





**Technische Universität Ilmenau**

Fakultät für Mathematik und Naturwissenschaften

Fachgebiet Stochastik

# **Nicht-asymptotische Konfidenzbereiche für extrinsische Erwartungswerte zirkulärer Daten**

Bachelorarbeit zur Erlangung des akademischen Grades Bachelor of Science

**Johannes Wieditz**

Verantwortlicher Hochschullehrer: Jun.-Prof. Dr. rer. nat. Thomas Hotz

Ilmenau, 12. September 2014

## Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich während der Anfertigung dieser Arbeit, insbesondere durch das Korrekturlesen, unterstützt haben. Ein besonderer Dank geht an Jun.-Prof. Dr. rer. nat. Thomas Hotz, der mir stets mit vielen hilfreichen Ratschlägen zur Seite stand und diese Arbeit somit ermöglichte.

# Zusammenfassung

In dieser Arbeit werden nicht-asymptotische Konfidenzbereiche für den extrinsischen Erwartungswert von Zufallsvariablen mit Werten auf dem Einheitskreis konstruiert. Dabei wird neben der Hoeffdingschen Ungleichung, der Tschebyscheffschen Ungleichung und Hoeffdings Theorem 3, eine verschärfte Version der Hoeffdingschen Ungleichung für die Konstruktion verwendet. Letztere wird dazu hergeleitet und bewiesen. Die Konfidenzbereiche werden dann für zwei Anwendungsbeispiele mit unterschiedlich stark konzentrierten Daten konstruiert und sowohl untereinander als auch mit asymptotischen Konfidenzbereichen verglichen.

## Abstract

In this Bachelor Thesis we will design non-asymptotic confidence regions for the extrinsic population mean of random variables with values on the unit circle. To this end we will consider the Hoeffding's and Chebychev's inequality as well as Hoeffdings Theorem 3 as well as a tightened version of the Hoeffding's inequality. The latter is first derived and proved, then used for the construction of confidence regions. These are determined for two datasets of different concentration. Finally, we compare the confidence regions among each other but also with asymptotic confidence regions.

---

# Inhaltsverzeichnis

<b>Notationen</b>	<b>III</b>
<b>1 Einleitung</b>	<b>1</b>
<b>2 Maßkonzentrationsungleichungen für beschränkte Zufallsgrößen</b>	<b>3</b>
2.1 Die Hoeffdingsche Ungleichung . . . . .	3
2.2 Ungleichungen zweiter Ordnung . . . . .	7
2.2.1 Die Tschebyscheffsche Ungleichung . . . . .	7
2.2.2 Hoeffdings Theorem 3 . . . . .	7
2.2.3 Hoeffdingsche Ungleichung zweiter Ordnung . . . . .	8
2.3 Vergleich der verschiedenen Ungleichungen . . . . .	13
<b>3 Die Konstruktion der Konfidenzbereiche</b>	<b>16</b>
3.1 Konfidenzbereiche mit der Hoeffdingschen Ungleichung . . . . .	17
3.1.1 Konstruktion – Schritt 1 . . . . .	18
3.1.2 Konstruktion – Schritt 2 . . . . .	22
3.2 Konfidenzbereiche mit Varianz . . . . .	27
3.2.1 Konfidenzbereiche mit der Tschebyscheffschen Ungleichung . . . . .	27
3.2.2 Konfidenzbereiche mit Hoeffdings Theorem 3 . . . . .	28
3.2.3 Konfidenzbereiche mit Hoeffdingscher Ungleichung zweiter Ordnung . . . . .	34
3.3 Asymptotische Konfidenzbereiche . . . . .	35
<b>4 Anwendung</b>	<b>38</b>
4.1 Strömungsrichtungen in Schrägschichtungen . . . . .	38
4.2 Bewegungsrichtungen von Ameisen . . . . .	41
4.3 Vergleich der Konfidenzbereiche . . . . .	43
<b>5 Fazit und Ausblick</b>	<b>44</b>
<b>A Anhang</b>	<b>46</b>
<b>Literaturverzeichnis</b>	<b>52</b>

---

# Notationen

## Mengen und Vektorräume

$\mathbb{N}$ .....	die natürlichen Zahlen $\{1, 2, 3, \dots\}$
$\mathbb{R}$ .....	die reellen Zahlen
$\mathbb{R}^n$ .....	$\{(v_1, \dots, v_n) \mid v_i \in \mathbb{R}\}$ der Vektorraum aller reellwertigen $n$ -Tupel, $n \in \mathbb{N}$
$\text{span } v$ .....	der von $v$ aufgespannte lineare Unterraum
$U^\perp$ .....	das orthogonale Komplement eines linearen Unterraums $U$
$(a, b) / [a, b]$ .....	das offene/abgeschlossene Intervall von $a$ bis $b$
$\mathbb{B}_\varepsilon(x) / \bar{\mathbb{B}}_\varepsilon(x)$ .....	der offene/abgeschlossene Ball um $x$ mit Radius $\varepsilon$
$S^1$ .....	$\{v \in \mathbb{R}^2 \mid \ v\  = 1\}$ der Einheitskreis
$\mathbb{1}_A$ .....	die charakteristische Funktion der Menge $A$

## Vektoren in der Ebene

$\langle \cdot, \cdot \rangle$ .....	das Standardskalarprodukt im $\mathbb{R}^2$
$\ v\ $ .....	die euklidische Norm eines Vektors $v \in \mathbb{R}^2$
$v^\top$ .....	der transponierte Vektor von $v \in \mathbb{R}^2$
$\angle(v_1, v_2)$ .....	der zwischen $v_1$ und $v_2$ eingeschlossene Winkel, $v_1, v_2 \in \mathbb{R}^2$

## Stochastische Begriffe

$\mathbb{P}$ .....	ein Wahrscheinlichkeitsmaß
$\mathbb{E}X$ .....	der Erwartungswert einer Zufallsvariable $X$
$\text{Var } X$ .....	die Varianz einer Zufallsvariable $X$
$\mathcal{N}(\mu, \sigma^2)$ .....	Normalverteilung mit Parametern $\mu$ und $\sigma^2$
$X_n \xrightarrow{\text{f.s.}} X$ .....	$X_n$ konvergiert fast sicher gegen $X$
$F_n \xrightarrow{\mathcal{D}} F$ .....	$F_n$ konvergiert in Verteilung gegen $F$

# 1 Einleitung

In vielen Bereichen der Naturwissenschaften treten in natürlicher Weise zirkuläre Daten, d.h. Daten in Form von Richtungen oder Achsen, auf. Insbesondere wenn es um Himmelsrichtungen, wie etwa den Flugrichtungen von Vögeln oder den Ausrichtungen von Sedimentschichten in Gesteinen geht, kann man dabei ohne Einschränkung die Daten auf dem Einheitskreis verteilt sehen.

Möchte man nun Erwartungswerte bestimmen oder Konfidenzbereiche konstruieren, tritt das Problem auf, dass der Einheitskreis  $S^1 = \{v \in \mathbb{R}^2 \mid \|v\| = 1\}$  nur ein topologischer Raum ohne Vektorraumstruktur ist. Insbesondere ist zunächst die Definition eines Erwartungswertes einer  $S^1$ -wertigen Zufallsvariable  $Z$  nicht unmittelbar klar. Man bettet daher die Kreislinie als Untermannigfaltigkeit in den Vektorraum  $\mathbb{R}^2$  ein. Dann kann man sich der Werkzeuge, die für die Ebene bekannt sind, bedienen und anschließend Projektionen auf den Kreis betrachten. In diesem Zusammenhang werden wir den *extrinsischen Erwartungswert* betrachten, welcher ein Punkt  $\varepsilon \in S^1$  ist, der den Abstand vom Erwartungswert  $\mathbb{E}Z$  im Vektorraum  $\mathbb{R}^2$  bzgl. der euklidischen Norm minimiert. Dieser liefert uns eine intuitive Definition des Erwartungswertes auf dem Einheitskreis. Ist  $\mathbb{E}Z \neq 0$ , so ist  $\varepsilon$  gegeben durch die Orthogonalprojektion von  $\mathbb{E}Z$  auf den Kreis. In diesem Fall gilt  $\varepsilon = \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|}$ . Andernfalls ist  $\varepsilon = S^1$ . Ziel dieser Arbeit ist es, nicht-asymptotische Konfidenzbereiche für den extrinsischen Erwartungswert zu konstruieren.

Hauptsächlich werden wir Abschätzungen mit Hilfe der Hoeffdingschen Ungleichung treffen, einer Ungleichung, welche für beschränkte Zufallsgrößen eine exponentiell gute Abschätzung für die positive Abweichung des Mittelwertes vom Erwartungswert liefert. Diese leiten wir in Abschnitt 2.1 her. Darüber hinaus betrachten wir in Abschnitt 2.2 drei weitere Ungleichungen für beschränkte Zufallsgrößen, welche die Varianz berücksichtigen. Insbesondere leiten wir eine verschärfte Variante der Hoeffdingschen Ungleichung her und beweisen diese. Abschließend werden wir einen Vergleich der Ungleichungen durchführen. Dieser wird uns darauf schließen lassen, dass sich mit den Ungleichungen aus Abschnitt 2.2 ggf. bessere Konfidenzbereiche konstruieren lassen, als lediglich unter Benutzung der Hoeffdingschen Ungleichung.

In Kapitel 3 konstruieren wir Konfidenzbereiche für den extrinsischen Erwartungswert. Dafür werden wir neben der Hoeffdingschen Ungleichung die drei Ungleichungen aus Abschnitt 2.2 benutzen, wobei die Tschebyscheffsche Ungleichung nur zu Vergleichszwecken dienen soll.

Genauer werden zunächst unter der Hypothese  $\mathbb{E}Z = q$  Mengen  $M(q)$  konstruiert, welche den Mittelwert  $\bar{Z}_n = \sum_{i=1}^n Z_i$  (im  $\mathbb{R}^2$  gebildet) mit großer Wahrscheinlichkeit enthalten. Mit Hilfe dieser Mengen und  $\bar{Z}_n$  betrachten wir dann die Menge aller  $q$ , für die  $\bar{Z}_n \in M(q)$  gilt, d.h. den Annahmehereich eines Tests der Hypothese  $\mathbb{E}Z = q$ . Diese Konstruktion werden wir auf zwei verschiedene Weisen durchführen – zuerst ohne und anschließend unter Einbeziehung der Varianz. Die Projektion des Annahmehereiches auf den Kreis liefert uns dann einen  $(1 - \alpha)$ -Konfidenzbereich für den extrinsischen Erwartungswert.

Darüber hinaus werden in Abschnitt 3.3 asymptotische Konfidenzbereiche konstruiert, welche wir später mit den nicht-asymptotischen Konfidenzbereichen vergleichen werden.

Alle konstruierten Konfidenzbereiche werden wir in Kapitel 4 für zwei Anwendungsbeispiele bestimmen. Diese werden dann untereinander, sowie mit den asymptotischen Konfidenzbereichen verglichen.

Kapitel 5 soll schließlich die Ergebnisse dieser Arbeit zusammenfassen und darüber hinaus Anregungen für weitere Konstruktionen bieten.

## 2 Maßkonzentrationsungleichungen für beschränkte Zufallsgrößen

In diesem Kapitel werden wir alle Ungleichungen vorstellen, die wir zur Konstruktion von Konfidenzbereichen in dieser Arbeit benötigten. Dabei beschäftigt sich Abschnitt 2.1 mit der Hoeffdingschen Ungleichung, welche wir für jeden Konstruktionsschritt benötigen werden. In Abschnitt 2.2 betrachten wir Ungleichungen mit zweiten Momenten, genauer die Tschebyscheffsche Ungleichung, Hoeffdings Theorem 3 und die Hoeffdingsche Ungleichung zweiter Ordnung. Diese werden wir in Abschnitt 3.2 zur Konstruktion der Konfidenzbereiche verwenden. Abschließend vergleichen wir letztere drei Ungleichungen aus Abschnitt 2.2 miteinander, um herauszufinden, ob es eine „beste“ Abschätzung gibt.

### 2.1 Die Hoeffdingsche Ungleichung

Dieser Abschnitt wird sich mit dem Ausgangspunkt dieser Arbeit befassen – der Hoeffdingschen Ungleichung. Der Beweis wurde in grundlegenden Zügen aus [Hoe63, Theorem 1] übernommen, wird jedoch in etwas allgemeinerer Form geführt. Im Fokus stehen dabei unabhängige, beschränkte Zufallsgrößen  $X_1, X_2, \dots, X_n$ , die (zumindest fast sicher) nur Werte aus dem kompakten Intervall  $[a, b]$  mit  $-\infty < a < b < \infty$  annehmen. Wir wollen zunächst einige Notationen einführen.

**Definition 2.1** Es seien  $X_1, X_2, \dots, X_n$  unabhängige Zufallsgrößen mit endlichen ersten und zweiten Momenten auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathfrak{A}, \mathbb{P})$ . Wir schreiben dann

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i, \quad \mu := \mathbb{E}\bar{X}_n = \frac{1}{n} \sum_{i=1}^n \mathbb{E}X_i, \quad \sigma^2 := n \operatorname{Var} \bar{X}_n = \frac{1}{n} \sum_{i=1}^n \operatorname{Var} X_i.$$

Bevor wir zur Aussage der Hoeffdingschen Ungleichung und deren Beweis kommen, müssen wir zunächst einige vorbereitende Aussagen treffen.

**Bemerkung 2.2** Es seien  $-\infty < a < b < \infty$  und  $X_1, X_2, \dots, X_n$  reelle Zufallsgrößen mit  $a \leq X_i \leq b$  für  $i = 1, 2, \dots, n$ . Dann gilt  $a \leq \mu \leq b$ . Gilt  $\mu = a$ , so folgt  $\mathbb{E}X_i = a$  und somit  $X_i = a$  fast sicher für  $i = 1, 2, \dots, n$ . Die Zufallsgrößen sind in diesem Fall einpunktverteilt auf  $a$  und es gilt  $\mathbb{P}(\bar{X}_n - \mu \geq t) = 0$  für jedes  $t > 0$ . Eine entsprechende Aussage gilt, falls  $\mu = b$  ist.

**Bemerkung 2.3** Es seien  $X_1, X_2, \dots, X_n$  unabhängige Zufallsgrößen und  $S_n = n\bar{X}_n$ . Dann ist

$$\mathbb{P}(S_n - \mathbb{E}S_n \geq nt) = \mathbb{E}\mathbb{1}_{\{S_n - \mathbb{E}S_n - nt \geq 0\}}.$$

Für jede Konstante  $h > 0$  ist  $\exp(h(S_n - \mathbb{E}S_n - nt))$  eine Majorante für  $\mathbb{1}_{\{S_n - \mathbb{E}S_n - nt \geq 0\}}$ . Folglich ist

$$\mathbb{P}(\bar{X}_n - \mu \geq t) = \mathbb{P}(S_n - \mathbb{E}S_n \geq nt) \leq \mathbb{E}e^{h(S_n - \mathbb{E}S_n - nt)}.$$

Aus der Unabhängigkeit der  $X_i$  folgt

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \mathbb{E}e^{h(S_n - \mathbb{E}S_n - nt)} = e^{-hnt} \prod_{i=1}^n \mathbb{E}e^{h(X_i - \mathbb{E}X_i)}. \quad (2.1)$$

Damit wir  $\mathbb{P}(\bar{X}_n - \mu \geq t)$  abschätzen können, gilt es eine obere Schranke für den Erwartungswert (2.1) in Abhängigkeit von  $h$  zu finden. Diese liefert das folgende Lemma.

**Lemma 2.4** *Es seien  $-\infty < a < b < \infty$  und es sei  $X$  eine reelle Zufallsgröße mit  $a \leq X \leq b$ . Dann gilt für jede reelle Zahl  $h \in \mathbb{R}$  die Ungleichung*

$$\mathbb{E}e^{hX} \leq \frac{b - \mathbb{E}X}{b - a} e^{ha} + \frac{\mathbb{E}X - a}{b - a} e^{hb}.$$

**Beweis** Da für alle  $h \in \mathbb{R}$  die Funktion  $t \mapsto e^{ht}$  konvex ist, gilt die Ungleichung

$$e^{ht} \leq \frac{b - t}{b - a} e^{ha} + \frac{t - a}{b - a} e^{hb}$$

für alle  $a \leq t \leq b$ . Aus der Monotonie des Erwartungswertes folgt dann die Behauptung.  $\blacksquare$

Wir werden nun die Aussage der Hoeffdingschen Ungleichung formulieren und beweisen.

**Theorem 2.5** (Hoeffdingsche Ungleichung) *Es seien  $X_1, X_2, \dots, X_n$  unabhängige Zufallsgrößen mit  $a \leq X_i \leq b$  für  $i = 1, 2, \dots, n$  und es sei  $\mu = \mathbb{E}\bar{X}_n$ . Dann gilt für  $t \in (0, b - \mu)$*

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ \left( \frac{\mu - a}{\mu - a + t} \right)^{\mu - a + t} \left( \frac{b - \mu}{b - \mu - t} \right)^{b - \mu - t} \right]^{\frac{n}{b - a}}. \quad (2.2)$$

**Beweis** Es sei  $h > 0$ . Aufgrund von Bemerkung 2.2 können wir o.B.d.A.  $a < \mu < b$  annehmen. Mit (2.1) folgt

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq e^{-hnt - hm\mu} \prod_{i=1}^n \mathbb{E}e^{hX_i}.$$

Mit Lemma 2.4 können wir dann die Abschätzung

$$\prod_{i=1}^n \mathbb{E}e^{hX_i} \leq \frac{1}{(b - a)^n} \prod_{i=1}^n \left( (b - \mathbb{E}X_i) e^{ha} + (\mathbb{E}X_i - a) e^{hb} \right)$$

treffen. Da das geometrische Mittel durch das arithmetische Mittel nach oben beschränkt ist, gilt dann

$$\begin{aligned} \left( \prod_{i=1}^n \left( (b - \mathbb{E}X_i)e^{ha} + (\mathbb{E}X_i - a)e^{hb} \right) \right)^{\frac{1}{n}} &\leq \frac{1}{n} \sum_{i=1}^n \left( (b - \mathbb{E}X_i)e^{ha} + (\mathbb{E}X_i - a)e^{hb} \right) \\ &= e^{ha} (b - \mu) + e^{hb} (\mu - a) \end{aligned} \quad (2.3)$$

und folglich

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ \frac{e^{-ht-h\mu}}{b-a} \left( e^{ha} (b - \mu) + e^{hb} (\mu - a) \right) \right]^n. \quad (2.4)$$

Um eine möglichst gute Abschätzung treffen zu können, minimieren wir nun die rechte Seite bzgl.  $h$ . Dazu betrachten wir die Funktion

$$f : (0, \infty) \rightarrow \mathbb{R}, \quad h \mapsto f(h) = \frac{e^{-h(t+\mu)}}{b-a} \left( e^{ha} (b - \mu) + e^{hb} (\mu - a) \right). \quad (2.5)$$

Die rechte Seite von (2.4) ist genau dann minimal, wenn  $f$  es ist. Für die Ableitung gilt

$$f'(h) = -\frac{e^{-h(t+\mu)}}{b-a} \left[ e^{ha} ((t + \mu)(b - \mu) - a(b - \mu)) + e^{hb} ((t + \mu)(\mu - a) - b(\mu - a)) \right].$$

Diese ist dann und nur dann gleich null, wenn

$$e^{ha} ((t + \mu - a)(b - \mu)) = -e^{hb} ((t + \mu - b)(\mu - a)),$$

d.h. falls

$$e^{h(b-a)} = -\frac{(t + \mu - a)(b - \mu)}{(t + \mu - b)(\mu - a)}$$

gilt. Die Nullstelle  $\bar{h}$  hat demnach die Gestalt

$$\bar{h} = \frac{1}{b-a} \ln \left( \frac{(b - \mu)(\mu - a + t)}{(\mu - a)(b - \mu - t)} \right) > 0.$$

In  $\bar{h}$  liegt dann ein Minimum vor. Wegen  $0 < t < b - \mu$  gilt nämlich

$$\begin{aligned} \lim_{h \rightarrow 0^+} f'(h) &= \frac{1}{b-a} \left( (a - t - \mu)(b - \mu) + (b - \mu - t)(\mu - a) \right) \\ &\leq \frac{b - \mu}{b - a} (a - t - \mu + \mu - a) \\ &= -\frac{t(b - \mu)}{b - a} < 0. \end{aligned}$$

Andererseits ist wegen  $a - t - \mu < 0$  und  $b - t - \mu > 0$

$$\lim_{h \rightarrow \infty} f'(h) = \lim_{h \rightarrow \infty} \frac{1}{b-a} \left[ e^{h(a-t-\mu)} (a-t-\mu)(b-\mu) + e^{h(b-\mu-t)} (b-t-\mu)(\mu-a) \right] = \infty.$$

Die Ableitung von  $f$  hat demnach in  $\bar{h}$  einen Vorzeichenwechsel von Minus nach Plus.  $\bar{h}$  ist also Minimalstelle.

