



 **Universität Trier**

# DISSERTATION

**Proximität von  
Binomial- und Poisson-Verteilungen  
und  
Konfidenzintervalle für durchschnittliche  
Erfolgswahrscheinlichkeiten**

vorgelegt von  
Christoph Tasto

zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften

Trier, im März 2015

Dissertation im	Fachbereich IV – Mathematik
Betreuer:	Prof. Dr. Lutz Mattner
Berichterstatter:	Apl. Prof. Dr. Helmut Finner



# Inhaltsverzeichnis

<b>Einleitung</b>	<b>5</b>
<b>1 Notationen und Definitionen</b>	<b>9</b>
<b>2 Metriken, uniforme Strukturen und Topologien</b>	<b>11</b>
2.1 Metriken . . . . .	11
2.2 Uniforme Strukturen und Topologien . . . . .	13
2.3 Konvergenzen auf BP . . . . .	18
<b>3 Proximität von Binomial- und Poisson-Verteilungen</b>	<b>23</b>
3.1 Approximationen von Binomialverteilungen . . . . .	23
3.1.1 Binomialverteilungen gleicher Länge . . . . .	24
3.1.2 Binomialverteilungen gleicher Erfolgswahrscheinlichkeit . . . . .	24
3.1.3 Binomialverteilungen mit gleichem Erwartungswert . . . . .	25
3.1.4 Weitere Binomialverteilungen . . . . .	25
3.2 Approximationen von Poisson-Verteilungen . . . . .	26
3.3 Approximationen von Binomial- und Poisson-Verteilungen . . . . .	28
3.4 Normalverteilungs-Approximationen . . . . .	29
3.5 Uniforme Struktur auf der Menge BP . . . . .	30
3.6 Offene Fragen . . . . .	33
3.7 Beweise der Ergebnisse . . . . .	33
3.7.1 Beweise der Abschnitte 3.1 – 3.3 . . . . .	33
3.7.2 Beweise des Abschnitts 3.5 . . . . .	43
3.8 Ergebnisse für BPBP-Faltungen . . . . .	53
3.8.1 Verwendete Ergebnisse aus Tasto (2011) . . . . .	53
3.8.2 Nicht verwendete Ergebnisse aus Tasto (2011) . . . . .	54
<b>4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten</b>	<b>57</b>
4.1 Konfidenzregionen und Binomialverteilungen . . . . .	57
4.2 Der Bernoulli-Faltungs-Fall . . . . .	59
4.3 Spezialisierung auf hypergeometrische Verteilungen . . . . .	65
4.4 Offene Fragen . . . . .	68
4.5 Beweise der Ergebnisse . . . . .	69
4.5.1 Beweise des Abschnitts 4.2 . . . . .	69
4.5.2 Beweise des Abschnitts 4.3 . . . . .	73
<b>Literaturverzeichnis</b>	<b>79</b>



# Einleitung

Die vorliegende Arbeit teilt sich in zwei Themengebiete:

- Proximität von Binomial- und Poisson-Verteilungen,
- Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten.

Inhalt des ersten Teils dieser Arbeit ist die Untersuchung der Proximität, also einer gewissen Messung der Nähe, von Binomial- und Poisson-Verteilungen. Aufgrund der vielfältigen Anwendungsgebiete der Verteilungsapproximation, insbesondere der Poisson-Approximation, ist dieses Forschungsgebiet das Objekt vieler Untersuchungen. Der Artikel von Zacharovas und Hwang (2010) gibt eine ausführliche Übersicht verschiedener Ergebnisse zur Poisson-Approximation seit den Anfängen in de Moivre (1718) bis zu seinem Erscheinen. Eine wichtige Methode in der Verteilungsapproximation ist die Stein-Chen-Methode, die z. B. von Barbour und Hall (1984) und Ehm (1991) verwendet wird, um untere und obere Abschätzungen für den Totalvariationsabstand (Supremumsabstand) einer Bernoulli-Faltung und einer Poisson-Verteilung bzw. Binomialverteilung, jeweils mit gleichem Erwartungswert, anzugeben. Diese Ergebnisse können wie folgt kombiniert werden (vgl. Seite 9 für die Notationen): Für zwei Folgen  $(P_n)$  und  $(Q_n)$  mit

$$(P_n, Q_n) \in \{(P, Q) : \mu(P) = \mu(Q), P \in \{BC_p, P_\mu : n \in \mathbb{N}_0, p \in [0, 1]^n, \mu \in [0, \infty[ \}, \\ Q \in \{B_{n,p}, P_\lambda : n \in \mathbb{N}_0, p \in [0, 1], \lambda \in [0, \infty[ \} \}$$

für alle  $n \in \mathbb{N}$  konvergiert der Totalvariationsabstand  $d_{TV}(P_n, Q_n)$  genau dann gegen Null, wenn der Quotient  $|\sigma^2(P_n) - \sigma^2(Q_n)| / (1 + \sigma^2(P_n) + \sigma^2(Q_n))$  gegen Null konvergiert.

Die Topologie des Totalvariationsabstandes auf der Menge der (unendlichen) Faltungen von Bernoulli- und Poisson-Verteilungen wird in Tasto (2011) mit Hilfe der beteiligten Parameter charakterisiert.

In der vorliegenden Arbeit wird die uniforme Struktur des Totalvariationsabstandes auf der abgeschlossenen Menge  $BP := \{B_{n,p}, P_\lambda : n \in \mathbb{N}_0, p \in [0, 1], \lambda \in [0, \infty[ \}$  aller Binomial- und Poisson-Verteilungen charakterisiert, wobei diese uniforme Struktur auch vom Kolmogorov-, Hellinger- und Prochorov-Abstand erzeugt wird.

Es wird also die uniforme Struktur auf einer Teilmenge der in Tasto (2011) behandelten Menge untersucht: Nach Spezialisierung auf Binomial- und Poisson-Verteilungen wird das obige, auf Barbour und Hall (1984) und Ehm (1991) zurückgehende Ergebnis verallgemeinert, indem die Bedingung der gleichen Erwartungswerte aufgegeben wird: Der Abstand zweier Folgen  $(P_n)$  und  $(Q_n)$  aus BP konvergiert genau dann gegen Null, wenn  $(\mu(P_n) - \mu(Q_n))^2 / (\sigma^2(P_n) + \sigma^2(Q_n))$  und  $|\sigma^2(P_n) - \sigma^2(Q_n)| / (1 + \sigma^2(P_n) + \sigma^2(Q_n))$

gegen Null konvergieren, wobei das Teilen durch Null auf Seite 9 wie üblich definiert wird.

Insbesondere wird eine universelle obere Abschätzung des Totalvariationsabstandes auf BP durch eine entsprechende Funktion der zugehörigen Erwartungswerte und Varianzen angegeben: Sind  $P$  und  $Q$  beliebige Binomial- oder Poisson-Verteilungen, so ist

$$d_{\text{TV}}(P, Q) \leq 5 \cdot \sqrt{\frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)}} + 12 \cdot \frac{|\sigma^2(P) - \sigma^2(Q)|}{1 + \sigma^2(P) + \sigma^2(Q)}.$$

Der zweite Teil der Arbeit widmet sich Konfidenzintervallen für durchschnittliche Erfolgswahrscheinlichkeiten, oder besser Konfidenzintervallen für Durchschnitte von Erfolgswahrscheinlichkeiten.

Die Ergebnisse dieses Teils, insbesondere die Spezialisierung auf hypergeometrische Verteilungen in Abschnitt 4.3, gehen über die Ergebnisse in Mattner und Tasto (2014) hinaus.

Eine der ersten und bekanntesten Arbeiten zu Konfidenzintervallen von Erfolgswahrscheinlichkeiten ist die von Clopper und Pearson (1934). Im Binomialmodell werden hier bei gegebenem Stichprobenumfang und Konfidenzniveau Konfidenzintervalle für die unbekannte Erfolgswahrscheinlichkeit entwickelt.

Betrachtet man bei festem Stichprobenumfang  $n$  statt einer Binomialverteilung  $B_{n,p}$ , also dem Bildmaß einer homogenen Bernoulli-Kette unter der Summationsabbildung, das entsprechende Bildmaß einer inhomogenen Bernoulli-Kette, so erhält man eine Bernoulli-Faltung  $BC_p$  mit Erfolgswahrscheinlichkeiten  $p \in [0, 1]^n$ .

Für das Schätzen der durchschnittlichen Erfolgswahrscheinlichkeit im größeren Bernoulli-Faltungs-Modell sind z. B. die einseitigen Clopper-Pearson-Intervalle im Allgemeinen nicht gültig, was schon in Agnew (1974) erkannt, aber später an einigen Stellen in der Literatur ignoriert wurde. In Agnew (1974) werden auch gültige aber nicht optimale einseitige Konfidenzintervalle angegeben.

Wir entwickeln hier (im Sinne von Buehler (1957)) optimale einseitige und gültige zweiseitige Konfidenzintervalle für die durchschnittliche Erfolgswahrscheinlichkeit im Bernoulli-Faltungs-Modell.

Die einseitigen Clopper-Pearson-Intervalle sind im Allgemeinen auch nicht gültig für das Schätzen der Erfolgswahrscheinlichkeit im hypergeometrischen Modell, das nach Vatutin und Mikhailov (1983) ein Teilmodell des Bernoulli-Faltungs-Modells ist. Für das hypergeometrische Modell mit festem Stichprobenumfang und bekannter Urnengröße sind die einseitigen Konfidenzintervalle in Katz (1953) optimal. Bei festem Stichprobenumfang und unbekannter Urnengröße entwickeln wir aus den im Bernoulli-Faltungs-Modell optimalen Konfidenzintervallen optimale Konfidenzintervalle für das hypergeometrische Modell. Wir betrachten außerdem den Fall, dass eine obere Schranke für die unbekannte Urnengröße gegeben ist.

Die vorliegende Arbeit gliedert sich wie folgt: In Kapitel 1 finden sich einige Notationsvereinbarungen und Definitionen. Kapitel 2 enthält die Definitionen und Eigenschaften verschiedener Metriken, uniformer Strukturen und Topologien, die im Kapitel 3 benötigt werden. Dieses Kapitel 3 enthält insbesondere die Untersuchung und Beschreibung

uniformer Strukturen auf der Menge der Binomial- und Poisson-Verteilungen. Die Ergebnisse zu Konfidenzintervallen finden sich in Kapitel 4. Die Beweise der Ergebnisse der letzten beiden Kapitel sind jeweils in einem eigenen Abschnitt zusammengefasst.

### **Danksagung**

Mein besonderer Dank gilt Herrn Professor Lutz Mattner für das ausgezeichnete Betreuungsverhältnis und die große Unterstützung während der letzten Jahre.

Herrn Professor Helmut Finner danke ich sehr herzlich dafür, dass er sich spontan bereit erklärt hat, diese Dissertation zu begutachten.

Weiter möchte ich meinen Kollegen an der Universität Trier für die angenehme und stimulierende Arbeitsatmosphäre danken. Insbesondere danke ich Todor Dinev, Yves Gessner und Jona Schulz für viele hilfreiche Diskussionen und das Korrekturlesen dieser Arbeit.

Nicht zuletzt möchte ich meiner Familie danken. Ohne ihre Unterstützung wäre die Erstellung dieser Arbeit nicht möglich gewesen.



