Übungsblatt 6 – Concentration Bounds

Randomisierte Algorithmik

Aufgabe 1 – Rechenregeln für Varianz

Seien X, Y unabhängige Zufallsvariablen deren Varianz existiert. Seien ferner $s, t > 0$. Zeige, dass gilt:

(a) $\mathrm{Var}(sX) = s^2\mathrm{Var}(X)$

(b) $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

(c) $\mathrm{Var}(sX + tY) = s^2\mathrm{Var}(X) + t^2\mathrm{Var}(Y)$

Hinweis: Nutze $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$. Wenn du Lust hast, kannst du das auch nochmal aus der Definition von Unabhängigkeit (für diskrete Zufallsvariablen) herleiten, die lediglich $\Pr[X = i \wedge Y = j] = \Pr[X = i] \cdot \Pr[Y = j]$ für alle i, j garantiert.

Lösung 1

Wir zeigen zunächst nochmal $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$. Seien hierfür $R_X, R_Y \subseteq \mathbb{R}$ abzählbare Mengen, die alle möglichen Werte für X und Y enthalten. Dann gilt:

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= \sum_{(x,y)\in R_X\times R_Y} x \cdot y \cdot \Pr[X = x \wedge Y = y] \\
&\overset{\text{Unabh.}}{=} \sum_{x\in R_X} \sum_{y\in R_Y} x \cdot y \cdot \Pr[X = x] \Pr[Y = y] \\
&= \sum_{x\in R_X} \Big(x \cdot \Pr[X = x] \sum_{y\in R_Y} y \cdot \Pr[Y = y]\Big) \\
&= \Big(\sum_{x\in R_X} x \cdot \Pr[X = x]\Big)\Big(\sum_{y\in R_Y} y \cdot \Pr[Y = y]\Big) \\
&= \mathbb{E}[X] \cdot \mathbb{E}[Y].
\end{aligned}$$

Wir beweisen nun die drei einfachen Rechenregeln (und machen das sehr ausführlich).

(a) Hier ist neben der Definition der Varianz nur die Einsicht $\mathbb{E}[sZ] = s\mathbb{E}[Z]$ nötig (folgt aus Linearität des Erwartungswertes). Letzteres gilt für jede Zufallsvariable Z, deren Erwartungswert existiert und jedes $s \in \mathbb{R}$. Also:

$$\begin{aligned}
\mathrm{Var}(sX) &= \mathbb{E}[(sX - \mathbb{E}[sX])^2] = \mathbb{E}[(sX - s\mathbb{E}[X])^2] \\
&= \mathbb{E}[s^2(X - \mathbb{E}[X])^2] = s^2\mathbb{E}[(X - \mathbb{E}[X])^2] = s^2\mathrm{Var}(X).
\end{aligned}$$

(b) Zunächst haben zentrierte Zufallsvariablen Erwartungswert 0:

$$\mathbb{E}[X - \mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[\mathbb{E}[X]] = \mathbb{E}[X] - \mathbb{E}[X] = 0. \tag{1}$$

Wenn zwei Zufallsvariablen X und Y unabhängig sind, dann sind auch $X - c$ und $Y - d$ unabhängig für beliebige Konstanten $c, d \in \mathbb{R}$. Wenn wir $c = \mathbb{E}[X]$ und $d = \mathbb{E}[Y]$ setzen folgt, dass $X - \mathbb{E}[X]$ unabhängig von $Y - \mathbb{E}[Y]$ ist. Also folgt:

$$\mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])] = \mathbb{E}[X - \mathbb{E}[X]] \cdot \mathbb{E}[Y - \mathbb{E}[Y]] \overset{(1)}{=} 0 \cdot 0 = 0. \tag{2}$$

Wir betrachten nun den Term $(X + Y - \mathbb{E}[X + Y])^2$, dessen Erwartungswert die gesuchte Varianz ist.

$$\begin{aligned}(X + Y - \mathbb{E}[X + Y])^2 &= (X + Y - \mathbb{E}[X] - \mathbb{E}[Y])^2 = \big((X - \mathbb{E}[X]) + (Y - \mathbb{E}[Y])\big)^2 \\ &= (X - \mathbb{E}[X])^2 + (Y - \mathbb{E}[Y])^2 + 2(X - \mathbb{E}[X])(Y - \mathbb{E}[Y]).\end{aligned}$$

