ETH Zürich
Institute of Theoretical Computer Science
Prof. Rasmus Kyng
Prof. Angelika Steger

FS 2024

Algorithmen und Wahrscheinlichkeit
Theorie-Aufgaben 7

Lösung zu Aufgabe 1

(a) Wir bezeichnen für jede Strasse $e_i = \{v_{i-1}, v_i\}$ für $i = 1, \ldots, k$ mit X_i die Indikatorvariable, dass an dieser Strasse Blumen spriessen.

Die Anzahl der Strassen an der Blumen spriessen ist dann $X = X_1 + \ldots + X_k$. Die erwartete Anzahl Strassen an denen Blumen spriessen können wir mithilfe der Linearität des Erwartungswerts berechnen als

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \ldots + \mathbb{E}[X_k] = \frac{k}{2},$$

wobei wir verwendet haben, dass $\mathbb{E}[X_i] = 1/2$ für alle $i = 1, \ldots, k$.

Die Wahrscheinlichkeit, dass mindestens an 3/4 der Strassen Blumen spriessen (also dass der Spaziergang kein Vergnügen ist), lässt sich dann mit Markovs Ungleichung wie folgt abschätzen:

$$\Pr\left[X \geq \frac{3k}{4}\right] \leq \frac{\mathbb{E}[X]}{3k/4} = \frac{k/2}{3k/4} = \frac{2}{3}.$$

(b) Um Chebychevs Ungleichung anwenden zu können, berechnen wir zunächst $\mathrm{Var}[X]$. Da die X_i unabhängig sind, folgt:

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_k] = \frac{k}{4},$$

wobei wir verwendet haben, dass $\mathrm{Var}[X_i] = 1/4$, da X_i eine Bernoulli Zufallsvariable mit Parameter 1/2 ist.

Wir können nun folgende Abshätzung vornehmen:

$$\Pr\left[X \geq \frac{3k}{4}\right] = \Pr\left[X - \mathbb{E}[X] \geq \frac{k}{4}\right] \leq \Pr\left[|X - \mathbb{E}[X]| \geq \frac{k}{4}\right] \leq \frac{\mathrm{Var}[X]}{(k/4)^2} = \frac{4}{k}.$$

Anmerkung: Wir haben hier die Abschätzung $\Pr[X - \mathbb{E}[X] \geq a] \leq \Pr[|X - \mathbb{E}[X]| \geq a]$ verwendet, die für alle Zufallsvariablen X zutrifft. In unseren Spezialfall ist $X \sim \mathrm{Bin}(n, 1/2)$ symmetrisch, und wir wissen sogar, dass $\Pr[X - \mathbb{E}[X] \geq a] = \frac{1}{2}\Pr[|X - \mathbb{E}[X]| \geq a]$, womit wir unsere finale Abschätzung sogar um einen Faktor 2 verbessern könnten.

(c) Wir kümmern uns zunächst nur um unseren eigenen Hund und überlegen uns, wie wahrscheinlich es ist, dass er schnieft. Das passiert, wenn an allen k Strassen unseres Spaziergangs Blumen spriessen. Das passiert mit Wahrscheinlichkeit $\left(\frac{1}{2}\right)^k$, da alle k Strassen entlang unseres Spaziergangs unabhängig voneinander Blumen aufweisen.

Sei nun S_i die Indikatorvariable, dass der Hund unseres i-ten Freundes schnieft und S_n die Indikatorvariable, dass unser Hund schnieft. Die Anzahl schniefender Hunde ist $S = S_1 + \ldots + S_n$, dank der Linearität des Erwartungswert ist die erwartete Anzahl schniefender Hunde $\mathbb{E}[S] = \mathbb{E}[S_1] + \ldots + \mathbb{E}[S_n]$.

Oben haben wir gesehen, dass $\mathbb{E}[S_n] = \left(\frac{1}{2}\right)^k$. Mit genau dem gleichen Argument sieht mann, dass $\mathbb{E}[S_i] = \left(\frac{1}{2}\right)^k$ für $i = 1, \ldots, n-1$. Daraus folgt $\mathbb{E}[S] = n \cdot \left(\frac{1}{2}\right)^k$.

Die Wahrscheinlichkeit, dass es einen schniefenden Hund gibt, ist $\Pr[S \geq 1]$, die wir wie folgt mit der Markov Ungleichung abschätzen können:

$$\Pr[S \geq 1] \leq \frac{\mathbb{E}[S]}{1} = n \cdot \left(\frac{1}{2}\right)^k .$$

Falls $k \geq \log_2 n + 1$, dann ist $\Pr[S \geq 1] \leq 1/2$. Also ist die Wahrscheinlichkeit, dass kein Hund shnieft mindestens $1/2$.

(d) Zuerst merken wir an, dass die Aufgabenstellungen für $n = 1$ nicht interessant ist (Spaziergänge der Länge $k = 0$ sind sicher kein Vergnügen) und nehmen deshalb $n \geq 2$ an.

Wir kümmern uns wieder zuerst um unseren eigenen Hund und versuchen eine noch bessere Schranke als in (a) und (b) zu finden, dass unser Spaziergang kein Vergnügen wird. Dafür verwenden wir die Chernoff-Ungleichung. Zunächst bemerken wir dass X eine Summe unabängiger Bernoulli-Variablen ist, dass wir also Chernoff anwenden können. Dann gilt

$$\begin{aligned}
\Pr\left[X \geq \mathbb{E}[X] + \frac{k}{4}\right] &= \Pr\left[X \geq \left(1 + \frac{1}{2}\right)\mathbb{E}[X]\right] && (1)\\
&\leq e^{-\frac{1}{3}\cdot\frac{1}{4}\,\mathbb{E}[X]} && (2)\\
&= e^{-k/24} && (3)\\
&= e^{-1000\log_2 n/24} && (4)\\
&\leq 2^{-1000\log_2 n/24} && (5)\\
&\leq n^{-1000/24} && (6)\\
&\leq n^{-40} && (7)
\end{aligned}$$

Sei P_i die Indikatorvariable, dass der Spaziergang unseres i-ten Freundes kein Vergnügen wird und $P = P_1 + \ldots + P_n$. Analog zu (c) erhalten wir

$$\Pr[P \geq 1] \leq \mathbb{E}[P] \leq n \cdot \mathbb{E}[P_n] \leq n \cdot n^{-40} = n^{-39}.$$

Wenn wir nun $n \geq 2$ verwenden, erhalten wir $\Pr[P \geq 1] \leq 2^{-39} \leq 2^{-7} \leq 0.01$, dass heisst mit Wahrscheinlichkeit mindestens 0.99 werden alle Spaziergänge ein Vergnügen.