Übung 5

5.1 Fixpunkte einer Permutation (2pt)

Ein *Fixpunkt* einer Permutation Π der Folge $1, 2, \ldots, n$ ist ein Index i für welchen $\Pi_i = i$. Finden Sie die Varianz der Anzahl Fixpunkte in einer uniform gewählten zufälligen Permutation. (Hinweis: Benutzen Sie eine Indikator-Zufallsvariable X_i für das Ereignis $\Pi_i = i$ und setzen Sie $X = \sum_{i=1}^{n} X_i$.)

5.2 Varianz (2pt)

Zeigen Sie, dass für jede Konstante $c > 0$ und jede diskrete Zufallsvariable X gilt, dass $\mathrm{Var}[cX] = c^2 \mathrm{Var}[X]$.

5.3 Einseitige Chebyshev-Ungleichung (4pt)

Zeigen Sie, dass für eine Zufallsvariable $X \in \mathbb{N}$ mit Standardabweichung σ und beliebiges $t > 0$ gilt

$$\mathrm{P}\big[\, X - \mathrm{E}[X] \geq t\sigma \,\big] \;\leq\; \frac{1}{1+t^2}.$$

Diese Beziehung ist auch bekannt als Chebyshev-Cantelli-Ungleichung. Sie ist verwandt mit der Chebyshev-Ungleichung, jedoch weder stärker noch schwächer.

Der Beweis erfolgt in drei Schritten.

1. Transformieren Sie X in eine Zufallsvariable Y, deren Erwartungswert 0 ist und für welche gilt $\mathrm{Var}[Y] = \mathrm{Var}[X]$.
2. Wenden Sie die Markov-Ungleichung an auf $\mathrm{P}[Y \geq t\sigma]$. Transformieren Sie vorher den Ausdruck so, dass der Erwartungswert der Zufallsvariablen grösser als Null wird. Erhöhen Sie dazu Y um ein beliebiges $a > 0$. Warum kann man die Markov-Ungleichung nicht direkt auf Y anwenden?
3. Setzen Sie jetzt $a = \frac{\sigma}{t}$ ein und leiten Sie die Ungleichung her.

*Bonusaufgabe: Leiten Sie daraus ebenfalls die folgende zweiseitge Abschätzung her:

$$\mathrm{P}\Big[\, \big|X - \mathrm{E}[X]\big| \geq t\sigma \,\Big] \;\leq\; \frac{2}{1+t^2}.$$

Transformieren Sie als ersten Schritt X in eine Zufallsvariable Z, für die gilt $\mathrm{E}[Z] = 0$ und $\mathrm{Var}[Z] = \mathrm{Var}[X]$ (wie für Y), die aber verschieden von Y ist. Im zweiten Schritt verwenden Sie den Union-Bound.

$\Longrightarrow$

5.4 Varianz im Sammlerproblem (2pt)

In Aufgabe 3.3 haben Sie die Verteilung der Zielorte von Fritz' Reisen analysiert. Dies ist eine Formulierung des *Sammlerproblems*: Es gibt n verschiedene Bilder und pro Experiment erhält man ein Bild mit uniformer Verteilung und unabhängig von allen anderen. Wie lange dauert es bis man von allen Bildern wenigstens eines hat? Sei dies die Zufallsvariable X. Der Erwartungswert von X ist $n\sum_{i=1}^{n}\frac{1}{i} = nH(n)$, wobei $H(n)$ die n-te *harmonische Zahl* ist, für die gilt $H(n) = \ln n + \Theta(1)$. Siehe auch [MU17, 2.4.1].

a) Berechnen Sie eine obere Schranke für Varianz von X. (Hinweis: $\sum_{i=1}^{\infty}\frac{1}{i^2} = \frac{\pi^2}{6}$.)

b) Schätzen Sie mit der Chebyshev-Ungleichung die Wahrscheinlichkeit ab, dass jemand mehr als viermal so lange wie erwartet auf alle Bilder warten muss.

Referenzen

[MU17] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, 2nd edition, 2017.