Setzen wir nun  $\bar{h}$  in (2.5) ein, so erhalten wir

$$\begin{aligned} f(\bar{h}) &= \frac{1}{b-a} \left[ \left( \frac{(b-\mu)(\mu-a+t)}{(\mu-a)(b-\mu-t)} \right)^{\frac{a-t-\mu}{b-a}} (b-\mu) + \left( \frac{(b-\mu)(\mu-a+t)}{(\mu-a)(b-\mu-t)} \right)^{\frac{b-t-\mu}{b-a}} (\mu-a) \right] \\ &= \frac{1}{b-a} \left[ \left( \frac{(\mu-a+t)}{(\mu-a)(b-\mu-t)} \right)^{\frac{a-t-\mu}{b-a}} (b-\mu)^{\frac{b-\mu-t}{b-a}} + \left( \frac{(b-\mu)(\mu-a+t)}{b-\mu-t} \right)^{\frac{b-t-\mu}{b-a}} (\mu-a)^{\frac{\mu+t-a}{b-a}} \right] \\ &= \frac{1}{b-a} \left( \frac{\mu-a+t}{\mu-a} \right)^{\frac{a-t-\mu}{b-a}} \left[ \frac{(b-\mu)^{\frac{b-\mu-t}{b-a}}}{(b-\mu-t)^{\frac{a-t-\mu}{b-a}}} + \left( \frac{b-\mu}{b-\mu-t} \right)^{\frac{b-t-\mu}{b-a}} (\mu-a+t) \right] \\ &= \frac{1}{b-a} \left( \frac{\mu-a}{\mu-a+t} \right)^{\frac{\mu-a+t}{b-a}} \left( \frac{b-\mu}{b-\mu-t} \right)^{\frac{b-\mu-t}{b-a}} [(b-\mu-t) + (\mu-a+t)] \\ &= \left[ \left( \frac{\mu-a}{\mu-a+t} \right)^{\mu-a+t} \left( \frac{b-\mu}{b-\mu-t} \right)^{b-\mu-t} \right]^{\frac{1}{b-a}}. \end{aligned}$$

Mit (2.4) erhalten wir die Ungleichung (2.2). ■

**Folgerung 2.6** *Es seien  $X_1, X_2, \dots, X_n$  unabhängige Zufallsgrößen mit  $-1 \leq X_i \leq 1$  für  $i = 1, 2, \dots, n$  und es sei  $\mu = \mathbb{E}\bar{X}_n$ . Dann gilt für  $0 < t < 1 - \mu$ :*

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ \left( \frac{\mu+1}{\mu+1+t} \right)^{\mu+1+t} \left( \frac{1-\mu}{1-\mu-t} \right)^{1-\mu-t} \right]^{\frac{n}{2}}. \quad (2.6)$$

**Bemerkung 2.7** In [Hoe63, Theorem 1] wird zudem gezeigt, dass die schwächere Abschätzung

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq e^{-\frac{n}{2} t^2} \quad (2.7)$$

gilt, wenn die Voraussetzungen aus Folgerung 2.6 erfüllt sind. Auf den Beweis werden wir an dieser Stelle jedoch nicht näher eingehen. Die Ungleichung (2.7) ist etwas leichter als (2.6) zu handhaben und wird sich für gewisse Abschätzungen als ausreichend erweisen.

**Bemerkung 2.8** Für den Spezialfall  $|\mu| = 1$ , d.h.  $\mu = \pm 1$ , liefert Folgerung 2.6 keine Abschätzung. In diesem Fall sind die Zufallsgrößen  $X_1, \dots, X_n$  einpunktverteilt auf  $\mu$  und es gilt  $\bar{X}_n = \mu$  fast sicher, wie in Bemerkung 2.2 bereits erwähnt wurde.

## 2.2 Ungleichungen zweiter Ordnung

Wünschenswert wäre es, wenn wir für unsere späteren Konstruktionsschritte auch noch die Varianz der Zufallsgröße  $Z$  ausnutzen können, um ggf. noch „bessere“, im Sinne von kleinere Konfidenzbereiche zum selben Niveau zu konstruieren. Dabei verstehen wir unter der *Varianz* einer vektorwertigen Zufallsvariable  $Z$  den Erwartungswert

$$\text{Var } Z := \mathbb{E} \|Z - \mathbb{E}Z\|^2,$$

sofern dieser existiert. Nimmt  $Z$  nur Werte auf dem Einheitskreis  $S^1$  an, so gilt

$$\text{Var } Z = \mathbb{E} \|Z - \mathbb{E}Z\|^2 = \mathbb{E} \|Z\|^2 - \|\mathbb{E}Z\|^2 = 1 - \|\mathbb{E}Z\|^2,$$

wobei  $\mathbb{E}Z$  der Erwartungswert im  $\mathbb{R}^2$  ist. Wir stellen im Folgenden drei Ungleichungen vor, welche wir anschließend in Kapitel 3 zur Konstruktion von Konfidenzbereichen verwenden werden.

### 2.2.1 Die Tschebyscheffsche Ungleichung

Die einfachste Möglichkeit, die Varianz mit einzubeziehen, bietet eine multivariate Version der Tschebyscheffschen Ungleichung. Diese hat für  $\varepsilon > 0$  die Gestalt

$$\mathbb{P}(\|Z - \mathbb{E}Z\| \geq \varepsilon) \leq \frac{\mathbb{E} \|Z - \mathbb{E}Z\|^2}{\varepsilon^2},$$

sofern das zweite Moment der Zufallsvariable  $Z$  existiert. Betrachten wir unabhängig und identisch verteilte Zufallsvariablen  $Z, Z_1, Z_2, \dots, Z_n$  mit Werten auf dem Einheitskreis, so gilt für  $\alpha \in (0, 1)$  und  $\bar{Z}_n = \frac{1}{n} \sum_{i=1}^n Z_i$  die Ungleichung

$$\mathbb{P}\left(\|\bar{Z}_n - \mathbb{E}Z\|^2 \geq \frac{1 - \|\mathbb{E}Z\|^2}{n\alpha}\right) \leq \alpha, \quad (2.8)$$

siehe [HK14, Theorem 1].

### 2.2.2 Hoeffdings Theorem 3

Die Tschebyscheffsche Ungleichung nutzt jedoch nicht aus, dass die Zufallsvariablen, welche wir später betrachten werden, beschränkt sind und ist darüber hinaus meist unscharf. Für eine bessere Abschätzung finden wir in [Hoe63] die folgende Aussage, welche für zentrierte, nach oben beschränkte Zufallsvariablen gilt.

**Theorem 2.9** (Hoeffdings Theorem 3) *Es seien  $X_1, X_2, \dots, X_n$  unabhängige, zentrierte Zufallsgrößen mit  $X_i \leq b$  für  $i = 1, 2, \dots, n$ . Dann gilt für  $t \in (0, b)$ :*

$$\mathbb{P}(\bar{X}_n \geq t) \leq \left[ \left(1 + \frac{bt}{\sigma^2}\right)^{-\frac{\sigma^2+bt}{b^2+\sigma^2}} \left(1 - \frac{t}{b}\right)^{-\frac{b^2-tb}{b^2+\sigma^2}} \right]^n.$$

**Beweis** Zum Beweis verweisen wir auf [Hoe63, Theorem 3]. ■

Für  $b = 1$  ergibt sich der für uns interessante Spezialfall:

**Folgerung 2.10** *Sind  $X_1, X_2, \dots, X_n$  unabhängige Zufallsgrößen mit  $X_i \leq 1$  und  $\mathbb{E}X_i = 0$  für  $i = 1, 2, \dots, n$ , so gilt für alle  $0 < t < 1$ :*

$$\mathbb{P}(\bar{X}_n \geq t) \leq \left[ \left(1 + \frac{t}{\sigma^2}\right)^{-\frac{\sigma^2+t}{1+\sigma^2}} (1-t)^{-\frac{1-t}{1+\sigma^2}} \right]^n. \quad (2.9)$$

**Bemerkung 2.11** Es gelten die folgenden Aussagen:

- (i) Für  $\sigma^2 \rightarrow 0$  konvergiert die rechte Seite von (2.9) gegen Null.
- (ii) Ist  $\sigma^2 = 1$ , so liefert Hoeffdings Theorem 3 dieselbe Aussage wie Folgerung 2.6, wenn wir dort  $\mu = 0$  annehmen.

### 2.2.3 Hoeffdingsche Ungleichung zweiter Ordnung

Ziel dieses Abschnittes ist es, die Aussage der Hoeffdingschen Ungleichung zu verschärfen, um damit bessere Abschätzungen treffen zu können. Insbesondere werden wir dann zweite Momente mit einbeziehen können. Für den Beweis verbessern wir dazu zuerst die Aussage von Lemma 2.4.

**Lemma 2.12** *Es seien  $-\infty < a < b < \infty$  reelle Zahlen. Für jede reelle Zahl  $h > 0$  genügt dann die quadratische Funktion  $p : \mathbb{R} \rightarrow \mathbb{R}$  mit*

$$(1) \quad p(a) = e^{ha},$$

$$(2) \quad p(b) = e^{hb},$$

$$(3) \quad p'(a) = he^{ha}$$

im Intervall  $[a, b]$  der folgenden Ungleichung:

$$e^{ht} \leq p(t) \leq \frac{b-t}{b-a} e^{ha} + \frac{t-a}{b-a} e^{hb}. \quad (2.10)$$

Ist  $\tilde{p}$  eine weitere quadratische Funktion, die (2.10) erfüllt, so gilt für  $t \in [a, b]$ :  $p(t) \leq \tilde{p}(t)$ . Die Funktion  $p$  ist damit die beste quadratische Funktion in dem Sinne, dass sie die beste Abschätzung unter allen Parabeln in (2.10) bietet.

**Beweis** O.B.d.A. seien  $a = 0$  und  $b = 1$ , andernfalls skaliert man mittels  $s = \frac{t-a}{b-a}$  um. Wir nehmen an, dass  $p(t) = p_2 t^2 + p_1 t + p_0$  mit reellen Koeffizienten  $p_0, p_1, p_2$  ist. Die Bedingungen (1)–(3) erzwingen dann für die Koeffizienten die folgende Form:

$$p_0 = 1, \quad p_1 = h, \quad p_2 = e^h - h - 1.$$

Die Funktion  $t \mapsto e^{ht}$  ist auf  $[0, 1]$  gleichmäßig konvex mit Konstante  $h^2 > 0$ . Dann gilt  $p_2 = e^h - h - 1 \geq h^2 > 0$ . Somit ist  $p$  konvex, d.h. eine nach oben geöffnete Parabel. Für  $t \in [0, 1]$  gilt demnach

$$p(t) \leq (1-t)p(0) + tp(1) = (1-t) + te^h,$$

was das zweite Ungleichheitszeichen in (2.10) beweist.

Betrachten wir nun die Funktion  $q$  mit  $q(t) = e^{ht} - p(t)$  etwas genauer. Wegen (1) und (2) gilt  $q(0) = q(1) = 0$ . Aus dem Satz von Rolle (siehe z.B. [AE10, IV.2 Theorem 2.3]) folgt nun, dass es ein  $\xi \in (0, 1)$  gibt mit  $q'(\xi) = 0$ . Wegen (3) gilt zudem  $q'(0) = 0$ . Dann sind 0 und  $\xi$  die einzigen kritischen Punkte von  $q$ , denn andernfalls gäbe es eine zweite Nullstelle der zweiten Ableitung, was wegen

$$q''(t) = h^2 e^{ht} - 2p_2 = 0 \iff t = \frac{1}{h} \ln \left( \frac{2p_2}{h^2} \right)$$

nicht möglich ist. Dann kann  $q$  aber neben 0 und 1 keine weiteren Nullstellen haben, sonst gäbe es neben 0 und  $\xi$  einen weiteren kritischen Punkt. Ferner gilt

$$\lim_{t \rightarrow \infty} q'(t) = \lim_{t \rightarrow \infty} h e^{ht} - 2p_2 t - p_1 = \infty.$$

Somit ist  $q'(t) > 0$  für alle  $t > \xi$ . Aus dem Mittelwertsatz folgt dann, dass es ein  $\xi' \in (\xi, 1)$  mit

$$q(\xi) = q(1) + q'(\xi')(\xi - 1) = q'(\xi')(\xi - 1)$$

gibt und wegen  $\xi \in (0, 1)$  gilt dann  $q(\xi) < 0$ . Da es in  $(0, 1)$  keine Nullstellen gibt und  $q$  stetig ist, gilt dann  $q(t) \leq 0$  für alle  $t \in [0, 1]$ . Somit ist  $e^{ht} \leq p(t)$  für  $t \in [0, 1]$ . Dies zeigt das erste Ungleichheitszeichen in (2.10).

Ist nun  $\tilde{p}$  eine weitere quadratische Funktion, die (2.10) erfüllt, so gilt in jedem Falle  $\tilde{p}(0) = p(0)$  und  $\tilde{p}(1) = p(1)$ ; sonst wäre das zweite Ungleichheitszeichen nicht erfüllt. Angenommen, es gilt  $\tilde{p}'(0) \neq h$ . Dann ist  $\tilde{p}'(0) > h$ , andernfalls gäbe es ein  $\varepsilon \in (0, 1)$ , sodass  $\tilde{p}(\varepsilon) < e^{h\varepsilon}$ , im Widerspruch zu (2.10). Dann gibt es aber ein  $\delta \in (0, 1)$  derart, dass  $\tilde{p}(\delta) > p(\delta)$  gilt. Folglich gilt  $\tilde{p}(t) \geq p(t)$  für alle  $t \in [0, 1]$ . ■

**Folgerung 2.13** Für Zufallsgrößen  $X$  mit  $a \leq X \leq b$  und  $h > 0$  gilt die Ungleichung

$$\mathbb{E}e^{hX} \leq \mathbb{E}p(X) \leq \frac{b - \mathbb{E}X}{b - a} e^{ha} + \frac{\mathbb{E}X - a}{b - a} e^{hb}, \quad (2.11)$$

wobei  $p$  wie in Lemma 2.12 gewählt sei.

**Beweis** Die Aussage folgt unmittelbar aus (2.10) durch Bildung des Erwartungswertes. ■

**Bemerkung 2.14** Für  $a = -1$  und  $b = 1$  haben die Koeffizienten der Parabel  $p$  in Lemma 2.12 die Gestalt

$$\begin{aligned} p_2 &= \frac{1}{4} (e^h - e^{-h} - 2he^{-h}), \\ p_1 &= \frac{1}{2} (e^h - e^{-h}), \\ p_0 &= \frac{1}{4} (e^h + 3e^{-h} + 2he^{-h}). \end{aligned}$$

**Bemerkung 2.15** Es sei  $h > 0$  und  $p$  die in Lemma 2.12 betrachtete Parabel mit Koeffizienten  $p_0, p_1, p_2$ . Dann ergibt sich mit (2.11) die folgende Ungleichung:

$$\mathbb{E}e^{hX} \leq p_2 \mathbb{E}X^2 + p_1 \mathbb{E}X + p_0.$$

Die Verwendung der Parabel in der Hoeffdingschen Ungleichung liefert demnach eine bessere Abschätzung als in (2.2). Betrachten wir nun den Beweis der Hoeffdingschen Ungleichung genauer, so bemerken wir, dass wir mit Lemma 2.12 eine verschärfte Aussage formulieren können.

**Theorem 2.16** (Hoeffdingsche Ungleichung zweiter Ordnung) *Es seien  $X_1, X_2, \dots, X_n$  unabhängig und identisch verteilte Zufallsgrößen mit  $-1 \leq X_i \leq 1$  für  $i = 1, 2, \dots, n$  und es sei  $\mu = \mathbb{E}X_1$  sowie  $\sigma^2 = \text{Var} X_1$ . Dann gilt für  $t \in (0, 1 - \mu)$  und  $h > 0$ :*

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ e^{-h(t+\mu)} (p_2 (\sigma^2 + \mu^2) + p_1 \mu + p_0) \right]^n. \quad (2.12)$$

Die Koeffizienten  $p_0, p_1, p_2$  haben dabei die Gestalt wie in Bemerkung 2.14. Für die Lösung  $\bar{h}$  der nichtlinearen Gleichung

$$e^{2h} = \frac{(t + \mu + 1) ((\sigma^2 + \mu^2)(2h + 1) - 2h + 2\mu - 3) - 2(\sigma^2 + \mu^2 - 1)}{(t + \mu - 1)(\sigma^2 + \mu^2 + 2\mu + 1)} \quad (2.13)$$

wird die rechte Seite von (2.12) minimal. Die beste Abschätzung liefert demnach

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ e^{-\bar{h}(t+\mu)} (\bar{p}_2 (\sigma^2 + \mu^2) + \bar{p}_1 \mu + \bar{p}_0) \right]^n, \quad (2.14)$$

wobei  $\bar{p}_i$ ,  $i = 0, 1, 2$ , die Koeffizienten entsprechend Bemerkung 2.14 mit  $h = \bar{h}$  seien.

**Beweis** Es sei  $h > 0$ . Mit (2.1) folgt dann

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq e^{-hnt-hn\mu} \prod_{i=1}^n \mathbb{E}e^{hX_i}.$$

Aus Bemerkung 2.15 folgt nun

$$\prod_{i=1}^n \mathbb{E}e^{hX_i} \leq \prod_{i=1}^n (p_2 \mathbb{E}X_i^2 + p_1 \mathbb{E}X_i + p_0).$$

Analog zu (2.3) können wir dann die Abschätzung

$$\begin{aligned} \left( \prod_{i=1}^n (p_2 \mathbb{E}X_i^2 + p_1 \mathbb{E}X_i + p_0) \right)^{\frac{1}{n}} &\leq \frac{1}{n} \sum_{i=1}^n (p_2 \mathbb{E}X_i^2 + p_1 \mathbb{E}X_i + p_0) \\ &= p_2 (\sigma^2 + \mu^2) + p_1 \mu + p_0 \end{aligned}$$

treffen. Folglich gilt:

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ e^{-ht-h\mu} (p_2 (\sigma^2 + \mu^2) + p_1 \mu + p_0) \right]^n. \quad (2.15)$$

Damit die Abschätzung möglichst gut wird, gilt es die rechte Seite der Ungleichung in Abhängigkeit vom Parameter  $h$  zu minimieren. Wir betrachten nur den Term innerhalb der eckigen Klammern, da die rechte Seite genau dann minimal ist, wenn dies für den inneren Term gilt. Dessen erste Ableitung nach  $h$  hat die Gestalt

$$\begin{aligned} -\frac{e^{-h(t+\mu)}}{4} &\left[ (t+\mu) \left( (e^h - e^{-h} - 2he^{-h})(\sigma^2 + \mu^2) + 2\mu(e^h - e^{-h}) + e^h + 2he^{-h} + 3e^{-h} \right) \right. \\ &\quad \left. - \left( (e^h - e^{-h} + 2he^{-h})(\sigma^2 + \mu^2) + 2\mu(e^h + e^{-h}) + e^h - 2he^{-h} - e^{-h} \right) \right] \end{aligned} \quad (2.16)$$

und wird genau dann Null, wenn

$$\begin{aligned} &e^h \left( (t+\mu)(\sigma^2 + \mu^2 + 2\mu + 1) - \sigma^2 - \mu^2 - 2\mu - 1 \right) \\ &= e^{-h} \left( (t+\mu) \left( (\sigma^2 + \mu^2)(2h+1) - 2h + 2\mu - 3 \right) + (\sigma^2 + \mu^2)(2h-1) + 2\mu - 2h - 1 \right), \end{aligned}$$

d.h. falls

$$e^{2h} = \frac{(t+\mu) \left( (\sigma^2 + \mu^2)(2h+1) - 2h + 2\mu - 3 \right) + (\sigma^2 + \mu^2)(2h-1) + 2\mu - 2h - 1}{(t+\mu-1)(\sigma^2 + \mu^2 + 2\mu + 1)}.$$

Dies ist genau dann der Fall, wenn  $h$  die Gleichung (2.13) löst. Die eckige Klammer in (2.15) konvergiert für  $h \rightarrow 0$  gegen 1; ihre erste Ableitung konvergiert gegen  $-t < 0$ . Letzteres sehen

wir mit (2.16) ein. Für  $h \rightarrow 0$  erhalten wir dort

$$-\frac{1}{4} \left( (t + \mu)(1 + 3) - (2\mu(1 + 1) + 1 - 1) \right) = -\frac{1}{4} (4t + 4\mu - 4\mu) = -t.$$

Zusammen mit der Tatsache, dass die rechte Seite von (2.15) wegen  $1 - t - \mu > 0$  für  $h \rightarrow \infty$  gegen  $\infty$  divergiert, sichert dies die Existenz eines (eindeutigen) Minimums. Die Minimalstelle ist dabei gegeben durch die Lösung  $\bar{h}$  der Gleichung (2.13). ■

**Bemerkung 2.17** Wir betrachten eine Zufallsvariable  $Z$  mit  $\|Z\| \leq 1$ . Ist  $\mu = \|\mathbb{E}Z\|$ , so gilt

$$\sigma^2 = \mathbb{E}\|Z - \mathbb{E}Z\|^2 = \mathbb{E}\|Z\|^2 - \|\mathbb{E}Z\|^2 \leq 1 - \mu^2.$$

Die Hoeffdingsche Ungleichung zweiter Ordnung in Theorem 2.16 liefert dann eine bessere Abschätzung als die Hoeffdingsche Ungleichung in Theorem 2.5, wenn die Varianz  $\sigma^2$  nicht maximal ist, d.h. falls  $\sigma^2 < 1 - \mu^2$  ist. Für  $\sigma^2 = 1 - \mu^2$  erhalten wir gerade die Hoeffdingsche Ungleichung zurück. Dann vereinfacht sich nämlich (2.13) zu

$$\begin{aligned} e^{2h} &= \frac{(t + \mu + 1)(2h + 1 - 2h + 2\mu - 3) - 2(1 - 1)}{(t + \mu - 1)(1 + 2\mu + 1)} \\ &= \frac{2(t + \mu + 1)(\mu - 1)}{2(t + \mu - 1)(\mu + 1)} = \frac{(1 - \mu)(1 + t + \mu)}{(1 + \mu)(1 - t - \mu)}. \end{aligned}$$

Somit löst

$$\bar{h} = \frac{1}{2} \ln \left( \frac{(1 - \mu)(1 + t + \mu)}{(1 + \mu)(1 - t - \mu)} \right)$$

die Gleichung (2.13). Betrachten wir nun (2.15) etwas genauer. Für  $\sigma^2 = 1 - \mu^2$  hat die rechte Seite die Gestalt

$$\left[ e^{-h(t+\mu)} (p_2 + p_0 + p_1\mu) \right]^n = \left[ \frac{e^{-h(t+\mu)}}{2} \left( e^h(1 + \mu) + e^{-h}(1 - \mu) \right) \right]^n.$$

Setzen wir dort  $h = \bar{h}$  ein, so erhalten wir

$$\begin{aligned} &\left[ \frac{1}{2} \left( \frac{(1 + \mu)(1 - t - \mu)}{(1 - \mu)(1 + t + \mu)} \right)^{\frac{1}{2}(t+\mu)} \left[ \left( \frac{(1 - \mu)(1 + t + \mu)}{(1 + \mu)(1 - t - \mu)} \right)^{\frac{1}{2}} (1 + \mu) + \left( \frac{(1 + \mu)(1 - t - \mu)}{(1 - \mu)(1 + t + \mu)} \right)^{\frac{1}{2}} (1 - \mu) \right] \right]^n \\ &= \left[ \frac{1}{2} \left( \frac{\mu + 1}{1 + t + \mu} \right)^{\frac{1}{2}(t+\mu+1)} \left( \frac{1 - t - \mu}{1 - \mu} \right)^{\frac{1}{2}(t+\mu)} \left[ \left( \frac{1 - \mu}{1 - t - \mu} \right)^{\frac{1}{2}} (1 + t + \mu) + ((1 - t - \mu)(1 - \mu))^{\frac{1}{2}} \right] \right]^n \\ &= \left[ \frac{1}{2} \left( \frac{\mu + 1}{1 + t + \mu} \right)^{\frac{1}{2}(t+\mu+1)} \left( \frac{1 - t - \mu}{1 - \mu} \right)^{\frac{1}{2}(t+\mu-1)} [1 + t + \mu + 1 - t - \mu] \right]^n \\ &= \left[ \left( \frac{\mu + 1}{1 + t + \mu} \right)^{t+\mu+1} \left( \frac{1 - \mu}{1 - t - \mu} \right)^{1-\mu-t} \right]^{\frac{n}{2}}. \end{aligned}$$

Für  $0 < t < 1 - \mu$  folgt somit insgesamt

$$\mathbb{P}(\bar{X}_n - \mu \geq t) \leq \left[ \left( \frac{\mu + 1}{\mu + 1 + t} \right)^{\mu + 1 + t} \left( \frac{1 - \mu}{1 - \mu - t} \right)^{1 - \mu - t} \right]^{\frac{n}{2}}.$$

Dies entspricht der Hoeffdingschen Ungleichung in Folgerung 2.6.

