1 Analyse der Abweichung vom Erwartungsswert

1.1 Die Markov'sche Ungleichung

Satz 1.1 *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. **qed.**

1.2 Die Chebyshev'sche Ungleichung

Satz 1.2 *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:

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

qed.

1.3 Die Chernoff-Schranken

Es seien $X_1, \ldots, X_n$ unabhängige ZV, die jeweils Bernoulli-verteilt mit Erfolgswahrscheinlichkeiten $p_1, \ldots, p_n$ sind (n unabhängige „Bernoulli-Experimente“) und $p := \frac{1}{n}(p_1 + \cdots + p_n)$. Weiter sei $Z := \frac{1}{n}(X_1 + \cdots + X_n)$ der resultierende „empirische Mittelwert“. Dann gelten für alle $0 \leq \delta \leq 1$ die folgenden Chernoff-Schranken der „additiven Form“:

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

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

Unter denselben Voraussetzungen wie eben gelte die folgenden Chernoff-Schranken der „multiplikativen Form“:

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

$$\Pr[Z \leq (1-\delta)p] \leq e^{-np\delta^2/2} \tag{2}$$

Die Ungleichungen (1) und (2) gelten nur unter Voraussetzung $0 \leq \delta \leq 1$. Für beliebige $\delta > 0$ gilt das folgende Pendant zu (1):

$$\Pr[Z \geq (1+\delta)p] \leq \frac{e^{np\delta}}{(1+\delta)^{(1+\delta)np}} = \left(e^{-((1+\delta)\cdot\ln(1+\delta)-\delta}\right)^{np} \tag{3}$$

Folgerung 1.3 *Für alle* $0 \leq \delta \leq B$ *und alle*

$$\lambda \leq \lambda(B) := \frac{(1+B)\ln(1+B) - B}{B^2} \tag{4}$$

gilt

$$\Pr[Z \geq (1+\delta)p] \leq e^{-\lambda np\delta^2} \quad . \tag{5}$$

Beweis Die rechte Seite von (3) legt nahe, den Ausdruck $(1+\delta)\cdot\ln(1+\delta) - \delta$ nach unten abzuschätzen. Wir wollen diesen Ausdruck zu δ^2 in Beziehung setzen und betrachten daher die Hilfsfunktion $\lambda(\delta)$ gemäß (4). Unter Anwendung der Regel von l'Hospital ergibt sich, dass λ in $\delta = 0$ stetig ergänzbar ist mit $\lambda(0) = 1/2$. Mit Kurvendiskussion stellt man fest, dass λ für alle $\delta \geq 0$ streng monoton fallend ist. Damit ergibt sich für alle $0 \leq \delta \leq B$:

$$\frac{(1+\delta)\cdot\ln(1+\delta) - \delta}{\delta^2} = \lambda(\delta) \geq \lambda(B)$$

und somit

$$(1+\delta)\cdot\ln(1+\delta) - \delta \geq \lambda(B)\cdot\delta^2 \quad .$$

Zusammen mit (3) ergibt sich die Behauptung. **qed.**

Es gilt zum Beispiel

$$\lambda(1) = 2\ln(2) - 1 \geq 1.368 > 1/3 \tag{6}$$

Hieran sieht man, dass (1) sich über Folgerung 1.3 aus (3) ableiten lässt. Weiterhin gilt

$$\lambda(4) = \frac{5\ln(5) - 4}{16} > \frac{1}{4} \ . \tag{7}$$

Folgerung 1.4 *Für alle $0 \leq \delta \leq 4$ gilt*

$$\Pr[Z \geq (1+\delta)p] \leq e^{-np\delta^2/4} \ . \tag{8}$$

Folgerung 1.5 *Für alle $\delta \geq B$ und alle*

$$\kappa \leq \kappa(B) := \ln(B+1) - 1 \tag{9}$$

gilt

$$\Pr[Z \geq (1+\delta)p] \leq e^{-\kappa np\delta} \tag{10}$$

Beweis Der Beweis ergibt sich mit Folgerung 1.3 unter Verwendung von $1 + \delta \geq 1 + B = \exp(\ln(1 + B))$ direkt aus folgender Rechnung:

$$\Pr[Z \geq (1+\delta)p] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{np} \leq (\exp(-((1+\delta)\ln(B+1) - \delta))^{np} < e^{-\kappa(B)np\delta}$$

qed.

Mit $\kappa(4) > 1/2$ erhalten wir direkt die

Folgerung 1.6 *Für alle $\delta \geq 4$ gilt:*

$$\Pr[Z \geq (1+\delta)p] \leq e^{-np\delta/2} \tag{11}$$

1.4 Die Hoeffding-Schranken

Die Hoeffding-Schranken sind Verallgemeinerungen der Chernoff-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 $0 \leq \delta \leq 1$:

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