Algorithmische Mathematik II

Dozent
Professor Dr. Patrik Ferrari

Mitschrift
Maximilian Kessler

Version
31. Mai 2021 10:16

Zusammenfassung

Bei folgenden Vorlesungsnotizen handelt es sich um (inoffizielle) Mitschriften zur Vorlesung 'Algorithmische Mathematik II', die im Sommersemester 2021 an der Universität Bonn gehalten wird. Ich garantiere weder für Korrektheit noch Vollständigkeit dieser Notizen, und bin dankbar für jegliche Art von Korrektur, sowohl inhaltlich, als auch Tippfehler.

Bemerkungen, die nicht zum eigentlichen Vorlesungsinhalt gehören, wurden mit einem * gekennzeichnet. Sie werden nach eigenem Ermessen hinzugefügt, um weitere Details oder evtl. mündliche Anmerkungen beizufügen.

Manche Umgebungen sind mit einem † versehen. Das ist dann der Fall, wenn ihr Inhalt so, oder zumindest in sehr ähnlicher Form, in der Vorlesung vorkam (unter Umständen auch mündlich), ich aber die Umgebung der Aussage geändert habe. Insbesondere also, wenn ich aus Aussagen, die einfach erwähnt werden, ein **Lemma**† mache, um sie hervorzuheben.

Weitere Informationen finden sich bei GitHub oder auf der Vorlesungshomepage

Inhaltsverzeichnis

Übersicht der Vorlesungen

Vorlesung 1
Mo 17 Mai 2021 10:15

0.0.1 Abschätzung von Abweichungen

Seien $X_1, X_2, \ldots, X_n \sim \operatorname{Ber}(p)$ unabhängige Zufallsvariablen. wobei $p \neq 0, 1$. Dann ist $\mathcal{S}_n \coloneqq X_1 + \ldots + X_n \sim \operatorname{Bin}(n, p)$. Bekannt ist für uns bereits

$$\mathbb{E}\left(\frac{\mathcal{S}_n}{n}\right) = p, \operatorname{Var}\left(\frac{\mathcal{S}_n}{n}\right) = \frac{1}{n^2}\operatorname{Var}(\mathcal{S}_n) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}.$$

Binomialverteilung für große n

Frage. Was ist $\mathbb{P}\left(\left|\frac{\mathcal{S}_n}{n} - p\right| \geqslant \varepsilon\right)$, bzw ist das $\leqslant$? für $n \gg 1$.

Satz 0.1 (Ungleichung von Tchebishev)**.** Sei X eine Zufallsvariable auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann gilt $\forall \varepsilon > 0$:

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\operatorname{Var}(X)}{\varepsilon^2}.$$

Beweis. Wir stellen fest, dass beide Seiten unabhängig vom Milltwert sind, also können wir $\mathbb{E}(X) = 0$ voraussetzen, indem wir $Y \coloneqq X - \mathbb{E}(X)$ als Zufallsvariable betrachten. Wir überlegen uns nun:

$$\begin{aligned}
\mathbb{P}(|X| \geqslant \varepsilon) &= \mathbb{E}\left(\mathbb{1}_{|X| \geqslant \varepsilon}\right) \\
&\underset{(1)}{\leqslant} \mathbb{E}\left(\mathbb{1}_{|X| \geqslant \varepsilon} \frac{X^2}{\varepsilon^2}\right) \\
&\leqslant \frac{\mathbb{E}(X^2)}{\varepsilon^2} \\
&= \frac{\operatorname{Var}(X)}{\varepsilon^2}
\end{aligned}$$

Man beachte, dass wir hier bei (1) einfach benutzen, dass $1 \leqslant \frac{X^2}{\varepsilon^2}$, denn $|X| \geqslant \varepsilon$. □

Wir lernen nun eine Verallgemeinerung kennen:

Lemma 0.2. Sei X enie Zufallsvariable, $f : \mathbb{R} \to \mathbb{R}_+$ eine monoton wachsende Zufallsvariable. Dann ist

$$\mathbb{P}(X \geqslant a) \leqslant \frac{\mathbb{E}(f(X))}{f(a)} \qquad \forall a \in \mathbb{R}.$$

Beweis. Völlig analog:

$$\begin{aligned}\mathbb{P}(X \geqslant a) &= \mathbb{E}(\mathbb{1}_{X \geqslant a}) \\ &\leqslant \mathbb{E}\left(\mathbb{1}_{X \geqslant a} \frac{f(X)}{f(a)}\right) \\ &\leqslant \frac{1}{f(a)} \mathbb{E}(f(X))\end{aligned}$$

