Die Markov'sche Ungleichung

Satz: Für eine ZV $X \geq 0$ und für alle $t > 0$ gilt die Ungleichung

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}$$

oder äquivalent dazu

$$\Pr[X \geq t \cdot \mathbb{E}[X]] \leq \frac{1}{t} \ .$$

Die Wahrscheinlichkeit für eine nicht-negativen ZV, ihren Erwartungswert um den Faktor t (oder stärker) zu überschreiten, ist also durch $1/t$ beschränkt.

Beweis: Wie wir wissen:

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}[X|X \geq t] \cdot \Pr[X \geq t] + \mathbb{E}[X|X < t] \Pr[X < t] \\ &\geq t \cdot \Pr[X \geq t] + 0 \end{aligned}$$

Auflösen der Ungleichung nach $\Pr[X \geq t]$ liefert die Behauptung.

Die Chebychev'sche Ungleichung

Satz: Für eine ZV X und für alle $t > 0$ gilt die Ungleichung

$$\Pr[|X - \mathbb{E}[X]| \geq t] \leq \frac{\mathrm{Var}[X]}{t^2}$$

oder äquivalent dazu

$$\Pr[|X - \mathbb{E}[X]| \geq t \cdot \sqrt{\mathrm{Var}[X]}] \leq \frac{1}{t^2} \ .$$

Die Wahrscheinlichkeit, dass die absolute Abweichung einer ZV von ihrem Erwartungswert das t-fache ihrer Standardabweichung beträgt (oder mehr), ist also durch $1/t^2$ beschränkt.

Beweis: Wir wenden die Markov'sche Ungleichung auf die ZV $(X - \mathbb{E}[X])^2$ an:

$$\begin{aligned} \Pr[|X - \mathbb{E}[X]| \geq t] &= \Pr[(X - \mathbb{E}[X])^2 \geq t^2] \\ &\leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{t^2} = \frac{\mathrm{Var}[X]}{t^2} \end{aligned}$$

Chernov-Schranken (additive Form)

Wir setzen folgendes voraus:

- $X_1, \ldots, X_n$ sind unabhängige ZV, die jeweils Bernoulli-verteilt mit Erfolgswahrscheinlichkeit p sind (n unabhängige „Bernoulli-Experimente").
- $Z := \frac{1}{n}(X_1 + \cdots + X_n)$ ist der resultierende „empirische Mittelwert".

Dann gilt für alle $0 \le \delta \le 1$:

$$\begin{aligned} \Pr[Z \ge p + \delta] &\le e^{-2\delta^2 n} \\ \Pr[Z \le p - \delta] &\le e^{-2\delta^2 n} \\ \Pr[|Z - p| \ge \delta] &\le 2e^{-2\delta^2 n} \end{aligned}$$

Da $e^{-2\delta^2 n}$ mit wachsendem n in mit exponentieller Geschwindigkeit gegen Null konvergiert, sind empirische Schätzungen für die Erfolgswahrscheinlichkeit p schon bei moderaten Stichprobengrößen n mit „überwältigender W.-keit" sehr genau.

Chernov-Schranken (multiplikative Form)

Unter denselben Voraussetzungen wie eben gilt:

$$\begin{aligned} \Pr[Z \geq (1+\delta)p] &\leq e^{-np\delta^2/3} \\ \Pr[Z \leq (1-\delta)p] &\leq e^{-np\delta^2/2} \end{aligned}$$

Hoeffding-Schranken

Die Hoeffding-Schranken sind Verallgemeinerungen der Chernov-Schranken auf beschränkte ZV (die nicht notwendig Bernoulli-verteilt sind). Wir beschränken uns hier auf die additive Form.

Wir setzen folgendes voraus:

- $X_1, \ldots, X_n$ sind unabhängige und gleichverteilte ZV mit Werten im Intervall $[a, b]$ und Erwartungswert μ.
- $Z := \frac{1}{n}(X_1 + \cdots + X_n)$ ist der resultierende „empirische Mittelwert“ (bei n unabhängigen Versuchen).

Dann gilt für alle $\delta \geq 0$:

$$
\begin{aligned}
\Pr[Z \geq \mu + \delta] &\leq e^{\frac{-2\delta^2 n}{(b-a)^2}} \\
\Pr[Z \leq \mu - \delta] &\leq e^{\frac{-2\delta^2 n}{(b-a)^2}} \\
\Pr[|Z - \mu| \geq \delta] &\leq 2e^{\frac{-2\delta^2 n}{(b-a)^2}}
\end{aligned}
$$

Hoeffding-Schranken (etwas allgemeiner)

$X_1, \ldots, X_n$ seien unabhängige Zufallsvariable mit $a_i \le X_i \le b_i$ und Erwartungswert μ_i. Dann gilt für alle $\delta \ge 0$:

$$\begin{aligned}
\Pr\left[\sum_{i=1}^{n}(X_i - \mu_i) \ge \delta\right] &\le e^{\frac{-2\delta^2}{\sum_{i=1}^{n}(b_i-a_i)^2}} \\
\Pr\left[\sum_{i=1}^{n}(X_i - \mu_i) \le -\delta\right] &\le e^{\frac{-2\delta^2}{\sum_{i=1}^{n}(b_i-a_i)^2}} \\
\Pr\left[\left|\sum_{i=1}^{n}(X_i - \mu_i)\right| \ge \delta\right] &\le 2e^{\frac{-2\delta^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}
\end{aligned}$$

Hoeffding-Schranken (noch allgemeiner)

Die bisherigen Hoeffding-Schranken sind alle aus der folgenden herleitbar.

Sei X eine Zufallsvariable mit $a \leq X \leq b$. Dann gilt für alle $s \in \mathbb{R}$:

$$s\mathbb{E}[X] \leq \ln \mathbb{E}[e^{sX}] \leq s\mathbb{E}[X] + \frac{s^2(b-a)^2}{8}$$

Zum Beweis s. (zum Bsp.) Anhang A.1 des Buches „Prediction, Learning, and Games“ von Cesa-Bianchi und Lugosi.

Anwendung

Sei $v \in \mathbb{R}^n$ und X eine in $\{-1,1\}^n$ uniform verteilte ZV. Dann gilt für alle $\delta \geq 0$:

$$
\begin{aligned}
\Pr[v^\top X \geq \delta] &\leq \exp\left(\frac{-\delta^2}{2\|v\|_2^2}\right) \\
\Pr[v^\top X \leq -\delta] &\leq \exp\left(\frac{-\delta^2}{2\|v\|_2^2}\right) \\
\Pr[|v^\top X| \geq \delta] &\leq 2\exp\left(\frac{-\delta^2}{2\|v\|_2^2}\right)
\end{aligned}
$$

Weitere exponentielle Ungleichung

Sei S^{n-1} die Einheitssphäre in $\mathbb{R}^n$, $v \in S^{n-1}$ und X eine in S^{n-1} uniform verteilte ZV. Dann gilt:

$$
\begin{aligned}
\Pr[v^\top X \geq \delta] &\leq \exp\left(-\frac{1}{2}n\delta^2\right) \\
\Pr[v^\top X \leq -\delta] &\leq \exp\left(-\frac{1}{2}n\delta^2\right) \\
\Pr[|v^\top X| \leq \delta] &\leq 2\exp\left(-\frac{1}{2}n\delta^2\right)
\end{aligned}
$$

Ein Beweis hierfür findet sich in dem Buch „Lectures on Discrete Geometry“ von J. Matousek.