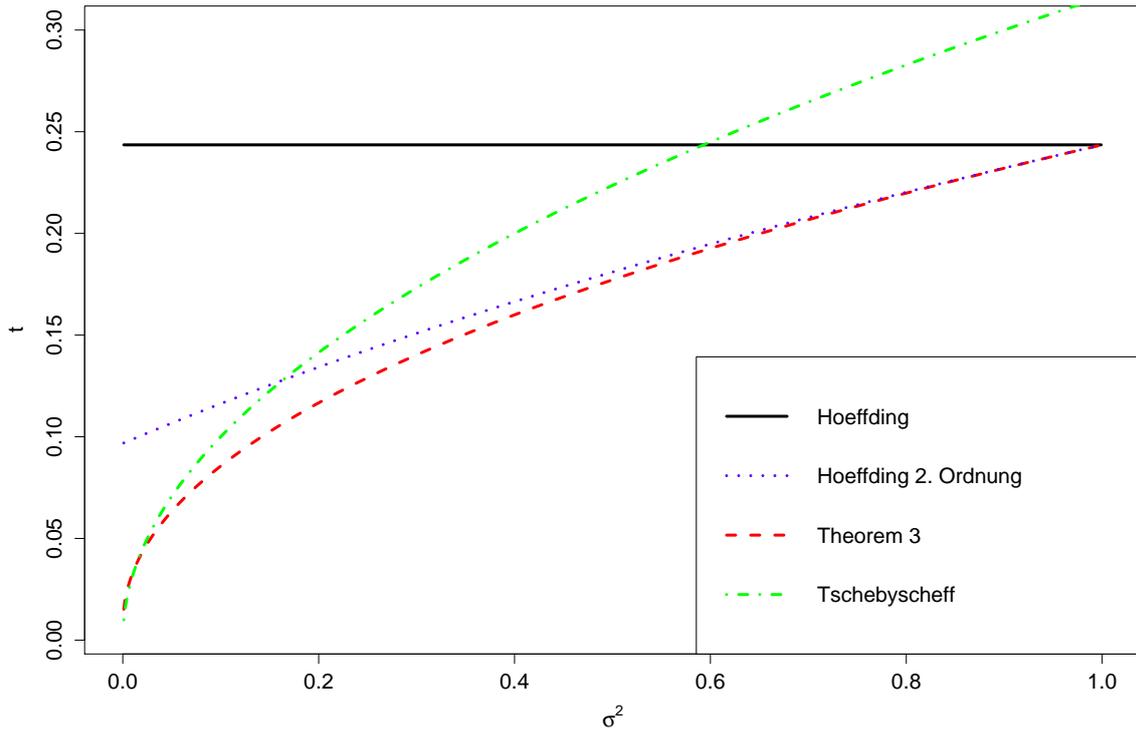
## 2.3 Vergleich der verschiedenen Ungleichungen

Wir wollen die in den Abschnitten 2.2.1–2.2.3 vorgestellten Ungleichungen nun miteinander vergleichen. Bei der Verwendung der Ungleichungen wird in der späteren Konstruktion stets  $\mu = 0$  gelten. Unter dieser Annahme werden wir, gegeben  $n \in \mathbb{N}$  und  $\alpha \in (0, 1)$ , die rechten Seiten der Ungleichungen jeweils gleich  $\alpha$  setzen. Dann bestimmen wir das entsprechende  $t = t(\sigma^2)$  und untersuchen das Verhalten von  $t$  in Abhängigkeit der Varianz  $\sigma^2$ . Für die Tschebyscheffsche Ungleichung verwenden wir zum Vergleich den Radius des  $\varepsilon$ -Balls. Für diesen gilt  $\varepsilon = \sqrt{\frac{\sigma^2}{n\alpha}}$  gemäß (2.8). Da die Tschebyscheffsche Ungleichung jedoch, im Gegensatz zu den anderen Ungleichungen, in beide Richtungen gleichzeitig abschätzt, werden wir dort für die rechte Seite  $2\alpha$  verwenden, um einen fairen Vergleich zu erhalten. Wir bemerken dazu, dass ein kleines  $t$  für die spätere Konstruktion stets besser ist.

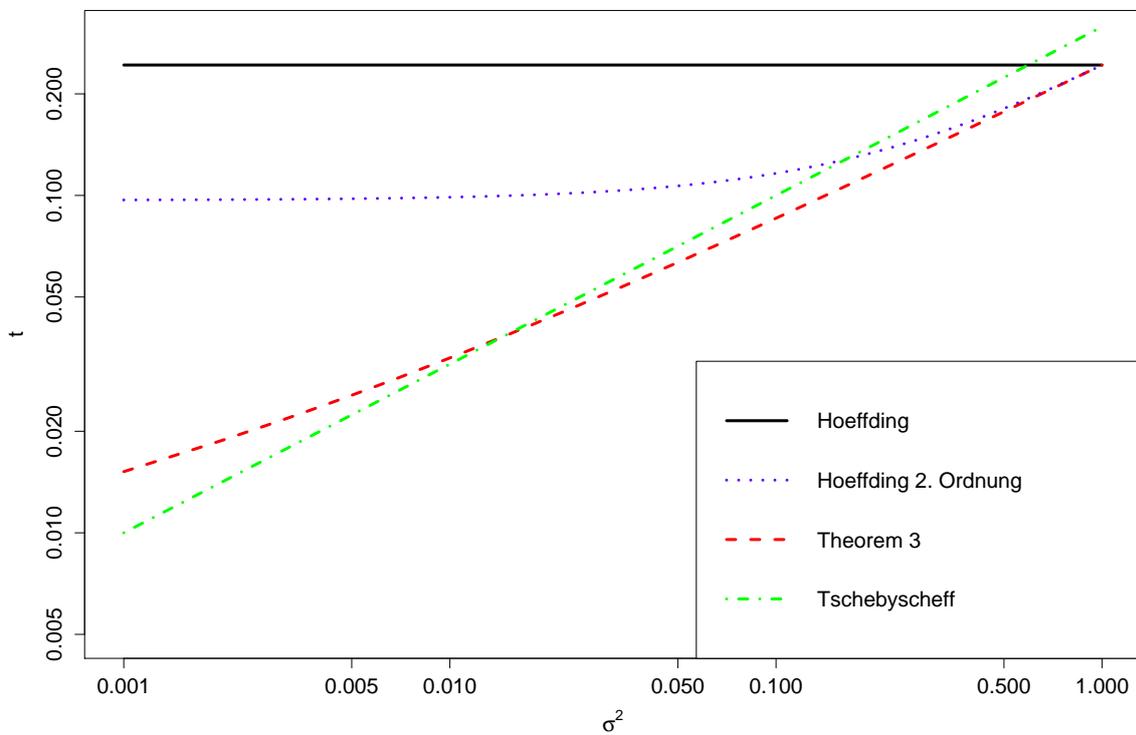
Zu Demonstrationszwecken haben wir  $t(\sigma^2)$  in Abbildung 2.1a für  $n = 100$  und  $\alpha = 0,05$  geplottet. Abbildung 2.1b zeigt selbigen Plot mit logarithmierten Achsen. Die in diesem Abschnitt betrachteten Ungleichungen sind farbig, im Vergleich zur Hoeffdingschen Ungleichung in schwarz zu sehen.

Es fällt auf, dass die Abschätzungen von Theorem 3 und der Hoeffdingschen Ungleichung zweiter Ordnung für  $\sigma^2 < 1$  stets besser sind als die Abschätzung mit Hilfe der Hoeffdingschen Ungleichung. Ferner wird die Aussage von Bemerkung 2.17 bestätigt, dass für maximale Varianz Hoeffdingsche Ungleichung und Hoeffdingsche Ungleichung zweiter Ordnung dieselbe Abschätzung liefern.

Weiter ist zu bemerken, dass die Abschätzungen nach Theorem 3 stets besser sind als jene nach Theorem 2.16, wenn auch für große  $\sigma^2$  nur marginal. Die Tschebyscheffsche Ungleichung ist erst für kleine  $\sigma^2$  besser als die Hoeffdingsche Ungleichung, liefert jedoch für sehr kleine  $\sigma^2$  sogar bessere Abschätzungen als die Hoeffdingsche Ungleichung zweiter Ordnung und ist schließlich sogar besser als die Abschätzung nach Theorem 3. Dies erkennt man in Abbildung 2.1b besonders gut. Zusammen mit Theorem 3 ist die Tschebyscheffsche Ungleichung zudem die einzigen Ungleichung, für die die entsprechende Lösung der rechten Seite für  $\sigma^2 \rightarrow 0$  gegen Null konvergiert.



(a) Achsen nicht logarithmiert



(b) Doppelt logarithmische Darstellung

Abbildung 2.1: Lösungen  $t(\sigma^2)$  in der Hoeffdingschen Ungl. (schwarz), Hoeffdingschen Ungl. zweiter Ordnung (blau), Theorem 3 (rot), sowie Tschebyscheffschen Ungl. (grün) für  $\alpha = 0,05$ ,  $\mu = 0$  und  $n = 100$ .

Zusammenfassend können wir sagen, dass für stark streuende Daten Hoeffdings Theorem 3 die beste Abschätzung liefert. Sind die Daten stark konzentriert, sollte man eher die Tschebyscheffsche Ungleichung verwenden. Ist die Varianz klein, bieten jedoch alle in Abschnitt 2.2 betrachteten Ungleichungen eine bessere Abschätzung als die Hoeffdingsche Ungleichung. Dies soll uns dazu motivieren, in Kapitel 3 auch Konfidenzbereiche mit diesen Ungleichungen zu konstruieren. Im Falle kleiner Varianz, werden diese Konfidenzbereiche dann vermutlich besser sein.

---

## 3 Die Konstruktion der Konfidenzbereiche

In diesem Kapitel wollen wir neben einem Konfidenzbereich mit Hilfe der Hoeffdingschen Ungleichung 2.5 in Abschnitt 3.1 auch weitere Konfidenzbereiche unter Einbeziehung der Varianz in Abschnitt 3.2 betrachten. Wenn wir von Konfidenzbereich sprechen, meinen wir dabei immer Konfidenzbereiche (zum Niveau  $1 - \alpha$ ) für den *extrinsischen Erwartungswert*

$$\varepsilon = \operatorname{argmin}_{\nu \in S^1} \|\mathbb{E}Z - \nu\|,$$

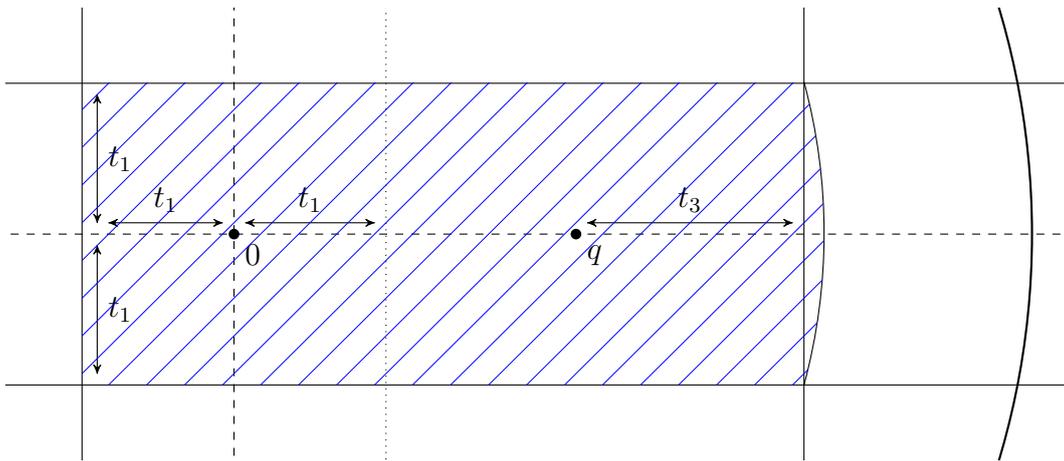
das ist ein Punkt  $\varepsilon \in S^1$ , welcher den Abstand vom Erwartungswert  $\mathbb{E}Z$  im Vektorraum  $\mathbb{R}^2$  bzgl. der euklidischen Norm in der Ebene minimiert. Dies liefert uns eine natürliche Definition des Erwartungswertes auf dem Einheitskreis. Die Existenz von  $\varepsilon$  wird dabei stets durch die Kompaktheit der Kreislinie und die Stetigkeit der Norm gewährleistet. Die Eindeutigkeit hingegen ist dann und nur dann gegeben, wenn  $\mathbb{E}Z \neq 0$  ist. In diesem Fall ist  $\varepsilon$  durch die Orthogonalprojektion von  $\mathbb{E}Z$  auf den Einheitskreis

$$\varepsilon = \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|}$$

eindeutig bestimmt, wie in [BP03, Abschnitt 3] gezeigt wird. Für den Fall  $\mathbb{E}Z = 0$  ist offenbar  $\varepsilon = S^1$ , d.h. jeder Punkt auf dem Einheitskreis besitzt denselben minimalen Abstand von  $\mathbb{E}Z$ . Wir sprechen im Folgenden stets so, als wäre der extrinsische Erwartungswert  $\varepsilon$  eindeutig gegeben und schließen dabei den Fall  $\varepsilon = S^1$  stillschweigend mit ein.

Gegeben Daten  $Z_1, Z_2, \dots, Z_n$  wollen wir nun einen *Konfidenzbereich* für  $\varepsilon$ , d.h. eine Menge, die jeden extrinsischen Erwartungswert mit vorgegebener Wahrscheinlichkeit enthält, konstruieren. Die genauen Ideen und Konstruktionsschritte werden wir dabei einzeln erläutern und in jedem Abschnitt abschließend in einem Algorithmus zusammenfassen.

---

Abbildung 3.1: Menge  $M(q)$  (blau schraffiert) unter der Hypothese  $\mu = q$ 

### 3.1 Konfidenzbereiche mit der Hoeffdingschen Ungleichung

Im folgenden Abschnitt seien  $Z, Z_1, Z_2, \dots, Z_n$  unabhängige, identisch verteilte Zufallsvariablen mit Werten auf dem Einheitskreis  $S^1$  und Erwartungswert  $\mu = \mathbb{E}Z = \mathbb{E}\bar{Z}_n$  für  $\bar{Z}_n = \frac{1}{n} \sum_{i=1}^n Z_i$ . Für festes  $q$  mit  $\|q\| \leq 1$  betrachten wir die Hypothese  $H_0: \mu \in \Theta_0 = \{q\}$  vs. die Alternative  $H_1: \mu \in \Theta_1 = \bar{\mathbb{B}}_1(0) \setminus \{q\}$ . Unter der Hypothese  $\mu = q$  konstruieren wir eine Menge  $M(q)$  der Form wie in Abbildung 3.1. Diese soll den Mittelwert  $\bar{Z}_n$  mindestens mit Wahrscheinlichkeit  $1 - \alpha$  enthalten.

Motiviert wird dies durch die folgende Idee: Gegeben  $M(q)$ , können wir einen Konfidenzbereich für den extrinsischen Erwartungswert  $\varepsilon$  konstruieren, indem wir den Annahmehereich  $A = \{q \in \bar{\mathbb{B}}_1(0) \mid \bar{Z}_n \in M(q)\}$ , das ist die Menge aller  $q$ , die bei Beobachtung von  $\bar{Z}_n$  bei einem Test zum Niveau  $\alpha$  nicht verworfen werden würden, betrachten, und diesen schließlich auf den Einheitskreis projizieren. Genauer ist dann durch

$$K = \begin{cases} S^1, & 0 \in A, \\ \left\{ \frac{q}{\|q\|} \mid q \in A \right\}, & 0 \notin A \end{cases}$$

ein  $(1 - \alpha)$ -Konfidenzbereich für den extrinsischen Erwartungswert  $\varepsilon$  gegeben.

Dies soll die im Weiteren verfolgte Vorgehensweise zur Konstruktion der Konfidenzbereiche sein. Wir werden die Konstruktion in zwei Schritte, 3.1.1 und 3.1.2, aufteilen, wobei wir im ersten Schritt die Mengen  $M(q)$  bestimmen. Im zweiten Schritt werden wir dann, gegeben  $\bar{Z}_n$ , den Konfidenzbereich  $K$  konstruieren.

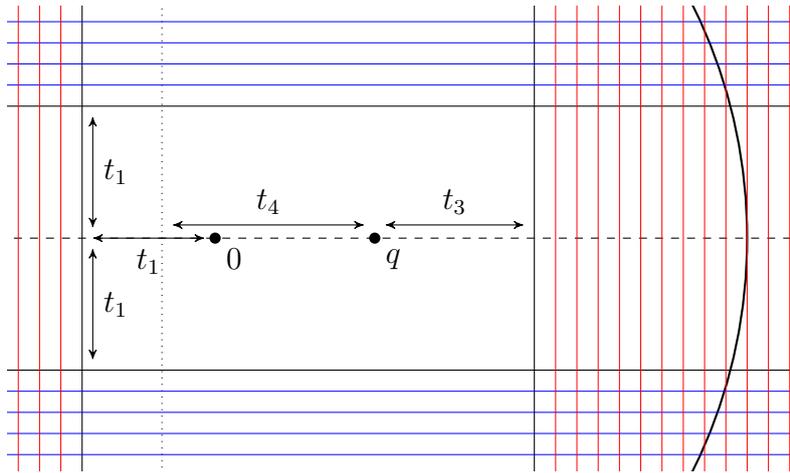


Abbildung 3.2: Konstruktion des Rechtecks mit Hoeffdingscher Ungleichung

### 3.1.1 Konstruktion – Schritt 1

Wir bestimmen zunächst die Längen  $t_1$  und  $t_3$  der in Abbildung 3.2 zu sehenden geometrischen Figur derart, dass das Rechteck den Mittelwert  $\bar{Z}_n$  mindestens mit Wahrscheinlichkeit  $1 - \alpha$  enthält. Dann gilt letzteres auch für die größere Menge  $M(q)$  aus Abbildung 3.1. Wir konstruieren das Rechteck symmetrisch in dem Sinne, dass wir jede Seite mit der Wahrscheinlichkeit  $\frac{\alpha}{4}$  gewichten.

