Seminararbeit in Finanz- und Versicherungsmathematik

Probability Inequalities

Valentine Zemanek
e01202674@student.tuwien.ac.at

Technische Universität Wien
Institute für Finanz- und Versicherungsmathematik

Juli 2017

Inhaltsverzeichnis

1 Elementare Ungleichungen von Wahrscheinlichkeiten von Ereignissen **3**

1.1 Ungleichung von unabhängige Ereignisse . . . 3

1.2 Feller-Chung Ungleichung . . . 3

1.3 Chung-Erdös Ungleichung . . . 5

2 Ungleichungen in Zusammenhang mit der Normalverteilung **6**

2.1 Ungleichung 1 in Zusammenhang mit der Normalverteilung . . . 6

2.2 Ungleichung 2 in Zusammenhang mit der Normalverteilung . . . 7

3 Ungleichungen in Zusammenhang mit der charakteristischen Funktion **8**

3.1 Ungleichung bezogen nur auf die Charakteristische Funktion . . . 8

3.2 Inkrement Ungleichung . . . 9

3.3 Ungleichung zusammenhängend mit Charakteristische Funktion und Verteilungsfunktion . . . 10

4 Wahrscheinlichkeitsungleichungen von Zufallsvariablen **11**

4.1 Ungleichungen bezogen auf zwei Zufallsvariablen . . . 11

4.2 Symmetrisierende Ungleichungen - Schwache symmetrisierende Ungleichungen . . 12

4.3 Symmetrisierende Ungleichungen - Allgemeine symmetrisierende Ungleichungen . 12

4.4 Levy Ungleichung . . . 13

4.5 Levy Ungleichung Korollar . . . 13

5 Wahrscheinlichkeitsschranken bezogen auf die Momente **15**

5.1 Allgemeine Form - Chebyshev-Markov Art von Ungleichungen . . . 15

5.2 Chebyshev Ungleichung . . . 15

5.3 Markov Ungleichung . . . 15

5.4 Verallgemeinerung der Chebyshev Ungleichung . . . 15

5.5 Kolmogorov Ungleichung . . . 16

6 Hoeffding Ungleichung **19**

Vorwort

Dieses Dokument besteht aus ausgearbeiteten Beweisen von Wahrscheinlichkeitsungleichungen. Die zugrunde liegenden Beweise wurden aus Lin und Bai 2010 aus den Kapiteln 1 bis 7 ausgewählt. Weiterführende Literatur kann unter anderem in Petrov 1995 und Kusolitsch 2014 gefunden werden.

Mathematische Schreibweise:
$\mathbb{E}X$ bzw. äquivalent dazu $\mathbb{E}(X)$ bezeichnet den Erwartungswert von X.
Analog dazu ist $\mathbb{V}X$ bzw. äquivalent dazu $\mathbb{V}(X)$ als die Varianz von X, sowie mX als der Median von X definiert.

1 Elementare Ungleichungen von Wahrscheinlichkeiten von Ereignissen

1.1 Ungleichung von unabhängige Ereignisse

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und Ereignisse $A, B, D \in \mathcal{F}$ mit A, B sind unabhängig, $A \cap B \subseteq D$ und $A^c \cap B^c \subseteq D^c$ Dann gilt:

$$\mathbb{P}(A \cap D) \geq \mathbb{P}(A)\mathbb{P}(D) \tag{1}$$

Beweis. Mittels direktem Beweis und Umformungen gilt:

$$\begin{aligned}
\mathbb{P}(A \cap D) &= \mathbb{P}(A \cap D \cap B) + \mathbb{P}(A \cap D \cap B^c) \\
&\overset{1}{=} \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) - \mathbb{P}(A \cap D^c \cap B^c) \\
&\overset{2}{=} \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A \cap B^c) - \mathbb{P}(B^c \cap D^c) + \mathbb{P}(A^c \cap D^c \cap B^c) \\
&\overset{3}{=} \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A \cap B^c) - \mathbb{P}(B^c \cap D^c) + \mathbb{P}(A^c \cap B^c) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(B^c) - \mathbb{P}(B^c \cap D^c) \\
&\overset{4}{=} \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(B^c \cap D) \geq \mathbb{P}(A)\mathbb{P}(B \cap D) + \mathbb{P}(A)\mathbb{P}(B^c \cap D) \\
&\overset{5}{=} \mathbb{P}(A)(\mathbb{P}(B \cap D) + \mathbb{P}(B^c \cap D)) = \mathbb{P}(A)\mathbb{P}(D)
\end{aligned}$$ □

Mit den verwendeten Umformungen:

1. Wegen $A \cap B \subseteq D$ gilt $A \cap B \cap D = A \cap B$.
   Zusätzlich gilt die Gleichheit $\mathbb{P}(A\cap D\cap B^c)+\mathbb{P}(A\cap D^c\cap B^c) = \mathbb{P}(A\cap B^c) \Leftrightarrow \mathbb{P}(A\cap D\cap B^c) = \mathbb{P}(A \cap B^c) - \mathbb{P}(A \cap D^c \cap B^c)$
2. Die Gleichung $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ gilt da A und B unabhängig sind.
   Zusätzlich gilt die Gleichung $\mathbb{P}(A \cap D^c \cap B^c) + \mathbb{P}(A^c \cap D^c \cap B^c) = \mathbb{P}(D^c \cap B^c) \Leftrightarrow -\mathbb{P}(A \cap D^c \cap B^c) = -\mathbb{P}(D^c \cap B^c) + \mathbb{P}(A^c \cap D^c \cap B^c)$
3. Wegen $A^c \cap B^c \subseteq D^c \Rightarrow A^c \cap B^c \cap D^c = A^c \cap B^c$
4. $\mathbb{P}(B^c \cap D^c) + \mathbb{P}(B^c \cap D) = \mathbb{P}(B^c) \Leftrightarrow \mathbb{P}(B^c \cap D) = \mathbb{P}(B^c) - \mathbb{P}(B^c \cap D^c)$
5. $B \supseteq B \cap D \Rightarrow \mathbb{P}(B) \geq \mathbb{P}(B \cap D)$ und $1 \geq \mathbb{P}(A)$

1.2 Feller-Chung Ungleichung

Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und $A_0 = \emptyset, (A_n), (B_n) \in \mathcal{F}$,wobei $(A_n)_{n\in\mathbb{N}}, (B_n)_{n\in\mathbb{N}}$ Mengenfolgen in $\mathcal{F}$. Angenommen, es gilt einer der folgenden Aussagen:

(i). (B_n) ist unabhängig von $A_n \cap A_{n-1}^c \cap \cdots \cap A_0^c$ für alle $n \in \mathbb{N}$

(ii). (B_n) ist unabhängig von $A_n \cap A_{n+1}^c \cap \cdots \cap A_{n+k}^c$ für alle $n \in \mathbb{N}$ und $k \in \mathbb{N} \cup \{0\}$

Dann gilt:

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty} (A_n \cap B_n)\Big) \geq \inf_{n \in \mathbb{N}} \mathbb{P}(B_n) \cdot \mathbb{P}\Big(\bigcup_{n=1}^{\infty} A_n\Big) \tag{2}$$

Beweis. Allgemein gilt für Mengenfolgen C_n:

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} \Big(C_n \setminus (\bigcup_{k=1}^{n-1} C_k)\Big) = \bigcup_{n=1}^{\infty} \Big(C_n \cap \bigcap_{k=1}^{n-1} C_k^c\Big)$$

"$\supseteq$": Die rechte Seite ist offensichtlich eine Teilmenge der linken Seite.
"$\subseteq$":
Sei $x \in \bigcup_{n=1}^{\infty} C_n \implies \exists n \in \mathbb{N} : x \in C_n$, wähle $n_0 \in \mathbb{N}$ als die kleinste natürliche Zahl mit $x \in C_{n_0}$. Dann gilt:

$$x \in C_{n_0} \setminus \Big(\bigcup_{k=1}^{n_0-1} C_k\Big) = C_{n_0} \cap \bigcap_{k=1}^{n_0-1} C_k^c \subseteq \bigcup_{n=1}^{\infty} \Big(C_n \cap \bigcap_{k=1}^{n-1} C_k^c\Big)$$

Daraus folgt:

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} \Big(C_n \cap \bigcap_{k=1}^{n-1} C_k^c\Big) \tag{3}$$

Angenommen (i) gilt:

$$\begin{aligned}
\mathbb{P}\Big(\bigcup_{n=1}^{\infty} \underbrace{(A_n \cap B_n)}_{C_n}\Big) &\overset{(3)}{=} \mathbb{P}\Big(\bigcup_{n=1}^{\infty} \underbrace{(A_n \cap B_n)}_{C_n} \cap \bigcap_{k=1}^{n-1} \underbrace{(A_k \cap B_k)^c}_{C_k^c}\Big) \\
&= \sum_{n=1}^{\infty} \mathbb{P}\Big(B_n \cap A_n \cap \bigcap_{k=1}^{n-1} \underbrace{(B_k \cap A_k)^c}_{B_k^c \cup A_k^c \supseteq A_k^c}\Big) \\
&\geq \sum_{n=1}^{\infty} \mathbb{P}\Big(B_n \cap A_n \cap \bigcap_{k=1}^{n-1} A_k^c\Big) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(B_n) \cdot \mathbb{P}(A_n \cap \bigcap_{k=1}^{n-1} A_k^c) \\
&\geq \inf_{n \in \mathbb{N}} \mathbb{P}(B_n) \cdot \sum_{n=1}^{\infty} \mathbb{P}\Big(A_n \cap \bigcap_{k=1}^{n-1} A_k^c\Big) \\
&= \inf_{n \in \mathbb{N}} \mathbb{P}(B_n) \cdot \mathbb{P}\Big(\bigcup_{n=1}^{\infty} (A_n \cap \bigcap_{k=1}^{n-1} A_k^c)\Big) \\
&\overset{(3)}{=} \inf_{n \in \mathbb{N}} \mathbb{P}(B_n) \cdot \mathbb{P}\Big(\bigcup_{n=1}^{\infty} A_n\Big)
\end{aligned}$$

Wodurch gezeigt wurde, dass auf (i) Ungleichung (2) folgt.