# 1 Notationen und Definitionen

Innerhalb dieser Arbeit definieren wir  $0/0 := 0$ ,  $\infty \cdot 0 := 0$  und  $c/0 := \infty$  für  $c > 0$ . Wir verwenden weiter die folgenden gebräuchlichen Notationen: Für  $x \in \mathbb{R}$  ist  $\lfloor x \rfloor := \max\{z \in \mathbb{Z}: z \leq x\}$  und  $\lceil x \rceil := \min\{z \in \mathbb{Z}: z \geq x\}$ . Für  $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$  ist  $]a, b] := \{x: a < x \leq b\}$ , und die weiteren Intervalle sind analog definiert. Es ist  $\mathbb{N} := \{1, 2, 3, \dots\}$  die Menge der natürlichen Zahlen und  $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ .

Ist  $(\mathcal{X}, \mathcal{A})$  ein Messraum, so ist  $\text{Prob}(\mathcal{X}, \mathcal{A})$  die Menge aller Wahrscheinlichkeitsmaße auf  $(\mathcal{X}, \mathcal{A})$ . Wenn klar ist, welche  $\sigma$ -Algebra betrachtet wird, so kann die Nennung dieser weggelassen werden: z. B. ist  $\text{Prob}(\mathbb{R}) := \text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . Ist  $P \in \text{Prob}(\mathbb{R})$ , so ist  $\check{P}$  definiert durch  $\check{P}(A) := P(\{-x: x \in A\})$  für alle  $A \in \mathcal{B}(\mathbb{R})$ .

Ist  $P \in \text{Prob}(\mathbb{R})$ , so bezeichnen  $\mu(P)$  und  $\sigma^2(P)$  den Erwartungswert bzw. die Varianz von  $P$ , falls sie existieren.

Wir betrachten die folgenden Verteilungen mit den jeweiligen Parametern: Für  $p \in [0, 1]$  ist  $B_p$  eine Bernoulli- und für  $n \in \mathbb{N}_0$  und  $p \in [0, 1]$  ist  $B_{n,p}$  eine Binomialverteilung. Für  $n \in \mathbb{N}_0$  und  $p = (p_1, \dots, p_n) \in [0, 1]^n$  ist  $BC_p := \ast_{j=1}^n B_{p_j}$  eine Bernoulli-Faltung, die in der Literatur auch als Poisson-Binomial-Verteilung oder Poissons Binomialverteilung (vgl. Wang (1993)) bezeichnet wird. Für  $\lambda \in [0, \infty[$  ist  $P_\lambda$  eine Poisson-Verteilung und für  $(\mu, \sigma^2) \in \mathbb{R} \times [0, \infty[$  ist  $N_{\mu, \sigma^2}$  eine Normalverteilung. Für  $n, N \in \mathbb{N}_0$  mit  $N \geq n$  und  $p \in \{j/N: j \in \{0, \dots, N\}\}$  bezeichnet  $H_{n,p,N}$  die hypergeometrische Verteilung der Anzahl roter Kugeln, die in einer einfachen Zufallsstichprobe der Größe  $n$  aus einer Urne mit  $N$  Kugeln, davon  $Np$  rot und  $N(1-p)$  blau, ohne Zurücklegen gezogen werden. Es ist  $BP := \{B_{n,p}, P_\lambda: n \in \mathbb{N}_0, p \in [0, 1], \lambda \in [0, \infty[ \}$  die Menge aller Binomial- und Poisson-Verteilungen. Für die Definition einer BPBP-Faltung vgl. Definition 3.24.



## 2 Metriken, uniforme Strukturen und Topologien

Das vorliegende Kapitel dient als Vorbereitung des Kapitels 3, in dem die Menge BP aller Binomial- und Poisson-Verteilungen untersucht wird, weshalb das Augenmerk verstärkt auf diese Menge gelegt wird.

Im folgenden Abschnitt 2.1 führen wir die im Weiteren benutzten Metriken ein und stellen einige Eigenschaften der Abstände und Zusammenhänge zwischen diesen dar. Im darauffolgenden Abschnitt 2.2 stellen wir die benötigten Definitionen und Resultate zu uniformen Strukturen und Topologien bereit. Schließlich werden in Abschnitt 2.3 verschiedene Topologien auf BP untersucht.

### 2.1 Metriken

Der *Totalvariationsabstand* (oder *Supremumsabstand*)  $d_{TV}$  auf  $\text{Prob}(\mathcal{X}, \mathcal{A})$  ist definiert durch

$$(P, Q) \mapsto d_{TV}(P, Q) = \sup\{|P(A) - Q(A)| : A \in \mathcal{A}\}$$

(vgl. Strasser (1985, S. 5)). Ist  $\nu$  ein  $P$  und  $Q$  dominierendes Maß auf  $\mathcal{A}$ , so ist (vgl. Strasser (1985, Lemma 2.4))

$$d_{TV}(P, Q) = \frac{1}{2} \int \left| \frac{dP}{d\nu} - \frac{dQ}{d\nu} \right| d\nu. \quad (2.1)$$

Der *Hellinger-Abstand*  $d_H$  auf  $\text{Prob}(\mathcal{X}, \mathcal{A})$  ist definiert durch

$$(P, Q) \mapsto d_H(P, Q) = \left( \frac{1}{2} \int (\sqrt{f} - \sqrt{g})^2 d\mu \right)^{1/2},$$

wobei  $\mu$  ein  $\sigma$ -endliches Maß auf  $\mathcal{A}$  mit  $(P, Q) = (f\mu, g\mu)$  ist. Der Hellinger-Abstand ist wohldefiniert, also unabhängig von der Wahl des Maßes  $\mu$  und der Dichten  $f, g$  (vgl. Strasser (1985, S. 8)).

Der *Kolmogorov-Abstand*  $d_K$  auf  $\text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  ist definiert durch

$$(P, Q) \mapsto d_K(P, Q) = \sup\{|P([-\infty, x]) - Q([-\infty, x])| : x \in \mathbb{R}\}.$$

Der *Prochorov-Abstand*  $d_P$  auf  $\text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  ist definiert durch

$$\begin{aligned} (P, Q) &\mapsto d_P(P, Q) \\ &= \inf \{ \varepsilon > 0 : P(A) \leq Q(A^\varepsilon) + \varepsilon, Q(A) \leq P(A^\varepsilon) + \varepsilon \text{ für alle } A \in \mathcal{B}(\mathbb{R}) \}, \end{aligned}$$

wobei  $A^\varepsilon := \{x : \inf \{|x - y| : y \in A\} < \varepsilon\}$  für  $A \in \mathcal{B}(\mathbb{R})$  und  $\varepsilon > 0$  (vgl. Elstrodt (2009, S. 402)).

Es gelten die folgenden Abschätzungen: Nach Strasser (1985, Lemma 2.15) gilt

$$d_{\mathbb{H}}^2 \leq d_{\text{TV}} \leq d_{\mathbb{H}} \cdot \sqrt{2 - d_{\mathbb{H}}^2}. \quad (2.2)$$

Außerdem ist auf  $\text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$$d_{\mathbb{K}} \leq d_{\text{TV}}, \quad (2.3)$$

und auf BP gilt weiter die Abschätzung  $d_{\text{TV}} \leq 2d_{\mathbb{K}}$ , die wir in Satz 2.1 analog zu Ehm (1991, Corollary 3) mit Hilfe von Klenke und Mattner (2010) beweisen, und die in Spezialfällen schon bekannt ist, z. B. nach Ehm (1991, Corollary 3) im Fall von Binomialverteilungen mit gleichem Erwartungswert.

**Satz 2.1.** *Auf BP ist*

$$d_{\mathbb{K}} \leq d_{\text{TV}} \leq 2d_{\mathbb{K}}. \quad (2.4)$$

*Beweis.* Die erste Ungleichung gilt auf ganz  $\text{Prob}(\mathbb{R})$  (vgl. (2.3)). Wir zeigen die zweite Ungleichung zuerst für Binomialverteilungen: Seien  $n, m \in \mathbb{N}_0$ ,  $p, q \in [0, 1]$ ,  $P := B_{n,p}$ ,  $Q := B_{m,q}$  und  $k_* := \min\{k \in \mathbb{N}_0 : (P + Q)(\{k\}) > 0\}$ ,  $k^* := \max\{k \in \mathbb{N}_0 : (P + Q)(\{k\}) > 0\}$ . Dann ist nach Klenke und Mattner (2010, Lemma 2.1)

$$l: \{k_*, \dots, k^*\} \rightarrow [0, \infty], \quad k \mapsto \frac{P(\{k\})}{Q(\{k\})}$$

für ein  $k_0 \in \{k_*, \dots, k^*\}$  monoton auf  $\{k_*, \dots, k_0\}$  und monoton auf  $\{k_0, \dots, k^*\}$ .

1. Fall:  $l$  ist wachsend auf  $\{k_*, \dots, k_0\}$  und fallend auf  $\{k_0, \dots, k^*\}$ . Dann ist  $L := \{k \in \{k_*, \dots, k^*\} : l(k) > 1\}$  ein Intervall in  $\mathbb{N}_0$ . Sei  $J := \{k_*, \dots, \min(L) - 1\}$ . Für  $M \subseteq \mathbb{N}_0$  sei  $\Delta(M) := \sum_{k \in M} (P(\{k\}) - Q(\{k\}))$ . Dann ist  $d_{\text{TV}}(P, Q) = \Delta(L)$  und

$$d_{\mathbb{K}}(P, Q) = \begin{cases} \Delta(J \cup L) = \Delta(J) + \Delta(L) \geq \frac{\Delta(L)}{2}, & \text{falls } \Delta(L) \geq -2\Delta(J), \\ -\Delta(J) \geq \frac{\Delta(L)}{2}, & \text{falls } \Delta(L) < -2\Delta(J). \end{cases}$$

Die anderen Fälle werden analog behandelt, und damit ist die Behauptung im Binomialfall bewiesen.

Wenn  $P$  eine Poisson-Verteilung und  $Q$  eine Binomialverteilung ist, sei  $P_n := B_{n, \frac{\mu(P)}{n}}$  für alle  $n \in \mathbb{N}$ . Nach Koopman (1950) gilt  $P_n(\{k\}) \rightarrow P(\{k\})$ ,  $n \rightarrow \infty$ , für alle  $k \in \mathbb{N}_0$ , und damit nach Gut (2005, Chapter 5, Theorem 6.4)  $d_{\text{TV}}(P_n, P) \rightarrow 0$ , also  $d_{\mathbb{K}}(P_n, P) \rightarrow 0$  für  $n \rightarrow \infty$ . Wie oben im Beweis gezeigt wurde, gilt  $d_{\text{TV}}(P_n, Q) \leq 2d_{\mathbb{K}}(P_n, Q)$  für alle  $n \in \mathbb{N}$ , und damit

$$d_{\text{TV}}(P, Q) \leq d_{\text{TV}}(P, P_n) + d_{\text{TV}}(P_n, Q) \leq d_{\text{TV}}(P, P_n) + 2d_{\mathbb{K}}(P_n, P) + 2d_{\mathbb{K}}(P, Q).$$

Also ist  $d_{\text{TV}}(P, Q) \leq 2d_{\mathbb{K}}(P, Q)$ .