□

Beispiel. Wir erhalten zum Beispiel die **Markov-Ungleichung**:

$$\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \frac{\mathbb{E}(|X|)}{\varepsilon}.$$

Korollar 0.3. Sei X eine Zufallsvariable auf $(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist

$$\mathbb{P}(X \geqslant a) \leqslant \inf_{\lambda \geqslant 0} e^{-\lambda a} \mathbb{E}(e^{\lambda x}).$$

Beweis. Verwende das Lemma mit $f(x) = e^{\lambda x}$ für alle $\lambda \geqslant 0$, dann sind wir schon fertig. □

Satz 0.4. Betrachte wieder das Setting zu Beginn des Kapitels. Für alle $\varepsilon > 0$ ist

$$\mathbb{P}\left(\frac{\mathcal{S}_n}{n} \geqslant p + \varepsilon\right) \leqslant e^{-c\varepsilon^2 n}.$$

und

$$\mathbb{P}\left(\frac{\mathcal{S}_n}{n} \leqslant p - \varepsilon\right) \leqslant e^{-c\varepsilon^2 n}.$$

für ein $0 < c = c(p)$, das nur von p abhängt. Also ist

$$\mathbb{P}\left(\left|\frac{\mathcal{S}_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant 2 \cdot e^{-c\varepsilon^2 n}.$$

Bemerkung. Man kann $c = 2$ zeigen.

Die Abschätzung konvergiert gegen 0 exponentiell schnell in n für $n \to \infty$. Wäs würden wir mit Tchebishev erhalten?

$$\mathbb{P}\left(\left|\frac{\mathcal{S}_n}{n} - p\right| \geqslant \varepsilon\right) \leqslant \frac{1}{\varepsilon^2} \operatorname{Var}\left(\frac{\mathcal{S}_n}{n}\right) = \frac{p(1-p)}{n \cdot \varepsilon^2}.$$

Das geht auch gegen 0, aber eben nicht exponentiell (sondern nur quadratisch / polynomiell), also nicht wirklich schnell.

Bemerkung. Es gilt sogar

$$\mathbb{P}\left(\left|\frac{\mathcal{S}_n}{n} - p\right| \geqslant \frac{X}{\sqrt{n}}\right) \to 2 \cdot \frac{1}{2\pi\gamma^2} \int_x^\infty e^{-y^2/(2\sigma^2)} \, \mathrm{dy}\,.$$

aber das soll nicht Teil dieser Vorlesung sein.

Beweis. Stellen wir zunächst fest, dass wir mit $X := \frac{\mathcal{S}_n}{n} - p$

$$\begin{aligned}\mathbb{P}\left(\frac{\mathcal{S}_n}{n} \geqslant p+\varepsilon\right) &= \mathbb{P}(X \geqslant \varepsilon) \\ &\leqslant \inf_{\lambda>0} e^{-\lambda\varepsilon}\mathbb{E}\left(e^{\lambda\left(\frac{\mathcal{S}_n}{n}-p\right)}\right) \\ &\overset{\lambda:=\mu n}{=} \inf_{\mu>0} e^{-\mu n(p+\varepsilon)} \underbrace{\mathbb{E}(e^{\mu\mathcal{S}_n})}_{=\psi(e^\mu)}\end{aligned}$$

Setzen wir nun $\psi(z) = \sum_{k=0}^n \binom{n}{k} p^k(1-p)^{n-k}z^k = (p\cdot z + 1 - p)^n$, so erhalten wir weiter:

$$\mathbb{P}\left(\frac{\mathcal{S}_n}{n} \geqslant p+\varepsilon\right) = \inf_{\mu>0} e^{-n[\underbrace{\mu(p+\varepsilon)-\ln[pe^\mu+1-p]}_{:=I(\mu,p)}]}$$

Wir untersuchen nun $I(\mu,p)$ weiter:

- $I(0,p) = 0$
- $\frac{d}{d\mu}I(\mu,p) = p+\varepsilon - \frac{pe^\mu}{1-p+pe^\mu}$ und diese ist 0 genau dann, wenn $e^\mu = \frac{(1-p)(p+\varepsilon)}{p(1-p-\varepsilon)} =: e^{\mu^*}$
- Es ist

  $$\begin{aligned}I(\mu^*,p) &= (p+\varepsilon)\ln\left(\frac{p+\varepsilon}{p}\right) + (1-p-\varepsilon)\ln\left(\frac{1-p-\varepsilon}{}\right) \\ &\overset{\text{Taylor}}{=} \frac{\varepsilon^2}{2p(1-p)} + O(\varepsilon^3)\end{aligned}$$

  Also erhalten wir für kleine ε und $p \in (0,1)$, dass

  $$I(\mu^*,p) \geqslant 2\varepsilon^2 + O(\varepsilon^3) \geqslant \varepsilon^2.$$

  und damit

  $$\mathbb{P}\left(\frac{\mathcal{S}_n}{n} \geqslant p+\varepsilon\right) \leqslant e^{-n\varepsilon^2}.$$

  Setze nun $X = -\frac{\mathcal{S}_n}{n} + p$, dann erhalten wir direkt die zweite Abschätzung, die ir zeigen wollten.