Angenommen (ii) gilt: Sei $n \in \mathbb{N}$ beliebig

$$\begin{aligned}
\mathbb{P}\Big(\bigcup_{j=1}^{\infty}(A_j \cap B_j)\Big) \geq \mathbb{P}\Big(\bigcup_{j=1}^{n}(A_j \cap B_j)\Big) &=^{(3)} \sum_{j=1}^{n} \mathbb{P}(A_j \cap B_j \cap \bigcap_{i=j+1}^{n} (A_i \cap B_i)^c\Big) \\
&\geq \sum_{j=1}^{n} \mathbb{P}(A_j \cap B_j \cap \bigcap_{i=j+1}^{n} A_i^c) \\
&= \sum_{j=1}^{n} \mathbb{P}(B_j) \cdot \mathbb{P}\Big(A_j \cap \bigcap_{i=j+1}^{n} A_i^c\Big) \\
&\geq \inf_{1 \leq j \leq n} \mathbb{P}(B_j) \sum_{j=1}^{n} \mathbb{P}\Big(A_j \cap \bigcap_{i=j+1}^{n} A_i^c\Big) \\
&= \inf_{1 \leq j \leq n} \mathbb{P}(B_j) \mathbb{P}\Big(\bigcup_{j=1}^{n} (A_j \cap \bigcap_{i=j+1}^{n} A_i^c)\Big) \\
&= \inf_{1 \leq j \leq n} \mathbb{P}(B_j) \mathbb{P}\Big(\bigcup_{j=1}^{n} (A_j \cap \Big(\bigcup_{i=j+1}^{n} A_i\Big)^c)\Big) \\
&= \inf_{1 \leq j \leq n} \mathbb{P}(B_j) \mathbb{P}\Big(\bigcup_{j=1}^{n} A_j\Big)
\end{aligned}$$

Die Aussage in Ungleichung 2 folgt für $n \to \infty$, da $n \in \mathbb{N}$ beliebig gewählt war. □

1.3 Chung-Erdös Ungleichung

Die Chung–Erdös Ungleichung gibt eine untere Schranke an für die Wahrscheinlichkeit, dass eine von vielen Ereignissen (möglicherweise abhängig) eintritt. Die untere Schranke ist gegeben durch die Wahrscheinlichkeit von Ereignis-Paaren.
Seien $A_1, \ldots, A_n$ Ereignisse auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ und $\mathbb{P}(A_i) > 0$ für einige i, so gilt:

$$\mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big) \geq \frac{\Big(\sum_{i=1}^{n} \mathbb{P}(A_i)\Big)^2}{\Big(\sum_{i=1}^{n} \mathbb{P}(A_i) + 2 \cdot \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j)\Big)} \tag{4}$$

Beweis. Definiere Zufallsvariablen $X_k(\omega)$, $\omega \in \Omega$, durch

$$X_i(\omega) = \begin{cases} 0 & \omega \in A_i \\ 1 & \omega \notin A_i \end{cases}$$

Wegen $X_i = X_i^2$, gilt $\mathbb{E}[X_i] = \mathbb{E}[X_i^2] = \mathbb{P}(A_i)$ und $\mathbb{P}(X_1 + \cdots + X_n > 0) = \mathbb{P}\Big(\bigcup_{i=1}^{n} A_i\Big)$ laut Definition. Weiters ist

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{P}(A_i \cap A_j) = \sum_{i=1}^{n} \mathbb{P}(A_i \cap A_i) + \sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n} \mathbb{P}(A_i \cap A_j) = \sum_{i=1}^{n} \mathbb{P}(A_i) + 2 \cdot \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j)$$

und

$$\sum_{i=1}^{n} \mathbb{P}(A_i) = \mathbb{E}[X_1^2] + \cdots + \mathbb{E}[X_n^2] = \mathbb{E}[X_1^2 + \cdots + X_n^2]$$

Es gilt dann, dass:

$$2 \cdot \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j) = \mathbb{E}[(X_1 + \cdots + X_n)^2] - \mathbb{E}[X_1^2 + \cdots + X_n^2]$$

Aus der Cauchy-Schwarz Ungleichung folgt für Zufallsvariablen X und Y:

$$\mathbb{E}|XY| \leq (\mathbb{E}(X^2)\mathbb{E}(Y^2))^{\frac{1}{2}} \Leftrightarrow (\mathbb{E}|XY|)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$$

sowie

$$\begin{aligned}(\mathbb{E}[X_1 + \cdots + X_n]) &\leq \mathbb{E}[(X_1 + \cdots + X_n) \cdot \mathbb{1}_{(X_1+\cdots+X_n>0)}] \\ &\leq (\mathbb{E}[(X_1 + \cdots + X_n)^2] \cdot \mathbb{E}[\mathbb{1}_{(X_1+\cdots+X_n>0)}])^{\frac{1}{2}} \end{aligned} \tag{5}$$

Wenn beide Seiten von Ungleichung (5) quadriert werden und die Gleichheit $\mathbb{E}(\mathbb{1}_A) = \int \mathbb{1}_A d\mathbb{P} = \mathbb{P}(A)$ eingesetzt wird, so erhält man

$$(\mathbb{E}[X_1 + \cdots + X_n])^2 \leq \mathbb{E}[(X_1 + \cdots + X_n)^2] \cdot \mathbb{P}(X_1 + \cdots + X_n > 0)$$

$$\mathbb{P}(X_1 + \cdots + X_n > 0) \geq \frac{(\mathbb{E}[X_1 + \cdots + X_n])^2}{\mathbb{E}[(X_1 + \cdots + X_n)^2]}$$

Was der zu zeigenden Behauptung in Ungleichung (4) entspricht. □

2 Ungleichungen in Zusammenhang mit der Normalverteilung

Sei Z eine Zufallsvariable, die Verteilungsfunktion von Z ist definiert durch $F(x) = \mathbb{P}(Z \leq x)$, falls die Wahrscheinlichkeitsdichtefunktion $p(x)$ existiert, dann ist sie messbar mit $F(x) := \int_{-\infty}^{x} p(y)dy$
Es werden in diesem Kapitel folgende Definitionen verwendet:
Die Dichtefunktion der Standardnormalverteilung ist $\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2}$
Die Verteilungsfunktion der Standardnormalverteilung: $\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$

2.1 Ungleichung 1 in Zusammenhang mit der Normalverteilung

Es gilt für $-\infty < a < b < \infty$:

$$\frac{1}{\sqrt{2\pi}}(b-a)(e^{-\frac{(a^2 \vee b^2)}{2}}) \leq \Phi(b) - \Phi(a) \leq \frac{1}{\sqrt{2\pi}}(b-a) \tag{6}$$

Beweis. Definitionsgemäß gilt:

$$\Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt$$

Die rechte Ungleichung von (6) folgt aus der Tatsache, dass

$$e^{-\frac{t^2}{2}} \leq 1 = e^0 \Rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt \leq \frac{1}{\sqrt{2\pi}} \int_a^b 1dt = \frac{1}{\sqrt{2\pi}}(b-a)$$

und die linke Ungleichung folgt mit $a^2 \vee b^2 = \max\{a^2, b^2\}$ und $x^2 \leq \max\{a^2, b^2\}$:

$$e^{-\frac{x^2}{2}} \geq e^{-\frac{(a^2 \vee b^2)}{2}} \Rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx \geq \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{(a^2 \vee b^2)}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{-\frac{(a^2 \vee b^2)}{2}} (b-a)$$

□

2.2 Ungleichung 2 in Zusammenhang mit der Normalverteilung

Für alle $x > 0$ gilt,

$$\left(\frac{1}{x} - \frac{1}{x^3}\right)\varphi(x) < \frac{x}{1+x^2}\varphi(x) < 1 - \Phi(x) < \frac{1}{x}\varphi(x) \tag{7}$$

Beweis.
(i) Die linke Ungleichung von (7) folgt aus elementaren Umformungen:

$$\frac{x^2-1}{x^3} < \frac{x}{x^2+1} \Leftrightarrow (x^2-1)\cdot(x^2+1) < x^4 \Rightarrow x^4 - 1 < x^4$$

(ii) Für den Beweis der mittleren Ungleichung von (7) definieren wir zunächst: $Q(x) = 1 - \Phi(x)$. Definitionsgemäß ergibt sich:

$$Q(x) = 1 - \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}dt - \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} e^{-\frac{t^2}{2}}dt = \frac{1}{\sqrt{2\pi}} \cdot \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt$$

Wodurch sich die zu zeigende Ungleichung

$$\frac{x}{x^2+1}\varphi(x) < Q(x) \tag{8}$$

als folgende Ungleichung darstellen lässt:

$$\frac{x}{x^2+1} \cdot e^{-\frac{x^2}{2}} < \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt$$

Es gilt:

$$\begin{aligned}
\left(1 + \frac{1}{x^2}\right)Q(x)\sqrt{2\pi} &= \int_{x}^{\infty} \left(1 + \frac{1}{x^2}\right)e^{-\frac{t^2}{2}}dt \\
&> \int_{x}^{\infty} \left(1 + \frac{1}{t^2}\right)e^{-\frac{t^2}{2}}dt \\
&= \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt + \int_{x}^{\infty} \frac{1}{t^2}e^{-\frac{t^2}{2}}dt \\
&= \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt + \left[-\frac{1}{t}e^{-\frac{t^2}{2}}\Big|_{x}^{\infty} - \int_{x}^{\infty} \frac{1}{t} \cdot te^{-\frac{t^2}{2}}dt\right] \\
&= \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt + \frac{1}{x}e^{-\frac{x^2}{2}} - \int_{x}^{\infty} e^{-\frac{t^2}{2}}dt \\
&= \frac{1}{x}e^{-\frac{x^2}{2}}
\end{aligned}$$

Wodurch die Ungleichungen folgen:

$$\begin{aligned}
\left(1 + \frac{1}{x^2}\right)Q(x)\sqrt{2\pi} &> \frac{1}{x}e^{-\frac{x^2}{2}} \\
\left(\frac{x^2+1}{x^2}\right)Q(x) &> \frac{1}{x\sqrt{2\pi}}e^{-\frac{x^2}{2}} = \frac{1}{x}\varphi(x)
\end{aligned}$$

Insgesamt ergibt sich

$$Q(x) > \frac{x^2}{x^2+1}\frac{1}{x}\varphi(x) = \frac{x}{x^2+1}\varphi(x)$$

Wodurch Ungleichung (8) gezeigt wurde.

(iii) Für die rechte Ungleichung von (7) betrachtet man das Integral: $\int_x^\infty e^{-\frac{t^2}{2}}dt$ für $t \geq x$ und $t \in [x, \infty] \Rightarrow \frac{t}{x} \geq 1$. Welches abgeschätzt werden kann durch:

$$\int_{x}^{\infty} e^{-\frac{t^2}{2}}dt < \int_{x}^{\infty} \frac{t}{x}e^{-\frac{t^2}{2}}dt$$

Durch Substitution von: $u = \frac{t^2}{2}$ mit $dt = \frac{1}{t}du$ ergibt sich

$$\int_x^{\infty} \frac{t}{x} e^{-\frac{t^2}{2}} dt = \int_x^{\infty} \frac{t}{x} e^{-u} \frac{1}{t} du = \frac{1}{x} \int_{x^2/2}^{\infty} e^{-u} du = -\frac{1}{x} e^{-\frac{t^2}{2}} \Big|_x^{\infty} = 0 + \frac{1}{x} e^{-\frac{x^2}{2}} = \frac{1}{x} e^{-\frac{x^2}{2}}$$

Wodurch zunächst

$$\int_x^{\infty} e^{-\frac{t^2}{2}} dt < \frac{1}{x} e^{-\frac{x^2}{2}}$$

gezeigt wurde, womit schlussendlich die zu zeigende rechte Ungleichung von (7) folgt:

$$\frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{t^2}{2}} dt < \frac{1}{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = \frac{1}{x} \cdot \varphi(x)$$

□

3 Ungleichungen in Zusammenhang mit der charakteristischen Funktion

Die Charakteristische Funktion (c.f.) oder Fouriertransformation spielt eine wichtige Rolle in der Wahrscheinlichkeitstheorie, besonders in der Theorie von Grenzwertsätzen über Summen von unabhängige Zufallsvariablen.