Auf der Menge der Poisson-Verteilungen gilt sogar  $d_{\text{TV}} = d_{\mathbb{K}}$  (vgl. z. B. Adell und Jodrá (2006, Proposition 2.1)).  $\square$

Bemerkung 2.2. Es ist

$$d_P \leq d_{TV}, \quad (2.5)$$

und auf  $\text{Prob}(\mathbb{Z})$  gilt

$$d_P = d_{TV}. \quad (2.6)$$

Seien nämlich  $P, Q \in \text{Prob}(\mathbb{R})$ . Um (2.5) zu zeigen, definieren wir  $\varepsilon := d_{TV}(P, Q)$ . Für alle  $A \in \mathcal{B}(\mathbb{R})$  ist dann  $P(A) \leq Q(A) + \varepsilon \leq Q(A^\varepsilon) + \varepsilon$  und  $Q(A) \leq P(A) + \varepsilon \leq P(A^\varepsilon) + \varepsilon$ , also ist  $d_P(P, Q) \leq d_{TV}(P, Q)$ .

Um (2.6) zu zeigen, nehmen wir an, dass  $P, Q \in \text{Prob}(\mathbb{Z})$ , und zeigen, dass  $d_{TV}(P, Q) \leq d_P(P, Q)$ . Wenn  $d_P(P, Q) = 1$ , ist  $d_{TV}(P, Q) \leq d_P(P, Q)$ . Wenn  $d_P(P, Q) < 1$ , gilt für  $1 > \delta > d_P(P, Q)$  und  $A \subseteq \mathbb{Z}$ , dass  $P(A^\delta \cap \mathbb{Z}) = P(A)$  und  $Q(A^\delta \cap \mathbb{Z}) = Q(A)$ , also  $d_{TV}(P, Q) \leq \delta$  und damit  $d_{TV}(P, Q) \leq d_P(P, Q)$ .

Die Gleichung (2.6) gilt nicht im Allgemeinen auf  $\text{Prob}(\mathbb{R})$ , denn für  $n \geq 2$  ist

$$d_{TV}(\delta_n, \delta_{n+\frac{1}{n}}) = 1 > \frac{1}{n} = d_P(\delta_n, \delta_{n+\frac{1}{n}}).$$

## 2.2 Uniforme Strukturen und Topologien

Im Folgenden stellen wir die für Kapitel 3 notwendigen Definitionen und Resultate zu uniformen Strukturen und Topologien bereit, wobei wir uns an Bourbaki (1989) orientieren. Weitere Einführungen in uniforme Strukturen findet man z. B. in Querenburg (2001), Schubert (1964) oder Kelley (1975). In all diesen Büchern werden auch die für die Behandlung uniformer Räume notwendigen topologischen Grundlagen bereitgestellt.

Durch eine uniforme Struktur ist eindeutig eine Topologie gegeben (vgl. Satz 2.17). Jeder uniforme Raum ist also insbesondere ein topologischer Raum. Die Idee der uniformen Struktur führt die der Topologie weiter: Die Interpretation „ $y$  liegt genügend nahe bei  $x$ “ im topologischen Kontext wird bei Betrachtung der uniformen Struktur weitergeführt zu „zwei Punkte  $x, y$  liegen genügend nahe beieinander“ (vgl. Schubert (1964, S.99)).

Andererseits ist jeder metrische Raum auch ein uniformer Raum, wobei die durch eine Metrik gegebene uniforme Struktur nach Beispiel 2.15 gebildet wird.

Die uniforme Struktur liegt also „zwischen“ der Metrik und der Topologie, in dem Sinne, dass eine Metrik eine uniforme Struktur definiert, welche eine Topologie definiert, die gleich der durch die Metrik induzierten Topologie ist.

Im Folgenden werden Filter und die Konvergenz entlang Filtern eingeführt, sowie auf einige in unserem Kontext wichtige Beispiele eingegangen, danach werden spezieller uniforme Strukturen eingeführt und untersucht.

**Definition 2.3.** Sei  $\mathcal{Y}$  eine Menge. Ein Mengensystem  $\mathcal{F} \subseteq 2^{\mathcal{Y}}$  heißt *Filter* auf  $\mathcal{Y}$ , falls gilt:

$$\text{Aus } F \in \mathcal{F} \text{ und } F \subseteq G \subseteq \mathcal{Y} \text{ folgt } G \in \mathcal{F}. \quad (2.7)$$

$$\mathcal{Y} \in \mathcal{F}. \text{ Sind } F_1, F_2 \in \mathcal{F}, \text{ so ist } F_1 \cap F_2 \in \mathcal{F}. \quad (2.8)$$

$$\emptyset \notin \mathcal{F}. \quad (2.9)$$

Das Paar  $(\mathcal{Y}, \mathcal{F})$  heißt *filtrierte Menge* und wird, wenn keine Verwechslung auftreten kann, mit  $\mathcal{Y}$  bezeichnet.

**Definition und Beispiel 2.4.** Sei  $(\mathcal{Y}, \mathcal{T})$  ein topologischer Raum und sei  $\emptyset \neq S \subseteq \mathcal{Y}$ . Dann ist

$$\{Y \subseteq \mathcal{Y} : \text{es existiert ein } T \in \mathcal{T} \text{ mit } S \subseteq T \subseteq Y\}$$

ein Filter auf  $\mathcal{Y}$ , der als *Umgebungsfilter* von  $S$  bezeichnet wird. Elemente des Umgebungsfilters heißen *Umgebungen* von  $S$ . Ist  $S = \{y\}$  für ein  $y \in \mathcal{Y}$ , so sprechen wir vom *Umgebungsfilter* von  $y$  (und schreiben  $\mathcal{U}(y)$ ) bzw. den *Umgebungen* von  $y$ . Ein Mengensystem  $\mathcal{B} \subseteq \mathcal{U}(y)$  heißt *Umgebungsbasis* von  $y$ , wenn zu jedem  $U \in \mathcal{U}(y)$  ein  $V \in \mathcal{B}$  mit  $V \subseteq U$  existiert.

**Definition 2.5.** Seien  $\mathcal{X}$  eine Menge und  $\mathcal{F}, \mathcal{F}'$  Filter auf  $\mathcal{X}$ . Wir sagen, dass  $\mathcal{F}'$  feiner als  $\mathcal{F}$  ist (oder  $\mathcal{F}$  gröber als  $\mathcal{F}'$ ), wenn  $\mathcal{F} \subseteq \mathcal{F}'$ . Ist weiter eine Topologie auf  $\mathcal{X}$  gegeben, so *konvergiert* der Filter  $\mathcal{F}$  gegen ein  $x \in \mathcal{X}$ , wenn  $\mathcal{F}$  feiner ist als der Umgebungsfilter von  $x$ .

**Definition 2.6.** Seien  $(\mathcal{X}, \mathcal{F})$  eine filtrierte Menge,  $(\mathcal{Y}, \mathcal{T})$  ein topologischer Raum,  $f : \mathcal{X} \rightarrow \mathcal{Y}$  eine Abbildung und  $y \in \mathcal{Y}$ .

Die Abbildung  $f$  *konvergiert* entlang  $\mathcal{F}$  gegen  $y$ , wenn für jede Umgebung  $V$  von  $y$  ein  $F \in \mathcal{F}$  mit  $f(F) \subseteq V$  existiert; und  $y$  heißt dann *Grenzwert* von  $f$  entlang  $\mathcal{F}$ . Wir schreiben dann  $\lim_{\mathcal{F}} f = y$ , oder  $\lim f = y$ , wenn klar ist, welcher Filter gemeint ist.

**Definition 2.7.** Auf  $\mathcal{X} = \mathbb{N}$  ist durch  $\mathcal{F} := \{F \subseteq \mathbb{N} : \mathbb{N} \setminus F \text{ endlich}\}$  der *Fréchet-Filter* definiert. Sei  $(\mathcal{Y}, \mathcal{T})$  ein topologischer Raum und  $y \in \mathcal{Y}$ . Eine Folge  $(y_n)_{n \in \mathbb{N}} : \mathbb{N} \rightarrow \mathcal{Y}$  *konvergiert* gegen  $y$ , falls sie entlang des Fréchet-Filters gegen  $y$  konvergiert. Wir benutzen dann die bekannten Schreibweisen  $\lim_{n \rightarrow \infty} y_n = y$  oder  $y_n \rightarrow y, n \rightarrow \infty$ .

**Definition und Beispiel 2.8.** Sei  $d$  eine Metrik auf  $\mathcal{X} = \text{Prob}(\mathbb{R})$  und

$$\{\mathcal{O} \subseteq \mathcal{X} : \text{für alle } P \in \mathcal{O} \text{ existiert ein } \varepsilon > 0 : \{Q \in \mathcal{X} : d(P, Q) < \varepsilon\} \subseteq \mathcal{O}\}$$

die durch  $d$  auf  $\mathcal{X}$  definierte Topologie und  $\mathcal{U}$  die durch die Betragsmetrik definierte Topologie auf  $[0, \infty[$ . Seien weiter  $P \in \mathcal{X}$  und  $\mathcal{F}$  der Umgebungsfilter von  $P$ . Dann konvergiert  $\mathcal{F}$  gegen  $P$  und es konvergiert  $f := d(\cdot, P) : \mathcal{X} \rightarrow [0, \infty[$  entlang  $\mathcal{F}$  gegen 0, denn für jedes  $\varepsilon > 0$  ist  $F := \{Q \in \mathcal{X} : d(Q, P) < \varepsilon\} \in \mathcal{F}$  und  $f(Q) < \varepsilon$  für alle  $Q \in F$ . Eine Folge  $(P_n)_{n \in \mathbb{N}}$  in  $\mathcal{X}$  *konvergiert in  $d$*  gegen  $P$ , wenn  $(P_n)_{n \in \mathbb{N}}$  bezüglich obiger

Topologie gegen  $P$  konvergiert, und das ist genau dann der Fall, wenn  $(d(P_n, P))_{n \in \mathbb{N}}$  gegen 0 konvergiert.

Wie in Abschnitt 2.1 zu sehen, betrachten wir in dieser Arbeit den Totalvariations-, den Kolmogorov-, den Hellinger- und den Prochorov-Abstand.

**Definition und Beispiel 2.9.** (vgl. Elstrodt (2009, S. 382)) Auf  $\text{Prob}(\mathbb{R})$  betrachten wir die *schwache Topologie*, also die größte Topologie auf  $\text{Prob}(\mathbb{R})$ , bezüglich der die Abbildung

$$\text{Prob}(\mathbb{R}) \ni P \mapsto \int_{\mathbb{R}} f dP$$

für alle  $f \in C_b(\mathbb{R}) := \{f \in \mathbb{R}^{\mathbb{R}} : f \text{ ist stetig und beschränkt}\}$  stetig ist. Eine Umgebungsbasis von  $P_0 \in \text{Prob}(\mathbb{R})$  bezüglich der schwachen Topologie ist

$$\left\{ U_{f_1, \dots, f_j; \varepsilon}(P_0) : j \in \mathbb{N}, f_1, \dots, f_j \in C_b(\mathbb{R}), \varepsilon > 0 \right\},$$

wobei für  $P_0 \in \text{Prob}(\mathbb{R})$ ,  $f_1, \dots, f_j \in C_b(\mathbb{R})$  und  $\varepsilon > 0$

$$U_{f_1, \dots, f_j; \varepsilon}(P_0) := \left\{ P \in \text{Prob}(\mathbb{R}) : \left| \int f_k dP - \int f_k dP_0 \right| < \varepsilon \text{ für alle } k = 1, \dots, j \right\}.$$

Eine Folge  $(P_n)$  in  $\text{Prob}(\mathbb{R})$  *konvergiert schwach* gegen ein  $P \in \text{Prob}(\mathbb{R})$ , wenn  $(P_n)$  bezüglich obiger Topologie gegen  $P$  konvergiert, wenn also  $\int f dP_n \rightarrow \int f dP$  für alle  $f \in C_b(\mathbb{R})$ . Als Schreibweise hierfür nutzen wir auch  $P_n \rightsquigarrow P$ .