Es sei dazu  $q \in \bar{\mathbb{B}}_1(0) \setminus \{0\}$  der Erwartungswert von  $Z$  unter der Hypothese, d.h.  $\mathbb{E}Z = q$ . Wir betrachten die Projektion von  $\bar{Z}_n$  auf den von  $q$  erzeugten linearen Unterraum  $\text{span } q$  bzw. auf dessen orthogonales Komplement  $(\text{span } q)^\perp$ . Genauer betrachten wir die (univariaten) Zufallsgrößen

$$\begin{aligned} (1) \quad X^{(1)} &:= \langle \bar{Z}_n, q' \rangle, & (3) \quad X^{(3)} &:= \langle \bar{Z}_n, \frac{q}{\|q\|} \rangle, \\ (2) \quad X^{(2)} &:= \langle \bar{Z}_n, -q' \rangle = -X^{(1)}, & (4) \quad X^{(4)} &:= \langle \bar{Z}_n, -\frac{q}{\|q\|} \rangle = -X^{(3)}, \end{aligned}$$

wobei  $q'$  ein Vektor mit  $\|q'\| = 1$  ist, der  $(\text{span } q)^\perp$  erzeugt. Dann gilt unter der Hypothese

$$\begin{aligned} \mathbb{E}X^{(1)} &= \langle \mathbb{E}\bar{Z}_n, q' \rangle = \langle q, q' \rangle = 0, & \mathbb{E}X^{(3)} &= \frac{1}{\|q\|} \langle \mathbb{E}Z_n, q \rangle = \frac{1}{\|q\|} \langle q, q \rangle = \|q\|, \\ \mathbb{E}X^{(2)} &= -\mathbb{E}X^{(1)} = 0, & \mathbb{E}X^{(4)} &= -\mathbb{E}X^{(3)} = -\|q\|. \end{aligned}$$

Es sei angemerkt, dass alle Zufallsgrößen Werte aus  $[-1, 1]$  annehmen. Aus der Hoeffdingschen Ungleichung 2.6 folgt dann für  $t_1 \in (0, 1)$ , sowie  $t_2 = -t_1$  die Abschätzung

$$\mathbb{P}(X^{(1)} \geq t_1) \leq \left[ \left( \frac{1}{1+t_1} \right)^{1+t_1} \left( \frac{1}{1-t_1} \right)^{1-t_1} \right]^{\frac{n}{2}}, \quad (3.1)$$

$$\mathbb{P}(X^{(1)} \leq t_2) = \mathbb{P}(X^{(2)} \geq t_1) \leq \left[ \left( \frac{1}{1+t_1} \right)^{1+t_1} \left( \frac{1}{1-t_1} \right)^{1-t_1} \right]^{\frac{n}{2}}. \quad (3.2)$$

Dabei entspricht  $t_1$  der Abweichung von  $\bar{Z}_n$  von  $\text{span } q$  in Richtung des Vektors  $q'$ . Für die Abschätzung in Richtung  $q$  gilt für  $t_3 \in (0, 1 - \|q\|)$  bzw.  $t_4 \in (0, 1 + \|q\|)$ :

$$\mathbb{P}(X^{(3)} - \|q\| \geq t_3) \leq \left[ \left( \frac{1 + \|q\|}{1 + \|q\| + t_3} \right)^{1 + \|q\| + t_3} \left( \frac{1 - \|q\|}{1 - \|q\| - t_3} \right)^{1 - \|q\| - t_3} \right]^{\frac{n}{2}}, \quad (3.3)$$

$$\begin{aligned} \mathbb{P}(X^{(3)} - \|q\| \leq -t_4) &= \mathbb{P}(X^{(4)} + \|q\| \geq t_4) \\ &\leq \left[ \left( \frac{1 - \|q\|}{1 - \|q\| + t_4} \right)^{1 - \|q\| + t_4} \left( \frac{1 + \|q\|}{1 + \|q\| - t_4} \right)^{1 + \|q\| - t_4} \right]^{\frac{n}{2}}. \end{aligned} \quad (3.4)$$

**Lemma 3.1** *Es sei  $-1 < \mu < 1$ ,  $n \in \mathbb{N}$  und  $\alpha \in (0, 1)$ . Wir betrachten für  $t \in (0, 1 - \mu)$  die Gleichung*

$$\left[ \left( \frac{\mu + 1}{\mu + 1 + t} \right)^{\mu + 1 + t} \left( \frac{1 - \mu}{1 - \mu - t} \right)^{1 - \mu - t} \right]^{\frac{n}{2}} = \alpha. \quad (3.5)$$

Dann gelten folgende Aussagen:

- (i) *Existiert eine Lösung  $t^*$  der Gleichung (3.5), so ist  $t^*$  die einzige Lösung.*
- (ii) *Ist  $\mu = 0$  und  $\alpha > 2^{-n}$ , so existiert eine Lösung  $t^*$  von (3.5).*

**Beweis** Die Gleichung (3.5) wird genau dann von  $t^*$  gelöst, wenn  $t^*$  die logarithmierte Gleichung

$$(\mu + 1 + t) \ln \left( \frac{\mu + 1}{\mu + 1 + t} \right) + (1 - \mu - t) \ln \left( \frac{1 - \mu}{1 - \mu - t} \right) = \frac{2}{n} \ln \alpha. \quad (3.6)$$

erfüllt. Wir betrachten die Funktion

$$\begin{aligned} f &: (0, 1 - \mu) \rightarrow \mathbb{R}, \\ t \mapsto f(t) &= (\mu + 1 + t) \ln \left( \frac{\mu + 1}{\mu + 1 + t} \right) + (1 - \mu - t) \ln \left( \frac{1 - \mu}{1 - \mu - t} \right) - \frac{2}{n} \ln \alpha. \end{aligned}$$

Eine Lösung  $t^*$  von (3.5) entspricht dann einer Nullstelle von  $f$ .

Um die Eindeutigkeit der Lösung  $t^*$  zu zeigen, betrachten wir die Ableitung  $f' = \frac{\partial f}{\partial t}$  von  $f$ . Diese hat die Form

$$\begin{aligned} f'(t) &= \ln \left( \frac{\mu + 1}{\mu + 1 + t} \right) + \frac{(\mu + 1 + t)^2}{\mu + 1} \left( -\frac{\mu + 1}{(\mu + 1 + t)^2} \right) - \ln \left( \frac{1 - \mu}{1 - \mu - t} \right) + \frac{(1 - \mu - t)^2}{1 - \mu} \left( \frac{1 - \mu}{(1 - \mu - t)^2} \right) \\ &= \ln \left( \frac{\mu + 1}{\mu + 1 + t} \right) - \ln \left( \frac{1 - \mu}{1 - \mu - t} \right), \end{aligned}$$

da  $\mu$  nicht von  $t$  abhängt.

Dieser Term ist genau dann gleich Null, wenn

$$\begin{aligned} \frac{\mu + 1}{\mu + 1 + t} = \frac{1 - \mu}{1 - \mu - t} &\iff (\mu + 1)(1 - \mu - t) = (1 - \mu)(\mu + 1 + t) \\ &\iff \mu - \mu^2 - \mu t + 1 - \mu - t = \mu + 1 + t - \mu^2 - \mu - \mu t \\ &\iff 2t = 0 \end{aligned}$$

und verschwindet demnach auf  $(0, 1 - \mu)$  nicht. Dann ist  $f$  injektiv und die Nullstelle  $t^*$  ist, im Falle ihrer Existenz, eindeutig bestimmt. Somit folgt die Aussage (i).

Es sei nun  $\mu = 0$ . Dann gilt  $\lim_{t \rightarrow 0} f(t) = -\frac{2}{n} \ln \alpha > 0$  und wegen  $\alpha > 2^{-n}$  ist

$$\lim_{t \rightarrow 1} f(t) = -2 \ln 2 - \frac{2}{n} \ln \alpha < -2 \ln 2 + 2 \ln 2 = 0.$$

Nach dem Zwischenwertsatz existiert somit, da  $f$  stetig ist, in  $(0, 1)$  eine Nullstelle  $t^*$ . Diese löst die Gleichung (3.5). Somit folgt die Aussage (ii) ■

**Bemerkung 3.2** Lemma 3.1 besagt, dass die Gleichung (3.5) für  $\mu = 0$  eine eindeutige Lösung  $t^*$  hat, sofern  $n$  hinreichend groß ist. Dies wollen wir im Folgenden stets annehmen.

Für den nächsten Konstruktionsschritt benötigen wir eine Monotonieaussage, welche uns das folgende Lemma liefert.

**Lemma 3.3** Für  $-1 < \mu < 1$ ,  $\alpha \in (0, 1)$  und  $n \in \mathbb{N}$  erfülle  $t \in (0, 1 - \mu)$  die Gleichung

$$\left[ \left( \frac{\mu + 1}{\mu + 1 + t} \right)^{\mu + 1 + t} \left( \frac{1 - \mu}{1 - \mu - t} \right)^{1 - \mu - t} \right]^{\frac{n}{2}} = \alpha. \quad (3.7)$$

Dann ist  $\mu + t$  strikt monoton wachsend in  $\mu$ .

**Beweis** Da  $t$  die Gleichung (3.7) erfüllt, gilt insbesondere

$$(\mu + 1 + t) \ln \left( \frac{\mu + 1}{\mu + 1 + t} \right) + (1 - \mu - t) \ln \left( \frac{1 - \mu}{1 - \mu - t} \right) = \frac{2}{n} \ln \alpha. \quad (3.8)$$

Dann ist  $t$  implizit als Lösung dieser Gleichung gegeben. Die Ableitung der linken Seite nach  $\mu$  und  $t$  existiert und ist stetig. Aus dem Beweis von Lemma 3.1 folgt, dass die Ableitung der linken Seite von (3.8) nach  $t$  nicht verschwindet. Nach dem Satz über implizite Funktionen (siehe z.B. [Aul12, Satz 4.5.2]) existiert dann die Ableitung  $t' = \frac{\partial t}{\partial \mu}$ .

Differenzieren wir (3.8) nach  $\mu$ , so erhalten wir die Gleichung

$$(1+t') \ln \left( \frac{\mu+1}{\mu+1+t} \right) + (\mu+1+t) \left( \frac{1}{\mu+t} - \frac{1+t'}{1+\mu+t} \right) \\ - (1+t') \ln \left( \frac{1-\mu}{1-\mu-1} \right) + (1-\mu-t) \left( -\frac{1}{1-\mu} + \frac{1+t'}{1-\mu-t} \right) = 0.$$

Diese ist äquivalent zu

$$(1+t') \left( \ln \left( \frac{\mu+1}{\mu+1+t} \right) - \ln \left( \frac{1-\mu}{1-\mu-t} \right) \right) = \frac{1-\mu-t}{1-\mu} - \frac{\mu+1+t}{\mu+t} \\ = -\frac{2t}{1-\mu^2}.$$

Man beachte, dass die rechte Seite wegen  $\mu \in (-1, 1)$  und  $t > 0$  negativ ist. Aufgrund von

$$\ln \left( \frac{\mu+1}{\mu+1+t} \right) < 0, \quad \ln \left( \frac{1-\mu}{1-\mu-t} \right) > 0 \quad (3.9)$$

folgt dann, dass  $1+t' = \frac{\partial}{\partial \mu}(\mu+t) > 0$  gelten muss. Somit ist  $\mu+t$  strikt monoton wachsend in  $\mu$ . ■

**Bemerkung 3.4** Für  $\|q\| = 0$  ist die rechte Seite von (3.4) mit der rechten Seite von (3.2) identisch. Mit Lemma 3.3 folgt dann, dass für  $\|q\| = 0$  die Länge  $\|q\| + t_4$  am größten wird; dann ist nämlich  $\mu = -\|q\|$  am größten. Für diesen Fall ist dann  $t_4 = t_1$ . Wir können demnach die Seitenlängen des Rechtecks konservativ durch  $2t_1$  (vertikal) und  $\|q\| + t_3 + t_1$  (horizontal) abschätzen.

Das in Abbildung 3.2 abgebildete weiße Rechteck enthält dann  $\bar{Z}_n$  wenigstens mit Wahrscheinlichkeit  $1-\alpha$ , wenn wir  $t_1$  und  $t_3$  in den Ungleichungen (3.1) bzw. (3.3) derart wählen, dass die rechte Seite jeweils gleich  $\frac{\alpha}{4}$  ist. Der blau schraffierte Bereich in Abbildung 3.2 enthält  $\bar{Z}_n$  nach Konstruktion höchstens mit Wahrscheinlichkeit  $\frac{\alpha}{2}$ , ebenso wie der rot schraffierte Bereich.

Für  $q = 0$  ist  $t_1 = t_3$ . Wir verwenden für  $M(q)$  einen Kreis mit Mittelpunkt 0 und Radius  $\sqrt{2}t_1$ , da wir keine Richtung auszeichnen können;  $M(q)$  ist dann konservativ gewählt.

**Bemerkung 3.5** Aufgrund von Lemma 3.1 folgt für  $\frac{\alpha}{4} > 2^{-n}$  die Existenz von  $t_1$  in Bemerkung 3.4. Man beachte jedoch, dass Lemma 3.1 nichts über die Existenz eines geeigneten  $t_3$ , wie in Bemerkung 3.4 gefordert, aussagt. Diese nehmen wir im Folgenden nichtsdestotrotz an. Wir zeigen später, dass für alle für die Konstruktion relevanten Fälle  $t_3$  existiert und darüber hinaus die Bestimmung des Konfidenzbereiches unabhängig von  $t_3$  erfolgen kann.

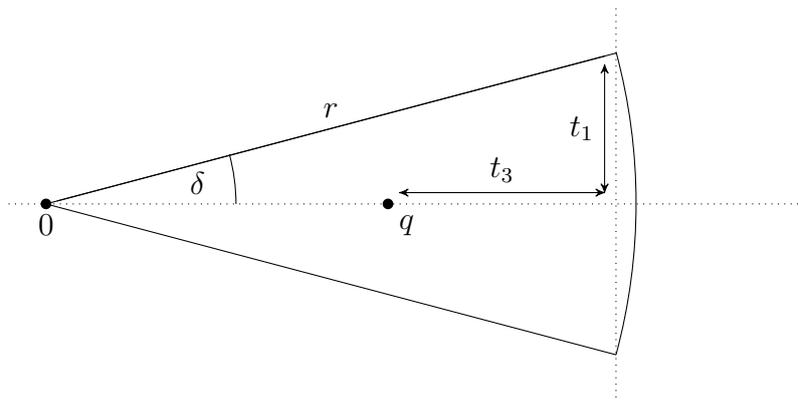


Abbildung 3.3: Kreissektor mit Radius  $r = \sqrt{(\|q\| + t_3)^2 + t_1^2}$

**Bemerkung 3.6** Wir betrachten die Skizze in Abbildung 3.3. Nach dem Satz des Pythagoras hat der Kreissektor den Radius  $r = \sqrt{(\|q\| + t_3)^2 + t_1^2}$ . Mit Lemma 3.3 folgt dann aufgrund der Tatsache, dass  $t_1$  nur von  $\alpha$  und  $n$  abhängt, dass  $r$  monoton wachsend in  $\|q\|$  ist.

### 3.1.2 Konstruktion – Schritt 2

Wir wollen nun, gegeben  $\bar{Z}_n$ , einen Annahmebereich  $A$  konstruieren, welcher all jene  $q$  enthält, für die bei Beobachtung von  $\bar{Z}_n$  bei einem Test zum Niveau  $\alpha$  die Hypothese  $\mathbb{E}Z = q$  nicht verworfen werden würde. Das sind gerade die  $q \in \bar{\mathbb{B}}_1(0)$ , für welche  $\bar{Z}_n$  in der in Schritt 1 konstruierten Menge  $M(q)$  enthalten ist. D.h. es gilt

$$A = \{q \in \bar{\mathbb{B}}_1(0) \mid \bar{Z}_n \in M(q)\}.$$

Die Konstruktion kommutiert mit Rotation in dem Sinne, dass für

- die nach Schritt 1 konstruierte und anschließend um den Ursprung mit Drehwinkel  $\varphi$  gedrehte Menge  $\tilde{M}(q)$  und
- die, für den aus  $q$  durch Drehung um den Ursprung mit Drehwinkel  $\varphi$  hervorgehenden Vektor  $\tilde{q}$ , konstruierte Menge  $M(\tilde{q})$

die Gleichheit  $\tilde{M}(q) = M(\tilde{q})$  gilt.

Daher reicht es aus, die für  $(\|q\|, 0)^\top$  konstruierten Mengen um den Ursprung rotieren zu lassen und zu beobachten, welche  $q$  die Bedingung  $\bar{Z}_n \in M(q)$  erfüllen, wie es in Abbildung 3.4 demonstriert wird.

Man beachte, dass bei der Rotation von  $M(0)$  um den Ursprung ein Ball mit Radius  $\sqrt{2}t_1$  entsteht. Gilt nun  $\|\bar{Z}_n\| \leq \sqrt{2}t_1$ , so ist  $\bar{Z}_n \in M(0)$ . Dann ist  $0 \in A$  und die Projektion ergibt die gesamte Kreislinie. In diesem Fall ist der Konfidenzbereich  $K$  gegeben durch  $K = S_1$ .

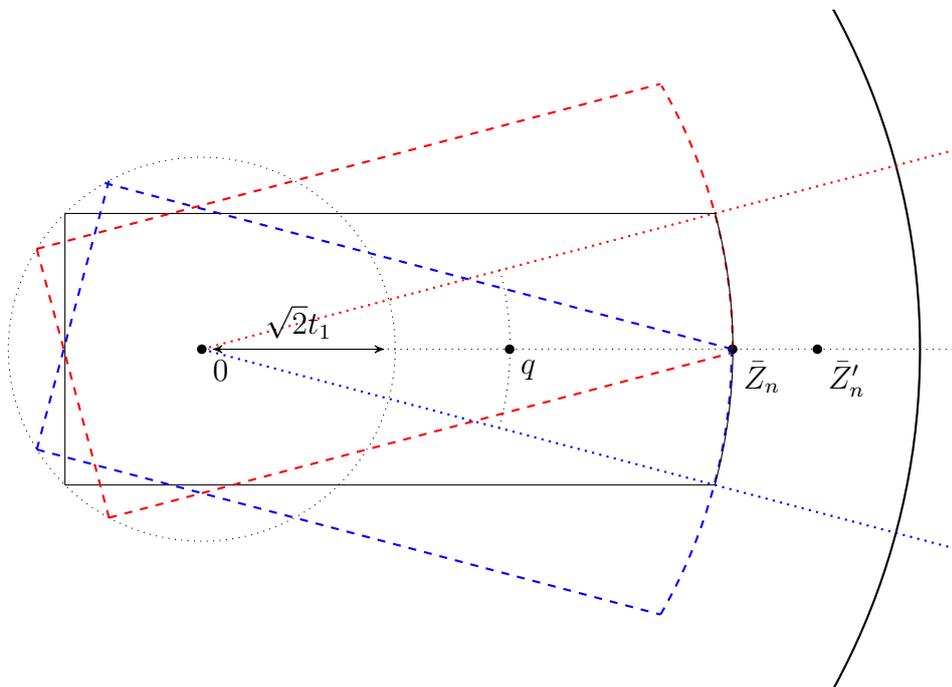


Abbildung 3.4: Rotation von  $M(q)$  (schwarz) um den Ursprung,  $\|\bar{Z}_n\| = r$ ,  $\|\bar{Z}'_n\| > r$

Wir betrachten nun den Fall  $\|\bar{Z}_n\| > \sqrt{2}t_1$  (und somit  $0 \notin A$ ). Ist  $\|\bar{Z}_n\| > r$ , wie es etwa für  $\bar{Z}'_n$  in Abbildung 3.4 der Fall ist, so gilt  $\bar{Z}_n \notin M(q)$ . Für diese  $q$  gilt dann  $q \notin A$ .

Wir müssen demnach nur die  $q$  betrachten, für die  $\|\bar{Z}_n\| \leq r$  gilt. Es reicht sogar nur die  $q$  zu betrachten, für die  $\|\bar{Z}_n\| = r$  gilt, da für alle  $q$  mit  $\|\bar{Z}_n\| < r$  der Öffnungswinkel  $\delta$  aufgrund von Lemma 3.3 höchstens kleiner wird. Wir suchen aber den größten Öffnungswinkel  $\delta$ , denn nur dieser ist für die Projektion von Bedeutung.

Gesucht sind nun die  $q$ , sodass  $\bar{Z}_n$  auf dem Kreisbogen mit Radius  $r$  liegt. Aufgrund obiger Überlegung können wir o.B.d.A. den Fall  $q = (\|q\|, 0)^\top$  betrachten. Dann gilt  $\bar{Z}_n \in M(q)$  mit  $\|\bar{Z}_n\| = r$ , wenn

$$\|q\| + t_3 = \sqrt{\|\bar{Z}_n\|^2 - t_1^2} \quad (3.10)$$

gilt. Da  $t_3$  die Gleichung

$$\left[ \left( \frac{1 + \|q\|}{1 + \|q\| + t_3} \right)^{1 + \|q\| + t_3} \left( \frac{1 - \|q\|}{1 - \|q\| - t_3} \right)^{1 - \|q\| - t_3} \right]^{\frac{n}{2}} = \frac{\alpha}{4}$$

löst, ist die Lösung  $\|q^*\|$  der Gleichung (3.10) somit ebenfalls als Lösung von

$$\left[ \left( \frac{1 + \|q\|}{1 + \sqrt{\|\bar{Z}_n\|^2 - t_1^2}} \right)^{1 + \sqrt{\|\bar{Z}_n\|^2 - t_1^2}} \left( \frac{1 - \|q\|}{1 - \sqrt{\|\bar{Z}_n\|^2 - t_1^2}} \right)^{1 - \sqrt{\|\bar{Z}_n\|^2 - t_1^2}} \right]^{\frac{n}{2}} = \frac{\alpha}{4} \quad (3.11)$$

bestimmt. Um Aussagen über die Existenz und Eindeutigkeit einer solchen Lösung  $\|q^*\|$  zu treffen, benötigen wir folgendes Lemma.

**Lemma 3.7** *Es seien  $n \in \mathbb{N}$ ,  $\alpha \in (2^{-n+2}, 1)$  und es sei  $t_1$  in der Ungleichung (3.1) derart gewählt, dass die rechte Seite gleich  $\frac{\alpha}{4}$  ist. Darüber hinaus sei  $\|\bar{Z}_n\| > \sqrt{2}t_1$  und  $\rho = \sqrt{\|\bar{Z}_n\|^2 - t_1^2}$ . Dann ist*

$$f : [0, \rho) \rightarrow \mathbb{R}, \quad \|q\| \mapsto f(\|q\|) = \left[ \left( \frac{1 + \|q\|}{1 + \rho} \right)^{1+\rho} \left( \frac{1 - \|q\|}{1 - \rho} \right)^{1-\rho} \right]^{\frac{n}{2}} \quad (3.12)$$

strikt monoton wachsend in  $\|q\|$ .

