$ und Y_x die Zufallsvariable mit

$$Y_x = \begin{cases} 1 & \text{falls } x \text{ durch Knoten } (0, 0, \ldots, 0) \text{ transportiert wird} \\ 0 & \text{sonst} \end{cases}$$

   Geben Sie eine Abschätzung für $\mathbf{E}[Y_x]$ an, in der keine Binomialkoeffizienten vorkommen.

4. Es sei $Z_k = \sum_{x \in S_k} Y_x$. Schätzen Sie $\mathbf{E}[Z_k]$ nach unten ab.

5. Geben Sie eine legale Wahl für k an, so dass $\mathbf{E}[Z_k]$ exponenziell in n ist. Was bedeutet das?

Lösung 4.4
Die folgende Argumentation findet man bei David Karger (Vorlesung Randomized Algorithms (6.856J/18.416J), `http://courses.csail.mit.edu/6.856/current/`)

1. (a) für $1 \le k \le n$ ist

$$\frac{n}{k} \cdot \frac{k-1}{n-1} = \frac{nk-n}{nk-k} \le 1$$

also $\frac{n}{k} \leq \frac{n-1}{k-1}$, und folglich

$$\begin{aligned}
\binom{n}{k} &= \frac{n \cdot (n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k} \\
&= \frac{n}{k} \cdot \frac{n-1}{k-1} \cdots \frac{n-k+1}{1} \\
&\geq \frac{n}{k} \cdots \frac{n}{k} = \left(\frac{n}{k}\right)^k
\end{aligned}$$

(b)

$$\begin{aligned}
\binom{n}{k} &= \frac{n \cdot (n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k} \\
&\leq \frac{n^k}{k!} \\
&\leq \frac{n^k \cdot e^k}{\sqrt{2\pi k} \cdot k^k (1 + h(k))} \\
&\leq \left(\frac{en}{k}\right)^k
\end{aligned}$$

2. $|S_k| = \binom{c}{k}$ also
$$\left(\frac{c}{k}\right)^k \leq |S_k| \leq \left(\frac{ec}{k}\right)^k$$

3. Insgesamt müssen beim Transport $2k$ Adressbits geändert werden. Es gibt $\binom{2k}{k}$ Möglichkeiten, welches diejenigen k sind, die als erste aktualisiert werden. Damit das Paket durch den Knoten $(0, \ldots, 0)$ transportiert wird, hat man nur eine Wahl: Es müssen die k Stellen in der linken Hälfte der Adresse sein. Also ist
$$\mathbf{E}\left[Y_x\right] = \frac{1}{\binom{2k}{k}} \geq \frac{1}{\left(\frac{2ek}{k}\right)^k} = \left(\frac{1}{2e}\right)^k$$

4.

$$\begin{aligned}
\mathbf{E}\left[Z_k\right] = \sum_{x \in S_k} \mathbf{E}\left[Y_x\right] &\geq \binom{n/2}{k} \left(\frac{1}{2e}\right)^k \geq \left(\frac{n}{2k}\right)^k \left(\frac{1}{2e}\right)^k \\
&= \left(\frac{n}{4ek}\right)^k
\end{aligned}$$

5. wähle $k = n/8e$; dann ist $\mathbf{E}\left[Z_k\right] \geq \left(\frac{n}{4en/8e}\right)^{n/8e} = 2^{n/8e}$