**Definition und Beispiel 2.10.** Auf  $\text{Prob}(\mathbb{R})$  betrachten wir die *Topologie der Konvergenz der absoluten Momente*, also die größte Topologie, bezüglich der die Abbildung

$$\text{Prob}(\mathbb{R}) \ni P \mapsto \int_{\mathbb{R}} |x|^k dP$$

für alle  $k \in \mathbb{N}$  stetig ist. Eine Umgebungsbasis von  $P_0 \in \text{Prob}(\mathbb{R})$  bezüglich dieser Topologie ist gegeben durch

$$\left\{ \left\{ P \in \text{Prob}(\mathbb{R}) : \left| \int |x|^k dP - \int |x|^k dP_0 \right| < \varepsilon \text{ für alle } k \in I \right\} : I \subseteq \mathbb{N} \text{ endlich}, \varepsilon > 0 \right\}.$$

Eine Folge  $(P_n)$  in  $\text{Prob}(\mathbb{R})$  *konvergiert in absoluten Momenten* gegen ein  $P \in \text{Prob}(\mathbb{R})$ , wenn  $(P_n)$  bezüglich obiger Topologie gegen  $P$  konvergiert, wenn also  $\int |x|^k dP_n \rightarrow \int |x|^k dP$  für alle  $k \in \mathbb{N}$ .

**Definition und Beispiel 2.11.** Auf  $\text{Prob}(\mathbb{N}_0)$  betrachten wir die *Topologie der Konvergenz im ersten und zweiten Moment*, also die größte Topologie, bezüglich der die Abbildungen

$$\text{Prob}(\mathbb{N}_0) \ni P \mapsto \int_{\mathbb{N}_0} x dP \quad \text{und} \quad \text{Prob}(\mathbb{N}_0) \ni P \mapsto \int_{\mathbb{N}_0} x^2 dP$$

stetig sind. Eine Umgebungsbasis von  $P_0 \in \text{Prob}(\mathbb{N}_0)$  bezüglich dieser Topologie ist gegeben durch

$$\left\{ \left\{ P \in \text{Prob}(\mathbb{N}_0) : \left| \int x^k dP - \int x^k dP_0 \right| < \varepsilon \text{ für } k \in \{1, 2\} \right\} : \varepsilon > 0 \right\}.$$

Eine Folge  $(P_n)$  in  $\text{Prob}(\mathbb{N}_0)$  *konvergiert im ersten und zweiten Moment* gegen ein  $P \in \text{Prob}(\mathbb{N}_0)$ , wenn  $(P_n)$  bezüglich obiger Topologie gegen  $P$  konvergiert, also genau dann, wenn  $\mu(P_n) \rightarrow \mu(P)$  und  $\sigma^2(P_n) \rightarrow \sigma^2(P)$ .

**Definition und Beispiel 2.12.** Auf  $\text{Prob}(\mathbb{N}_0)$  betrachten wir die Topologie der *punktweisen Konvergenz*, also die größte Topologie, bezüglich der die Abbildung  $\text{Prob}(\mathbb{N}_0) \ni P \mapsto P(\{x\})$  für alle  $x \in \mathbb{N}_0$  stetig ist. Eine Umgebungsbasis von  $P_0 \in \text{Prob}(\mathbb{N}_0)$  bezüglich dieser Topologie ist gegeben durch

$$\{ \{ P \in \text{Prob}(\mathbb{N}_0) : |P(\{x\}) - P_0(\{x\})| < \varepsilon \text{ für alle } x \in E \} : E \subseteq \mathbb{N}_0 \text{ endlich, } \varepsilon > 0 \}.$$

Eine Folge  $(P_n)$  in  $\text{Prob}(\mathbb{N}_0)$  *konvergiert punktweise* gegen ein  $P \in \text{Prob}(\mathbb{N}_0)$ , wenn  $(P_n)$  bezüglich obiger Topologie gegen  $P$  konvergiert, wenn also  $P_n(\{x\}) \rightarrow P(\{x\})$  für alle  $x \in \mathbb{N}_0$ .

Um uniforme Strukturen zu definieren, führen wir die folgende Schreibweise für das *Relationenprodukt* ein: Für eine Menge  $\mathcal{X}$  und  $U, V \subseteq \mathcal{X} \times \mathcal{X}$  definieren wir

$$U \circ V := \{ (x, z) \in \mathcal{X} \times \mathcal{X} : \text{es gibt ein } y \in \mathcal{X} \text{ mit } (x, y) \in U \text{ und } (y, z) \in V \},$$

wobei hier keine Verwechslung mit der Komposition von Funktionen zu befürchten ist.

**Definition 2.13.** Sei  $\mathcal{X}$  eine Menge. Eine Menge  $\mathcal{U} \subseteq 2^{\mathcal{X} \times \mathcal{X}}$  heißt *uniforme Struktur* auf  $\mathcal{X}$ , falls  $\mathcal{U}$  (2.7) und (2.8) erfüllt und weiter gilt:

$$\text{Für jedes } U \in \mathcal{U} \text{ gilt } \Delta := \{ (x, x) : x \in \mathcal{X} \} \subseteq U. \quad (2.10)$$

$$\text{Ist } U \in \mathcal{U}, \text{ so ist } \{ (y, x) : (x, y) \in U \} \in \mathcal{U}. \quad (2.11)$$

$$\text{Ist } U \in \mathcal{U}, \text{ so existiert ein } V \in \mathcal{U} \text{ mit } V \circ V \subseteq U. \quad (2.12)$$

Das Paar  $(\mathcal{X}, \mathcal{U})$  heißt *uniformer Raum* und wird auch mit  $\mathcal{X}$  bezeichnet. Die Elemente von  $\mathcal{U}$  heißen *Nachbarschaften* in  $\mathcal{X}$ .

**Definition und Bemerkung 2.14.** Wir betrachten nur  $\mathcal{X} \neq \emptyset$ , so dass eine uniforme Struktur  $\mathcal{U}$  wegen (2.10) auch (2.9) erfüllt, und somit  $\mathcal{U}$  immer ein Filter auf  $\mathcal{X} \times \mathcal{X}$  ist. Wir sprechen daher auch vom *Nachbarschaftsfilter*  $\mathcal{U}$ . Eine Menge  $\mathcal{B} \subseteq \mathcal{U}$  heißt *Basis* des Nachbarschaftsfilters  $\mathcal{U}$ , wenn zu jedem  $U \in \mathcal{U}$  ein  $B \in \mathcal{B}$  mit  $B \subseteq U$  existiert.

**Beispiel 2.15.** Metriken und uniforme Strukturen hängen wie folgt zusammen:

(a) Sei  $(\mathcal{X}, d)$  metrischer Raum. Dann bildet

$$\mathcal{B} := \{ \{ (x, y) \in \mathcal{X}^2 : d(x, y) < \varepsilon \} : \varepsilon > 0 \}$$

eine Basis des Nachbarschaftsfilters

$$\mathcal{U} := \{ V \subseteq \mathcal{X}^2 : \text{es existiert ein } B \in \mathcal{B} \text{ mit } V \supseteq B \}$$

auf  $\mathcal{X}$ .

- (b) Verschiedene Metriken können die gleiche uniforme Struktur erzeugen: Auf der Menge aller Binomial- und Poisson-Verteilungen BP gilt nach (2.2), (2.4) und (2.6)

$$\max \{d_K, d_H^2\} \leq d_{TV} = d_P \leq \min \left\{ d_H \cdot \sqrt{2 - d_H^2}, 2d_K \right\}.$$

Also definieren alle vier Metriken die gleiche uniforme Struktur und die gleiche Topologie auf BP.

- (c) Ist eine uniforme Struktur  $\mathcal{U}$  auf einer Menge  $\mathcal{X}$  wie in (a) durch eine Metrik  $d$  definiert, so lässt sich die uniforme Struktur vollständig durch konvergente Folgen charakterisieren: Eine Menge  $U \subseteq \mathcal{X}^2$  liegt genau dann in  $\mathcal{U}$ , wenn für alle Folgen  $(x_n, y_n)_n$  aus  $\mathcal{X}^2$  mit  $\lim_{n \rightarrow \infty} d(x_n, y_n) = 0$  schließlich  $(x_n, y_n) \in U$  gilt.

**Beispiel 2.16.** Auf  $[0, \infty[$  bildet  $\mathcal{U} = \{U_\varepsilon : \varepsilon > 0\}$ , mit

$$U_\varepsilon := \{(x, y) \in [0, \infty[^2 : (x - y)^2 < \varepsilon\} \text{ für } \varepsilon > 0,$$

eine Basis des Nachbarschaftsfilters

$$\mathcal{F} := \{V \subseteq [0, \infty[^2 : \text{es existiert ein } \varepsilon > 0 \text{ mit } V \supseteq U_\varepsilon\}.$$

*Beweis.* Da (2.7), (2.8), (2.10) und (2.11) direkt nach der Definition von  $\mathcal{F}$  gelten, muss nur (2.12) für  $\mathcal{F}$  nachgewiesen werden, um zu zeigen, dass  $\mathcal{F}$  eine uniforme Struktur auf  $[0, \infty[$  ist.

Sei  $V \in \mathcal{F}$ . Dann existiert ein  $\varepsilon > 0$  mit  $V \supseteq U_\varepsilon$ . Also reicht zu zeigen: Für alle  $\varepsilon > 0$  existiert ein  $\delta > 0$  mit  $U_\delta \circ U_\delta \subseteq U_\varepsilon$ . Seien  $\varepsilon > 0$ ,  $\delta := \frac{\varepsilon}{4}$  und  $(x, z) \in U_\delta \circ U_\delta$ , so dass ein  $y \in [0, \infty[$  mit  $\max\{(x - y)^2, (y - z)^2\} < \delta$  existiert. Dann ist  $(x, z) \in U_\varepsilon$ , denn  $(x - z)^2 \leq 2((x - y)^2 + (y - z)^2) < 4\delta = \varepsilon$ .  $\square$

Die Verbindung zwischen uniformen Strukturen und Topologien wird im folgenden Satz deutlich (vgl. Bourbaki (1989, Chapter II, § 1, no. 2, Proposition 1)).

**Satz 2.17.** Sei  $(\mathcal{X}, \mathcal{U})$  ein uniformer Raum, und für jedes  $x \in \mathcal{X}$  sei

$$\mathcal{B}(x) := \{\{y \in \mathcal{X} : (x, y) \in V\} : V \in \mathcal{U}\}.$$

Dann existiert genau eine Topologie auf  $\mathcal{X}$ , bezüglich der für jedes  $x \in \mathcal{X}$  die Menge  $\mathcal{B}(x)$  der Umgebungsfilter von  $x$  ist.

**Beispiel 2.18.** Ist  $\mathcal{F}$  wie in Beispiel 2.16 definiert, so ist für jedes  $x \in [0, \infty[$

$$\mathcal{B}(x) = \left\{ A \subseteq [0, \infty[ : \text{es gibt ein } \varepsilon > 0 : ]x - \sqrt{\varepsilon}, x + \sqrt{\varepsilon}[ \cap [0, \infty[ \subseteq A \right\},$$

also ist die nach Satz 2.17 eindeutig gegebene Topologie gleich der durch die Betragsmetrik definierten Topologie auf  $[0, \infty[$ .