Die charakteristische Funktion $\varphi(x)$ einer Zufallsvariable x mit Verteilungsdichte $f(x)$ ist definiert durch

$$\varphi(x) = \mathbb{E}[e^{iXt}] = \mathbb{E}[cos(Xt)] + i\mathbb{E}[sin(Xt)] = \int_{-\infty}^{+\infty} e^{itx} d\mathbb{P}(x) = \int_{-\infty}^{+\infty} e^{itx} f(x) dx$$

3.1 Ungleichung bezogen nur auf die Charakteristische Funktion

Für beliebige $t \in \mathbb{R}$,

$$1 - |f(2t)|^2 \leq 4(1 - |f(t)|^2) \tag{9}$$

Beweis.

Sei $G(x)$ eine beliebige Verteilungsfunktion und $g(t)$ die zugehörige charakteristische Funktion. Definitionsgemäß gilt $1 = \int_{-\infty}^{+\infty} 1 dG(x) = \int_{-\infty}^{+\infty} f(x)dx$. Dadurch ergibt sich

$$\begin{aligned} Re(1 - g(t)) &= Re\Big(1 - \int_{-\infty}^{+\infty} e^{itx} dG(x)\Big) \\ &= Re\Big(1 - \int_{-\infty}^{+\infty} cos(tx) dG(x) - \int_{-\infty}^{+\infty} i \cdot sin(tx) dG(x)\Big) \\ &= \Big(1 - \int_{-\infty}^{+\infty} cos(tx) dG(x)\Big) \\ &= \int_{-\infty}^{+\infty} \Big(1 - cos(tx)\Big) dG(x) \end{aligned} \tag{10}$$

An dieser Stelle werden wir zunächst den Hilfssatz

$$1 - cos(tx) \geq \frac{1}{4}(1 - cos(2tx)) \tag{11}$$

definieren und beweisen.

Dazu werden folgende Gleichheiten verwendet. Zunächst für die linke Seite gilt

$$\begin{aligned} cos(2x) &= cos^2 x - sin^2 x = 1 - 2sin^2 x \\ 1 - cos(2x) &= 2 \cdot sin^2(x) \\ 1 - cos(tx) &= 2 \cdot sin^2\Big(\frac{tx}{2}\Big) \end{aligned}$$

und die rechte Seite kann vereinfacht werden auf

$$1 - cos(2tx) = 2 \cdot sin^2(tx)$$
$$\frac{1}{4}(1 - cos(2tx)) = \frac{1}{2}sin^2(tx)$$

zusammengefasst folgt die Ungleichung

$$2 \cdot sin^2(\frac{tx}{2}) \geq \frac{1}{2}sin^2(tx)$$
$$4 \cdot sin^2(\frac{tx}{2}) \geq sin^2(tx)$$

Wendet man nun die Gleichung $sin(a+b) = sin(a)cos(b) + cos(a)sin(b)$ an so folgt:

$$sin(tx) = sin\left(\frac{tx}{2} + \frac{tx}{2}\right) = 2sin\left(\frac{tx}{2}\right)cos\left(\frac{tx}{2}\right)$$
$$sin^2(tx) = 4sin^2\left(\frac{tx}{2}\right)cos^2\left(\frac{tx}{2}\right)$$
$$4 \cdot sin^2(\frac{tx}{2}) \geq 4sin^2\left(\frac{tx}{2}\right)\underbrace{cos^2\left(\frac{tx}{2}\right)}_{\leq 1}$$

Nun kann Ungleichung (11) in Gleichung (10) auf der vorherigen Seite eingesetzt werden

$$Re(1 - g(t)) = \int_{-\infty}^{+\infty} \left(1 - cos(tx)\right)dG(x)$$
$$\geq \frac{1}{4}\int_{-\infty}^{+\infty} \left(1 - cos(2tx)\right)dG(x)$$

Was nun der Ungleichung

$$4Re(1 - g(t)) \geq Re(1 - g(2t))$$

für alle reelle t entspricht.
Die zu zeigende Ungleichung (9) folgt in dem man $g(t) = |f(t)|^2$ setzt. □

3.2 Inkrement Ungleichung

Für beliebige t und h gilt

$$|f(t) - f(t+h)|^2 \leq 2(1 - Ref(h)) \tag{12}$$

wobei $f(h)$ die zugehörige c.f von $F(x)$

Beweis.
Wegen der Dreiecksungleichung und der Cauchy-Schwarz Ungleichung

$(\mathbb{E}|XY|)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$ folgt unmittelbar

$$
\begin{aligned}
|f(t) - f(t+h)|^2 &= \left| \int_{-\infty}^{+\infty} e^{itx} - e^{i(t+h)x} dF(x) \right|^2 \\
&= \left| \int_{-\infty}^{+\infty} e^{itx}(1 - e^{ihx}) dF(x) \right|^2 \\
&\leq \left(\int_{-\infty}^{+\infty} |e^{itx}|^2 dF(x) \right) \cdot \left(\int_{-\infty}^{+\infty} |(1 - e^{ihx})|^2 dF(x) \right) \\
&\leq \int_{-\infty}^{+\infty} 1 dF(x) \cdot \left(\int_{-\infty}^{+\infty} |(1 - e^{ihx})|^2 dF(x) \right) \\
&= 1 \cdot \left(\int_{-\infty}^{+\infty} |(1 - cos(hx) - i\ sin(hx))|^2 dF(x) \right) \\
&= \left(\int_{-\infty}^{+\infty} ((1 - cos(hx))^2 + sin^2(hx)) dF(x) \right) \\
&= \left(\int_{-\infty}^{+\infty} (1 - 2cos(hx) + cos^2(hx) + sin^2(hx)) dF(x) \right) \\
&= \left(\int_{-\infty}^{+\infty} (2 - 2cos(hx)) dF(x) \right) \\
&= \left(2 \int_{-\infty}^{+\infty} (1 - cos(hx)) dF(x) \right) \\
&= 2 \cdot Re(1 - f(h)) = 2 \cdot (1 - Ref(h))
\end{aligned}
$$

□

3.3 Ungleichung zusammenhängend mit Charakteristische Funktion und Verteilungsfunktion

Sei $F(x)$ eine Verteilungsfunktion und $f(h)$ die zugehörige charakteristische Funktion, dann gilt

$$
\int \frac{x^2}{1+x^2} dF(x) \leq \int_0^{\infty} e^{-t} |1 - f(t)| dt \tag{13}
$$

Beweis.
Wir beginnen den Beweis mit der Definition und dem Beweis des Hilfssatz:

$$
\int_0^{\infty} e^{-t} cos(xt) dt = \frac{1}{x^2+1} \tag{14}
$$

Durch Partielle Integration der linken Seite erhält man:

$$
\begin{aligned}
\int_0^{\infty} e^{-t} cos(xt) dt &= \frac{sin(xt)}{x} e^{-t} \Big|_0^{\infty} + \int_0^{\infty} \frac{sin(xt)}{x} e^{-t} dt \\
&= \frac{sin(xt)}{x} e^{-t} \Big|_0^{\infty} + \frac{1}{x} \left(\frac{-cos(xt)}{x} e^{-t} \Big|_0^{\infty} - \frac{1}{x} \int_0^{\infty} e^{-t} cos(xt) dt \right) \\
&= \frac{sin(xt)}{x} e^{-t} - \frac{1}{x^2} cos(xt) e^{-t} \Big|_0^{\infty} - \frac{1}{x^2} \int_0^{\infty} e^{-t} cos(xt) dt \\
&= 0 + \frac{1}{x^2} - \frac{1}{x^2} \int_0^{\infty} e^{-t} cos(xt) dt
\end{aligned}
$$

Daraus folgt

$$
\frac{x^2+1}{x^2} \int_0^{\infty} e^{-t} cos(xt) dt = \frac{1}{x^2}
$$

und schlussendlich folgt mit einer Multiplikation von $\frac{x^2}{x^2+1}$ auf beiden Seiten die zu zeigende Hilfssatzgleichung (14).

Mit Hilfe von Gleichung (10) auf Seite 8: und der Äquivalenz $\int_0^\infty e^{-t}dt = 1$ folgen die Gleichungen:

$$\begin{aligned}
1 - \int_0^\infty e^{-t} cos(xt)dt &= 1 - \frac{1}{1+x^2} \\
\int_0^\infty e^{-t}dt - \int_0^\infty e^{-t} cos(xt)dt &= \frac{x^2}{1+x^2} \\
\int_0^\infty e^{-t}(1 - cos(xt))dt &= \frac{x^2}{1+x^2}
\end{aligned}$$

Nun wird auf beiden Seiten nach dF(x) integriert und anschließend der Satz von Fubini angewendet;

$$\begin{aligned}
\int_{-\infty}^{+\infty}\int_0^\infty e^{-t}(1 - cos(xt))dtdF(x) &= \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2} dF(x) \\
\int_0^\infty e^{-t}\int_{-\infty}^{+\infty}(1 - cos(xt))dF(x)dt &= \cdots \\
\int_0^\infty e^{-t}\Big(Re(1 - f(t))\Big)dt &= \cdots \\
\int_0^\infty e^{-t}\Big(1 - Ref(t)\Big)dt &= \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2} dF(x)
\end{aligned}$$

Es fehlt noch zu zeigen, dass

$$|1 - f(t)| \geq (1 - Ref(t))$$

gilt, indem man die linke Seite der Ungleichung umformt:

$$\begin{aligned}
|1 - f(t)| &= |1 - Ref(t) - Imf(t)| \\
&= \sqrt{(1 - Ref(t))^2 + \underbrace{Imf(t)^2}_{\geq 0}} \\
&\geq \sqrt{(1 - Ref(t))^2} = 1 - Ref(t)
\end{aligned}$$

womit schlussendlich die Behauptung (13) gezeigt werden kann:

$$\int_0^\infty e^{-t}|1 - f(t)|dt \geq \int_0^\infty e^{-t}\Big(1 - Ref(t)\Big)dt = \int_{-\infty}^{+\infty} \frac{x^2}{1+x^2} dF(x)$$

□

4 Wahrscheinlichkeitsungleichungen von Zufallsvariablen

Wahrscheinlichkeitsungleichungen von Zufallsvariablen besonders die von Summen von Zufallsvariablen sind wichtig in der analytischen Wahrscheinlichkeitstheorie, genauer bei den Grenzwertsätzen.
Sofern nicht anders gegeben bezeichnet $X_1, X_2, \ldots X_n$ eine Folge von Zufallsvariablen und S_k die Summe $S_k := \sum_{j=1}^k X_i$.