**Beweis** Wir betrachten lediglich den Term innerhalb der eckigen Klammer, da dieser genau dann in  $\|q\|$  wächst, wenn die Aussage für (3.12) gilt. Dessen Ableitung nach  $\|q\|$  hat die Form

$$\begin{aligned} & \left( \frac{1 + \|q\|}{1 + \rho} \right)^\rho \left( \frac{1 - \|q\|}{1 - \rho} \right)^{1-\rho} - \left( \frac{1 + \|q\|}{1 + \rho} \right)^{1+\rho} \left( \frac{1 - \|q\|}{1 - \rho} \right)^{-\rho} \\ &= \left( \frac{1 + \|q\|}{1 + \rho} \right)^\rho \left( \frac{1 - \|q\|}{1 - \rho} \right)^{-\rho} \left( \frac{1 - \|q\|}{1 - \rho} - \frac{1 + \|q\|}{1 + \rho} \right). \end{aligned}$$

Wegen  $0 \leq \|q\| < \rho$  folgt dann  $\frac{1-\|q\|}{1-\rho} > 1$ , sowie  $0 < \frac{1+\|q\|}{1+\rho} < 1$  und somit

$$\left( \frac{1 + \|q\|}{1 + \rho} \right)^\rho \left( \frac{1 - \|q\|}{1 - \rho} \right)^{-\rho} \left( \frac{1 - \|q\|}{1 - \rho} - \frac{1 + \|q\|}{1 + \rho} \right) > 0.$$

Folglich ist  $\frac{\partial f}{\partial \|q\|} > 0$ . Dies zeigt die Behauptung. ■

**Bemerkung 3.8** Mit den Bezeichnungen aus Lemma 3.7 betrachten wir erneut die Gleichung (3.11). Dann gilt:

$$(i) \quad \lim_{\|q\| \rightarrow \rho} f(\|q\|) = 1,$$

$$(ii) \quad \lim_{\|q\| \rightarrow 0} f(\|q\|) = \left( \frac{1}{1 + \rho} \right)^{1+\rho} \left( \frac{1}{1 - \rho} \right)^{1-\rho}.$$

Der Term  $(1 + t)^{-(1+t)} (1 - t)^{-(1-t)}$  ist monoton fallend in  $t$  (vgl. den Beweis aus Lemma 3.1 für  $\mu = 0$ ) und nimmt für  $t = t_1$  den Wert  $\frac{\alpha}{4}$  an. Wegen  $\|\bar{Z}_n\| > \sqrt{2}t_1$  ist  $\rho > t_1$ . Somit ist der Ausdruck aus (ii) höchstens  $\frac{\alpha}{4}$ . Da  $f$  stetig und nach Lemma 3.7 strikt wachsend in  $\|q\|$  ist, folgt dann mit dem Zwischenwertsatz, dass es genau ein  $\|q^*\| \in [0, \rho)$ , sodass  $f(\|q^*\|) = \frac{\alpha}{4}$  gilt. Damit existiert stets genau eine Lösung  $\|q^*\|$  der Gleichung (3.11).

Man beachte, dass die Bestimmung von  $\|q^*\|$  als Lösung von (3.11) unabhängig von  $t_3$  erfolgt. Insbesondere folgt damit aus (3.10) die Existenz von  $t_3$  für  $\|q^*\|$  und somit für alle  $q$  mit  $\|q\| \leq \|q^*\|$ .

Zur Bestimmung des Konfidenzbereiches benötigen wir abschließend den Begriff des Winkels zwischen zwei Vektoren in der Ebene.

**Definition 3.9** Es seien  $x, y \in \mathbb{R}^2 \setminus \{0\}$  zwei Vektoren. Dann definieren wir durch

$$\angle(x, y) := \arccos \left( \frac{\langle x, y \rangle}{\|x\| \|y\|} \right) \in [0, \pi]$$

den *zwischen  $x$  und  $y$  eingeschlossenen Winkel*. Für  $x = 0$  und  $y \in \mathbb{R}^2$  definieren wir  $\angle(x, y) := 0$ .

Gegeben  $\|q^*\|$  können wir den Konfidenzbereich  $K$  als Projektion der Menge

$$\{q \in \bar{\mathbb{B}}_1(0) \mid \|q\| = \|q^*\|, \angle(q, \bar{Z}_n) \leq \delta\}$$

bestimmen, wobei

$$\delta = \arctan \left( \frac{t_1}{\sqrt{\|\bar{Z}_n\|^2 - t_1^2}} \right) = \arcsin \left( \frac{t_1}{\|\bar{Z}_n\|} \right)$$

gilt (vgl. Abbildung 3.3). Dann ergibt sich der Konfidenzbereich zu

$$K = \{x \in S_1 \mid \angle(x, \bar{Z}_n) \leq \delta\}. \quad (3.13)$$

**Bemerkung 3.10** Man beachte, dass für die Bestimmung des Konfidenzbereiches  $K$  eine Berechnung der Lösung  $\|q^*\|$  der Gleichung (3.11) nicht erforderlich ist.

Der folgende Algorithmus fasst die Konstruktion zusammen.

**Algorithmus 3.11** (Konfidenzbereich mit Hoeffdingscher Ungleichung) Gegeben seien  $n \in \mathbb{N}$  und  $\bar{Z}_n$ , sowie  $\alpha \in (2^{-n+2}, 1)$ .

(S1) Bestimme (numerisch)  $t_1$  als Lösung von

$$\left[ \left( \frac{1}{1+t_1} \right)^{1+t_1} \left( \frac{1}{1-t_1} \right)^{1-t_1} \right]^{\frac{n}{2}} = \frac{\alpha}{4}.$$

(S2) Ist  $\|\bar{Z}_n\| \leq \sqrt{2}t_1$ , so ist  $K = S_1$ . Andernfalls gehe zu (S3).

(S3) Berechne  $\delta = \arcsin(t_1 \|\bar{Z}_n\|^{-1})$ . Der  $(1 - \alpha)$ -Konfidenzbereich ist dann gegeben durch

$$K = \{x \in S_1 \mid \angle(x, \bar{Z}_n) \leq \delta\}.$$

**Bemerkung 3.12** Aus Algorithmus 3.11 folgt, dass der Öffnungswinkel  $\delta$  eines mit Hilfe der Hoeffdingschen Ungleichung konstruierten Konfidenzbereiches  $K$  höchstens  $\arcsin\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{4}$  ist, sofern  $\|\bar{Z}_n\| > \sqrt{2}t_1$  gilt. In diesem Fall ist der Konfidenzbereich  $K$  damit nie größer als ein Viertelkreis.

Abschließend in Abschnitt 3.1 wollen wir unsere Resultate im folgenden Satz festhalten.

**Satz 3.13** *Es seien  $n \in \mathbb{N}$ ,  $\alpha \in (2^{-n+2}, 1)$  und es seien  $Z, Z_1, Z_2, \dots, Z_n$  unabhängig und identisch verteilte  $S^1$ -wertige Zufallsvariablen. Dann ist die nach Algorithmus 3.11 konstruierte Menge  $K$  ein  $(1 - \alpha)$ -Konfidenzbereich für den extrinsischen Erwartungswert  $\varepsilon$  der Zufallsvariable  $Z$ . Darüber hinaus gelten die folgenden Aussagen:*

(i) *Ist  $\|\bar{Z}_n\| > \sqrt{2}t_1$ , so ist*

$$K = \left\{ x \in S^1 \mid \angle(x, \bar{Z}_n) \leq \arcsin\left(\frac{t_1}{\|\bar{Z}_n\|}\right) \right\}.$$

*Andernfalls ist  $K = S^1$ .*

(ii) *Ist  $\mathbb{E}Z = 0$ , so gilt  $\|\bar{Z}_n\| > \sqrt{2}t_1$  höchstens mit Wahrscheinlichkeit  $\alpha$ . Es gilt folglich  $K = S^1$  mindestens mit Wahrscheinlichkeit  $1 - \alpha$ .*

(iii) *Ist  $\mathbb{E}Z \neq 0$ , so gilt*

$$\mathbb{P}\left(\|\bar{Z}_n\| \leq \sqrt{2}t_1\right) \leq e^{-\frac{\alpha}{2}(\|\mathbb{E}Z\| - \sqrt{2}t_1)^2}.$$

*Die Wahrscheinlichkeit in diesem Fall den trivialen Konfidenzbereich  $K = S^1$  zu erhalten, konvergiert demnach mit exponentieller Rate gegen Null für  $n \rightarrow \infty$ .*

**Beweis** Die Beweisidee ist aus [HK14, Proposition 5] übernommen. Die Aussagen (i) und (ii) folgen unmittelbar aus den Konstruktionsschritten 3.1.1 und 3.1.2. Für (iii) sei  $Y_i := \left\langle \bar{Z}_i, \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|} \right\rangle$  für  $i = 1, 2, \dots, n$  und  $\bar{Y}_n := \left\langle \bar{Z}_n, \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|} \right\rangle$ . Die Zufallsgrößen  $Y_i$  sind dann unabhängig und identisch verteilt. Dann geht die Aussage (iii) aus der folgenden Ungleichungskette hervor:

$$\begin{aligned} \mathbb{P}\left(\|\bar{Z}_n\| \leq \sqrt{2}t_1\right) &\leq \mathbb{P}\left(|\bar{Y}_n| \leq \sqrt{2}t_1\right) \\ &\leq \mathbb{P}\left(\bar{Y}_n \leq \sqrt{2}t_1\right) \\ &= \mathbb{P}\left(-\bar{Y}_n + \|\mathbb{E}Z\| \geq \|\mathbb{E}Z\| - \sqrt{2}t_1\right) \\ &\leq e^{-\frac{\alpha}{2}(\|\mathbb{E}Z\| - \sqrt{2}t_1)^2}. \end{aligned}$$

Dabei folgt die erste Ungleichung aus der Cauchy-Schwarzschen Ungleichung:

$$\|\bar{Z}_n\| \geq \frac{\left| \left\langle \bar{Z}_n, \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|} \right\rangle \right|}{\left\| \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|} \right\|} = \left| \left\langle \bar{Z}_n, \frac{\mathbb{E}Z}{\|\mathbb{E}Z\|} \right\rangle \right| = |\bar{Y}_n|. \quad (3.14)$$

Für die letzte Abschätzung überlege man sich, dass  $\bar{Y}_n$  eine Zufallsvariable mit Werten aus  $[-1, 1]$  und Erwartungswert  $\|\mathbb{E}Z\|$  ist; ersteres ergibt sich mit (3.14). Somit gilt  $-1 \leq -\bar{Y}_n \leq 1$  und  $\mathbb{E}(-\bar{Y}_n) = -\|\mathbb{E}Z\|$ . Nun folgt mit Bemerkung 2.7 die Abschätzung

$$\mathbb{P}\left(-\bar{Y}_n + \mathbb{E}\bar{Y}_n \geq \|\mathbb{E}Z\| - \sqrt{2}t_1\right) \leq e^{-\frac{\alpha}{2}(\|\mathbb{E}Z\| - \sqrt{2}t_1)^2}.$$

Dies zeigt die dritte Ungleichung. Da  $\|\mathbb{E}Z\| > 0$  und  $t_1 \xrightarrow{n \rightarrow \infty} 0$  gilt, folgt somit die Aussage (iii). ■

## 3.2 Konfidenzbereiche mit Varianz

Ziel dieses Abschnittes ist es, kleinere Konfidenzbereiche als in Abschnitt 3.1 zu konstruieren, in dem wir den Öffnungswinkel  $\delta$  verkleinern. Dazu wollen wir für  $t_1$  bessere Abschätzungen mit Hilfe der in Abschnitt 2.2 vorgestellten Ungleichungen erzielen. Diese berücksichtigen zusätzlich die Varianz. Wie in Abschnitt 2.3 bereits bemerkt, erhoffen wir uns davon, dass die so konstruierten Konfidenzbereiche kleiner werden. Die grundlegende Idee des vorangegangenen Abschnittes wird jedoch beibehalten.

### 3.2.1 Konfidenzbereiche mit der Tschebyscheffschen Ungleichung

Wir betrachten zunächst die einfachste Konstruktionsvorschrift, welche wir unter Einbezug der Varianz verfolgen können – mit Hilfe der Tschebyscheffschen Ungleichung. Diese soll uns später zu Vergleichszwecken dienen. Insbesondere wird hier die Hoeffdingsche Ungleichung nicht verwendet. Die Konstruktionen und Resultate in diesem Abschnitt sind dem Artikel [HK14, Abschnitt 2] entnommen.

Betrachten wir die Ungleichung (2.8), so erhalten wir in natürlicher Weise Tests für die Hypothese  $\mathbb{E}Z = q$  zum Niveau  $\alpha$ . Der Annahmebereich

$$A = \left\{ q \in \mathbb{B}_1(0_2) \mid \|\bar{Z}_n - q\|^2 < \frac{1 - \|q\|^2}{n\alpha} \right\}$$

des Tests ist dabei ein  $(1 - \alpha)$ -Konfidenzbereich für den Erwartungswert  $\mathbb{E}Z$  in der Ebene. Durch Projektion auf den Einheitskreis  $S^1$  erhalten wir dann einen  $(1 - \alpha)$ -Konfidenzbereich für den extrinsischen Erwartungswert. Genauer ist dieser gegeben durch die Projektion eines Balls um  $\bar{Z}_n$  mit Radius

$$\varepsilon = \sqrt{\frac{1}{n\alpha} \left( 1 - \|\bar{Z}_n\|^2 + \frac{1}{n\alpha} \right)},$$

wie in [HK14, Korollar 3] gezeigt wird. Wir erhalten das folgende Resultat:

**Satz 3.14** *Es sei  $\alpha \in (0, 1)$ . Dann ist durch die Projektion  $K$  des Annahmebereiches  $A$  auf den Einheitskreis ein  $(1 - \alpha)$ -Konfidenzbereich für den extrinsischen Erwartungswert  $\varepsilon$  gegeben. Darüber hinaus gelten die folgenden Aussagen:*

(i) *Ist  $\|\bar{Z}_n\|^2 \geq \frac{1}{n\alpha}$ , so ist der Konfidenzbereich gegeben durch*

$$K = \left\{ x \in S^1 \mid \angle(x, \bar{Z}_n) \leq \arcsin \left( \sqrt{\frac{1 - \|\bar{Z}_n\|^2 + \frac{1}{n\alpha}}{n\alpha\|\bar{Z}_n\|^2}} \right) \right\}.$$

*Andernfalls ist  $K = S^1$ .*

(ii) *Ist  $\mathbb{E}Z = 0$ , so gilt  $\|\bar{Z}_n\|^2 \geq \frac{1}{n\alpha}$  mit Wahrscheinlichkeit höchstens  $\alpha$ . Folglich erhalten wir den Konfidenzbereich  $K = S^1$  in diesem Fall mindestens mit Wahrscheinlichkeit  $1 - \alpha$ .*

(iii) *Ist  $\mathbb{E}Z \neq 0$ , so gilt*

$$\mathbb{P} \left( \|\bar{Z}_n\| \leq \frac{1}{n\alpha} \right) \leq e^{-\frac{n}{2} (\|\mathbb{E}Z\| - \frac{1}{n\alpha})^2}.$$

*Die Wahrscheinlichkeit den trivialen Konfidenzbereich  $K = S^1$  zu erhalten konvergiert somit mit exponentieller Rate gegen Null für  $n \rightarrow \infty$ .*

**Beweis** Zum Beweis verweisen wir auf [HK14, Proposition 5]. ■

**Bemerkung 3.15** Die Berechnung des Konfidenzbereiches gestaltet sich in dieser Konstruktion besonders einfach, da sich gegeben  $n$ ,  $\bar{Z}_n$  und  $\alpha$ , der Öffnungswinkel  $\delta$  explizit berechnen lässt.

### 3.2.2 Konfidenzbereiche mit Hoeffdings Theorem 3

**Bemerkung 3.16** Benutzen wir für die Bestimmung von  $t_1$  in Konstruktion 3.1.1 anstatt der Hoeffdingschen Ungleichung, die Ungleichung (2.9) aus Hoeffdings Theorem 3, so können wir zusätzlich die Varianz der Zufallsgröße ausnutzen. Ist  $Z = (Z_1, Z_2)$  eine  $S_1$ -wertige Zufallsvariable, so gilt

$$\text{Var } Z_2 \leq \text{Var } Z_1 + \text{Var } Z_2 = \text{Var } Z = 1 - \|\mathbb{E}Z\|^2.$$

Da wir unter der Hypothese den Erwartungswert  $\mathbb{E}Z = q$  kennen, können wir  $1 - \|q\|^2$  als obere Schranke für  $\sigma^2$  benutzen. Es ist zu vermuten, dass  $t_1$  dann groß wird, wenn  $\|q\|$  sehr klein ist, d.h. falls die Varianz von  $Z$  besonders groß ist. Dies zeigt folgendes Lemma.

**Lemma 3.17** *Für  $\sigma^2 \in (0, 1)$ ,  $\alpha \in (0, 1)$  und  $n \in \mathbb{N}$  erfülle  $t \in (0, 1)$  die Gleichung*

$$\left[ \left( 1 + \frac{t}{\sigma^2} \right)^{-\frac{\sigma^2+t}{1+\sigma^2}} (1-t)^{-\frac{1-t}{1+\sigma^2}} \right]^n = \alpha. \quad (3.15)$$

*Dann ist  $t$  monoton wachsend in  $\sigma^2$ .*

**Beweis** Wir substituieren  $\tau = \sigma^2$  und betrachten die logarithmierte Gleichung

$$\frac{\tau + t}{1 + \tau} \ln \left( 1 + \frac{t}{\tau} \right) + \frac{1 - t}{1 + \tau} \ln(1 - t) = -\frac{\ln \alpha}{n}. \quad (3.16)$$

Dann ist  $t$  implizit als Lösung dieser Gleichung gegeben. Da die Ableitung der linken Seite nach  $t$  und  $\tau$  existiert und stetig ist, sichert uns der Satz über implizite Funktionen die Existenz von  $\frac{\partial t}{\partial \tau}$ , sofern die Ableitung der linken Seite nach  $t$  nicht verschwindet.

Diese hat die Gestalt

$$\frac{1}{1 + \tau} \left[ \ln \left( 1 + \frac{t}{\tau} \right) + \frac{\tau + t}{\tau \left( 1 + \frac{t}{\tau} \right)} - \ln(1 - t) - 1 \right] = \frac{1}{1 + \tau} \left[ \ln \left( 1 + \frac{t}{\tau} \right) - \ln(1 - t) \right],$$

da  $\tau$  nicht von  $t$  abhängt. Dann verschwindet die Ableitung der linken Seite nach  $t$  genau dann, wenn

$$\begin{aligned} 1 + \frac{t}{\tau} = 1 - t &\iff \frac{t}{\tau} + t = 0 \\ &\iff t \left( 1 + \frac{1}{\tau} \right) = 0. \end{aligned}$$

Dies ist wegen  $t, \tau > 0$  aber nicht möglich. Dann folgt mit dem Satz über implizite Funktionen die Existenz von  $t' = \frac{\partial t}{\partial \tau}$ .

Für die Ableitung von (3.16) nach  $\tau$  gilt

$$\frac{(1 + t')(1 + \tau) - (\tau + t)}{(1 + \tau)^2} \ln \left( 1 + \frac{t}{\tau} \right) + \frac{t'\tau - t}{\tau(1 + \tau)} - \frac{t'(1 + \tau) + (1 - t)}{(1 + \tau)^2} \ln(1 - t) - \frac{t'}{1 + \tau} = 0. \quad (3.17)$$

Lösen wir die Gleichung (3.17) nach  $t'$  auf, so ergibt sich nach Multiplikation mit  $(1 + \tau)$ :

$$\begin{aligned} t' \left[ \ln \left( 1 + \frac{t}{\tau} \right) - \ln(1 - t) \right] &= \frac{\tau + t}{1 + \tau} \ln \left( 1 + \frac{t}{\tau} \right) + \frac{t}{\tau} + \frac{1 - t}{1 + \tau} \ln(1 - t) - \ln \left( 1 + \frac{t}{\tau} \right) \\ &= \frac{t}{\tau} - \frac{1 - t}{1 + \tau} \left[ \ln \left( 1 + \frac{t}{\tau} \right) - \ln(1 - t) \right] \\ &= \frac{t}{\tau} - \frac{1 - t}{1 + \tau} \ln \left( \frac{\tau + t}{\tau(1 - t)} \right). \end{aligned}$$

Wegen  $0 < \tau, t < 1$  gilt

$$\ln \left( 1 + \frac{t}{\tau} \right) - \ln(1 - t) > 0.$$

Dann ist  $t' \geq 0$  genau dann, wenn

$$\frac{t}{\tau} \geq \frac{1-t}{1+\tau} \ln \left( \frac{\tau+t}{\tau(1-t)} \right)$$

gilt. Wegen  $\ln(1+x) \leq x$  für  $x \in \mathbb{R}$  folgt

$$\ln \left( \frac{\tau+t}{\tau(1-t)} \right) = \ln \left( 1 + \frac{t(1+\tau)}{\tau(1-t)} \right) \leq \frac{t(1+\tau)}{\tau(1-t)}$$

und somit

$$\frac{1-t}{1+\tau} \ln \left( \frac{\tau+t}{\tau(1-t)} \right) \leq \frac{1-t}{1+\tau} \cdot \frac{t(1+\tau)}{\tau(1-t)} = \frac{t}{\tau}.$$

Da  $\frac{\partial t}{\partial \sigma^2} \geq 0$  ist, gilt somit, dass  $t$  monoton wachsend in  $\sigma^2$  ist. ■

**Bemerkung 3.18** Für  $\sigma^2 \in (0, 1)$ ,  $\alpha \in (0, 1)$  und  $n \in \mathbb{N}$  betrachte man für  $t \in (0, 1)$  die Gleichung

$$\left[ \left( 1 + \frac{t}{\sigma^2} \right)^{-\frac{\sigma^2+t}{1+\sigma^2}} (1-t)^{-\frac{1-t}{1+\sigma^2}} \right]^n = \alpha. \quad (3.18)$$

Dann existiert die Lösung  $t^*$  dieser Gleichung, falls  $\left( 1 + \frac{1}{\sigma^2} \right)^{-n} < \alpha$  gilt und ist im Falle ihrer Existenz eindeutig bestimmt. Insbesondere folgt die Existenz aus der Bedingung  $2^{-n} < \alpha$ . Mit der Beweisidee aus Lemma 3.1 kann man diese Aussage entsprechend zeigen. Wir werden im Folgenden davon ausgehen, dass  $n$  hinreichend groß ist, damit die Existenz der Lösung  $t^*$  gewährleistet ist.