## 2.3 Konvergenzen auf BP

Wir richten nun unseren Blick spezieller auf verschiedene Topologien auf der Menge der Binomial- und Poisson-Verteilungen, die z. B. in der Topologie der punktweisen Konvergenz auf  $\text{Prob}(\mathbb{N}_0)$  abgeschlossen ist:

**Satz 2.19.** *Die Menge  $\text{BP} := \{B_{n,p}, P_\lambda : n \in \mathbb{N}_0, p \in [0, 1], \lambda \in [0, \infty[ \}$  bildet den Abschluss der Menge der Binomialverteilungen in der Topologie der punktweisen Konvergenz auf  $\text{Prob}(\mathbb{N}_0)$ .*

Um Satz 2.19 zu beweisen, nutzen wir das folgende Ergebnis, dessen Beweisidee schon in Tasto (2011) verwendet wurde. Dieses Lemma gilt sogar für eine Obermenge von BP, die BPBP-Faltungen, die wie in Definition 3.24 unten definiert sind.

**Lemma 2.20.** *Seien  $(P_k)$  eine Folge von BPBP-Faltungen und  $P \in \text{Prob}(\mathbb{Z})$  mit  $P_k(\{z\}) \rightarrow P(\{z\}), k \rightarrow \infty$ , für alle  $z \in \mathbb{Z}$ . Dann konvergieren auch  $P_k \rightsquigarrow P$  und  $\mu(P_k) \rightarrow \mu(P), k \rightarrow \infty$ .*

*Beweis.* Nach Voraussetzung existiert ein  $z \in \mathbb{Z}$  mit  $\liminf_{k \rightarrow \infty} P_k(\{z\}) > 0$ . Also ist  $(\sigma^2(P_k))_k$  beschränkt (vgl. Barbour und Jensen (1989, Lemma 1) oder Baillon et al. (2013, Corollary 2)). Damit ist  $(P_k)_k$  nach der Jensen-Ungleichung zentriert  $L^1$ -beschränkt. Da  $P_k(\{z\}) \rightarrow P(\{z\}), k \rightarrow \infty$ , für alle  $z \in \mathbb{Z}$ , ist  $(P_k)_k$  straff. Aus der Straffheit und der zentrierten  $L^1$ -Beschränktheit folgt nach Lemma 3.26 unten die  $L^1$ -Beschränktheit von  $(P_k)_k$ . Damit ist  $\sup\{\int x^2 dP_k(x) : k \in \mathbb{N}\} = \sup\{\sigma^2(P_k) + (\mu(P_k))^2 : k \in \mathbb{N}\} < \infty$ . Also ist  $(P_k)_k$  gleichgradig integrierbar (vgl. z. B. Gut (2005, Chapter 5, Theorem 4.2.)). Nach Gut (2005, Chapter 5, Theorem 6.4.) gilt  $P_k \rightsquigarrow P$ , und aufgrund der gleichgradigen Integrierbarkeit folgt  $\mu(P_k) \rightarrow \mu(P), k \rightarrow \infty$ .  $\square$

Damit können wir Satz 2.19 beweisen:

*Beweis von Satz 2.19.* Seien  $(p) := (p_k) \in [0, 1]^{\mathbb{N}}$ ,  $(n) := (n_k) \in \mathbb{N}_0^{\mathbb{N}}$  und  $P \in \text{Prob}(\mathbb{N}_0)$  mit  $B_{n,p}(\{z\}) \rightarrow P(\{z\})$  für alle  $z \in \mathbb{N}_0$ , so dass nach Lemma 2.20  $n p \rightarrow \mu(P)$ . Wenn  $\mu(P) = 0$ , ist  $P = \delta_0 \in \text{BP}$ . Wenn  $\mu(P) \neq 0$ , sei  $(k_l)_l$  eine Teilfolge.

Wenn entlang der Teilfolge  $p \rightarrow 0$ , konvergiert nach Koopman (1950)  $B_{n,p}(\{z\}) \rightarrow P_{\mu(P)}(\{z\}) = P(\{z\})$  für alle  $z \in \mathbb{N}_0$ .

Wenn  $p \not\rightarrow 0$ , existieren ein  $j \in \mathbb{N}$  und eine Teilteilfolge  $(k_{l_m})_m$  mit  $(n, p) \rightarrow (j, \frac{\mu(P)}{j})$  und damit  $B_{n,p}(\{z\}) \rightarrow B_{j, \frac{\mu(P)}{j}}(\{z\}) = P(\{z\})$  für alle  $z \in \mathbb{N}_0$ .  $\square$

Als Vorbereitung für Satz 2.22 formulieren wir den folgenden Satz 2.21, wobei  $\mu_k(P) := \int x^k dP$  für  $P \in \text{BP}$  und  $k \in \mathbb{N}$ .

**Satz 2.21.** *Für jedes  $k \in \mathbb{N}$  ist*

$$\mu_k(P) \leq (k-1)! \mu_1(P) (1 + \mu_1(P))^{k-1} \quad \text{für alle } P \in \text{BP}. \quad (2.13)$$

*Beweis.* Wir zeigen induktiv für alle  $k \in \mathbb{N}$ : es gilt (2.13) und es existieren

$$\text{ein } \alpha_k \in [0, \infty[^k \text{ mit } \mu_k(P_\lambda) = \sum_{j=1}^k \alpha_{k,j} \lambda^j \text{ für alle } \lambda \in [0, \infty[, \quad (2.14)$$

$$\text{für } n \in \mathbb{N}_0 \text{ ein } \beta_{(n,k)} \in [0, \infty[^k \text{ mit } \mu_k(B_{n,p}) = \sum_{j=1}^k \beta_{(n,k),j} p^j \text{ für alle } p \in [0, 1]. \quad (2.15)$$

Induktions-Anfang: Für  $k = 1$  gelten (2.13), (2.14) und (2.15) mit  $\alpha_1 = 1$  und  $\beta_{(n,1)} = n$  für  $n \in \mathbb{N}_0$ .

Induktions-Schritt: Es gelten (2.13), (2.14) und (2.15) für ein  $k \in \mathbb{N}$ .

Dann ist

$$\sum_{j=1}^k \beta_{(n,k),j} p^j = \mu_k(B_{n,p}) = \sum_{j=1}^n j^k \binom{n}{j} p^j (1-p)^{n-j}$$

für alle  $n \in \mathbb{N}_0$  und  $p \in [0, 1]$ . Aufgrund der Gleichheit dieser Polynome ist  $\beta_{(n,k),j} = 0$  für alle  $n \in \mathbb{N}_0$  mit  $j \geq n + 1$ . Unter Ausnutzung von Riordan (1937, (3.7)) in der ersten Gleichung ist für alle  $n \in \mathbb{N}_0$  und  $p \in [0, 1]$

$$\begin{aligned} \mu_{k+1}(B_{n,p}) &= \mu_1(B_{n,p}) \mu_k(B_{n,p}) + p(1-p) \frac{d\mu_k(B_{n,p})}{dp} \\ &= \mu_1(B_{n,p}) \mu_k(B_{n,p}) + (1-p) \sum_{j=1}^k j \beta_{(n,k),j} p^j. \end{aligned}$$

Analog ist unter Ausnutzung von Riordan (1937, (3.8)) in der ersten Gleichung für alle  $\lambda \in [0, \infty[$

$$\mu_{k+1}(P_\lambda) = \lambda \int x^k dP_\lambda + \lambda \frac{d\mu_k(P_\lambda)}{d\lambda} = \mu_1(P_\lambda) \mu_k(P_\lambda) + \sum_{j=1}^k j \alpha_{k,j} \lambda^j.$$

Also ist für alle  $P \in \text{BP}$

$$\mu_{k+1}(P) \leq (\mu_1(P) + k) \mu_k(P) \leq (1 + \mu_1(P)) k \mu_k(P) \leq k! \mu_1(P) (1 + \mu_1(P))^k, \quad (2.16)$$

und es gilt (2.13) für  $k + 1$ .

Mit  $\alpha_{k,0} := \alpha_{k,k+1} := \beta_{(n,k),0} := \beta_{(n,k),k+1} := 0$  für alle  $n \in \mathbb{N}_0$  ist  $\mu_{k+1}(B_{n,p}) = \sum_{j=1}^{k+1} (\beta_{(n,k),j-1} (n-j+1) + \beta_{(n,k),j}) p^j$  für alle  $n \in \mathbb{N}_0$  und  $p \in [0, 1]$  und  $\mu_{k+1}(P_\lambda) = \sum_{j=1}^{k+1} \lambda^j (\alpha_{k,j-1} + \alpha_{k,j})$  für alle  $\lambda \in [0, \infty[$ . Also gelten (2.14) und (2.15) für  $k + 1$ .  $\square$

**Satz 2.22.** *Es seien  $(P_n) \in \text{BP}^{\mathbb{N}}$  und  $P \in \text{Prob}(\mathbb{N}_0)$  mit  $P_n(\{z\}) \rightarrow P(\{z\})$ ,  $n \rightarrow \infty$ , für alle  $z \in \mathbb{N}_0$ . Dann ist  $P \in \text{BP}$  und für alle  $k \in \mathbb{N}$  gilt*

$$\int x^k dP_n \rightarrow \int x^k dP, \quad n \rightarrow \infty.$$

## 2 Metriken, uniforme Strukturen und Topologien

*Beweis.* Nach Satz 2.19 ist  $P \in \text{BP}$ . Sei  $k \in \mathbb{N}$ . Nach Lemma 2.20 konvergieren  $P_n \rightsquigarrow P$  und  $\int x dP_n \rightarrow \int x dP, n \rightarrow \infty$ . Also existiert ein  $C_1 < \infty$  mit  $\sup\{\int x dP_n : n \in \mathbb{N}\} < C_1$ . Also existiert nach Satz 2.21 zu jedem  $k \in \mathbb{N}$  ein  $C_k < \infty$  mit  $\sup\{\int x^k dP_n : n \in \mathbb{N}\} < C_k$ . Damit folgt die Behauptung mit Gut (2005, Chapter 5, Theorem 4.2. und 5.9.)  $\square$

**Satz 2.23.** Sei  $(P_n) \in \text{BP}^{\mathbb{N}}$  mit

$$\mu(P_n) \rightarrow \mu \in [0, \infty[ \quad \text{und} \quad \sigma^2(P_n) \rightarrow \sigma^2, \quad n \rightarrow \infty.$$

Dann existiert genau ein  $P \in \text{BP}$  mit  $\mu(P) = \mu, \sigma^2(P) = \sigma^2$  und

$$P_n(\{k\}) \rightarrow P(\{k\}), \quad n \rightarrow \infty, \quad \text{für alle } k \in \mathbb{N}_0.$$

*Beweis.* Da  $0 \leq \sigma^2(P_n) \leq \mu(P_n)$  für alle  $n \in \mathbb{N}$ , ist  $0 \leq \sigma^2 \leq \mu < \infty$ , und wir untersuchen die Fälle  $(\alpha) \sigma^2 = 0$ ,  $(\beta) 0 < \sigma^2 = \mu$  und  $(\gamma) 0 < \sigma^2 < \mu$ .

Falls  $(\alpha)$  gilt, konvergiert nach der Chebyshev-Ungleichung  $P_n(\{\mu\}) \rightarrow 1, n \rightarrow \infty$ , und  $\mu \in \mathbb{N}_0$ , also  $P_n(\{k\}) \rightarrow \delta_\mu(\{k\}), n \rightarrow \infty$ , für alle  $k \in \mathbb{N}_0$ , und  $\delta_\mu \in \text{BP}$ .