4.1 Ungleichungen bezogen auf zwei Zufallsvariablen

Für zwei Zufallsvariablen X und Y gilt,

$$\begin{gathered}
\mathbb{P}(X + Y \geq x) \leq \mathbb{P}(X \geq x/2) + \mathbb{P}(Y \geq x/2); \\
\mathbb{P}(|X + Y| \geq x) \leq \mathbb{P}(|X| \geq x/2) + \mathbb{P}(|Y| \geq x/2); \\
|\mathbb{P}(X < x_1, Y < y_1) - \mathbb{P}(X < x_2, Y < y_2)| \leq |\mathbb{P}(X < x_1, X < x_2)| + |\mathbb{P}(Y < y_1, Y < y_2)|
\end{gathered}$$

Angenommen X und Y sind unabhängig. Dann gilt

$$\begin{aligned}\mathbb{P}(X+Y \leq x) &\geq \mathbb{P}(X \leq x/2) + \mathbb{P}(Y \leq x/2) \\ \mathbb{P}(|X+Y| \leq x) &\geq \mathbb{P}(|X| \leq x/2) + \mathbb{P}(|Y| \leq x/2)\end{aligned}$$

und für $x > 0$ groß genug,

$$\begin{aligned}\mathbb{P}(|X| > x) &\leq 2\mathbb{P}(|X| > x, |Y| < x/2) \\ &\leq 2\mathbb{P}(|X+Y| > x/2)\end{aligned}$$

4.2 Symmetrisierende Ungleichungen - Schwache symmetrisierende Ungleichungen

Seien X und X' zwei unabhängig identisch verteilte Zufallsvariablen und X^s definiert durch $X^s = X - X'$. Des Weiteren bezeichnet mX den Median von X, sprich eine Zahl welche die Bedingung $\mathbb{P}(X \geq mX) \geq \frac{1}{2} \leq \mathbb{P}(X \leq mX)$ erfüllt.
Für beliebige Zahlen x und a gilt

$$\frac{1}{2}\mathbb{P}(X - mX \geq x) \leq \mathbb{P}(X^s \geq x) \tag{15}$$

$$\frac{1}{2}\mathbb{P}(|X - mX| \geq x) \leq \mathbb{P}\Big(|X^s| \geq x\Big) \leq 2\mathbb{P}\Big(|X - a| \geq \frac{x}{2}\Big) \tag{16}$$

Beweis.
Wir beginnen mit dem Beweis von Ungleichung (15) indem die rechte Seite umgeformt wird

$$\begin{aligned}\mathbb{P}(X^s \geq x) &= \mathbb{P}\{(X - mX) - (X' - mX') \geq x\} \\ &\geq \mathbb{P}\{X - mX \geq x, X' - mX' \leq 0\} \\ &= \mathbb{P}(X - mX \geq x)\mathbb{P}(X' - mX' \leq 0) \\ &\geq \frac{1}{2}\mathbb{P}(X - mX \geq x)\end{aligned}$$

Womit die Ungleichung (15) bewiesen wäre.
Der Beweis von Ungleichung (16) widmet sich der linken Seite mit Hilfe einer Substitution von X durch $-X$. Die rechte Ungleichung folgt dann durch:

$$\begin{aligned}\mathbb{P}(|X^s| \geq x) &= \mathbb{P}\Big\{\Big|(X-a) - (X'-a)\Big| \geq x\Big\} \\ &\leq \mathbb{P}\Big\{|(X-a)| + |(X'-a)| \geq x\Big\} \\ &\leq \mathbb{P}\Big(|X-a| \geq \frac{x}{2} \dot{\cup} |X'-a| \geq \frac{x}{2}\Big) \\ &= \mathbb{P}\Big(|X-a| \geq \frac{x}{2}\Big) + \mathbb{P}\Big(|X'-a| \geq \frac{x}{2}\Big) \\ &= 2\mathbb{P}\Big(|X-a| \geq \frac{x}{2}\Big)\end{aligned}$$

□

4.3 Symmetrisierende Ungleichungen - Allgemeine symmetrisierende Ungleichungen

Sei $\{X_n, n \geq 1\}$, eine Folge von Zufallsvariablen. Dann gilt für beliebige $x > 0$ und für eine beliebige Folge $\{C_n, n \geq 1\}$ von Zahlen:

$$\frac{1}{2}\mathbb{P}\Big\{\sup_{n\geq 1}(X_n - mX_n) \geq x\Big\} \leq \mathbb{P}(\sup_{n\geq 1} X_n^s \geq x); \tag{17}$$

$$\frac{1}{2}\mathbb{P}\Big\{\sup_{n\geq 1}|X_n - mX_n| \geq x\Big\} \leq \mathbb{P}\Big\{\sup_{n\geq 1}|X_n^s| \geq x\Big\} \leq 2\mathbb{P}\Big\{\sup_{n\geq 1}|X_n - C_n| \geq \frac{x}{2}\Big\} \tag{18}$$

Beweis.
Zunächst definieren wir eine Folge X_n^s mit $X_n^s = X_n - X_n'$, sowie die Ereignisse A_n, B_n und C_n mit $A_n = \{X_n - mX_n \geq x\}, B_n = \{X_n' - mX_n' \leq 0\}, C_n = \{X_n^s \geq x\}$, dadurch folgt, dass $A_n \cap B_n \subset C_n$

Man erhält Ungleichung 17, indem Ungleichung (2) für $\inf_n \mathbb{P}(B_n) \geq \frac{1}{2}$ verwendet wird. Die Ungleichung (18) folgt durch äquivalente Argumentationen wie der Beweis von Ungleichung (16). □

4.4 Levy Ungleichung

Seien $X_1, X_2, \cdots, X_n$ unabhängige Zufallsvariablen, $x > 0$. Dann gilt:

$$\mathbb{P}\Big\{\max_{1\leq j\leq n}(S_j - m(S_j - S_n)) \geq x\Big\} \leq 2\mathbb{P}(S_n \geq x) \tag{19}$$

$$\mathbb{P}\Big\{\max_{1\leq j\leq n}|S_j - m(S_j - S_n)| \geq x\Big\} \leq 2\mathbb{P}(|S_n| \geq x) \tag{20}$$

Beweis.
Sei $S_0 = 0, S_k^* = \max\limits_{1\leq j\leq k}(S_j - m(S_j - S_n))$ und

$$\begin{aligned} A_k &= \{S_{k-1}^* < x, S_k - m(S_k - S_n) \geq x\} \\ B_k &= \{S_n - S_k - m(S_n - S_k) \geq 0\}. \end{aligned}$$

Nach Definition des Median gilt: $m(S_n - S_k) = -m(S_k - S_n)$. Des Weiteren gilt aus den obigen Definitionen, dass $S_0^* = 0$ und $\mathbb{P}(B_k) = \mathbb{P}(S_n - S_k \geq m(S_n - S_k)) \geq \frac{1}{2}$.
Außerdem sind die Mengen A_k disjunkt, die Folge S_k^* monoton steigend da $S_{k-1}^* < x$ und $S_k^* \geq x$. Wegen der Unabhängigkeit der Ereignisse A_k und B_k folgt, dass

$$\begin{aligned} \mathbb{P}(S_n \geq x) &\geq \mathbb{P}\Big(\underbrace{\bigcup_{k=1}^{n} A_k \cap B_k}_{\{S_{n-1}^* < x, S_n \geq x\}}\Big) \\ &= \sum_{k=1}^{n} \mathbb{P}(A_k \cap B_k) \\ &= \sum_{k=1}^{n} \mathbb{P}(A_k)\mathbb{P}(B_k) \\ &\geq \frac{1}{2}\sum_{k=1}^{n} \mathbb{P}(A_k) \\ &= \frac{1}{2}\mathbb{P}(S_n^* \geq x) \end{aligned}$$

womit Ungleichung (19) folgt. Durch Substitution von X_j durch $-X_j$ für $1 \leq j \leq n$ in Ungleichung (19) folgt unmittelbar auch Ungleichung (20). □

4.5 Levy Ungleichung Korollar

Angenommen $\mathbb{E}X_j = 0$, $\mathbb{E}X_j^2 < \infty$, $j = 1, \cdots, n$. Setze $B_n = \sum_{j=1}^{n} \mathbb{E}X_j^2$ Dann gilt:

$$\mathbb{P}\Big\{\max_{1\leq j\leq n} S_j \geq x\Big\} \leq 2\mathbb{P}(S_n \geq x - \sqrt{B_n}); \tag{21}$$

Beweis.
Wir verwenden ein Spezialfall von der Chebyshev Ungleichung (26) und setzen statt X, $S_n - S_j$

ein. Daraus folgt mit $m(X) \leq \mathbb{E}X + \sigma$ aus der Definition des Median: $P(X \geq m(X)) \geq \frac{1}{2} \geq P(X \geq \mathbb{E}X + \sigma)$

$$m(S_n - S_j) \leq \mathbb{E}(S_n - S_j) + \sqrt{\mathbb{V}(S_n - S_j)} = \sqrt{\mathbb{V}(S_n - S_j)}$$

weil $\mathbb{E}\Big(\sum_{k=j+1}^n X_k\Big) = \sum_{k=j+1}^n \mathbb{E}(X_k) = 0$
wegen $m(S_j - S_n) = -m(S_n - S_j)$ gilt die gleiche Ungleichung auch für $S_j - S_n$ d.h.

$$|m(S_j - S_n)| \leq \sqrt{\mathbb{V}(S_n - S_j)} = \sqrt{\mathbb{V}(S_j - S_n)}$$

und schließlich gilt

$$\begin{aligned}\mathbb{V}(S_n - S_j) &= \mathbb{E}\Big((\sum_{k=j+1}^n X_k)^2\Big) - \Big(\underbrace{\mathbb{E}\Big(\sum_{k=j+1}^n X_k\Big)}_{=0}\Big)^2 = \mathbb{E}\Big(\sum_{k=j+1}^n X_k^2 + \sum_{\substack{k=j+1\\k\neq l}}^n \sum_{l=j+1}^n X_k X_l\Big)\\ &= \sum_{k=j+1}^n \mathbb{E}X_k^2 + \underbrace{\sum_{\substack{k=j+1\\k\neq l}}^n \sum_{l=j+1}^n \mathbb{E}(X_k)\mathbb{E}(X_l)}_{=0} = \sum_{k=j+1}^n \mathbb{E}X_k^2\end{aligned}$$

Daraus folgt

$$m(S_n - S_j) \leq \sqrt{\mathbb{V}(S_n - S_j)} = \sqrt{\sum_{k=j+1}^n \mathbb{E}X_k^2} \leq \sqrt{\sum_{k=1}^n \mathbb{E}X_k^2} = \sqrt{B_n}$$

Damit gilt

$$\mathbb{P}\Big(\max_{1\leq j\leq n} S_j \geq x\Big) = \mathbb{P}\Big(\max_{1\leq j\leq n} S_j - \max_{k\in\{1,\cdots,n\}}(m(S_k - S_n)) \geq x - \max_{k\in\{1,\cdots,n\}}(m(S_k - S_n))\Big) \tag{22}$$

wegen $|m(S_k - S_n)| \leq \sqrt{B_n}$ gilt

$$x - m(S_k - S_n) \geq x - |m(S_k - S_n)| \geq x - \sqrt{B_n}$$

insgesamt bekommt man

$$x - \max_{k\in\{1,\cdots,n\}}(m(S_k - S_n)) \geq x - \sqrt{B_n}$$

Wodurch sich Gleichung (22) weiter abschätzen lässt mit

$$\cdots \leq \mathbb{P}\Big(\max_{1\leq j\leq n} S_j - \max_{k\in\{1,\cdots,n\}}(m(S_k - S_n)) \geq x - \sqrt{B_n}\Big) \leq \mathbb{P}\Big(\max_{1\leq j\leq n}(S_j - m(S_j - S_n)) \geq x - \sqrt{B_n}\Big)$$