**Konstruktion 3.19** Im Folgenden sei o.B.d.A. stets  $q = (\|q\|, 0)^\top$  und es gelte  $\frac{\alpha}{4} > 2^{-n}$ . Wir betrachten erneut den Konstruktionsschritt 3.1.2. Dabei setzen wir  $\sigma^2 = 1 - \|q\|^2$  und wählen  $t_1$  gemäß (3.15) mit rechter Seite  $\frac{\alpha}{4}$ . Wegen  $\text{Var } Z_2 \leq 1 - \|q\|^2$  führt dies zu einer konservativen Abschätzung. Nach Lemma 3.17 ist  $t_1$  fallend in  $\|q\|$ . Somit tritt das Problem auf, dass  $\|q\| + t_3$  in  $\|q\|$  zwar wächst, dafür  $t_1$  jedoch fällt. Das Verhalten von  $r = \sqrt{(\|q\| + t_3)^2 + t_1^2}$  ist daher ungewiss.

Um Aussagen über den Radius  $r$  treffen zu können, betrachten wir eine Zerlegung von  $[0, 1)$  in  $m$  Teilintervalle  $I_j := \left[ \frac{j-1}{m}, \frac{j}{m} \right)$ ,  $j = 1, 2, \dots, m$ . Auf jedem Teilintervall  $I_j$  schätzen wir  $t_1$  als Lösung von (3.15) mit rechter Seite  $\frac{\alpha}{4}$  durch seinen Wert an der linken Intervallgrenze ab; dort ist  $t_1$  nach Lemma 3.17 am größten. Für jedes Intervall  $I_j$  können wir nun mit der in Abschnitt 3.1 beschriebenen Konstruktion einen Konfidenzbereich  $K_j$  bestimmen. Als Konfidenzbereich für  $\varepsilon$  betrachten wir die Vereinigung  $K = \bigcup_{j=1}^m K_j$  der einzelnen Konfidenzbereiche. Dabei ist  $K$  gerade der Konfidenzbereich  $K_j$  mit dem größten Öffnungswinkel  $\delta_j$ . Es stellt sich nun die Frage, in welchem Intervall  $I_j$  der Öffnungswinkel  $\delta_j$  maximal wird.

Der Winkel

$$\delta = \arctan \left( \frac{t_1}{\|q\| + t_3} \right)$$

wächst monoton im Argument. Dabei ist  $t_1$  monoton fallend in  $\|q\|$ , während  $\|q\| + t_3$  monoton wachsend in  $\|q\|$  ist. Wegen  $t_1, \|q\| + t_3 > 0$  ist dann  $\frac{t_1}{\|q\| + t_3}$  monoton fallend in  $\|q\|$ . Wir suchen demnach das kleinste  $\|q\|$ , so dass  $\bar{Z}_n \in M(q)$ . Dann ist der Winkel  $\delta$  am größten.

Es sei  $t_{1,j}$  für  $j \in \{1, 2, \dots, m\}$  diejenige Lösung von (3.15) für  $\sigma^2 = 1 - \left(\frac{j-1}{m}\right)^2$  und rechter Seite  $\frac{\alpha}{4}$ . Nach Bemerkung 3.18 existiert dann  $t_{1,j}$  und ist eindeutig bestimmt. Dieses ist zudem der maximale Wert von  $t_1$  im Intervall  $I_j$ . Verwenden wir im Intervall  $I_j$  anstatt  $t_1$  den Wert  $t_{1,j}$ , so führt dies zu einer konservativen Abschätzung.

Man beachte, dass wir für diese Konstruktion  $\|\bar{Z}_n\| > \max\{\sqrt{2}t_{1,1}, \sqrt{2}t_1\}$  fordern müssen, wobei  $t_1$  Lösung von (3.1) mit rechter Seite  $\frac{\alpha}{4}$  ist. Andernfalls ist  $\bar{Z}_n \in M(0)$  und die Projektion ergibt die gesamte Kreislinie als Konfidenzbereich. Wegen Bemerkung 2.11 (ii) ist  $t_1 = t_{1,1}$ . Die Menge  $M(0)$  ist demnach ein Kreis um 0 mit Radius  $\sqrt{2}t_{1,1}$ .

Es sei nun  $\|\bar{Z}_n\| > \sqrt{2}t_{1,1}$ . Nach obigen Überlegungen ist der Öffnungswinkel  $\delta_j$  im Intervall  $I_j$  mit dem kleinsten linken Endpunkt, in dem  $\bar{Z}_n \in M(q)$  gilt, am größten. Letzteres ist nach dem Konstruktionsschritt 3.1.2 dann der Fall, wenn  $\|\bar{Z}_n\|$  auf einem Kreisbogen mit entsprechendem Radius  $r_j = \sqrt{(\|q\| + t_3)^2 + t_{1,j}^2}$  liegt.

Daher prüfen wir für  $j = 1, 2, \dots, m$ , ob es ein  $\|q_j^*\| \in I_j$  gibt, das die Gleichung

$$\|q\| + t_3 = \sqrt{\|\bar{Z}_n\|^2 - t_{1,j}^2} \quad (3.19)$$

löst und suchen das kleinste dieser Art. Existiert in den Intervallen  $I_1, I_2, \dots, I_j$  keine Lösung  $\|q_k^*\| \in I_k$ ,  $k = 1, 2, \dots, j$ , so ist  $\|\bar{Z}_n\| > r_j$ , d.h.  $\|\bar{Z}_n\|$  ist größer als der maximale Radius im Intervall  $I_j$ . In diesem Fall gilt  $\bar{Z}_n \notin M(q)$  für alle  $q \in I_j$  (vgl. Abbildung 3.4) und somit  $\bar{Z}_n \notin M(q)$  für alle  $q \in [0, \frac{j}{m})$ .

Analog zum Konstruktionsschritt 3.1.2 können wir  $\|q_j^*\| \in [0, \rho_j)$  statt als Lösung von (3.19) auch als Lösung von

$$\left[ \left( \frac{1 + \|q\|}{1 + \rho_j} \right)^{1+\rho_j} \left( \frac{1 - \|q\|}{1 - \rho_j} \right)^{1-\rho_j} \right]^{\frac{n}{2}} = \frac{\alpha}{4} \quad (3.20)$$

mit  $\rho_j = \sqrt{\|\bar{Z}_n\|^2 - t_{1,j}^2}$  bestimmen. Die Existenz der Lösung  $\|q_j^*\| \in [0, \rho_j)$  in (3.20) folgt nach Bemerkung 3.8. Somit existiert auch der entsprechende Wert für  $t_3$  in (3.19) und ist eindeutig bestimmt.

Ist  $\|q_j^*\| \notin I_j$ , so gilt  $\|q_j^*\| \geq \frac{j}{m}$ , sofern es in den Intervallen  $I_1, I_2, \dots, I_{j-1}$  keine Lösung  $\|q_k^*\| \in I_k$ ,  $k < j$  gegeben hat. Dies folgt, da  $\rho_j$  auf dem Intervall  $I_j$  konstant ist und die linke Seite von (3.20) somit nach Lemma 3.7 strikt monoton in  $\|q\|$  wächst. Es reicht somit für  $j = 1, 2, \dots, m$  zu prüfen, ob  $\|q_j^*\| < \frac{j}{m}$  gilt. Ist dies der Fall, so ist  $\|q_j^*\| \in \left[\frac{j-1}{m}, \frac{j}{m}\right)$ , falls es für  $k < j$  keine entsprechende Lösung  $\|q_k^*\| \in I_k$  gegeben hat.

Wir können unsere Überlegungen wie folgt zusammenfassen: Der Öffnungswinkel  $\delta_j$  ist im Intervall mit dem kleinsten linken Endpunkt  $I_\ell$ , in dem die Gleichung (3.19) eine Lösung  $\|q_\ell^*\| \in I_\ell$  besitzt am größten. Daher prüfen wir für  $j = 1, 2, \dots, m$ , ob für die Gleichung (3.20) eine Lösung  $\|q_j^*\| < \frac{j}{m}$  existiert, solange diese Bedingung nicht erfüllt ist. Das entsprechende Intervall, in dem die Gleichung (3.20) zum ersten Mal eine Lösung  $\|q_\ell^*\| \in \left[\frac{\ell-1}{m}, \frac{\ell}{m}\right)$  besitzt, sei  $I_\ell$ .

Der größte Öffnungswinkel  $\delta = \delta_\ell$  hat dann die Form

$$\delta = \arctan \left( \frac{t_{1,\ell}}{\sqrt{\|\bar{Z}_n\|^2 - t_{1,\ell}^2}} \right) = \arcsin \left( \frac{t_{1,\ell}}{\|\bar{Z}_n\|} \right).$$

Analog zu (3.13) ist der Konfidenzbereich  $K$  von der Gestalt

$$K = \left\{ x \in S_1 \mid \angle(x, \bar{Z}_n) \leq \delta \right\}.$$

Zur Bestimmung können wir den folgenden Algorithmus verwenden.

**Algorithmus 3.20** (Konfidenzbereich mit Hoeffdings Theorem 3) Gegeben seien  $n \in \mathbb{N}$  und  $\bar{Z}_n$ , sowie  $\alpha \in (2^{-n+2}, 1)$  und  $m \in \mathbb{N}$  die Anzahl der Intervalle.

(S1) Berechne  $t_{1,1}$  als Lösung von

$$\left[ (1+t)^{1+t} (1-t)^{1-t} \right]^{-\frac{n}{2}} = \frac{\alpha}{4}.$$

Ist  $\|\bar{Z}_n\| \leq \sqrt{2}t_{1,1}$ , so ist  $K = S_1$ . Andernfalls setze  $j = 1$  und gehe zu (S2).

(S2) Bestimme  $t_{1,j}$  als Lösung von

$$\left[ \left( 1 + \frac{t}{\sigma^2} \right)^{-\frac{\sigma^2+t}{1+\sigma^2}} (1-t)^{-\frac{1-t}{1+\sigma^2}} \right]^n = \frac{\alpha}{4} \quad (3.21)$$

für  $\sigma^2 = 1 - \left(\frac{j-1}{m}\right)^2$  und  $\|q_j^*\| \in \left[\frac{j-1}{m}, \rho_j\right)$  als Lösung von

$$\left[ \left( \frac{1+\|q\|}{1+\rho_j} \right)^{1+\rho_j} \left( \frac{1-\|q\|}{1-\rho_j} \right)^{1-\rho_j} \right]^{\frac{n}{2}} = \frac{\alpha}{4}$$

für  $\rho_j = \sqrt{\|\bar{Z}_n\|^2 - t_{1,j}^2}$ . Prüfe, ob  $\|q_j^*\| < \frac{j}{m}$  gilt. Falls ja gehe zu (S3), andernfalls setze  $j \leftarrow j + 1$  und iteriere (S2).

(S3) Es sei  $t_{1,\ell}$  die Lösung von (3.21) für das kleinste  $\ell$ , sodass  $\|q_\ell^*\| < \frac{\ell}{m}$  erfüllt ist. Berechne

$$\delta = \arcsin\left(\frac{t_{1,\ell}}{\|\bar{Z}_n\|}\right).$$

Dann ist der  $(1 - \alpha)$ -Konfidenzbereich gegeben durch

$$K = \left\{x \in S^1 \mid \angle(x, \bar{Z}_n) \leq \delta\right\}.$$

**Bemerkung 3.21** Aufgrund von Bemerkung 2.11 (ii) sind die Bedingungen in (S1) von Algorithmus 3.11 bzw. Algorithmus 3.20 identisch. Wir erhalten demnach den trivialen Konfidenzbereich  $K = S^1$  in Algorithmus 3.20 genau dann, wenn der nach Algorithmus 3.11 konstruierte Konfidenzbereich trivial ist.

**Bemerkung 3.22** Die Aussagen aus Bemerkung 3.10 und Bemerkung 3.12 gelten für einen mit Algorithmus 3.20 konstruierten Konfidenzbereich entsprechend.

Zusammenfassend können wir folgendes Resultat formulieren.

**Satz 3.23** *Es seien  $n \in \mathbb{N}$ ,  $\alpha \in (2^{-n+2}, 1)$  und es seien  $Z, Z_1, Z_2, \dots, Z_n$  unabhängige, identisch verteilte,  $S^1$ -wertige Zufallsvariablen. Dann ist die mit Algorithmus 3.20 konstruierte Menge  $K$  ein  $(1 - \alpha)$ -Konfidenzbereich für den extrinsischen Erwartungswert  $\varepsilon$  der Zufallsvariable  $Z$ . Darüber hinaus gelten die folgenden Aussagen:*

(i) *Ist  $\|\bar{Z}_n\| > \sqrt{2}t_{1,1}$ , so ist*

$$K = \left\{x \in S^1 \mid \angle(x, \bar{Z}_n) \leq \arcsin\left(\frac{t_{1,\ell}}{\|\bar{Z}_n\|}\right)\right\},$$

*wobei  $t_{1,\ell}$  entsprechend (S3) in Algorithmus 3.20 gewählt werde. Andernfalls ist  $K = S^1$ .*

(ii) *Ist  $\mathbb{E}Z = 0$ , so gilt  $\|\bar{Z}_n\| > \sqrt{2}t_{1,1}$  höchstens mit Wahrscheinlichkeit  $\alpha$ . Es gilt folglich  $K = S^1$  mindestens mit Wahrscheinlichkeit  $1 - \alpha$ .*

(iii) *Ist  $\mathbb{E}Z \neq 0$ , so gilt*

$$\mathbb{P}\left(\|\bar{Z}_n\| \leq \sqrt{2}t_{1,1}\right) \leq e^{-\frac{n}{2}(\|\mathbb{E}Z\| - \sqrt{2}t_{1,1})^2}.$$

*Folglich konvergiert die Wahrscheinlichkeit, in diesem Fall den trivialen Konfidenzbereich  $K = S^1$  zu erhalten, mit exponentieller Rate gegen Null für  $n \rightarrow \infty$ .*

**Beweis** Die Aussagen (i) und (ii) folgen aus den Konstruktionsschritten 3.19. Die Aussage (iii) ist identisch mit Satz 3.13 (iii). ■

### 3.2.3 Konfidenzbereiche mit Hoeffdingscher Ungleichung zweiter Ordnung

Abschließend wollen wir in diesem Abschnitt einen Konfidenzbereich mit der in Abschnitt 2.2.3 gezeigten Abschätzung konstruieren. Für die Konstruktion benötigen wir jedoch noch die folgende Monotonieaussage.

**Lemma 3.24** *Es seien  $\sigma^2 \in (0, 1)$ ,  $\alpha \in (0, 1)$  und  $n \in \mathbb{N}$ . Mit den Bezeichnungen von Theorem 2.16 sei  $t \in (0, 1)$  Lösung der Gleichung*

$$\left[ e^{-\bar{h}t} \left( \bar{p}_2 \sigma^2 + \bar{p}_0 \right) \right]^n = \alpha.$$

Dann hängt  $t$  monoton wachsend von  $\sigma^2$  ab.

Man beachte, dass hier  $\bar{h}$  und somit die Koeffizienten  $\bar{p}_2$  und  $\bar{p}_0$  von  $\sigma^2$  abhängen.

**Beweis** Für jedes  $h > 0$  und  $t \in (0, 1)$  gilt nach Theorem 2.16 die Ungleichung

$$\begin{aligned} \mathbb{P} \left( \bar{X}_n \geq t \right) &\leq \left[ e^{-ht} \left( p_2 \sigma^2 + p_0 \right) \right]^n \\ &= \left[ \frac{e^{-h(t+1)}}{4} \left( \left( e^{2h} - 2h - 1 \right) \sigma^2 + e^{2h} + 2h + 3 \right) \right]^n =: f_t(h, \sigma^2). \end{aligned} \quad (3.22)$$

Dann ist die rechte Seite wegen  $e^{2h} - 2h - 1 > 0$  strikt wachsend in  $\sigma^2$ , d.h. es gilt  $f_t(h, \sigma^2) < f_t(h, \tau^2)$  für  $\sigma^2 < \tau^2$ . Somit folgt

$$\inf_{h>0} f_t(h, \sigma^2) \leq \inf_{h>0} f_t(h, \tau^2). \quad (3.23)$$

Setzen wir  $g(\sigma^2, t) := \inf_{h>0} f_t(h, \sigma^2)$ , so ist  $g$  wegen (3.23) monoton wachsend in  $\sigma^2$ . Insbesondere gilt dann für  $\sigma^2 \leq \tau^2$  die Ungleichung  $g(\sigma^2, t) \leq g(\tau^2, t)$  für alle  $t \in (0, 1)$ . Für festes  $\sigma^2$  ist zudem  $g(\sigma^2, t)$  strikt fallend in  $t$ , wie man mit (3.22) einsieht.

Es seien nun  $t_{\sigma^2}, t_{\tau^2}$  Lösungen von  $g(\sigma^2, t_{\sigma^2}) = \alpha$  bzw.  $g(\tau^2, t_{\tau^2}) = \alpha$  für  $\sigma^2 < \tau^2$ . Da  $g$  fallend in  $t$  ist, können wir dann  $t_{\sigma^2} \leq t_{\tau^2}$  folgern. Folglich ist  $t$  monoton wachsend in  $\sigma^2$ . ■

**Bemerkung 3.25** Für  $\sigma^2 \in (0, 1)$ ,  $\alpha \in (0, 1)$  und  $n \in \mathbb{N}$  betrachte man für  $t \in (0, 1)$  die Gleichung

$$\left[ e^{-\bar{h}t} \left( \bar{p}_2 \sigma^2 + \bar{p}_0 \right) \right]^n = \alpha. \quad (3.24)$$

Dann lassen sich entsprechende Existenz- und Eindeutigkeitsaussagen für deren Lösung  $t^*$  wie in Bemerkung 3.18 treffen, die insbesondere für  $2^{-n} < \alpha$  erfüllt sind. Wir gehen im Folgenden stets davon aus, dass  $n$  derart groß ist, dass die Lösung  $t^*$  von (3.24) existiert.

**Konstruktion 3.26** Wir gehen wie in Konstruktion 3.19 vor, wobei wir zur Bestimmung von  $t_1$  statt der Ungleichung aus Hoeffdings Theorem 3 die Hoeffdingsche Ungleichung zweiter Ordnung 2.16 benutzen. Modifizieren wir den Algorithmus 3.20 wie folgt, so erhalten wir einen Konfidenzbereich mit der Hoeffdingschen Ungleichung zweiter Ordnung.

**Algorithmus 3.27** (Konfidenzbereich mit Hoeffdingscher Ungleichung zweiter Ordnung) Gegeben seien  $n \in \mathbb{N}$  und  $\bar{Z}_n$ , sowie  $\alpha \in (2^{-n+2}, 1)$  und  $m \in \mathbb{N}$  die Anzahl der Intervalle.

(S1) Berechne  $t_{1,1}$  als Lösung von

$$\left[ e^{-\bar{h}t} (\bar{p}_2 + \bar{p}_0) \right]^n = \frac{\alpha}{4}$$

mit  $\bar{h}$ ,  $\bar{p}_2$  und  $\bar{p}_0$  entsprechend Theorem 2.16. Ist  $\|\bar{Z}_n\| \leq \sqrt{2}t_{1,1}$ , so ist  $K = S_1$ . Andernfalls setze  $j = 1$  und gehe zu (S2).

(S2) Verwende in Algorithmus 3.20  $t_{1,j}$  als Lösung von

$$\left[ e^{-\bar{h}t} (\bar{p}_2 \sigma^2 + \bar{p}_0) \right]^n = \frac{\alpha}{4}$$

statt (3.21) mit  $\sigma^2 = 1 - \left(\frac{j-1}{m}\right)^2$ . Verfahre sonst analog zu den dort beschriebenen Schritten.

**Bemerkung 3.28** Mit Bemerkung 2.17 folgt, dass der mit Algorithmus 3.27 genau dann trivial ist, wenn wir in Algorithmus 3.11 den trivialen Konfidenzbereich  $K = S^1$  erhalten.

Darüber hinaus gelten für den mit Algorithmus 3.27 konstruierten Konfidenzbereich die Aussagen aus Bemerkung 3.22 und Satz 3.23 entsprechend.