Falls  $(\beta)$  gilt, existiert für jede Teilfolge von  $(P_n)$  eine Teilteilfolge, entlang der  $(P_n) \in \{B_{m,p} : m \in \mathbb{N}_0, p \in [0, 1]\}^{\mathbb{N}}$  oder  $(P_n) \in \{P_\lambda : \lambda \in [0, \infty]\}^{\mathbb{N}}$ . Im Binomial-Fall gilt  $\frac{\mu(P_n)}{\sigma^2(P_n)} \rightarrow \frac{\mu}{\sigma^2} = 1, n \rightarrow \infty$ , nach Koopman (1950) also  $P_n(\{k\}) \rightarrow P_\mu(\{k\}), n \rightarrow \infty$ , für alle  $k \in \mathbb{N}_0$ . Im Poisson-Fall gilt auch  $P_n(\{k\}) \rightarrow P_\mu(\{k\}), n \rightarrow \infty$ , für alle  $k \in \mathbb{N}_0$ . Es ist  $P_\mu \in \text{BP}$ , und nach dem Teilteilfolgenargument gilt  $P_n(\{k\}) \rightarrow P_\mu(\{k\}), n \rightarrow \infty$ , für alle  $k \in \mathbb{N}_0$ .

Falls  $(\gamma)$  gilt, ist schließlich  $P_n = B_{m_n, p_n}$  für  $(m_n, p_n) \in (\mathbb{N}_0 \times [0, 1])^{\mathbb{N}}$ , und  $p_n \rightarrow 1 - \frac{\sigma^2}{\mu} =: q \in ]0, 1[$ , also  $m_n \rightarrow \frac{\mu}{q}$ , und damit  $B_{m_n, p_n}(\{k\}) \rightarrow B_{\frac{\mu}{q}, q}(\{k\}), n \rightarrow \infty$ , und  $B_{\frac{\mu}{q}, q} \in \text{BP}$ .  $\square$

Betrachten wir nun in den Definitionen und Beispielen 2.8 – 2.12 statt  $\text{Prob}(\mathbb{R})$  bzw.  $\text{Prob}(\mathbb{N}_0)$  die Menge  $\text{BP}$  aller Binomial- und Poisson-Verteilungen mit den entsprechenden Relativtopologien, so gelten die folgenden (mehr oder weniger bekannten) Äquivalenzen, deren Aussagen wegen (2.21) auch äquivalent zu den Aussagen im Portmanteau-Theorem sind.

**Satz 2.24.** Für  $(P_n) \in \text{BP}^{\mathbb{N}}$  sind die Aussagen

$$P_n \text{ konvergiert in Totalvariation gegen } P, \quad (2.17)$$

$$P_n \text{ konvergiert im Prochorov-Abstand gegen } P \in \text{BP}, \quad (2.18)$$

$$P_n \text{ konvergiert im Kolmogorov-Abstand gegen } P \in \text{BP}, \quad (2.19)$$

$$P_n \text{ konvergiert im Hellinger-Abstand gegen } P, \quad (2.20)$$

$$P_n \text{ konvergiert schwach gegen } P, \quad (2.21)$$

$$P_n \text{ konvergiert in absoluten Momenten gegen } P \in \text{BP}, \quad (2.22)$$

$$P_n \text{ konvergiert im ersten und zweiten Moment gegen } P \in \text{BP}, \quad (2.23)$$

$$P_n \text{ konvergiert punktweise gegen } P \quad (2.24)$$

äquivalent, und aus jeder folgt  $P \in \text{BP}$ .

*Beweis.* Aus (2.17) folgt (2.24), nach Satz 2.19 also  $P \in \text{BP}$ . Aus (2.24) folgt nach Satz 2.22 (2.22). Aus (2.22) folgt (2.23). Aus (2.23) folgt nach Satz 2.23 (2.24). Aus (2.24) folgt nach Gut (2005, Chapter 5, Theorem 6.4.) (2.17).

Aus (2.17) folgt nach obigem  $P \in \text{BP}$ , also sind (2.17), (2.18), (2.19) und (2.20) nach (2.2), (2.4) und (2.6) äquivalent.

Für die Äquivalenz zwischen (2.21) und (2.18) vgl. z. B. Billingsley (1999, S. 72).  $\square$



# 3 Proximität von Binomial- und Poisson-Verteilungen

Gegenstand dieses Kapitels ist die Untersuchung und Beschreibung uniformer Strukturen auf der Menge der Binomial- und Poisson-Verteilungen. Dies vorbereitend werden in den folgenden Abschnitten 3.1 – 3.4 Abschätzungen für Abstände verschiedener Binomial-, Poisson- und Normalverteilungen angegeben. Die Ungleichungen der Abschnitte 3.1 – 3.3 ergeben zusammen eine obere Abschätzung für den Totalvariationsabstand auf der Menge BP. In Abschnitt 3.5 wird die uniforme Struktur des Totalvariationsabstandes auf der Menge BP untersucht und in Satz 3.19 gezeigt, wie sie mit Hilfe der Erwartungswerte und Varianzen der Verteilungen beschrieben werden kann. Insbesondere wird eine obere Abschätzung des Totalvariationsabstandes durch eine geeignete Funktion der Erwartungswerte und Varianzen angegeben, die sich aus den Ergebnissen der Abschnitte 3.1 – 3.3 ergibt.

In Abschnitt 3.6 werden offene Forschungsfragen zu diesem Thema aufgezeigt und es schließt sich ein Abschnitt mit den Beweisen der Ergebnisse an. Schließlich werden in Abschnitt 3.8 einige Ergebnisse der unveröffentlichten Arbeit Tasto (2011) bereitgestellt, die in diesem Themenkreis von Interesse sind.

Für den im Umgang mit uniformen Strukturen ungeübten Leser werden die für dieses Kapitel benötigten Aussagen über uniforme Strukturen in Abschnitt 2.2 bereitgestellt.

Spezialfälle der Ergebnisse dieses Kapitels gehen auf Barbour und Hall (1984), Ehm (1991) und Roos und Pfeifer (2003a) zurück, wobei insbesondere die ersten beiden die in Stein (1970) entwickelte und in Chen (1975) weiterentwickelte Stein-Chen-Methode verwenden, siehe z. B. Barbour und Chen (2005) für eine Einführung in die Methode.

Einführungen in und Übersichten über die Themen Proximität von Verteilungen und uniforme Strukturen auf Verteilungen finden sich in D'Aristotile et al. (1988), Davydov und Rotar (2009) und Dudley (2002, Abschnitt 11.7).

## 3.1 Approximationen von Binomialverteilungen

Inhalt dieses Abschnitts ist die Berechnung oberer Schranken des Totalvariationsabstandes verschiedener Binomialverteilungen. Die Kombination dieser Ergebnisse findet in Satz 3.20 statt.

### 3.1.1 Binomialverteilungen gleicher Länge

**Satz 3.1.** Seien  $n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$ . Dann ist

$$d_{\text{TV}}(B_{n,p}, B_{n,q}) \leq \sqrt{2} \sqrt{\frac{(np - nq)^2}{np(1-p) + nq(1-q)}}. \quad (3.1)$$

**Lemma 3.2.** In der Situation von Satz 3.1 ( $n \in \mathbb{N}_0, p, q \in [0, 1]$ ) gilt außerdem

$$\exp\left(-\frac{2}{(\sqrt{np} - \sqrt{nq})^2 + (\sqrt{n(1-p)} - \sqrt{n(1-q)})^2}\right) \leq d_{\text{H}}^2(B_{n,p}, B_{n,q}). \quad (3.2)$$

Dies liefert die folgenden Grenzwertaussagen:

**Satz 3.3.** Seien  $(n) := (n_k) \in \mathbb{N}_0^{\mathbb{N}}$  und  $(p) := (p_k), (q) := (q_k) \in [0, 1]^{\mathbb{N}}$ . Für  $k \rightarrow \infty$  folgen dann aus

$$\frac{(np - nq)^2}{np(1-p) + nq(1-q)} \rightarrow 0 \quad (3.3)$$

die äquivalenten Aussagen

$$d_{\text{H}}(B_{n,p}, B_{n,q}) \rightarrow 0, \quad (3.4)$$

$$|\sqrt{np} - \sqrt{nq}| + \left| \sqrt{n(1-p)} - \sqrt{n(1-q)} \right| \rightarrow 0, \quad (3.5)$$

$$|\sqrt{np} - \sqrt{nq}| + \left| \sqrt{np(1-p)} - \sqrt{nq(1-q)} \right| \rightarrow 0. \quad (3.6)$$

### 3.1.2 Binomialverteilungen gleicher Erfolgswahrscheinlichkeit

Der Spezialfall der Abschätzung des Abstandes zweier Binomialverteilungen mit gleicher Erfolgswahrscheinlichkeit  $p$  lässt sich mit Hilfe eines Ergebnisses für Zufallssummen aus Roos und Pfeifer (2003a) behandeln:

**Satz 3.4.** Seien  $m, n \in \mathbb{N}_0$  und  $p \in [0, 1]$ . Dann ist

$$d_{\text{TV}}(B_{n,p}, B_{m,p}) \leq \frac{|np - mp|}{\max\{1, \sqrt{np(1-p) + mp(1-p)}\}} \leq \sqrt{\frac{(np - mp)^2}{np(1-p) + mp(1-p)}}.$$

*Bemerkung 3.5.* Als Vorbereitung auf eine mögliche, aber hier nicht durchgeführte Abschätzung des Totalvariationsabstandes in (3.18) nach unten gegen die Ausdrücke in (3.19) sei angemerkt, dass in der Situation von Satz 3.4

$$\frac{|np(1-p) - mp(1-p)|}{1 + np(1-p) + mp(1-p)} = \frac{(1-p)|np - mp|}{1 + np(1-p) + mp(1-p)} \leq \sqrt{\frac{(np - mp)^2}{np(1-p) + mp(1-p)}}$$

gilt.

### 3.1.3 Binomialverteilungen mit gleichem Erwartungswert

Der Spezialfall der Approximation zweier Binomialverteilungen mit gleichem Erwartungswert ergibt sich z. B. durch Anwendung von Ergebnissen aus Ehm (1991):

**Satz 3.6.** Seien  $n \in \mathbb{N}_0$  und  $p, p_1, \dots, p_n \in [0, 1]$  mit  $np = \sum_{j=1}^n p_j$ . Dann ist  $\Sigma^2 :=$

$$\sigma^2(\mathbb{B}_{n,p}) \geq \sigma^2\left(\bigstar_{j=1}^n \mathbb{B}_{p_j}\right) =: \sigma^2 \text{ und}$$

$$C \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2} \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \bigstar_{j=1}^n \mathbb{B}_{p_j}) \leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2},$$

wobei  $C \geq 1/124$  eine universelle Konstante ist.

Inbesondere ist für  $m, n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $np = mq$

$$C \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \leq 3 \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}. \quad (3.7)$$

Roos und Pfeifer (2003b, Example 1, dort Verteilungen  $\mathbb{B}_{m,p}$  und  $\mathbb{B}_{n,q}$ ) entwickeln Abschätzungen für Zufallssummen, aus denen man durch die hier vorgenommene Spezialisierung die folgende obere Abschätzung erhält, die die gleiche Konvergenzrate wie in (3.7) nach Ehm (1991) ergibt, aber eine größere Konstante enthält:

**Satz 3.7.** Es seien  $m, n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $np = mq$ . Dann ist

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \leq \sqrt{2}(\sqrt{2} + 1) \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

### 3.1.4 Weitere Binomialverteilungen

**Satz 3.8.** Seien  $m \leq n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $q \leq p$ . Dann ist

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \leq (\sqrt{2} + 1) \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}}. \quad (3.8)$$

*Bemerkung 3.9.* Als Vorbereitung auf eine mögliche, aber hier nicht durchgeführte Abschätzung des Totalvariationsabstandes in (3.18) nach unten gegen die Ausdrücke in (3.19) sei angemerkt, dass in der Situation von Satz 3.8

$$\frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} \leq \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}} \quad (3.9)$$

gilt.

