Ungleichung von Markov

Sei X eine Zufallsvariable, die nur nicht-negative Werte annimmt. Dann gilt für alle $t \in \mathbb{R}$ mit $t > 0$, dass

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

Beweis: Wir rechnen direkt nach, dass

$$\begin{aligned} \mathbb{E}[X] &= \sum_{x \in W_X} x \cdot \Pr[X = x] \; \geq \sum_{x \in W_X,\; x \geq t} x \cdot \Pr[X = x] \\ &\geq t \cdot \sum_{x \in W_X,\; x \geq t} \Pr[X = x] \; = \; t \cdot \Pr[X \geq t]. \end{aligned}$$

Man erkennt, dass die Markov-Ungleichung im Wesentlichen durch Weglassen einiger Summanden aus der Definition des Erwartungswerts entsteht.

Ungleichung von Chebyshev

Sei X eine Zufallsvariable und $t \in \mathbb{R}$ mit $t > 0$. Dann gilt

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

Beweis:

Die Zufallsvariable $Y := (X - \mathbb{E}[X])^2$ hat nach Definition der Varianz den Erwartungswert $\mathbb{E}[Y] = \mathrm{Var}[X]$. Damit folgt die Behauptung durch Anwendung der Markov-Ungleichung:

$$\Pr[|X - \mathbb{E}[X]| \geq t] \;=\; \Pr[(X - \mathbb{E}[X])^2 \geq t^2] \;=\; \Pr[Y \geq t^2] \;\leq\; \frac{\mathbb{E}[Y]}{t^2} \;=\; \frac{\mathrm{Var}[X]}{t^2}.$$

Gesetz der großen Zahlen

Gegeben sei eine Zufallsvariable X und $\varepsilon, \delta > 0$. Dann gilt für alle $n \geq \mathrm{Var}[X]/(\varepsilon\delta^2)$:

Sind $X_1, \ldots, X_n$ unabhängige Zufallsvariablen mit derselben Verteilung wie X und setzt man

$$Z := \frac{X_1 + \ldots + X_n}{n},$$

so gilt

$$\Pr[|Z - \mathbb{E}[X]| \geq \delta] \;\leq\; \varepsilon.$$

Anschaulich: Der beobachtete Mittelwert nähert sich dem Erwartungswert der zugrunde liegenden Zufallsvariablen immer besser an.

Beweis: Für Z gilt

$$\mathbb{E}[Z] = \frac{1}{n} \cdot (\mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n]) = \frac{1}{n} \cdot n \cdot \mathbb{E}[X] = \mathbb{E}[X],$$

sowie

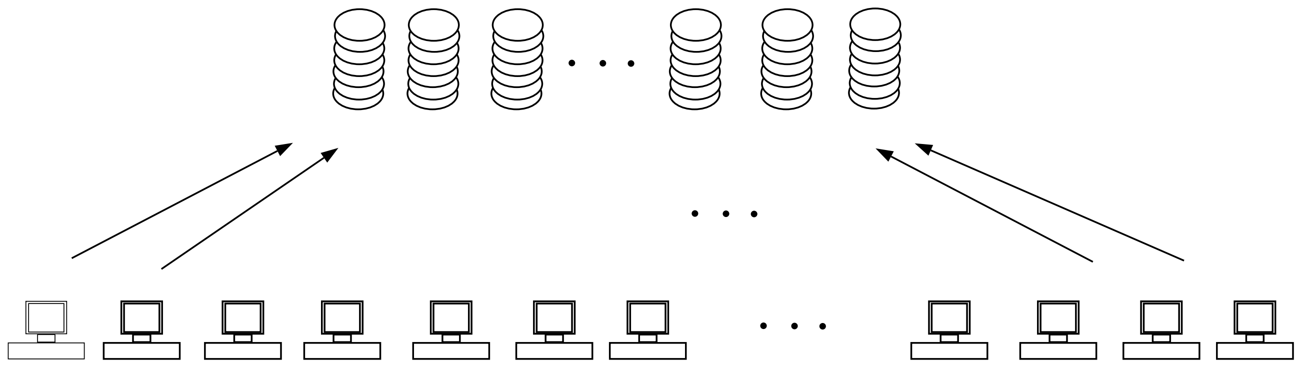
$$\mathrm{Var}[Z] = \frac{1}{n^2} \cdot (\mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]) = \frac{1}{n^2} \cdot n \cdot \mathrm{Var}[X] = \frac{\mathrm{Var}[X]}{n}.$$