$\square$

Details des Beweises aufschreiben.

Bemerkung. Wir hätten das ganze auch so schreiben können:

$$\begin{aligned}\mathbb{E}(e^{\mu\mathcal{S}_n}) &= \mathbb{E}(e^{\mu X_1}\cdot\ldots\cdot e^{\mu X_n}) \\ &\overset{zz.}{=} \prod_{k=1}^n \mathbb{E}(e^{\mu X_k})\end{aligned}$$

Lemma 0.5. Seien $X_1 : \Omega \to \mathcal{S}_1$, $X_2 : \Omega \to \mathcal{S}_2$ zwei Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$. Falls X_1 und X_2 unabhängig sind, so sind für alle $f_1 : \mathcal{S} \to \mathbb{R}$ und $f_2 : \mathcal{S}2 \to \mathbb{R}$ auch $f_1(X_1)$ und $f_2(X_2)$ zwei unabhängige Zufallsvariablen.

Beweis.

$$\begin{aligned}\mathbb{P}(f_1(X_1) = a_1, f_2(X_2) = a_2) &= \mathbb{P}(X_1 \in f_1^{-1}(a_1), X_2 \in f_2^{-1}(a_2)) \\ &\overset{X_1,X_2 \text{ unabhängig}}{=} \mathbb{P}(X_1 \in f_1^{-1}(a_1)) \cdot \mathbb{P}(X_2 \in f_2^{-1}(a_2)) \\ &= \mathbb{P}(f_1(X_1) = a_1) \cdot \mathbb{P}(f_2(X_2) = a_2)\end{aligned}$$

□

0.0.2 Das schwache Gesetz der großen Zahlen

Seien X_1,X_2,... Ergebnisse von Experimenten mit $X_k \sim \mu$ für jedes k. Setze

$$M_n(\omega) := \frac{1}{n}\sum_{k=1}^{n} X_k(\omega).$$

als den empirischen Mittelwert.

> **Frage.** Unter welchen Bedingungen gilt für große n, dass $M_n - \mathbb{E}(M_n)$ klein ist?

Extremfälle:

(a) Sei $X_1 = X_2 = X_3 = \ldots$ für alle ω. Dann ist $M_n - \mathbb{E}(M_n) = X_1 - \mathbb{E}(X_1)$, und das geht nicht gegen 0.

(b) Falls die Zufallsvariablen paarweise unabhängig (aber gleich verteilt) sind, so erhalten wir

$$\mathbb{P}\left(|M_n - \mathbb{E}(M_n)| \geqslant \varepsilon\right) \leqslant \frac{\operatorname{Var}(X_1)}{n\varepsilon^2} \to 0.$$

Ein allgemeineres Resultat bietet uns:

> **Satz 0.6** (Schwaches Gesetz der großen Zahlen)**.** Seien $X_1, X_2, \ldots$ Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}(X_k^2) < \infty$ für alle k, sodass
> (a) $\operatorname{Cov}(X_k, X_l) = 0$ für $k \neq l$
> (b) $\nu := \sup_{k \geqslant 1} \operatorname{Var}(X_k) < \infty$
> Sei $M_n = X_1 + \ldots + X_n$, so ist für jedes $\varepsilon > 0$:
>
> $$\mathbb{P}\left(\left|\frac{\mathcal{M}_n}{n} - \frac{\mathbb{E}(\mathcal{M}_n)}{n}\right| \geqslant \varepsilon\right) \leqslant \frac{\nu}{\varepsilon^2 n} \to 0.$$

> **Notation†.** Wir schreiben auch $X_k \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ für $\mathbb{E}(X_k)^2 < \infty$.

Beweis. Mit Tchebishev erhalten wir, dass

$$\begin{aligned}\text{LHS} &\leqslant \frac{\operatorname{Var}(M_n)}{\varepsilon^2 n^2} \\ &\overset{lemma20}{=} \frac{\operatorname{Var}(X_1) + \ldots + \operatorname{Var}(X_n)}{\varepsilon^2 n^2} \\ &\leqslant \frac{\nu}{e^2 n}\end{aligned}$$

□

Namen und label für Sätze hinzufügen

C Stichwortverzeichnis