



Einführung in die Mathematische Statistik

6. Tutorium

Im Zentrum dieses Tutoriums steht die folgende Variante der Hoeffdingschen Ungleichung für Bernoulli-verteilte Zufallsvariablen.

Satz 1. (Hoeffdingsche Ungleichung) Sei $p \in]0, 1[$ und $(X_i)_{i \in \mathbb{N}}$ eine unabhängige Folge von identisch verteilten Zufallsvariablen mit $X_1 \sim \mathbf{B}(1, p)$. Weiterhin sei $S_n = \sum_{i=1}^n X_i$ für $n \in \mathbb{N}$. Dann gilt für alle $\varepsilon > 0$ und alle $n \in \mathbb{N}$

$$P(\{1/n \cdot S_n - p \geq \varepsilon\}) \leq \exp(-2\varepsilon^2 \cdot n).$$

Zum Beweis benötigen wir die folgende Variante der Tschebyscheff-Ungleichung.

Proposition 1. Sei $X \in \mathcal{L}_1$ und $\varepsilon > 0$. Dann gilt

$$P(\{|X| \geq \varepsilon\}) \leq \frac{1}{\varepsilon} \cdot \mathbf{E}|X|.$$

Weiterhin ist folgendes Lemma nützlich.

Lemma 1. Sei $r > 0$, $n \in \mathbb{N}$ und S_n wie in Satz 1. Dann gilt

$$\mathbf{E} \exp(rS_n) = (1 - p + p \exp(r))^n \leq \exp\left(rnp + \frac{nr^2}{8}\right).$$

Aufgabe 1. Beweisen Sie Proposition 1 und Lemma 1. Verwenden Sie für den Beweis der Abschätzung in Lemma 1 eine Taylorentwicklung der Funktion $L(r) = n \ln(1 - p + p \exp(r))$, $r > 0$ bis zur Ordnung 2.

Aufgabe 2. Beweisen Sie Satz 1. Leiten Sie dazu mit Hilfe von Proposition 1 und Lemma 1 zuerst eine von $r > 0$ abhängige Abschätzung für $P(\{1/n \cdot S_n - p \geq \varepsilon\})$ her und minimieren Sie dann über $r > 0$.