Die rechte Ungleichung gilt, weil an jedem Punkt ω $\exists j : S_j(\omega) \geq S_i(\omega)$ $\forall 1 \leq i \leq n$. Damit folgt, dass $S_j(\omega) - \max_{k\in\{1,\cdots,n\}}(m(S_k - S_n)) \leq S_j(\omega) - m(S_j - S_n) \leq \max_{1\leq j\leq n}(S_j(\omega) - m(S_j - S_n))$
Nun verwende Levy Ungleichung (19) und ersetze x durch $x - \sqrt{B_n}$ dann folgt die zu zeigende Ungleichung

$$\mathbb{P}\Big(\max_{1\leq j\leq n}(S_j - m(S_j - S_n)) \geq x - \sqrt{B_n}\Big) \leq 2\mathbb{P}(S_n \geq x - \sqrt{B_n})$$

Diese Ungleichung wird verwendet um das Gesetz der großen Zahlen zu beweisen. Siehe Petrov 1995 □

5 Wahrscheinlichkeitsschranken bezogen auf die Momente

5.1 Allgemeine Form - Chebyshev-Markov Art von Ungleichungen

Sei X eine Zufallsvariable und $g(x)$ eine nichtfallende Funktion auf $\mathbb{R}$ Dann gilt für alle x

$$\mathbb{P}(X \geq x) \leq \frac{\mathbb{E}\Big(g(X)\Big)}{g(x)} \tag{23}$$

Beweis.

$$g(x) \cdot \mathbb{P}(X \geq x) = g(x) \cdot \int_\Omega \mathbb{1}_{[X \geq x]} d\mathbb{P} = \int_\Omega g(x) \mathbb{1}_{[X \geq x]} d\mathbb{P} \leq \int_\Omega g(x) d\mathbb{P} = \mathbb{E}\Big(g(x)\Big)$$

□

5.2 Chebyshev Ungleichung

Für beliebiges $x > 0$, gilt

$$\mathbb{P}(|X - \mathbb{E}X| \geq x) \leq \frac{\mathbb{V}\Big(X\Big)}{x^2}; \tag{24}$$

5.3 Markov Ungleichung

Für beliebige Zahlen $r > 0$ und $x > 0$, gilt

$$\mathbb{P}(|X| \geq x) \leq \frac{\mathbb{E}|X|^r}{x^r}; \tag{25}$$

5.4 Verallgemeinerung der Chebyshev Ungleichung

Setze $\sigma^2 = \mathbb{V}(X)$. Dann gilt für alle x und a

$$\mathbb{P}(X - \mathbb{E}X \geq x) \leq \frac{\sigma^2 + a^2}{(x+a)^2}; \tag{26}$$

Beweis.

$$\mathbb{P}(X - \mathbb{E}X \geq x) = \mathbb{P}(X - \mathbb{E}X + a \geq x + a) \leq \mathbb{P}(|X - \mathbb{E}X + a| \geq |x + a|) \leq \frac{\mathbb{E}[(X - \mathbb{E}X + a)^2]}{(x+a)^2}$$

Die ganz rechte Ungleichung gilt wegen der Ungleichung (23) von der allgemeinen Form von der Chebyshev - Markov Art von Ungleichungen. Weiters gilt

$$\begin{aligned}
\mathbb{E}[(X - \mathbb{E}X + a)^2] &= \mathbb{E}[(X - \mathbb{E}X)^2 + 2a(X - \mathbb{E}X) + a^2] \\
&= \mathbb{E}[X^2 - 2X\mathbb{E}X + (\mathbb{E}(X))^2] + 2a\mathbb{E}(X - \mathbb{E}X) + \mathbb{E}a^2 \\
&= \mathbb{E}[X^2] - 2\mathbb{E}X\mathbb{E}X + (\mathbb{E}(X))^2 + 2a(\mathbb{E}X - \mathbb{E}X) + a^2 \\
&= \mathbb{E}[X^2] - 2(\mathbb{E}X)^2 + (\mathbb{E}(X))^2 + a^2 = \mathbb{E}[X^2] - (\mathbb{E}X)^2 + a^2 \\
&= \mathbb{V}(X) + a^2 = \sigma^2 + a^2
\end{aligned}$$

was zu zeigen ist. □

Spezialfall: wähle $a = \frac{\sigma^2}{x}$, so folgt:

$$\frac{\sigma^2 + a^2}{(x+a)^2} = \frac{\sigma^2 + \frac{\sigma^4}{x^2}}{\left(x + \frac{\sigma^2}{x}\right)^2} = \frac{\frac{\sigma^2 x^2 + \sigma^4}{x^2}}{\left(\frac{x^2+\sigma^2}{x}\right)^2} = \frac{\frac{\sigma^2 x^2 + \sigma^4}{x^2}}{\frac{(x^2+\sigma^2)^2}{x^2}} = \frac{\sigma^2(x^2+\sigma^2)}{(x^2+\sigma^2)^2} = \frac{\sigma^2}{x^2+\sigma^2}$$

Daraus ergibt sich

$$\mathbb{P}(X - \mathbb{E}X \geq x) \leq \frac{\sigma^2}{x^2+\sigma^2}$$

wenn $x = \sigma$ dann ist

$$\mathbb{P}(X - \mathbb{E}X \geq \sigma) \leq \frac{\sigma^2}{\sigma^2+\sigma^2} = \frac{1}{2}$$

also

$$\mathbb{P}(X \geq \mathbb{E}X + \sigma) \leq \frac{1}{2}$$

nach Definition von Median $\mathbb{P}(X \geq m(X)) \geq \frac{1}{2}$ folgt damit $m(X) \leq \mathbb{E}X + \sigma$. Symmetrisch betrachtet erhält man $|\mathbb{E}X - m(X)| \leq \sigma$

5.5 Kolmogorov Ungleichung

Seien $X_1, \cdots, X_n$ unabhängige Zufallsvariablen mit $\mathbb{E}X_i = 0$ für $j = 1, \cdots, n$, sowie $x > 0$ beliebig und S_k definiert durch $S_k := \sum_{j=1}^k X_j$.
Dann gilt

$$\mathbb{P}(\max_{1\leq j \leq n} |S_j| \geq x) \leq \frac{\mathbb{V}(S_n)}{x^2} \tag{27}$$

und falls $|X_j| \leq c$ für $1 \leq j \leq n$ mit einem $c > 0$, dann gilt

$$\mathbb{P}(\max_{1\leq j \leq n} |S_j| \geq x) \geq 1 - \frac{(x+c)^2}{\mathbb{V}(S_n)} \tag{28}$$

Beweis.
Sei $S_0 := 0$, $A_k := \{\omega : \max_{1 \leq j < k} |S_j(\omega)| < x \leq |S_k(\omega)|\}$ d.h. A_k ist die Menge aller Punkte ω, sodass $|S_j(\omega)|$ beim Index k erstmals größer oder gleich x ist. Dann folgt

$$\begin{aligned}
\int_{A_k} S_n^2 d\mathbb{P} &= \int_{A_k} (S_k + S_n - S_k)^2 d\mathbb{P} \\
&= \int_{A_k} S_k^2 d\mathbb{P} + 2 \cdot \int S_k \mathbb{1}_{A_k} \cdot (S_n - S_k) d\mathbb{P} + \int_{A_k} (S_n - S_k)^2 d\mathbb{P} \\
&\geq \int_{A_k} S_k^2 d\mathbb{P} + 2 \cdot \mathbb{E}[\underbrace{\mathbb{1}_{A_k} S_k}_{\text{sind}} \cdot \underbrace{(S_n - S_k)}_{\text{unabhängig}}] \\
&= \int_{A_k} S_k^2 d\mathbb{P} + 2\mathbb{E}[\mathbb{1}_{A_k} S_k]\mathbb{E}[(S_n - S_k)]
\end{aligned}$$

Es gilt $\mathbb{E}(S_n - S_k) = \mathbb{E}(\sum\limits_{j=k+1}^{n} X_j) = \sum\limits_{j=k+1}^{n} \mathbb{E}X_j = 0$, wodurch weiters folgt:

$$\int_{A_k} S_n^2 d\mathbb{P} \geq \int_{A_k} S_k^2 d\mathbb{P} \geq \int_{A_k} x^2 d\mathbb{P} = x^2 \cdot \mathbb{P}(A_k)$$

Durch Aufsummieren über $k = 1, \cdots, n$ ergibt sich:

$$\begin{aligned}\mathbb{V}(S_n) &= \mathbb{E}S_n^2 - (\mathbb{E}S_n)^2 = \mathbb{E}S_n^2 - (\mathbb{E}(\sum_{k=1}^{n} X_k))^2 = \mathbb{E}S_n^2 = \int S_n^2 d\mathbb{P} \geq \int_{\dot\bigcup_{k=1}^{n} A_k} S_n^2 d\mathbb{P} = \\ &= \sum_{k=1}^{n} \int_{A_k} S_n^2 d\mathbb{P} \geq x^2 \sum_{k=1}^{n} \mathbb{P}(A_k) = x^2 \mathbb{P}(\bigcup_{k=1}^{n} A_k) = x^2 \mathbb{P}(\max_{1\leq j<n} |S_j| \geq x)\end{aligned}$$

Woraus Ungleichung (27) folgt.