### 3.3 Asymptotische Konfidenzbereiche

In diesem Abschnitt wollen wir asymptotische Konfidenzbereiche für unabhängig und identisch verteilte Zufallsgrößen  $Z, Z_1, Z_2, \dots, Z_n$  mit Werten auf dem Einheitskreis konstruieren. Diese vergleichen wir später mit den nicht-asymptotischen Konfidenzbereichen. Für die Konstruktion verwenden wir die Aussage eines zentralen Grenzwertsatzes. Daher betrachten wir den *vorzeichenbehafteten Winkel*, welcher für  $x, y \in \mathbb{R}^2$  durch  $\text{sgn}(\langle x, y \rangle) \angle(x, y)$  gegeben ist. Dabei sei  $\text{sgn}(\langle x, y \rangle)$  das *Vorzeichen* von  $\langle x, y \rangle \in \mathbb{R}$ . Es gilt dann der folgende zentrale Grenzwertsatz

$$\sqrt{n} \text{sgn}(\langle \bar{Z}_n, e_1 \rangle) \angle(\bar{Z}_n, e_1) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\mathbb{E} \sin^2 \angle(Z, e_1)}{\|\mathbb{E} Z\|}\right), \quad (3.25)$$

sofern der extrinsische Erwartungswert  $\varepsilon$  eindeutig mit  $\angle(\varepsilon, e_1) = 0$  und  $\bar{Z}_n \neq 0$  ist, siehe z.B. [Hot13a, Theorem 5]. Dabei sei  $e_1 = (1, 0)^\top$  der erste kanonische Einheitsvektor.

**Satz 3.29** *Es seien  $Z, Z_1, Z_2, \dots, Z_n$  unabhängige, identisch verteilte,  $S^1$ -wertige Zufallsvariablen und für den Erwartungswert  $\mathbb{E}Z$  in der Ebene gelte  $\mathbb{E}Z \neq 0$ . Dann gilt der folgende zentrale Grenzwertsatz:*

$$\sqrt{n} \operatorname{sgn}(\langle \bar{Z}_n, e_1 \rangle) \angle(\bar{Z}_n, e_1) \left( \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n)}{\|\bar{Z}_n\|}} \right)^{-1} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

**Beweis** Wir führen den Beweis in mehreren Schritten:

(1) Wir zeigen zunächst die Konvergenz von

$$\frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n) \xrightarrow{\text{f.s.}} \mathbb{E} \sin^2 \angle(Z, \mathbb{E}Z).$$

Wegen  $\mathbb{E}Z \neq 0$  ist  $\bar{Z}_n \neq 0$  fast sicher für  $n \rightarrow \infty$ . Für  $i = 1, 2, \dots, n$  gilt dann:

$$\begin{aligned} \sin^2 \angle(Z_i, \bar{Z}_n) &= 1 - \cos^2 \angle(Z_i, \bar{Z}_n) \\ &= 1 - \left( \frac{\langle Z_i, \bar{Z}_n \rangle}{\|Z_i\| \|\bar{Z}_n\|} \right)^2 \\ &= 1 - \frac{1}{\|\bar{Z}_n\|^2} \left( (Z_i)_1^2 (\bar{Z}_n)_1^2 + (Z_i)_2^2 (\bar{Z}_n)_2^2 + 2(Z_i)_1 (Z_i)_2 (\bar{Z}_n)_1 (\bar{Z}_n)_2 \right) \end{aligned} \quad (3.26)$$

Mit [Hot13b, Proposition 3.35] können wir

$$\frac{1}{n} \sum_{i=1}^n (Z_i)_1 \xrightarrow{\text{f.s.}} \mathbb{E}(Z)_1, \quad \frac{1}{n} \sum_{i=1}^n (Z_i)_1^2 \xrightarrow{\text{f.s.}} \mathbb{E}(Z)_1^2, \quad \frac{1}{n} \sum_{i=1}^n (Z_i)_1 (Z_i)_2 \xrightarrow{\text{f.s.}} \mathbb{E}((Z)_1 (Z)_2)$$

folgern; dabei sei  $(Z_i)_j$  die  $j$ -te Komponente von  $Z_i$ . Somit folgt mit  $\bar{Z}_n \xrightarrow{\text{f.s.}} \mathbb{E}Z$  die fast sichere Konvergenz von

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (Z_i)_1^2 (\bar{Z}_n)_1^2 &\xrightarrow{\text{f.s.}} \mathbb{E}(Z)_1^2 (\mathbb{E}Z)_1^2, \\ \frac{1}{n} \sum_{i=1}^n (Z_i)_1 (Z_i)_2 (\bar{Z}_n)_1 (\bar{Z}_n)_2 &\xrightarrow{\text{f.s.}} \mathbb{E}((Z)_1 (Z)_2) (\mathbb{E}Z)_1 (\mathbb{E}Z)_2 \end{aligned}$$

aus dem starken Gesetz der großen Zahlen. Mit (3.26) schließen wir dann auf die fast sichere Konvergenz von

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n) &\xrightarrow{\text{f.s.}} 1 - \frac{1}{\|\mathbb{E}Z\|^2} \left( \mathbb{E}(Z)_1^2 (\mathbb{E}Z)_1^2 + \mathbb{E}(Z)_2^2 (\mathbb{E}Z)_2^2 + 2 \mathbb{E}((Z)_1 (Z)_2) (\mathbb{E}Z)_1 (\mathbb{E}Z)_2 \right) \\ &= 1 - \frac{\mathbb{E} \langle Z, \mathbb{E}Z \rangle^2}{\|\mathbb{E}Z\|^2} \\ &= \mathbb{E} \sin^2 \angle(Z, \mathbb{E}Z). \end{aligned}$$

(2) Mit (1) und [Hot13b, Proposition 3.35] können wir nun folgern, dass durch

$$\frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n) / \|\bar{Z}_n\|$$

ein stark konsistenter Schätzer für  $\mathbb{E} \sin^2 \angle(Z, \mathbb{E}Z) / \|\mathbb{E}Z\|$  gegeben ist.

(3) Sind  $X_n, Y_n$  zwei Folgen von Zufallsvariablen mit  $X_n \xrightarrow{\mathcal{D}} X$  und  $Y_n \xrightarrow{\mathcal{D}} c$ , wobei  $X$  eine Zufallsvariable und  $c \in \mathbb{R}$  eine Konstante ist, so folgt aus Slutskys Theorem [Leh99, Theorem 2.3.3] die Konvergenz  $Y_n X_n \xrightarrow{\mathcal{D}} cX$ .

(4) Dann folgt mit (2), (3) und (3.25) der folgende zentrale Grenzwertsatz:

$$\sqrt{n} \operatorname{sgn}(\langle \bar{Z}_n, e_1 \rangle) \angle(\bar{Z}_n, e_1) \left( \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n)}{\|\bar{Z}_n\|}} \right)^{-1} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

■

**Bemerkung 3.30** Durch

$$\left\{ x \in \bar{\mathbb{B}}_1(0) \mid \angle(x, \bar{Z}_n) \leq q_{1-\frac{\alpha}{2}} \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n)}{n \|\bar{Z}_n\|}} \right\}$$

ist somit nach Satz 3.29 ein asymptotisches zweiseitiges  $(1-\alpha)$ -Konfidenzintervall für  $\angle(\varepsilon, e_1)$  gegeben; dabei sei  $q_{1-\frac{\alpha}{2}}$  das  $(1-\frac{\alpha}{2})$ -Quantil der Standardnormalverteilung. Für den Öffnungswinkel  $\delta$  gilt demnach

$$\delta = q_{1-\frac{\alpha}{2}} \sqrt{\frac{\frac{1}{n} \sum_{i=1}^n \sin^2 \angle(Z_i, \bar{Z}_n)}{n \|\bar{Z}_n\|}}.$$

## 4 Anwendung

In diesem Kapitel bestimmen wir für zwei Beispiele die in Kapitel 3 konstruierten Konfidenzbereiche. Diese vergleichen wir sowohl untereinander als auch mit den asymptotischen Konfidenzbereichen aus Abschnitt 3.3. Anhand der Konfidenzbereiche schließen wir dann entsprechend der Problemstellung.

### 4.1 Strömungsrichtungen in Schrägschichtungen

In der Geologie versucht man die Strömungsrichtung von Flüssen in der geologischen Vergangenheit zu bestimmen. Dazu kann man z.B. die Verlaufsrichtung von Schrägschichtungen betrachten.

Als *Schrägschichtungen* (engl. cross-bedding) bezeichnet man Ablagerungen von Sediment, wie Sand oder Kies in schräg einfallenden Schichten. Sind die übereinander liegenden Schichten unterschiedlich oder gar senkrecht orientiert, so spricht man auch von *Kreuzschichtung*, siehe Abbildung 4.1. Als Ursache dafür werden wechselnde Strömungsverhältnisse bei der Ablagerung genannt. Man kann demnach die Orientierung einer solchen Aufschüttung dazu benutzen, die Richtung der verursachenden Strömung zu ermitteln. Dazu betrachtet man



Abbildung 4.1: Schrägschichtung von Sandsteinsediment in den Vermilion Cliffs, Arizona. Quelle: [Geo14]

Geländequerschnitte und notiert die Orientierung des Flussbettes, wobei man normalerweise die Flussrichtung flussabwärts richtet. Es handelt sich also um orientierte Daten.

Bei der Bestimmung allgemeiner Landschaftsbilder vergangener geologischer Zeitalter ist von Interesse, in welche Richtung das Wasser im Mittel geflossen ist. Man sucht also einen Konfidenzbereich für den extrinsischen Erwartungswert der Strömungsrichtungen.

Wir betrachten einen Datensatz mit  $n = 104$  Messungen in Schrägschichtungen von Sedimenten in Pakistan. Die Daten sind unter [Fis93, S. 250, Beispiel B.16, Set 2] einsehbar und unter dem Namen `cross.beds2` im Paket `NPCirc` in R [R C14] verfügbar. Wir werden nach den Abschnitten 3.1 bzw. 3.2 jeweils einen Konfidenzbereich konstruieren. Dazu wurden die oben aufgeführten Algorithmen in der Programmiersprache R implementiert und auf dem Datensatz `cross.beds2` getestet. Der dazugehörige Quellcode ist im Anhang A zu finden.

Wir konstruieren die Konfidenzbereiche jeweils zum Niveau 95% und 90%. Zusätzlich zum Öffnungswinkel des Konfidenzbereiches  $2\delta$  in Grad werden wir die Bogenlänge  $\ell$  des konstruierten Konfidenzbereiches angeben, welche sich zu

$$\ell = \frac{4\pi\delta}{360^\circ} \quad (4.1)$$

ergibt. Die Winkel sind mit zweistelliger Genauigkeit angegeben. Eine höhere Genauigkeit hätte allenfalls numerische Signifikanz, würde aber die Längen der Konfidenzbereiche nur marginal verändern und ist somit aus anwendungsorientierter Sicht vernachlässigbar. Für die Konfidenzbereiche mit Theorem 3 bzw. der Hoeffdingschen Ungleichung zweiter Ordnung wurde die Diskretisierung  $m = 1000$  verwendet.

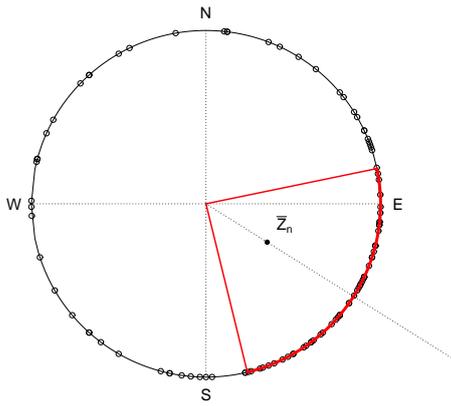
	$\alpha = 0,05$		$\alpha = 0,1$	
	Winkel $2\delta$	Länge $\ell$	Winkel $2\delta$	Länge $\ell$
Hoeffding	88,03°	1,536	79,32°	1,384
Tschebyscheff	360°	6,283	91,88°	1,604
Theorem 3	88,02°	1,536	79,19°	1,382
Hoeffding 2. Ordnung	88,02°	1,536	79,19°	1,382
Asymptotisch	20,87°	0,364	17,52°	0,306

Tabelle 4.1: Öffnungswinkel und Längen der konstruierten Konfidenzbereiche

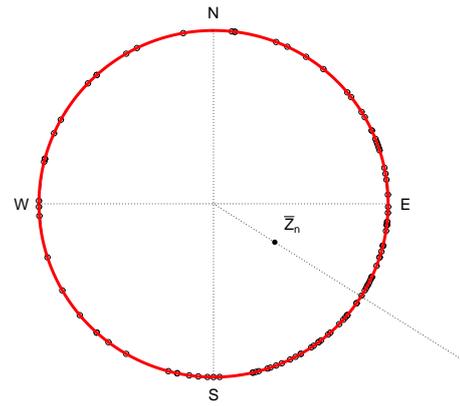
Für alle mit Hilfe von Varianten der Hoeffdingschen Ungleichung konstruierten Konfidenzbereiche ist  $\|\bar{Z}_n\|$  groß genug, um keinen trivialen Konfidenzbereich zu erhalten. Da jedoch  $\|\bar{Z}_n\|^2 = 0,172 < 0,192 = \frac{1}{n\alpha}$  gilt, liefert die Tschebyscheffsche Ungleichung die gesamte Kreislinie  $S^1$  als Konfidenzbereich. Die Konstruktion mit der Tschebyscheffschen Ungleichung benötigt demnach mehr Daten für einen nicht-trivialen Konfidenzbereich, als die anderen Varianten.

Die konstruierten Konfidenzbereiche sowie die Daten wurden geplottet und sind in Abbildung 4.2 zu sehen. Dabei wurden die Konfidenzbereiche auf der Kreislinie und der gesamte Kreissektor farbig markiert.

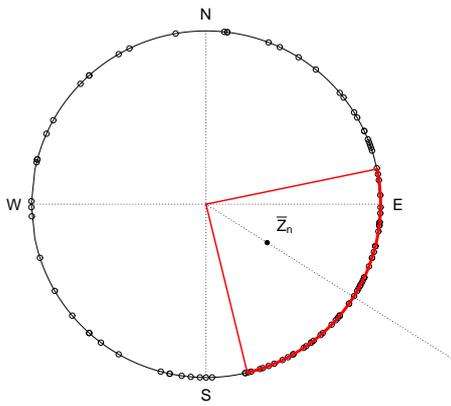
Man kann mit großer Sicherheit schließen, dass die mittlere Strömungsrichtung in süd-östlicher Richtung liegt.



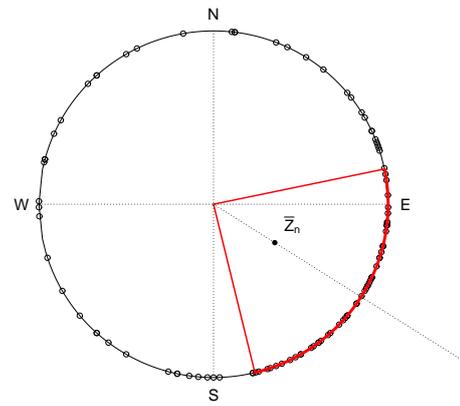
(a) Hoeffdingsche Ungleichung



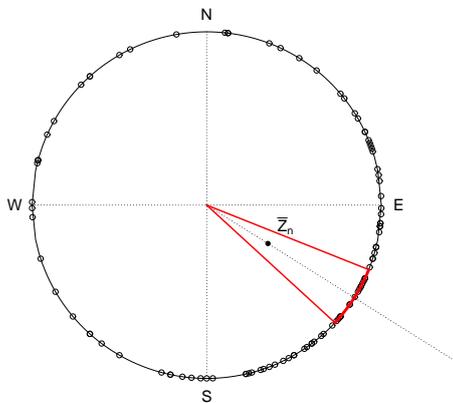
(b) Tschebyscheffsche Ungleichung



(c) Theorem 3



(d) Hoeffdingsche Ungleichung zweiter Ordnung



(e) Asymptotisch

Abbildung 4.2: Datensatz `cross.beds2` ( $\circ$ ) und mit verschiedenen Ungleichungen konstruierte Konfidenzbereiche zum Niveau 95% (jeweils rot markiert)

## 4.2 Bewegungsrichtungen von Ameisen

In einem Experiment betrachtet man die Bewegungsrichtungen von  $n = 100$  Ameisen. Dabei ist man an deren Reaktion interessiert, die sie auf ein gleichmäßig beleuchtetes, schwarzes Ziel zeigen. Wir definieren Norden als  $0^\circ$  und messen die Winkel im Uhrzeigersinn, d.h. Osten entspricht  $90^\circ$  usw. Dann ist das Ziel zwischen  $170^\circ$  und  $190^\circ$  aufgestellt, vgl. Abbildung 4.3. Es stellt sich die Frage, ob sich der extrinsische Erwartungswert in diesem Intervall befindet. Dann könnte man darauf schließen, dass die Ameisen eine Reaktion auf das Ziel zeigen und sich in dessen Richtung bewegen.

Wir betrachten dazu einen Datensatz aus [Fis93, S. 243, Beispiel B.7]. Dieser ist ebenfalls im Paket `CircNNTSR` unter dem Namen `Ants_radians` zu finden. Wir konstruieren jeweils einen Konfidenzbereich zu den Niveaus 95% und 90% mit Hilfe des in Anhang A zu findenden Programms. Auch hier werden wir für die Ungleichungen mit Theorem 3 bzw. der Hoeffdingschen Ungleichung zweiter Ordnung die Diskretisierung  $m = 1000$  verwenden. Wir erhalten:

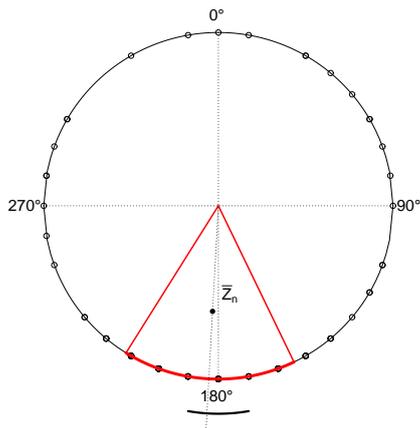
	$\alpha = 0,05$		$\alpha = 0,1$	
	Winkel $2\delta$	Länge $\ell$	Winkel $2\delta$	Länge $\ell$
Hoeffding	$57,59^\circ$	1,005	$52,52^\circ$	0,917
Tschebyscheff	$83,67^\circ$	1,460	$52,49^\circ$	0,916
Theorem 3	$55,59^\circ$	0,970	$50,13^\circ$	0,875
Hoeffding 2. Ordnung	$55,60^\circ$	0,970	$50,15^\circ$	0,875
Asymptotisch	$15,01^\circ$	0,262	$12,59^\circ$	0,220

Tabelle 4.2: Öffnungswinkel und Längen der konstruierten Konfidenzbereiche

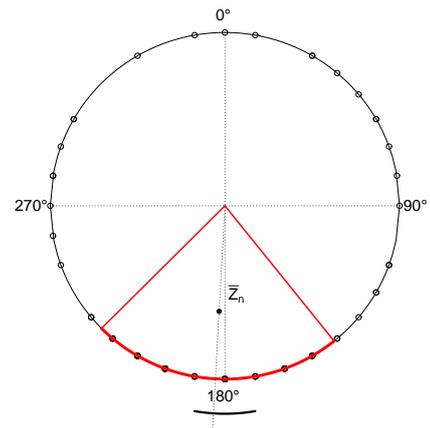
In diesem Beispiel sind die Daten deutlich konzentrierter als in Abschnitt 4.1. Es gilt  $\|\bar{Z}_n\| = 0,610$ , während im vorherigen Abschnitt  $\|\bar{Z}_n\| = 0,415$  galt. Alle in diesem Beispiel konstruierten Konfidenzbereiche sind außerdem nicht-trivial.

Der Mittelwert  $\bar{Z}_n$  schließt mit der Nordrichtung den Winkel  $\angle(\bar{Z}_n, e_2) = 183,14^\circ$  (im Uhrzeigersinn) ein. Die Bewegungsrichtungen (bzw. die Winkel der Richtungen) der Ameisen streuen dann um  $\pm\delta$  um diesen Winkel.

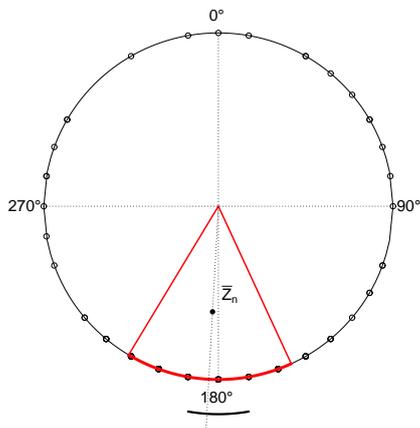
Die Daten haben Werte aus  $\{0^\circ, 10^\circ, \dots, 350^\circ\}$ . Um Bindungen besser sichtbar zu machen, wurden die Winkel für den Plot mit der Funktion `jitter` leicht gestört. Die Skizzen der Konfidenzbereiche in Abbildung 4.3 legen nun nahe, dass die Ameisen eine Reaktion auf das beleuchtete, markierte Ziel zeigen und sich in dessen Richtung bewegen.