**Satz 3.10.** Seien  $m \leq n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $r_1 := \frac{mp}{m} < \min\{q, 1/2\}$ . Dann ist

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \leq 5 \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}} + 3 \left( \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} \right).$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

**Satz 3.11.** Seien  $m, n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $m \leq n$  und  $p \in [\frac{m}{n}q, 1 - \frac{m}{n}q] \cup [1 - \frac{m}{n}q, \frac{m}{n}q]$ . Dann ist

$$d_{\text{TV}}(B_{n,p}, B_{m,q}) \leq \sqrt{2} \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}} + 3 \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

**Lemma 3.12.** Seien  $1 \leq k \in \mathbb{N}$  und  $p, q \in [0, 1]$  mit  $\frac{1}{2} \leq p \leq q$  und  $k(1-p) \leq (k-1)q \leq kp$ . Dann ist

$$d_{\text{TV}}(B_{k,p}, B_{k-1,q}) \leq 2 \sqrt{\frac{(kp - (k-1)q)^2}{kp(1-p) + (k-1)q(1-q)}} + 3 \frac{|kp(1-p) - (k-1)q(1-q)|}{1 + kp(1-p) + (k-1)q(1-q)}.$$

**Satz 3.13.** Seien  $m, n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $m \leq n$  und  $\max\{p, \frac{n}{m}(1-p)\} < q < \frac{n}{m}p$ . Dann ist

$$d_{\text{TV}}(B_{n,p}, B_{m,q}) \leq 3 \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}} + 6 \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

**Satz 3.14.** Seien  $m, n \in \mathbb{N}_0$  und  $p, q \in [0, 1]$  mit  $m \leq n$  und  $\frac{m}{n}\frac{1}{2} \leq p < \frac{m}{n}q$ . Dann ist

$$d_{\text{TV}}(B_{n,p}, B_{m,q}) \leq \sqrt{2} \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}} + 12 \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

## 3.2 Approximationen von Poisson-Verteilungen

In diesem Abschnitt werden nicht nur Abschätzungen für den Totalvariationsabstand auf der Menge der Poisson-Verteilungen angegeben, sondern auch die uniforme Struktur eingehender untersucht, da diese relativ anschaulich zu beschreiben ist.

Da für Poisson-Verteilungen Erwartungswert und Varianz übereinstimmen, vereinfachen sich die die Konvergenz beschreibenden Ausdrücke im Poisson-Fall:

Für  $\lambda, \mu \in [0, \infty[$  ist

$$\frac{|\lambda - \mu|}{1 + \lambda + \mu} = |\sqrt{\lambda} - \sqrt{\mu}| \cdot \frac{\sqrt{\lambda} + \sqrt{\mu}}{1 + \lambda + \mu} \leq |\sqrt{\lambda} - \sqrt{\mu}|$$

und

$$\begin{aligned} \frac{(\lambda - \mu)^2}{\lambda + \mu} &= \begin{cases} \max\{\lambda, \mu\}, & \text{falls } \lambda\mu = 0, \\ (\sqrt{\lambda} - \sqrt{\mu})^2 \left(1 + 2 / \left(\sqrt{\frac{\mu}{\lambda}} + \sqrt{\frac{\lambda}{\mu}}\right)\right), & \text{falls } \lambda\mu > 0, \end{cases} \\ &\leq 2(\sqrt{\lambda} - \sqrt{\mu})^2, \end{aligned}$$

also

$$\frac{1}{2} \left( \frac{|\lambda - \mu|}{1 + \lambda + \mu} \right)^2 + \frac{1}{6} \frac{(\lambda - \mu)^2}{\lambda + \mu} \leq (\sqrt{\lambda} - \sqrt{\mu})^2 \leq \frac{(\lambda - \mu)^2}{\lambda + \mu}.$$

Für zwei Poisson-Verteilungen mit Parametern  $\lambda, \mu \in [0, \infty[$  lässt sich der Hellinger-Abstand explizit ausrechnen als

$$d_H(P_\lambda, P_\mu) = \sqrt{1 - \exp\left(-\frac{1}{2}(\sqrt{\lambda} - \sqrt{\mu})^2\right)}, \quad (3.10)$$

und man erhält mit obigem die Abschätzungen

$$\begin{aligned} & \sqrt{1 - \exp\left(-\left(\frac{1}{12} \frac{(\lambda - \mu)^2}{\lambda + \mu} + \frac{1}{4} \left(\frac{|\lambda - \mu|}{1 + \lambda + \mu}\right)^2\right)\right)} \\ & \leq d_H(P_\lambda, P_\mu) \leq \sqrt{1 - \exp\left(-\frac{1}{2} \frac{(\lambda - \mu)^2}{\lambda + \mu}\right)}. \end{aligned} \quad (3.11)$$

Für zwei Folgen  $(P_{\lambda_n})$  und  $(P_{\mu_n})$  von Poisson-Verteilungen sind nach obigem also die folgenden Konvergenzen für  $n \rightarrow \infty$  äquivalent:

$$\begin{aligned} & \text{(i) } d_{TV}(P_{\lambda_n}, P_{\mu_n}) \rightarrow 0, \quad \text{(ii) } d_K(P_{\lambda_n}, P_{\mu_n}) \rightarrow 0, \\ & \text{(iii) } d_H(P_{\lambda_n}, P_{\mu_n}) \rightarrow 0, \quad \text{(iv) } d_P(P_{\lambda_n}, P_{\mu_n}) \rightarrow 0, \\ & \text{(v) } |\sqrt{\lambda_n} - \sqrt{\mu_n}| \rightarrow 0, \quad \text{(vi) } \frac{(\lambda_n - \mu_n)^2}{\lambda_n + \mu_n} \rightarrow 0, \\ & \text{(vii) } \frac{|\lambda_n - \mu_n|}{1 + \lambda_n + \mu_n} + \frac{(\lambda_n - \mu_n)^2}{\lambda_n + \mu_n} \rightarrow 0. \end{aligned}$$

Aus (3.10) erhält man außerdem nach (2.2) als obere Schranke für den Totalvariationsabstand

$$d_{TV}(P_\lambda, P_\mu) \leq \sqrt{1 - \exp\left(-(\sqrt{\lambda} - \sqrt{\mu})^2\right)},$$

die aber z. B. für  $\mu \approx 0$  nicht kleiner als die beste obere Schranke in (3.12) unten ist. Ein vollständiger Vergleich scheint hier noch sinnvoll.

Einige weitere obere Schranken für den Totalvariationsabstand zweier Poisson-Verteilungen findet man in Adell und Jodrá (2006, (2.2)):

$$\begin{aligned} d_{TV}(P_\lambda, P_\mu) & \leq \min \left\{ 1 - \exp(-|\lambda - \mu|), \int_{\min\{\lambda, \mu\}}^{\max\{\lambda, \mu\}} P_u(\{\lfloor u \rfloor\}) du \right\} \\ & \leq \int_{\min\{\lambda, \mu\}}^{\max\{\lambda, \mu\}} P_u(\{\lfloor u \rfloor\}) du \\ & \leq \min \left\{ |\lambda - \mu|, \sqrt{\frac{2}{e}} \cdot |\sqrt{\lambda} - \sqrt{\mu}| \right\}, \end{aligned} \quad (3.12)$$

wobei die erste Schranke in Adell und Lekuona (2005, Corollary 3.1) gegeben ist, die zweite in Ruzankin (2004, Beweis von Lemma 1), und die dritte in Roos (2003, S. 469).

### 3 Proximität von Binomial- und Poisson-Verteilungen

Ruzankin (2004) weist außerdem die letzte Ungleichung in (3.12) nach, wobei nicht klar ist, ob das Ergebnis aus Roos (2003) bekannt sein konnte.

Aus (3.12) ergibt sich für  $\lambda, \mu \in [0, \infty[$ , da  $(\sqrt{\lambda} - \sqrt{\mu})^2 \leq (\lambda - \mu)^2 / (\lambda + \mu)$ , die obere Schranke

$$d_{\text{TV}}(P_\lambda, P_\mu) \leq \sqrt{\frac{2}{e}} \cdot \sqrt{\frac{(\lambda - \mu)^2}{\lambda + \mu}}. \quad (3.13)$$

Eine weitere Darstellung des Totalvariationsabstandes, die für unsere Zwecke aber nicht geeignet ist, findet man in Adell und Lekuona (2005, Proposition 2.1).

Damit können wir die uniforme Struktur, z. B. des Totalvariationsabstandes (vgl. Beispiel 2.15), auf der Menge  $\{P_\lambda: \lambda \in [0, \infty[ \}$  der Poisson-Verteilungen charakterisieren: Eine Umgebungsbasis ist nach (3.10) gegeben durch

$$\begin{aligned} \mathcal{B} &= \left\{ \left\{ (P_\lambda, P_\mu) : \sqrt{1 - \exp\left(-\frac{1}{2}(\sqrt{\lambda} - \sqrt{\mu})^2\right)} < \varepsilon \right\} : \varepsilon \in ]0, 1] \right\} \\ &= \{U_\delta : \delta > 0\}, \end{aligned}$$

wobei

$$U_\delta := \{(P_\lambda, P_\mu) : |\sqrt{\lambda} - \sqrt{\mu}| < \delta\} \quad \text{für } \delta > 0.$$

In Abbildung 3.1 ist die uniforme Struktur auf der Menge der Poisson-Verteilungen, parametrisiert durch den Erwartungswert, durch die Darstellung der Elemente  $U_1$ ,  $U_{1/2}$  und  $U_{1/10}$  der Umgebungsbasis  $\mathcal{B}$  und der Menge  $\{\lambda = \mu\}$  angedeutet:

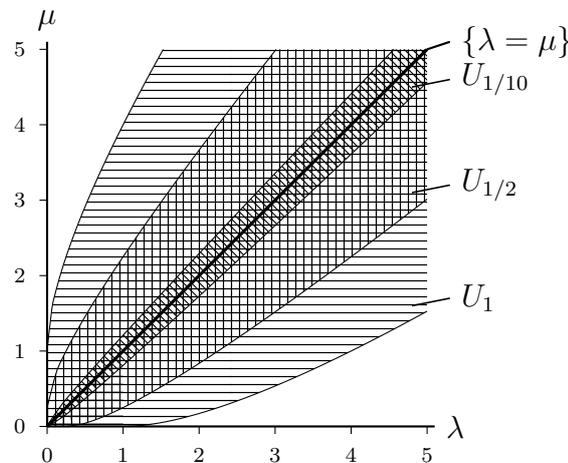


Abbildung 3.1: Uniforme Struktur auf der Menge der Poisson-Verteilungen

## 3.3 Approximationen von Binomial- und Poisson-Verteilungen

Wir nutzen Barbour und Hall (1984) für die folgende Abschätzung für den Abstand einer Binomial- und einer Poisson-Verteilung mit gleichem Erwartungswert:

**Satz 3.15.** Seien  $n \in \mathbb{N}_0$ ,  $p_1, \dots, p_n \in [0, 1]$  und  $\lambda := \sum_{j=1}^n p_j$ . Dann ist  $\Sigma^2 := \sigma^2(\mathbb{P}_\lambda) \geq \sigma^2(\bigstar_{j=1}^n \mathbb{B}_{p_j}) =: \sigma^2$  und

$$\frac{1}{32} \cdot \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2} \leq d_{\text{TV}}(\bigstar_{j=1}^n \mathbb{B}_{p_j}, \mathbb{P}_\lambda) \leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}. \quad (3.14)$$

Insbesondere ist für  $n \in \mathbb{N}_0$ ,  $p \in [0, 1]$  und  $\lambda := np$

$$\frac{p \cdot \min\{1, np\}}{32} \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{P}_\lambda) \leq p \cdot \min\{1, np\} \quad (3.15)$$

und

$$\frac{1}{32} \cdot \frac{|\lambda - np(1-p)|}{1 + \lambda + np(1-p)} \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{P}_\lambda) \leq 3 \frac{|\lambda - np(1-p)|}{1 + \lambda + np(1-p)}. \quad (3.16)$$

Eine Abschätzung des Abstandes beliebiger Binomial- und Poisson-Verteilungen liefert der folgende Satz, wobei die obere Schranke nicht in jedem Fall die schärfste bekannte ist (vgl. z. B. (3.33) und (3.12)), aber einfach und für die Bestimmung der uniformen Struktur auf BP ausreichend.