Nun folgt der Beweis von Ungleichung (28)
Angenommen, $|X_j| \leq c$, $j = 1, \cdots, n$ Sei $B_0 := \Omega$. Wenn X_j Funktionen von $(\Omega, \mathbf{S}, \mathbb{P})$ nach $\mathbb{R}$ sind, $B_k := \{\omega : \max_{1\leq j\leq k} |S_j(\omega)| < x\}$ d.h. B_k ist die Menge aller Punkte ω, wo alle $|S_j(\omega)|$ kleiner x sind. Dann ergibt sich folgende Gleichheit:

$$S_{k-1}\mathbb{1}_{B_{k-1}} + X_k\mathbb{1}_{B_{k-1}} = S_k\mathbb{1}_{B_{k-1}} = S_k\mathbb{1}_{B_k} + S_k\mathbb{1}_{A_k}$$

weil nach Definition $B_{k-1} = B_k \dot\cup A_k$, da

$$\begin{aligned}\omega \in B_{k-1} \Leftrightarrow \max_{1\leq j<k-1} |S_j(\omega)| < x \Leftrightarrow \left(\max_{1\leq j<k-1} |S_j(\omega)| < x \text{ und } |S_k(\omega)| \geq x\right) \text{ oder } |S_k(\omega)| < x \\ \Leftrightarrow \omega \in A_k \text{ oder } \omega \in B_k\end{aligned}$$

Aus der Definition gelten folgende Aussagen: $S_{k-1}\mathbb{1}_{B_{k-1}} = \sum_{j=1}^{k-1} X_j\mathbb{1}_{B_{k-1}}$, X_k sind unabhängig und $A_k \cap B_k = \emptyset$
Daraus folgend ergibt sich

$$\begin{aligned}\mathbb{E}(S_{k-1}\mathbb{1}_{B_{k-1}})^2 + \mathbb{E}(X_k^2)\mathbb{P}(B_{k-1}) &= \mathbb{E}(S_{k-1}\mathbb{1}_{B_{k-1}} + X_k\mathbb{1}_{B_{k-1}})^2 \\ &= \mathbb{E}(S_k\mathbb{1}_{B_k} + S_k\mathbb{1}_{A_k})^2 = \mathbb{E}(S_k\mathbb{1}_{B_k})^2 + \mathbb{E}(S_k\mathbb{1}_{A_k})^2\end{aligned}$$

Auf A_k ist $|S_{k-1}| < x$, daraus resultiert

$$|S_k\mathbb{1}_{A_k}| = |S_{k-1}\mathbb{1}_{A_k} + X_k\mathbb{1}_{A_k}| \leq |S_{k-1}\mathbb{1}_{A_k}| + |\underbrace{X_k}_{\leq c}\mathbb{1}_{A_k}| \leq (x+c)\mathbb{1}_{A_k}$$

Zusammen mit der Ungleichung $\mathbb{P}(B_n) \leq \mathbb{P}(B_{k-1})$ (weil $B_n \subseteq B_{k-1}$) folgt

$$\begin{aligned}\mathbb{E}(S_{k-1}\mathbb{1}_{B_{k-1}})^2 + \mathbb{E}(X_k^2)\mathbb{P}(B_n) &\leq \mathbb{E}(S_{k-1}\mathbb{1}_{B_{k-1}})^2 + \mathbb{E}(X_k^2)\mathbb{P}(B_{k-1}) \\ &= \mathbb{E}(S_k\mathbb{1}_{B_k})^2 + \mathbb{E}(S_k\mathbb{1}_{A_k})^2 \\ &\leq \mathbb{E}(S_k\mathbb{1}_{B_k})^2 + \mathbb{E}(\underbrace{((x+c)\mathbb{1}_{A_k})^2}_{(x+c)^2\mathbb{1}_{A_k}}) \\ &= \mathbb{E}(S_k\mathbb{1}_{B_k})^2 + (x+c)^2\mathbb{P}(A_k)\end{aligned}$$

Aufsummieren über $k = 1, \cdots, n$ ergibt:

$$\sum_{k=1}^{n} \mathbb{E}(S_{k-1}\mathbb{1}_{B_{k-1}})^2 + (\sum_{k=1}^{n} \mathbb{E}(X_k^2))\mathbb{P}(B_n) \leq \sum_{k=1}^{n} \mathbb{E}(S_k\mathbb{1}_{B_k})^2 + (x+c)^2\mathbb{P}(\bigcup_{k=1}^{n} A_k)$$

Man sieht, dass es sich hier um eine Teleskopsumme handelt. Auf der linken Seite bleibt $\mathbb{E}(S_0\mathbb{1}_{B_0})^2 = \mathbb{E}(0) = 0$ über während auf der rechten Seite $\mathbb{E}(S_n\mathbb{1}_{B_n})^2$ stehen bleibt.

$$\Big(\sum_{k=1}^{n} \mathbb{E}(X_k^2)\Big)\mathbb{P}(B_n) \leq \mathbb{E}(S_n\mathbb{1}_{B_n})^2 + (x+c)^2\mathbb{P}\Big(\bigcup_{k=1}^{n} A_k\Big) \leq x^2\mathbb{P}(B_n) + (x+c)^2\mathbb{P}(B_n^c) \leq (x+c)^2$$

Daraus folgt die Ungleichung

$$\mathbb{P}(B_n) \leq \frac{(x+c)^2}{(\sum\limits_{k=1}^{n} \mathbb{E}(X_k^2))}$$

Nun muss nur noch die Gleichung

$$\mathbb{V}(S_n) = \sum_{k=1}^{n} \mathbb{E}(X_k^2))$$

gezeigt werden.
Es gilt:

$$\begin{aligned}
\mathbb{V}(S_n) &= \mathbb{E}\Big((\sum_{k=1}^{n} X_k)^2\Big) - \Big(\mathbb{E}(\sum_{k=1}^{n} X_k)\Big)^2 \\
&= \mathbb{E}\Big(\sum_{k=1}^{n}\sum_{j=1}^{n} X_j X_k\Big) - \underbrace{\Big(\sum_{k=1}^{n} \underbrace{\mathbb{E}(X_k)}_{=0}\Big)^2}_{=0} \\
&= \mathbb{E}\Big(\sum_{k=1}^{n}\sum_{j=1}^{n} X_j X_k\Big)
\end{aligned}$$

Ist $j \neq k$, dann gilt wegen der Unabhängigkeit der X_i, $\mathbb{E}(X_j X_k) = \mathbb{E}(X_j)\mathbb{E}(X_k) = 0$. Also gilt

$$\begin{aligned}
\mathbb{V}(S_n) &= \mathbb{E}\Big(\sum_{k=1}^{n} X_k X_k\Big) \\
&= \sum_{k=1}^{n} \mathbb{E}(X_k^2)
\end{aligned}$$

Daraus folgt

$$\mathbb{P}(B_n) \leq \frac{(x+c)^2}{\mathbb{V}(S_n)}$$

Bzw. Äquivalent dazu

$$\begin{aligned}
\mathbb{P}(B_n^c) &= 1 - \mathbb{P}(B_n) \\
&\geq 1 - \frac{(x+c)^2}{\mathbb{V}(S_n)} \\
&= \mathbb{P}(\max_{1\leq j\leq n} |S_j| \geq x)
\end{aligned}$$

□

Verallgemeinerung der Kolmogorov Ungleichung

$$\mathbb{P}(\max_{1\leq j\leq n} S_j \geq x) \leq \frac{\mathbb{V}(S_n)}{x^2 + \mathbb{V}(S_n)} \tag{29}$$

Beweis.
Sei $a > 0$ beliebig: Weil S_k und $S_n - S_k$ unabhängig sind gilt

$$\begin{aligned}
\int_{A_k} 2(S_k + a)(S_n - S_k) d\mathbb{P} &= 2\Big(\int (S_k \mathbb{1}_{A_k})(S_n - S_k) d\mathbb{P} + \int a\mathbb{1}_{A_k}(S_n - S_k) d\mathbb{P}\Big) \\
&= 2\Big(\mathbb{E}(S_k \mathbb{1}_{A_k})\mathbb{E}(S_n - S_k) + \mathbb{E}(a\mathbb{1}_{A_k})\mathbb{E}(S_n - S_k)\Big) \\
&= 0
\end{aligned}$$

da $\mathbb{E}(S_n - S_k) = \sum_{i=k+1}^n \mathbb{E}(X_i) = 0$, weiters gilt

$$\begin{aligned}
\int_{A_k} (S_n + a)^2 d\mathbb{P} &= \int_{A_k} (S_k + a + S_n - S_k)^2 d\mathbb{P} \\
&= \int_{A_k} (S_k + a)^2 d\mathbb{P} + \underbrace{\int_{A_k} 2(S_k + a)(S_n - S_k) d\mathbb{P}}_{=0} + \int_{A_k} (S_n - S_k)^2 d\mathbb{P} \\
&\geq \int_{A_k} (S_k + a)^2 d\mathbb{P} \\
&\geq \int_{A_k} (x + a)^2 d\mathbb{P} \\
&= (x + a)^2 \int_{A_k} d\mathbb{P} \\
&= (x + a)^2 \mathbb{P}(A_k)
\end{aligned}$$

Ergibt mit einer Division durch $(x + a)^2$

$$\begin{aligned}
\mathbb{P}(A_k) &\leq \frac{1}{(x + a)^2} \int_{A_k} (S_n + a)^2 d\mathbb{P} \\
&\leq \frac{1}{(x + a)^2} \int (S_n + a)^2 d\mathbb{P} \\
&= \frac{1}{(x + a)^2} \Big(\int S_n^2 d\mathbb{P} + 2a \underbrace{\int S_n d\mathbb{P}}_{=\mathbb{E}(S_n)=0} + \int a^2 d\mathbb{P} \Big) \\
&= \frac{1}{(x + a)^2} \Big(\mathbb{V}(S_n) + a^2 \Big)
\end{aligned}$$

Mit $a = \frac{\mathbb{V}(S_n)}{x}$ ergibt sich die zu zeigende Behauptung. $\square$

6 Hoeffding Ungleichung

Seien X_j , $j = 1, \cdots, n$ unabhängig voneinander und $0 \leq X_j \leq 1$. Weiters sei $\mu_j = \mathbb{E}X_j$, $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$ und $\mu = \mathbb{E}\bar{X}$. Dann gilt für $0 < x < 1 - \mu$,

$$\mathbb{P}(\bar{X} - \mu \geq x) \leq \mathbb{P}(\max_{1 \leq j \leq n} (S_j - \mathbb{E}S_j) \geq nx) \leq \Big\{ \Big(\frac{\mu}{\mu + x} \Big)^{\mu + x} \Big(\frac{1 - \mu}{1 - \mu - x} \Big)^{1 - \mu - x} \Big\}^n \tag{30}$$

$$\leq e^{-g(\mu) n x^2} \leq e^{-2nx^2} \tag{31}$$

wobei

$$g(\mu) = \begin{cases} \frac{1}{1-2\mu} \log \frac{1-\mu}{\mu} & , \text{ für } 0 < \mu < \frac{1}{2} \\ \frac{1}{2\mu(1-\mu)} & , \text{ für } \frac{1}{2} \leq \mu < 1 \end{cases}$$

Wenn $a_j \leq b_j$ existiert, sodass $a_j \leq X_j \leq b_j$, $j = 1, \cdots, n$ dann gilt für beliebiges $x > 0$

$$\mathbb{P}(\bar{X} - \mu \geq x) \leq \mathbb{P}(\max_{1 \leq j \leq n} (S_j - \mathbb{E}S_j) \geq nx) \leq \exp \Big\{ -2n^2 x^2 / \sum_{j=1}^n (b_j - a_j)^2 \Big\}. \tag{32}$$

Beweis.
Es wird zuerst die erste Ungleichung in (31) gezeigt, da die Zweite (32) analog verläuft.
Sei $t > 0$. Da e^{tx} eine konvexe Funktion von x ist, dann erhält man mit

$$\begin{aligned}
e^{tx} &\leq (1 - x) + xe^t && \text{für } 0 \leq x \leq 1 \\
M_j(t) &= \mathbb{E}[e^{tX_j}] && \cdots \text{ Momentenerzeugende Funktion von } X_j \\
M_j(t) &= \mathbb{E}e^{tX_j} \leq \mathbb{E}(1 - X_j + X_j e^t) = 1 - \mu_j + \mu_j e^t
\end{aligned}$$

Wir verwenden erstens, die Tatsache, dass der geometrische Mittelwert kleiner ist, als der arithmetische Mittelwert also $\sqrt[n]{\prod_{j=1}^{n} X_j} \leq \frac{1}{n}\sum_{j=1}^{n} X_j$ und zweitens, die Doob Ungleichung auf das Submartingal $e^{t(S_j-\mathbb{E}S_j)}$