Wenn wir nun den Erwartungswert auf beiden Seiten nehmen erhalten wir:

$$\begin{aligned}\mathrm{Var}(X + Y) &= \mathbb{E}[(X + Y - \mathbb{E}[X + Y])^2] \\ &= \mathbb{E}[(X - \mathbb{E}[X])^2 + (Y - \mathbb{E}[Y])^2 + 2(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[(X - \mathbb{E}[X])^2] + \mathbb{E}[(Y - \mathbb{E}[Y])^2] + 2\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 0\end{aligned}$$

wobei wir im letzten Schritt die Definition von Varianz sowie Gleichung (2) verwenden.

(c) Wir verwenden, dass für unabhängige Zufallsvariablen X und Y auch sX und tY unabhängig sind. Damit folgt unmittelbar aus (a) und (b):

$$\mathrm{Var}(sX + tY) \overset{(b)}{=} \mathrm{Var}(sX) + \mathrm{Var}(tY) \overset{(a)}{=} s^2\mathrm{Var}(X) + t^2\mathrm{Var}(Y).$$

Aufgabe 2 – Chernoff in noch einfacher für starke Abweichung

Sei $X = X_1 + \cdots + X_n$ eine Summe von unabhängigen Bernoulli-Zufallsvariablen mit $\mu = \mathbb{E}[X]$ und sei $b \geq 6\mu$. Zeige

$$\Pr[X \geq b] \leq 2^{-b}.$$

Hinweis: Benutze die Chernoff Schranke $\Pr[X \geq (1 + \delta)\mu] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$.

Lösung 2

Um dem Hinweis zu folgen setzen wir $\delta = b/\mu - 1$. Es ergibt sich:

$$\begin{aligned}\Pr[X \geq b] = \Pr[X \geq (1 + \delta)\mu] &\leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\mu \\ &\leq \left(\frac{e^{b/\mu}}{(b/\mu)^{b/\mu}}\right)^\mu = \left(\frac{e}{b/\mu}\right)^b \leq \left(\frac{e}{6}\right)^b \leq 2^{-b}.\end{aligned}$$

Aufgabe 3 – Vergleich von Konzentrationsschranken

Für $n \in \mathbb{N}$ sei X_n die Anzahl der Sechsen bei n-maligem Würfeln eines gewöhnlichen Würfels. Sei p_n die Wahrscheinlichkeit, dass X_n seinen Erwartungswert um mindestens 10% übersteigt. Finde jeweils eine Schranke an p_n mit…

(a) … der Markov-Ungleichung.

(b) … der Tschebyscheffschen Ungleichung (engl. „Chebyshev's Inequality").

(c) … der Chernoff-Ungleichung (oder einer Variante davon).

(d) Vergleiche die asymptotische Stärke der Schranken.

Lösung 3

Vorbereitungen:

- Es gilt $\mu := \mathbb{E}[X] = n/6$.
- Es gilt $\mathrm{Var}(X) = n \cdot (\frac{1}{6} \cdot (\frac{5}{6})^2 + \frac{5}{6} \cdot (\frac{1}{6})^2) = \frac{5}{36}n$.
- Gesucht ist eine Schranke an $p_n = \Pr[X \geq 1.1\mu] \leq \Pr[|X - \mu| \geq 0.1\mu]$.

(a) $p_n = \Pr[X \geq 1.1\mu] \leq \mu/(1.1\mu) = \frac{10}{11} = \Theta(1)$.

(b) $p_n \leq \Pr[|X - \mu| \geq 0.1\mu] \leq \frac{\mathrm{Var}(X)}{(0.1\mu)^2} = \frac{5n/36}{0.01 \cdot n^2/36} = \frac{500}{n} = \Theta(1/n)$.

(c) $p_n = \Pr[X \geq (1 + 0.1)\mu] \leq \exp(-\frac{0.1^2}{2+0.1}\mu) = \exp(-\frac{1}{210}\mu) = \exp(-\frac{1}{1260}n) = \exp(-\Theta(n))$.

(d) Asympototisch ist die Chernoff-Schranke die stärkste und die Markov-Schranke die schwächste.
Bermerkung: Während Markov und Chernoff für alle $n \in \mathbb{N}$ eine Schranke mit Wert < 1 liefert, greift Tschebyscheff erst ab $n = 501$.