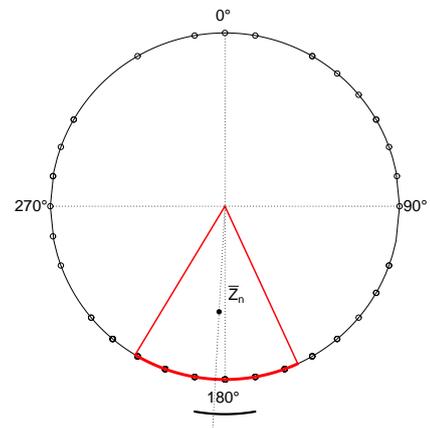
(a) Hoeffdingsche Ungleichung



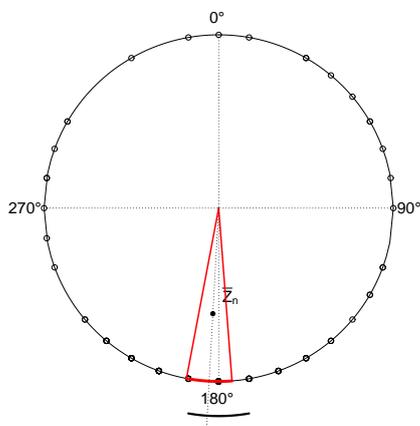
(b) Tschebyscheffsche Ungleichung



(c) Theorem 3



(d) Hoeffdingsche Ungleichung zweiter Ordnung



(e) Asymptotisch

Abbildung 4.3: Datensatz `Ants_radians` ( $\circ$ ) und mit verschiedenen Ungleichungen konstruierte Konfidenzbereiche zum Niveau 95% (jeweils rot markiert), Ziel bei  $180^\circ$  ist schwarz markiert

### 4.3 Vergleich der Konfidenzbereiche

Wir betrachten zunächst die Konfidenzbereiche zum Niveau 95%. Dabei fällt im ersten Beispiel 4.1 auf, dass die Norm des arithmetischen Mittels  $\bar{Z}_n$  zu klein ist, um mittels der Tschebyscheffschen Ungleichung einen nicht-trivialen Konfidenzbereich zu erhalten, während alle anderen konstruierten Konfidenzbereiche nicht-trivial sind. Die Tschebyscheffsche Ungleichung benötigt demnach für einen nicht-trivialen Konfidenzbereich mehr Daten. Die Konfidenzbereiche, welche mit Hilfe der Hoeffdingschen Ungleichung, der Hoeffdingschen Ungleichung zweiter Ordnung bzw. Theorem 3 konstruiert wurden, unterscheiden sich nur marginal. Dies suggerierte bereits Abbildung 2.1a in Kapitel 2, da  $1 - \|\bar{Z}_n\|^2 = 0,828$  verhältnismäßig groß ist.

Ein ähnliches Ergebnis liefert das in Abschnitt 4.2 betrachtete Beispiel. Hier sind die Daten allerdings stärker konzentriert. Die Konstruktion mit Tschebyscheffscher Ungleichung ergibt keinen trivialen Konfidenzbereich, dieser ist jedoch immer noch mindestens  $26^\circ$  größer als die anderen Konfidenzbereiche. Die Hoeffdingsche Ungleichung zweiter Ordnung bietet diesmal einen wesentlich besseren Konfidenzbereich als die Hoeffdingsche Ungleichung. Die Konstruktion mit Theorem 3 ergibt das beste Resultat.

Betrachten wir nun die Konfidenzbereiche zum Niveau 90%. Qualitativ hat sich im ersten Beispiel an den Ergebnissen nichts verändert. Lediglich der mittels Tschebyscheffscher Ungleichung konstruierte Konfidenzbereich ist nun nicht-trivial, jedoch immer noch erheblich größer als die anderen drei.

Im zweiten Beispiel ist der mit Hilfe der Tschebyscheffschen Ungleichung konstruierte Konfidenzbereich nun sogar besser, als jener mit der Hoeffdingschen Ungleichung konstruierte. Dies hat sich in Abbildung 2.1b bereits angedeutet, da die Daten hier ziemlich stark konzentriert sind. Nichtsdestotrotz sind die, mit der Hoeffdingschen Ungleichung zweiter Ordnung und mit Theorem 3 konstruierten Konfidenzbereiche kleiner, letzterer ist der beste.

Es ist zudem zu bemerken, dass alle Konfidenzbereiche, welche mit Varianten der Hoeffdingschen Ungleichung konstruiert wurden, einen Öffnungswinkel  $2\delta < 90^\circ$  besitzen. Für die Konstruktion mit Hilfe der Tschebyscheffschen Ungleichung ist dies nicht der Fall, wie Beispiel 4.1 für  $\alpha = 0,1$  zeigte. Die Aussagen aus den Bemerkungen 3.12, 3.22 und 3.28 werden bestätigt.

Die asymptotischen Konfidenzbereiche sind in allen Beispielen um ein Vielfaches kleiner als die nicht-asymptotischen Konfidenzbereiche. Diese enthalten den extrinsischen Erwartungswert aber nur in etwa mit Wahrscheinlichkeit  $1 - \alpha$ . Eine genauere Aussage ist aufgrund der Tatsache, dass der in Bemerkung 3.30 konstruierte Konfidenzbereich das Niveau  $1 - \alpha$  nur asymptotisch einhält, wir aber nur  $n$  Daten zur Verfügung haben, nicht möglich.

## 5 Fazit und Ausblick

Die verschiedenen Varianten der Hoeffdingschen Ungleichung haben sich als sehr nützlich bei der Konstruktion von nicht-asymptotischen Konfidenzbereichen für den extrinsischen Erwartungswert  $\varepsilon$  erwiesen. Allerdings entfalten sie ihre volle Wirkung erst für große Anzahlen von Daten  $n$  oder wenn diese stark konzentriert sind. Sind nur wenige Daten gegeben bzw.  $\alpha$  zu klein, können wir die Konstruktion ggf. nicht anwenden, wie wir in Aussagen wie Lemma 3.1 festgestellt haben. Ist  $\|\bar{Z}_n\| \lll 1$ , so können wir für kleine  $n$  meist den Ursprung nicht ausschließen, da  $t_1$  (bzw.  $t_{1,1}$ ) in den Algorithmen zu groß ist, sodass wir den trivialen Konfidenzbereich auch für  $\mathbb{E}Z \neq 0$  erhalten.

Des Weiteren ist zu bemerken, dass der Konfidenzbereich, welcher mit der Hoeffdingschen Ungleichung zweiter Ordnung konstruiert wurde, besser im Sinne von kleiner ist, als der mit der Hoeffdingschen Ungleichung konstruierte Konfidenzbereich, wenn auch im ersten Beispiel nur marginal. Die Konstruktion nach Hoeffdings Theorem 3 lieferte in allen Beispielen jeweils das beste Ergebnis. Dies ließen uns bereits die Grafiken aus Abschnitt 2.3 schließen. Es ist davon auszugehen, dass für stark konzentrierte Daten sich noch größere Unterschiede zwischen den einzelnen Konstruktionen aufzeigen werden.

Mit einem entsprechend feinen Gitter, d.h. für eine hinreichend große Anzahl an Intervallen  $m$ , kann man in (S3) der Algorithmen 3.20 bzw. 3.27 das Intervall  $I_\ell$  (bzw.  $\|q_\ell^*\|$ ) und damit  $t_{1,\ell}$  sowie  $\delta_\ell$  beliebig genau bestimmen.

Es ist zudem zu bemerken, dass die Konfidenzbereiche, die mit Hilfe von Varianten der Hoeffdingschen Ungleichung konstruiert wurden, eine Länge von  $\ell = \frac{\pi}{2}$  nicht überschreiten, sofern sie nicht-trivial sind.

Noch kleiner sind lediglich die asymptotischen Konfidenzbereiche. Bei den in Kapitel 4 betrachteten Beispielen waren diese im Vergleich zu den nicht-asymptotischen Konfidenzbereichen nur in etwa ein Viertel so groß. Da wir für die Konstruktion nur endlich viele Daten zur Verfügung haben, halten erstere das Niveau jedoch nicht exakt ein und enthalten den extrinsischen Erwartungswert  $\varepsilon$  nur ungefähr mit Wahrscheinlichkeit  $1 - \alpha$ .

---

Man beachte, dass alle konstruierten Konfidenzbereiche, insbesondere die Konfidenzbereiche mit Varianz, konservativ sind. Durch eine andere Gestalt der Mengen  $M(q)$  kann man möglicherweise exaktere Konfidenzbereiche konstruieren.

Ferner kann man auch höhere Momente in Betracht ziehen und versuchen, Lemma 2.12 zu verbessern, indem man kubische, biquadratische oder Polynome höheren Grades betrachtet, welche zwischen der Exponentialfunktion und der Parabel liegen. Dann kann man die Abschätzung der Hoeffdingschen Ungleichung, wie in Theorem 2.16, weiter verschärfen. Für die Konstruktion eines Konfidenzbereiches müsste man dann entsprechende Monotonieeigenschaften zeigen.

Weiterführend kann man auch Daten in höheren Dimensionen, etwa  $(k + 1)$ -dimensionale Daten, verteilt auf der  $k$ -Sphäre  $S^k = \{v \in \mathbb{R}^{k+1} \mid \|v\| = 1\}$ , betrachten. Hierzu verweisen wir auf [HK14]. Dort werden Konfidenzbereiche mit Hilfe der Tschebyscheffschen Ungleichung konstruiert, wie in Abschnitt 3.2.1 bereits angesprochen wurde, und darüber hinaus Daten in projektiven Räumen  $\mathbb{RP}^k$  bzw.  $\mathbb{CP}^k$  betrachtet.

---

# A Anhang

```

1 # Konfidenzbereich mit Hoeffdingscher Ungleichung
2 #
3 ##### Definition von Funktionen #####
4
5 # Bestimme t in der entsprechenden rechten Seite
6 t.alpha <- function(f, eps = 1e-6) Vectorize(function(mu, alpha, ...) if(abs
  (mu) == 1) 0 else tryCatch(
7   uniroot(function(t) f(t, mu, ...) - alpha/4, lower = eps, upper = 1 - mu -
  eps, f.upper = -alpha)$root,
8   error = function(e) {warning(e$message); print(c(mu = mu, alpha = alpha));
  NA}))
9
10 # Einfache Hoeffdingsche Ungleichung
11 hoeff <- function(t, mu, n) {
12   ( ( (mu+1)/(mu+1+t) )^(mu+1+t) * ( (1-mu)/(1-mu-t) )^(1-mu-t) )^(n/2)
13 }
14
15 # Hoeffdingsche Ungleichung zweiter Ordnung
16 hoeff2 <- function(t, mu, n, sigmaq, inf = 1e2) {
17   rlhs <- function(h) ( ( 1 + t ) * ( (sigmaq)*(2*h + 1) - 2*h -3 ) - 2 *
  (sigmaq - 1) ) / (( t - 1 ) * ( sigmaq + 1 )) - exp(2 * h)
18   h <- tryCatch(uniroot(rlhs, lower = 0, upper = inf)$root)
19   ( exp(-h*(1 + t))/4 * ( ( exp(2*h)-2*h-1) * sigmaq + exp(2*h)+2*h+3 ) )^n
20 }
21
22 # Hoeffdings Theorem 3
23 theorem3 <- function(t, mu, n, sigmaq) {
24   ((1+t/sigmaq)^(-(sigmaq+t)/(1+sigmaq)) * (1-t)^(-(1-t)/(1+sigmaq)))^n
25 }
26
27 # Bestimme q
28 getq <- function(rho){
29
30   qstern <- uniroot(function(q){
31     (((1+q)/(1+rho))^(1+rho) * ((1-q)/(1-rho))^(1-rho))^n - alpha/4
32   }, lower=0, upper=rho)$root
33   return(qstern)
34 }

```

```
35 # Plot-Funktion
36 zeichnen <- function(delta, title=" ", sub=" ") {
37
38   # Zeichnen der Datenpunkte
39   plot(kart, xlab=" ", ylab=" ", pch=1, bty="n", xaxt="n", yaxt="n", xlim=c
40         (-1.2,1.2), ylim=c(-1.2,1.2), main=title, sub=sub, asp=1)
41
42   # Kreis zeichnen
43   x <- seq(-1, 1, 1e-3)
44   y <- seq(-1, 1, 1e-3)
45
46   lines(x,y, xlim = c(-1,1), ylim=c(-1,1), type='n', asp = 1)
47   curve(( 1 * (1 - x^2)^0.5 ), add=TRUE, from=-1 , to =1)
48   curve((-1 * (1 - x^2)^0.5 ), add=TRUE, from=-1 , to =1)
49
50   # Konfidenzbereich zeichnen auf der Kreislinie
51   k <- matrix(nrow = round(2*delta/1e-3)+1, ncol = 2)
52   j<-1
53
54   for (i in seq(from=-delta, to=delta, by=1e-3)){
55     k[j,1] <- sin(i)
56     k[j,2] <- cos(i)
57     j<- j+1
58   }
59
60   M <- matrix(nrow = nrow(k), ncol = 2)
61   for(i in seq(from=1, to=nrow(k))){
62     M[i,1]<- sin(arg)*k[i,1]+cos(arg)*k[i,2]
63     M[i,2]<- -cos(arg)*k[i,1]+sin(arg)*k[i,2]
64   }
65
66   points(M, col="red", type="l", lwd=4)
67
68   # Einzeichnen von Znquer
69   mwert <- matrix(nrow = 2, ncol = 2)
70   mwert[1,1] <- 0
71   mwert[1,2] <- 0
72   mwert[2,1] <- 1/znquerq*znquer[1]
73   mwert[2,2] <- 1/znquerq*znquer[2]
74   points(mwert, type="l", lty=3)
75   points(znquer[1],znquer[2], pch=16)
76   text((znquer[1]+0.1),(znquer[2]+0.1), expression(bar(Z)[n]))
77
78   # Plot der Kreisradien
79   axes1 <- matrix(nrow = 2, ncol = 2)
80   axes1[1,1] <- 0
81   axes1[1,2] <- 0
82   axes1[2,1] <- sin(arg)*sin(-delta)+cos(arg)*cos(-delta)
```

```
81 axes1[2,2] <- -cos(arg)*sin(-delta)+sin(arg)*cos(-delta)
82 points(axes1, col="red", type="l", lty=1, lwd=2)
83
84 axes2 <- matrix(nrow = 2, ncol = 2)
85 axes2[1,1] <- 0
86 axes2[1,2] <- 0
87 axes2[2,1] <- sin(arg)*sin(delta)+cos(arg)*cos(delta)
88 axes2[2,2] <- -cos(arg)*sin(delta)+sin(arg)*cos(delta)
89 points(axes2, col="red", type="l", lty=1, lwd=2)
90
91 # Himmelsrichtungen
92 text(0,1.1, "N")
93 text(0,-1.1, "S")
94 text(1.1,0, "E")
95 text(-1.1,0, "W")
96
97 # Einzeichnen der Achsen
98 ns <- matrix(nrow = 2, ncol = 2)
99 ns[1,1] <- 0
100 ns[1,2] <- 1
101 ns[2,1] <- 0
102 ns[2,2] <- -1
103 points(ns, type="l", lty=3)
104
105 we <- matrix(nrow = 2, ncol = 2)
106 we[1,1] <- -1
107 we[1,2] <- 0
108 we[2,1] <- 1
109 we[2,2] <- 0
110 points(we, type="l", lty=3)
111
112 }
113
114 ##### Anwendung #####
115 # Beispiel – A dataset of cross-beds measurements from Himalayan molasse in
116 # Pakistan presented in Fisher (1993). This dataset collects 104
117 # measurements of Chaudan Zam large bedforms. From R-Package NPCirc
118
119 # install.packages("NPCirc")
120 # library(NPCirc)
121 # data(cross.beds2)
122 # n <- nrow(cross.beds2)
123 # angles <- t(cross.beds2)
124
125 ##### Anwendung #####
126 # Beispiel – Directions chosen by 100 ants in response to an evenly
127 # illuminated black target. From R-Package CircNNTSR
```

```
125 # install.packages("CircNNTSR")
126 library(CircNNTSR)
127 data(Ants_radians)
128 n <- length(Ants_radians)
129 angles <- Ants_radians
130
131 m <- 1000
132 # alpha <- 0.05
133 alpha <- 0.1
134 eps <- 1e-6
135
136 # Berechnung von Znquer
137
138 kart <- matrix(nrow=length(angles), ncol=2)
139 for (i in 1:length(angles)) {
140   kart[i,1] <- sin(angles[i])
141   kart[i,2] <- cos(angles[i])
142 }
143
144 znquer <- c(mean(kart[,1]), mean(kart[,2]))
145 znquerq <- sum(abs(znquer)^2)
146
147 # Argument des Mittelwertes
148 arg <- atan2(znquer[2], znquer[1])
149
150 ##### Normaler Hoeffding #####
151
152 # Berechnung von t_1
153 t1 <- t.alpha(hoeff)(mu=0, alpha=alpha, n=n)
154
155 if(sqrt(znquerq) >=sqrt(2)*t1) {
156
157   # Berechnung des (halben) Oeffnungswinkels
158   deltaH <- asin(t1/sqrt(znquerq))
159 } else deltaH <- pi
160 zeichnen(deltaH)
161
162 ##### Tschebyscheff #####
163
164 if (znquerq >= 1/(n*alpha)) {
165   deltaT <- asin(sqrt( (1-znquerq+(1/(n*alpha))) / (n*alpha*znquerq) ))
166 } else deltaT <- pi
167
168 zeichnen(deltaT)
169
170
171
```

```
172 ##### Theorem 3 #####
173
174 # Berechnung von t11
175 t11 <- t.alpha(theorem3)(mu=0, alpha=alpha, n=n, sigmaq=1)
176
177 if (sqrt(znquerq) >= sqrt(2)*t11) {
178
179   # Berechnung des passenden tij
180   intervalT3 <- 1
181   for (j in 1:m) {
182
183     sigmaq <- 1- ((j-1)/m)^2
184
185     t1j <- t.alpha(theorem3)(mu=0, alpha=alpha, n=n, sigmaq=sigmaq)
186     rhoj <- sqrt(znquerq-t1j^2)
187
188     critqj <- getq(rhoj)
189
190     if(critqj < j/m) {
191       intervalT3 <- j
192       t1 <- t1j
193       break
194     }
195
196   }
197
198   # Berechnung des (halben) Oeffnungswinkels
199   deltaT3 <- asin(t1/sqrt(znquerq))
200 } else deltaT3 <- pi
201
202 zeichnen(deltaT3)
203
204 ##### Hoeffding zweiter Ordnung #####
205
206 # Berechnung von t11
207 t11 <- t.alpha(hoeff2)(mu=0, alpha=alpha, n=n, sigmaq=1)
208
209 if (sqrt(znquerq) >= sqrt(2)*t11) {
210
211   # Berechnung des passenden tij
212   intervalH2 <- 1
213   for (j in 1:m) {
214
215     sigmaq <- 1- ((j-1)/m)^2
216
217     t1j <- t.alpha(hoeff2)(mu=0, alpha=alpha, n=n, sigmaq=sigmaq)
218     rhoj <- sqrt(znquerq-t1j^2)
```

```
219   critqj <- getq(rhoj)
220
221   if(critqj < j/m) {
222     intervalH2 <- j
223     t1 <- t1j
224     break
225   }
226
227 }
228
229 # Berechnung des (halben) Oeffnungswinkels
230 deltaH2 <- asin(t1/sqrt(znquerq))
231
232 } else deltaH2 <- pi
233 zeichnen(deltaH2)
234
235 ##### Winkel und Laengen #####
236 delta <- c(deltaH, deltaT, deltaT3, deltaH2)
237 # Oeffnungswinkel
238 delta*360/pi
239 # Bogenlaengen
240 delta*2
```

# Literaturverzeichnis

- [AE10] AMANN, Herbert ; ESCHER, Joachim: *Analysis*. Bd. 1. 3. Aufl., 2. Nachdr. Basel [u.a.] : Birkhäuser, 2010.
- [Aul12] AULBACH, Bernd: *Analysis I und Analysis II : verständlich, anschaulich, kompakt ; ein Lehrbuch für das Bachelorstudium*. Berlin : Pro Business Verl., 2012.
- [BP03] BHATTACHARYA, Rabi ; PATRANGENARU, Vic: Large sample theory of intrinsic and extrinsic sample means on manifolds. In: *The Annals of Statistics* 31 (2003), Nr. 1.
- [Fis93] FISHER, Nicholas I.: *Statistical analysis of circular data*. Cambridge [u.a.] : Cambridge Univ. Press, 1993.
- [Geo14] GEOSCIENCE, Department of: *Rocks and the Rock Cycle*. [http://geoscience.wisc.edu/~chuck/Classes/Mtn\\_and\\_Plates/rock\\_cycle.html](http://geoscience.wisc.edu/~chuck/Classes/Mtn_and_Plates/rock_cycle.html). Version: 2014. – Datei: vermilion.cliffs.jpeg (zuletzt aufgerufen am 12. September 2014).
- [HK14] HOTZ, Thomas ; KELMA, Florian: *Non-asymptotic Confidence Regions for Extrinsic Means*. 2014 – unveröffentlicht.
- [Hoe63] HOEFFDING, Wassily: Probability Inequalities for Sums of Bounded Random Variables. In: *Journal of the American Statistical Association* 58 (1963), Nr. 301, S. 13–30.
- [Hot13a] HOTZ, Thomas: Extrinsic vs intrinsic means on the circle. In: *Geometric science of information. First international conference, GSI 2013, Paris, France, August 28–30, 2013. Proceedings*. Berlin : Springer, 2013, S. 433–440.
- [Hot13b] HOTZ, Thomas: *Mathematische Statistik*. 2013. – Vorlesungsskript zu „Mathematische Statistik“, WS 2013/14, TU Ilmenau, unveröffentlicht.
- [Leh99] LEHMANN, Erich L.: *Elements of Large-Sample Theory*. New York [u.a.] : Springer, 1999.
- [R C14] R CORE TEAM ; R FOUNDATION FOR STATISTICAL COMPUTING (Hrsg.): *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2014.
-

Ich erkläre, dass ich die vorliegende Bachelorarbeit selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe und dass ich alle Stellen, die ich wörtlich oder sinngemäß aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe. Die Arbeit hat bisher in gleicher oder ähnlicher Form oder auszugsweise noch keiner Prüfungsbehörde vorgelegen.

Ilmenau, den 12. September 2014

---

Johannes Wieditz