Für $S_k = \sum_{j=1}^{k} X_j$ folgt, dass $(e^{t(S_k-\mathbb{E}S_k)})_{k\in\mathbb{N}}$ ein Submartingal ist mit der Filtration

$$\mathcal{A}_k = \mathcal{A}_\sigma(\{e^{t(S_j-\mathbb{E}S_j)} : j \leq k\})$$

denn natürlich ist $e^{t(S_k-\mathbb{E}S_k)}\mathcal{A}_k$ messbar und es gilt

$$\begin{aligned}
\mathbb{E}\Big[e^{t(S_{k+1}-\mathbb{E}S_{k+1})}|\mathcal{A}_k\Big] &= \mathbb{E}\Big[e^{t(S_k-\mathbb{E}S_k)+t(X_{k+1}-\mathbb{E}X_{k+1})}|\mathcal{A}_k\Big] \\
&= \mathbb{E}\Big[e^{t(S_k-\mathbb{E}S_k)} \cdot \underbrace{e^{t(X_{k+1}-\mathbb{E}X_{k+1})}}_{*}|\mathcal{A}_k\Big] \\
&= \mathbb{E}\Big[e^{t(X_{k+1}-\mathbb{E}X_{k+1})}\Big]\mathbb{E}\Big[\underbrace{e^{t(S_k-\mathbb{E}S_k)}}_{**}|\mathcal{A}_k\Big] \\
&= \underbrace{\mathbb{E}[e^{tX_{k+1}}]}_{\geq e^{\mathbb{E}tX_{k+1}}***} e^{-\mathbb{E}tX_{k+1}}e^{t(S_k-\mathbb{E}S_k)} \geq e^{t(S_k-\mathbb{E}S_k)} \qquad \Rightarrow \text{ist Submartingal}
\end{aligned}$$

* $e^{t(X_{k+1}-\mathbb{E}X_{k+1})}$ ist unabhängig von $\mathcal{A}_k$ und von $e^{t(S_k-\mathbb{E}S_k)}$
** $e^{t(S_k-\mathbb{E}S_k)}$ ist $\mathcal{A}_k$ messbar
*** gilt wegen Jensen da exp- Fkt konvex und integrierbar ist

Die Doob- Ungleichung: Für $\{Y_n, \mathcal{A}_n, n \geq 1\}$ Submartingal und $x > 0$ ist geben durch

$$x \cdot \mathbb{P}\{\max_{1\leq j\leq n} Y_j \geq x\} \leq \mathbb{E}|Y_n| \tag{33}$$

liefert mit $\mathbb{P}\{\max_{1\leq j\leq n}(S_j - \mathbb{E}S_j) \geq nx\} = \mathbb{P}\{\max_{1\leq j\leq n} e^{t(S_j-\mathbb{E}S_j)} \geq e^{nxt}\}$, dass

$$\begin{aligned}
e^{nxt} \cdot \mathbb{P}\{\max_{1\leq j\leq n} e^{t(S_j-\mathbb{E}S_j)} \geq e^{nxt}\} &\leq \mathbb{E}|e^{t(S_n-\mathbb{E}S_n)}| = \mathbb{E}e^{t(S_n-\mathbb{E}S_n)} \\
\mathbb{P}\{\max_{1\leq j\leq n} e^{t(S_j-\mathbb{E}S_j)} \geq e^{nxt}\} &\leq e^{-nxt} \cdot \mathbb{E}e^{t(S_n-\mathbb{E}S_n)}
\end{aligned}$$

Aus $S_n = \sum_{j=1}^{n} X_j$ folgt $\mathbb{E}S_n = \sum_{j=1}^{n} \mathbb{E}X_j = \sum_{j=1}^{n} \mu_j$ wegen der Unabhängigkeit der X_j und $\mu = \mathbb{E}\bar{X} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\ X_j = \frac{1}{n}\sum_{j=1}^{n} \mu_j$ bzw. $n\mu = \sum_{j=1}^{n} \mu_j = \mathbb{E}S_n$ Damit gilt:

$$\begin{aligned}
\mathbb{P}\{\max_{1\leq j\leq n} e^{t(S_j-\mathbb{E}S_j)} \geq e^{nxt}\} \leq e^{-nxt} \cdot \mathbb{E}e^{t(S_n-\mathbb{E}S_n)} &= e^{-nxt}\mathbb{E}e^{tS_n}e^{-t\mathbb{E}S_n} \\
&= e^{-nxt}\mathbb{E}e^{t\sum_{j=1}^{n} X_j}e^{-t\sum_{j=1}^{n}\mu_j} \\
&= e^{-nxt}\mathbb{E}\prod_{j=1}^{n} e^{t\cdot X_j}e^{-tn\mu} \\
&= e^{-n(x+\mu)t}\prod_{j=1}^{n}\mathbb{E}[e^{t\cdot X_j}] = e^{-n(x+\mu)t}\prod_{j=1}^{n} M_j(t)
\end{aligned}$$

$$\begin{aligned}
&\leq e^{-n(x+\mu)t}\prod_{j=1}^{n}(1-\mu_j(1-e^t)) \leq e^{-n(x+\mu)t}\Big(\frac{1}{n}\sum_{j=1}^{n}(1-\mu_j(1-e^t))\Big)^n = e^{-n(x+\mu)t}\Big(1-(1-e^t)\frac{1}{n}\sum_{j=1}^{n}\mu_j\Big)^n \\
&= e^{-n(x+\mu)t}\Big(1-\mu(1-e^t)\Big)^n = \Big(e^{-(x+\mu)t}(1-\mu(1-e^t))\Big)^n
\end{aligned}$$

Um das Minimum der rechte Seite für $t > 0$ zu bestimmen setzten wir

$$\frac{\partial\Big(e^{-(x+\mu)t}(1-\mu(1-e^t))\Big)^n}{\partial t} = 0 \tag{34}$$

Diese Gleichung kann vereinfacht werden indem Faktoren die niemals 0 werden weggelassen werden. Dadurch ergeben sich Schrittweise folgende zur Aussage (34) äquivalente Gleichungen:

$$\begin{aligned} n\cdot\Big(e^{-(x+\mu)t}(1-\mu(1-e^t))\Big)^{n-1}\Big(-(\mu+x)e^{-(x+\mu)t}(1-\mu(1-e^t))+e^{-(x+\mu)t}\mu e^t\Big) &= 0\\ e^{-(x+\mu)t}\Big(\mu e^t-(\mu+x)(1-\mu(1-e^t))\Big) &= 0\\ \mu e^t-(\mu+x)(1-\mu(1-e^t)) &= 0\\ \mu e^t-(\mu+x)+\mu(\mu+x)-\mu(\mu+x)e^t &= 0 \end{aligned}$$

Umstellen auf e^t ergibt

$$\begin{aligned} e^t &= \frac{(\mu+x)-\mu(\mu+x)}{\mu(1-(\mu+x))} \Leftrightarrow\\ t &= \ln\Big(\frac{(\mu+x)(1-\mu)}{\mu(1-(\mu+x))}\Big) =: t_0 \end{aligned}$$

Es ist $t_0 > 0$ da $t_0 = \ln\Big(\frac{x+\mu-\mu^2-\mu x}{\mu-\mu^2-\mu x}\Big) = \ln\Big(1+\frac{x}{\mu(1-(\mu+x))}\Big) > \ln 1 = 0$
Nun wird das Minimum t_0 in die ursprüngliche Gleichung eingesetzt:

$$\begin{aligned} \Big(e^{-(\mu+x)t_0}(1-\mu(1-e^{t_0}))\Big)^n &= \Big((e^{t_0})^{-(\mu+x)}(1-\mu(1-e^{t_0}))\Big)^n\\ &= \Big[\Big(\frac{(\mu+x)(1-\mu)}{\mu(1-(\mu+x))}\Big)^{-(\mu+x)}\Big(1-\mu(1-\Big(1+\frac{x}{\mu(1-(\mu+x))}\Big))\Big)\Big]^n\\ &= \Big[\Big(\frac{(\mu+x)}{\mu}\frac{(1-\mu)}{(1-(\mu+x))}\Big)^{-(\mu+x)}\Big(1-\mu\Big(\frac{-x}{\mu(1-(\mu+x))}\Big)\Big)\Big]^n\\ &= \Big[\Big(\frac{\mu}{\mu+x}\Big)^{(\mu+x)}\Big(\frac{(1-\mu)}{(1-(\mu+x))}\Big)^{-(\mu+x)}\Big(\frac{1-\mu}{(1-(\mu+x))}\Big)\Big]^n\\ &= \Big[\underbrace{\Big(\frac{\mu}{\mu+x}\Big)^{(\mu+x)}\Big(\frac{(1-\mu)}{(1-(\mu+x))}\Big)^{1-(\mu+x)}}_{***}\Big]^n = \exp\Big\{n\cdot\ln\Big(***\Big)\Big\}\\ &= \exp\Big\{-n\cdot\ln\Big(\frac{1}{***}\Big)\Big\} = \exp\Big\{-nx^2\cdot\Big(\frac{1}{x^2}\ln\Big(\frac{1}{***}\Big)\Big)\Big\}\\ &= \exp\Big\{-nx^2\cdot\Big(\underbrace{\frac{1}{x^2}\Big[(\mu+x)\ln\Big(\frac{\mu+x}{\mu}\Big)+(1-(\mu+x))\ln\Big(\frac{(1-(\mu+x))}{1-\mu}\Big)\Big]}_{=:G(x,\mu)}\Big)\Big\} \end{aligned}$$

Als nächsten Schritt wird $G(x,\mu)$ nach x minimiert und das resultierende Minimum wird bezeichnet mit $g(\mu)$

$$\frac{\partial G(x,\mu)}{\partial x} =$$

$$\begin{aligned} &= \frac{-2}{x^3}\Big[(\mu+x)\ln\Big(\frac{\mu+x}{\mu}\Big)+(1-(\mu+x))\ln\Big(\frac{(1-(\mu+x))}{1-\mu}\Big)\Big]+\frac{1}{x^2}\Big[\ln\Big(\frac{\mu+x}{\mu}\Big)-\ln\Big(\frac{(1-(\mu+x))}{1-\mu}\Big)\Big]\\ &= \frac{1}{x^2}\Big[\ln\Big(\frac{\mu+x}{\mu}\Big)-\ln\Big(\frac{(1-(\mu+x))}{1-\mu}\Big)-\frac{2}{x}(\mu+x)\ln\Big(\frac{\mu+x}{\mu}\Big)-\frac{2(1-(\mu+x))}{x}\ln\Big(\frac{(1-(\mu+x))}{1-\mu}\Big)\Big]\\ &= \frac{1}{x^2}\Big[\ln\Big(\frac{\mu+x}{\mu}\Big)\Big(1-\frac{2(\mu+x)}{x}\Big)-\ln\Big(\frac{(1-\mu-x)}{1-\mu}\Big)\Big(1+\frac{2(1-\mu-x)}{x}\Big)\Big] \end{aligned}$$