Mit der Chebyshev-Ungleichung erhalten wir

$$\Pr[|Z - \mathbb{E}[X]| \geq \delta] = \Pr[|Z - \mathbb{E}[Z]| \geq \delta] \leq \frac{\mathrm{Var}[Z]}{\delta^2} = \frac{\mathrm{Var}[X]}{n\delta^2} \leq \varepsilon.$$

BSP.: Fallbeispiel 1: Zuordnung von Datenbankanfragen

Szenario:



Aufgabe: Entwerfe ein möglichst einfaches Verfahren, das jede Anfrage genau einem Server zuordnet — und zwar so, daß alle Server möglichst gleich viel Arbeit erhalten.

Idealisiertes Modell: Wir werfen n Bälle *zufällig* in m Urnen.

$X_i :=$ Anzahl Bälle in $i$ter Urne

Frage: Wie groß ist $\max\limits_{1 \leq i \leq n} X_i$?

[Wir betrachten im folgenden nur den Fall $n = m$.]

$X_i :=$ Anzahl Bälle in $i$ter Urne

Frage: Wie groß ist $\max\limits_{1 \le i \le n} X_i$?

Antwort: Mit hoher Wahrscheinlichkeit etwa $\dfrac{\ln n}{\ln \ln n}$.

Beweisidee: Wir zeigen

$$\Pr[\max_{1 \le i \le n} X_i \ge \alpha \frac{\ln n}{\ln \ln n}] = \begin{cases} o(1) & , \text{ falls } \alpha > 1 \\ 1 - o(1) & , \text{ falls } \alpha < 1. \end{cases}$$

Setze

$$Y_i := \begin{cases} 1 & , \text{ falls } X_i \ge \alpha \frac{\ln n}{\ln \ln n} \\ 0 & , \text{ sonst} \end{cases}$$

Dann gilt für $Y := \sum_{i=1}^n Y_i$:

$$Y = 0 \qquad \Longleftrightarrow \qquad \max_{1 \le i \le n} X_i < \alpha \frac{\ln n}{\ln \ln n}$$

Markov-Ungleichung: $\Pr[Y \ge 1] \le \mathbb{E}[Y]$

Chebychev-Ungleichung: $\Pr[Y = 0] \le \dfrac{\mathrm{Var}[Y]}{(\mathbb{E}[Y])^2}$

Etwas Rechnung (weggelassen):

$$\begin{aligned}\mathbb{E}[Y] &\approx n^{1-\alpha}\\ \mathrm{Var}[Y] &\approx \mathbb{E}[Y]\end{aligned}$$

Also

$$\begin{aligned}\Pr[\max_{1\le i\le n} X_i \ge \alpha\frac{\ln n}{\ln\ln n}] &= \Pr[Y \ge 1] \le \mathbb{E}[Y]\\ &= o(1), \qquad \text{falls } \alpha > 1\end{aligned}$$

$$\begin{aligned}\Pr[\max_{1\le i\le n} X_i < \alpha\frac{\ln n}{\ln\ln n}] &= \Pr[Y = 0]\\ &\le \frac{\mathrm{Var}[Y]}{(\mathbb{E}[Y])^2}\\ &\approx \frac{1}{\mathbb{E}[Y]}\\ &= o(1), \qquad \text{falls } \alpha < 1\end{aligned}$$