Durch einsetzen folgender Gleichheiten $\ln\left(\frac{\mu+x}{\mu}\right) = -\ln\left(\frac{\mu}{\mu+x}\right) = -\ln\left(1 - \frac{x}{\mu+x}\right)$ und $\ln\left(\frac{(1-\mu-x)}{1-\mu}\right) = \ln\left(1 - \frac{x}{1-\mu}\right)$ lässt sich $\frac{\partial G(x,\mu)}{\partial x}$ weiter vereinfachen:

$$
\begin{aligned}
&= \frac{1}{x^2}\Big[-\ln\Big(1 - \frac{x}{\mu+x}\Big)\Big(1 - \frac{2(\mu+x)}{x}\Big) - \ln\Big(1 - \frac{x}{1-\mu}\Big)\Big(1 + \frac{2(1-\mu)}{x} - 2\Big)\Big] \\
&= \frac{1}{x^2}\Big[\ln\Big(1 - \frac{x}{1-\mu}\Big)\Big(1 - \frac{2\cdot(1-\mu)}{x}\Big) - \ln\Big(1 - \frac{x}{\mu+x}\Big)\Big(1 - \frac{2\cdot(\mu+x)}{x}\Big)\Big] \\
&= \frac{1}{x^2}\Big[H\Big(\frac{x}{1-\mu}\Big) - H\Big(\frac{x}{\mu+x}\Big)\Big]
\end{aligned}
$$

wobei aus $0 < x < 1-\mu$, $\mu \in (0,1)$ folgt

(i). $0 < \frac{x}{1-\mu} < 1$

(ii). $0 < x < 1-\mu$, durch Addition von μ ergibt sich
$0 < \mu < x + \mu < 1$ bzw.
$0 < x + \mu < 1$

(iii). $1 = \frac{\mu+x}{\mu+x} = \frac{\mu}{\mu+x} + \frac{x}{\mu+x} > \frac{x}{\mu+x} > 0$

Die Funktion H ist auf $(0,1)$ definiert als:

$$H(s) := \Big(1 - \frac{2}{s}\Big)\ln(1-s)$$

Es gilt:

$$H'(s) = \frac{2}{s^2}\ln(1-s) + \Big(1 - \frac{2}{s}\Big)\cdot\frac{-1}{1-s} = \frac{2(s-1)\ln(1-s) + s(s-2)}{(s-1)s^2}$$

Behauptung: $H(s)$ ist auf $(0,1)$ monoton wachsend, das ist gleichbedeutend mit $H'(s) \geq 0$ auf $(0,1)$.
Beweis: Der Nenner $(s-1)s^2$ ist offensichtlich < 0 auf $(0,1)$. Damit $H'(s) \geq 0$ ist, muss der Zähler auch < 0 sein.
Der Zähler $Z(s) = 2(s-1)\ln(1-s) + s(s-2)$ erfüllt $Z(0) = 0$ und
$Z'(s) = 2\ln(1-s) + 2 + 2s - 2 = 2(s + \ln(1-s))$

Auf ganz $\mathbb{R}$ gilt $e^x \geq x + 1 \Leftrightarrow x \geq \ln(x+1) \Rightarrow -s \geq \ln(1-s)$
Daraus ergibt sich

$$Z'(s) \leq 2(s-s) = 0$$

und damit ist $Z(s) \leq 0$ auf $(0,1)$ also ist $H'(s) \geq 0$ auf $(0,1)$, sprich es wurde gezeigt, dass $H(s)$ auf $(0,1)$ monoton steigend ist.

Zurückkommend zur Berechnung von $\frac{\partial G(x,\mu)}{\partial x}$ folgt dadurch:

$$\frac{\partial G(x,\mu)}{\partial x} > 0 \text{ genau dann, wenn } \frac{x}{1-\mu} > \frac{x}{\mu+x} \text{ bzw. } x > 1 - 2\mu \text{ ist}$$

Für $1 - 2\mu > 0$ wählen wir also $x = 1 - 2\mu$, andernfalls für $1 - 2\mu \leq 0$ wird $G(x,\mu)$ minimal durch die Wahl von Limes $x \to 0$
Einsetzen in $G(x,\mu)$ ergibt:

1. Fall: $(1 - 2\mu) > 0$

aus $1 > 2\mu$, d.h. $\frac{1}{2} > \mu$ und $\mu \in (0,1)$ folgt $0 < \mu < \frac{1}{2}$, damit:

$$\begin{aligned}
g(\mu) &= \frac{1}{(1-2\mu)^2}\Big[(\mu+1-2\mu)\ln\Big(\frac{\mu+1-2\mu}{\mu}\Big) + (1-\mu-1+2\mu)\ln\Big(\frac{1-\mu-1+2\mu}{1-\mu}\Big)\Big] \\
&= \frac{1}{(1-2\mu)^2}\Big[(1-\mu)\ln\Big(\frac{1-\mu}{\mu}\Big) + \mu\ln\Big(\frac{\mu}{1-\mu}\Big)\Big] \\
&= \frac{1}{(1-2\mu)^2}\Big[(1-\mu)\ln\Big(\frac{1-\mu}{\mu}\Big) - \mu\ln\Big(\frac{1-\mu}{\mu}\Big)\Big] = \frac{1}{(1-2\mu)^2}\Big[(1-2\mu)\ln\Big(\frac{1-\mu}{\mu}\Big)\Big] \\
&= \frac{1}{(1-2\mu)} \cdot \ln\Big(\frac{1-\mu}{\mu}\Big)
\end{aligned}$$

2. Fall: $(1-2\mu) \leq 0$

aus $1 \leq 2\mu$, d.h. $\frac{1}{2} \leq \mu$ und $\mu \in (0,1)$ folgt $\frac{1}{2} \leq \mu < 1$, damit:

$$\begin{aligned}
g(\mu) &= \lim_{x\to 0} G(x,\mu) = \lim_{x\to 0}\frac{1}{x^2}\Big[(\mu+x)\ln\Big(\frac{\mu+x}{\mu}\Big) + (1-(\mu+x))\ln\Big(\frac{(1-(\mu+x))}{1-\mu}\Big)\Big] \\
&= \lim_{x\to 0}\frac{\Big[\ln\Big(\frac{\mu+x}{\mu}\Big) + \frac{\mu+x}{\mu}\frac{1}{\mu}(\mu+x) - \ln\Big(\frac{(1-\mu-x)}{1-\mu}\Big) - (1-\mu-x)\frac{(1-\mu-x)}{1-\mu}\frac{1}{1-\mu}\Big]}{2x} \\
&= \lim_{x\to 0}\frac{\ln\Big(\frac{\mu+x}{\mu}\Big) - \ln\Big(\frac{(1-\mu-x)}{1-\mu}\Big)}{2x} = \lim_{x\to 0}\frac{\frac{\mu}{\mu+x}\frac{1}{\mu} + \frac{1-\mu}{1-\mu-x}\frac{1}{1-\mu}}{2} = \lim_{x\to 0}\frac{\frac{1}{\mu+x} + \frac{1}{1-\mu-x}}{2} \\
&= \lim_{x\to 0}\frac{1}{2\Big(\mu+x-\mu^2-2\mu x-x^2\Big)} = \frac{1}{2\Big(\mu(1-\mu)\Big)}
\end{aligned}$$

Also gilt:

$$g(\mu) = \begin{cases} \frac{1}{1-2\mu}\log\frac{1-\mu}{\mu} & \text{, für } 0 < \mu < \frac{1}{2} \\ \frac{1}{2\mu(1-\mu)} & \text{, für } \frac{1}{2} \leq \mu < 1 \end{cases}$$

Nun folgt noch die Ungleichung zu zeigen dass für $\mu \in (0,1)$ gilt

$$g(\mu) \geq g(\frac{1}{2}) = 2 \tag{35}$$

Beweis von Ungleichung (35):

$$\begin{aligned}
\lim_{\mu\to\frac{1}{2}-} g(\mu) &= \lim_{\mu\to\frac{1}{2}-}\Big(\frac{1}{1-2\mu}\ln\Big(\frac{1-\mu}{\mu}\Big)\Big) = \cdots \text{ de l' H } = \lim_{\mu\to\frac{1}{2}-}\frac{\frac{\mu}{1-\mu}\frac{-\mu-(1-\mu)}{\mu^2}}{-2} \\
&= \lim_{\mu\to\frac{1}{2}-}\frac{1}{2\mu(1-\mu)} = 2
\end{aligned}$$

(i) Für $0 < \mu < \frac{1}{2}$ ist $g(\mu) \geq 2$ äquivalent zu $\ln(1-\mu) - \ln(\mu) \geq 2(1-2\mu)$ $\Leftrightarrow \ln(1-\mu) - \ln(\mu) + 4\mu - 2 \geq 0$. Für $\mu = \frac{1}{2}$ ist die linke Seite dieser Ungleichung $= 0$. Diese Funktion abgeleitet ergibt:

$$\frac{-1}{1-\mu} - \frac{1}{\mu} + 4 = \frac{\mu-(\mu-1)}{\mu(\mu-1)} + 4 = \frac{1}{\mu(\mu-1)} + 4 \stackrel{!}{=} 0 \Leftrightarrow \frac{1}{\mu^2-\mu} = -4 \Leftrightarrow -(2\mu-1)^2 = 0$$

Daraus folgt, dass diese Funktion für $0 < \mu < \frac{1}{2}$ bzw. $g(\mu) \geq 2$ für $\mu \in (0, \frac{1}{2})$ monoton fällt.
(ii) Für $\frac{1}{2} \leq \mu < 1$ betrachte $g(\mu) = \frac{1}{2\mu(1-\mu)} = \frac{1}{-2\mu^2+2\mu}$. Der Nenner $-2\mu^2 + 2\mu$ abgeleitet ergibt $-4\mu + 2$, ist auf $[\frac{1}{2}, 1)$ monoton fallend. Daraus folgt $g(\mu)$ ist auf $[\frac{1}{2}, 1)$ monoton steigend $\Rightarrow g(\mu) \geq g(\frac{1}{2}) = 2$ auf $[\frac{1}{2}, 1)$. Insgesamt folgt, dass Ungleichung (35) auf dem gesamten Intervall $\mu \in (0,1)$ gilt.
Schlussendlich folgt die zu zeigende Aussage:

$$\mathbb{P}(\max_{1\leq j\leq n}(S_j - \mathbb{E}S_j) \geq nx) \leq e^{-g(\mu)nx^2} \leq e^{-2nx^2}$$

$\square$

Literatur

[1] Norbert Kusolitsch. *Maß- und Wahrscheinlichkeitstheorie; Eine Einführung*. Springer-Lehrbuch. Berlin, Heidelberg: Springer Berlin Heidelberg, Berlin, Heidelberg, 2014. ISBN: 9783642453861.

[2] Zhengyan Lin und Zhidong Bai. *Probability Inequalities*. Beijing: Science Press, 2010. ISBN: 9787030255624.

[3] Valentin V. Petrov. *Limit theorems of probability theory; Sequences of Independent Random Variables*. Oxford studies in probability ; 4. Oxford Science Publications: Clarendon Press, 1995. ISBN: 9780198534990.