**Satz 3.16.** Seien  $n \in \mathbb{N}_0$ ,  $p \in [0, 1]$  und  $\lambda \in [0, \infty[$ . Dann ist

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{P}_\lambda) \leq \left( \sqrt{\frac{2}{e}} + 3 \right) \sqrt{\frac{(np - \lambda)^2}{np(1-p) + \lambda}} + 3 \frac{|np(1-p) - \lambda|}{1 + np(1-p) + \lambda}.$$

## 3.4 Normalverteilungs-Approximationen

Für die Berechnung des Hellinger-Abstandes unten benötigen wir die folgende Identität: Für  $a, b, c \in \mathbb{R}$  mit  $a > 0$  ist  $\int_{\mathbb{R}} \exp(-ax^2 + bx + c) d\lambda(x) = \sqrt{\pi/a} \exp(b^2/(4a) + c)$ . Für  $\mu \in \mathbb{R}$  und  $\sigma^2 \in ]0, \infty[$  sei  $f_{\mu, \sigma^2}: \mathbb{R} \rightarrow [0, 1]$ ,  $x \mapsto \exp(-(x - \mu)^2/(2\sigma^2)) / (\sqrt{2\pi}\sigma)$ .

Damit ist für  $\mu, \lambda \in \mathbb{R}$  und  $\sigma^2, \tau^2 \in ]0, \infty[$

$$\begin{aligned} & d_{\text{H}}(\mathbb{N}_{\mu, \sigma^2}, \mathbb{N}_{\lambda, \tau^2}) \\ &= \left( 1 - \int_{\mathbb{R}} \sqrt{f_{\mu, \sigma^2} f_{\lambda, \tau^2}} d\lambda \right)^{1/2} \\ &= \left( 1 - \frac{1}{\sqrt{2\pi\sigma\tau}} \int_{\mathbb{R}} \exp\left(-\frac{\tau^2 + \sigma^2}{4\sigma^2\tau^2} x^2 + \frac{\mu\tau^2 + \lambda\sigma^2}{2\sigma^2\tau^2} x - \frac{\lambda^2\sigma^2 + \mu^2\tau^2}{4\sigma^2\tau^2}\right) d\lambda(x) \right)^{1/2} \\ &= \left( 1 - \sqrt{\frac{2\sigma\tau}{\sigma^2 + \tau^2}} \exp\left(-\frac{1}{4} \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2}\right) \right)^{1/2} \end{aligned}$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

(vgl. auch Liese und Miescke (2008, S. 37, dort  $D := \sqrt{2}d_H$ )). Da für  $\mu, \lambda \in \mathbb{R}$  und  $0 < \sigma^2 \leq \tau^2$   $\{f_{\mu, \sigma^2} \geq f_{\lambda, \tau^2}\} \subseteq \mathbb{R}$  ein Intervall ist, folgt analog zu Satz 2.1 für  $\mu, \lambda \in \mathbb{R}$  und  $\sigma^2, \tau^2 \in ]0, \infty[$

$$d_K(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}) \leq d_{TV}(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}) \leq 2d_K(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}).$$

In Verbindung mit (2.2) ergibt sich

$$\begin{aligned} \frac{1}{2} \left( 1 - \exp \left( -\frac{1}{4} \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2} \right) \right) &\leq \frac{1}{2} \left( 1 - \sqrt{\frac{2\sigma\tau}{\sigma^2 + \tau^2}} \exp \left( -\frac{1}{4} \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2} \right) \right) \\ &= \frac{1}{2} d_H^2(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}) \\ &\leq d_K(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}) \\ &\leq d_H(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}) \sqrt{2 - d_H^2(N_{\mu, \sigma^2}, N_{\lambda, \tau^2})} \\ &= \sqrt{1 - \frac{2\sigma\tau}{\sigma^2 + \tau^2} \exp \left( -\frac{1}{2} \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2} \right)}. \end{aligned} \tag{3.17}$$

## 3.5 Uniforme Struktur auf der Menge BP

Ziel der vorliegenden Arbeit war es, die uniforme Struktur des Totalvariationsabstandes auf der Menge der Binomialverteilungen zu charakterisieren, also zu untersuchen, wann zwei Binomialverteilungen „nahe beieinander“ liegen, wobei diese Nähe durch den Totalvariationsabstand gemessen wird. Da wir die uniforme Struktur des Totalvariationsabstandes, also einer Metrik, betrachten, kann die uniforme Struktur durch die Charakterisierung der Konvergenz  $\lim_{n \rightarrow \infty} d_{TV}(P_n, Q_n) = 0$  für Folgen  $(P_n), (Q_n)$  von Binomialverteilungen charakterisiert werden.

Durch Beschränkung auf die Untersuchung von  $\lim_{n \rightarrow \infty} d_{TV}(P_n, P) = 0$  für eine Folge  $(P_n)$  von Binomialverteilungen erhält man eine Charakterisierung der Topologie des Totalvariationsabstandes, so dass es sinnvoll ist, die Untersuchung nicht auf Binomialverteilungen zu beschränken, sondern den Abschluss der Binomialverteilungen in der Topologie des Totalvariationsabstandes zu untersuchen. Dieser besteht nach Satz 2.24 aus der Menge der Binomial- und Poisson-Verteilungen

$$BP := \{B_{n,p}, P_\lambda : n \in \mathbb{N}_0, p \in [0, 1], \lambda \in [0, \infty[ \}.$$

Zur Illustration der Menge BP nutzen wir, dass sie sich vermöge der Abbildung

$$BP \ni P \mapsto (\mu(P), \sigma^2(P)) \in \{(\mu(P), \sigma^2(P)) : P \in BP\}$$

mit letzterer Menge, die in Abbildung 3.2 skizziert ist, identifizieren lässt:

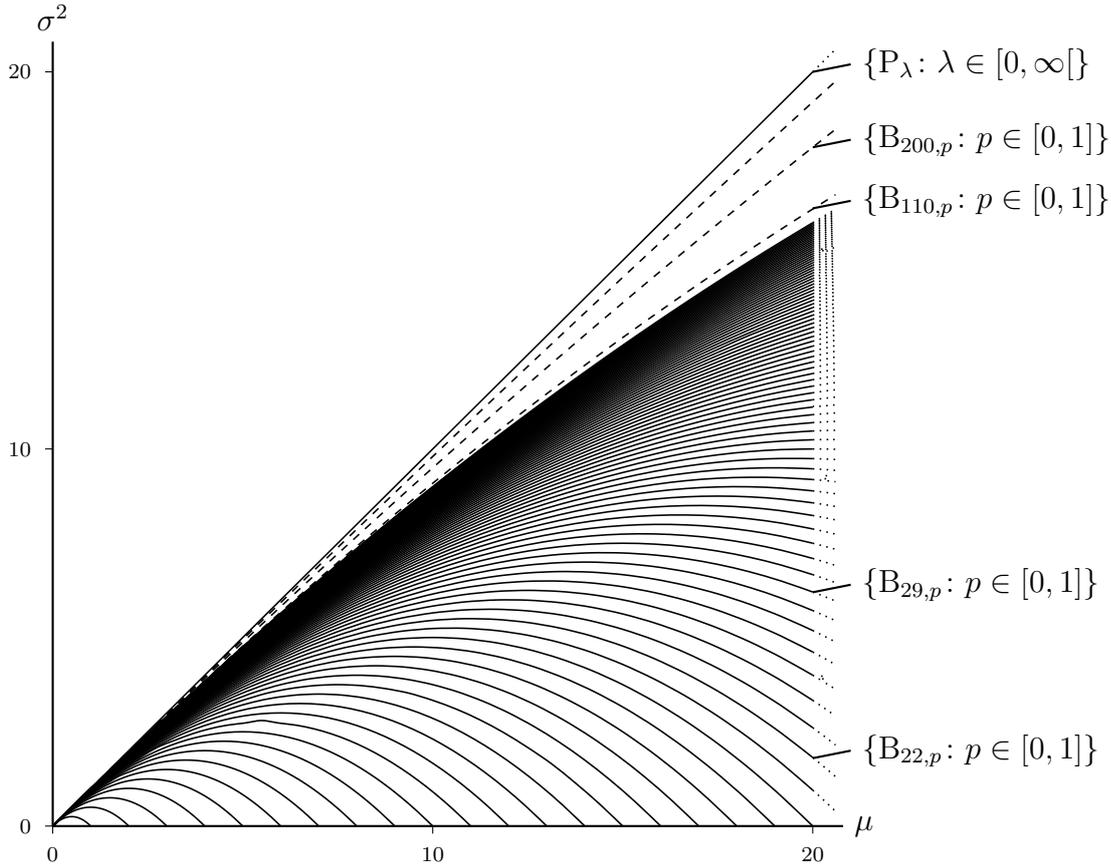


Abbildung 3.2: Die Menge BP, identifiziert durch Erwartungswerte und Varianzen

Wie in Beispiel 2.15 zu sehen, ist auf BP durch  $d_{TV}$ ,  $d_K$ ,  $d_P$  und  $d_H$  die gleiche uniforme Struktur gegeben.

Der folgende Satz 3.17 zeigt, dass eine geeignete Funktion der Erwartungswerte und Varianzen, die die Elemente in BP eindeutig identifizieren, eine uniforme Struktur auf BP erzeugt. Satz 3.19 zeigt, dass diese uniforme Struktur gleich der uniformen Struktur des Totalvariationsabstandes ist. In Satz 3.20 wird diese Aussage durch die Angabe einer oberen Abschätzung des Totalvariationsabstandes durch eine geeignete Funktion der Erwartungswerte und Varianzen konkretisiert. Eine entsprechende untere Abschätzung ist Gegenstand aktueller Forschung. Erste Schritte hierfür wurden unternommen, unter anderem gilt die untere Abschätzung in (3.21).

**Satz 3.17.** Auf  $\mathcal{X} := BP$  bildet  $\mathcal{U} = \{U_\varepsilon : \varepsilon > 0\}$ , mit

$$U_\varepsilon := \left\{ (P, Q) \in \mathcal{X} \times \mathcal{X} : \frac{(\mu(P) - \mu(Q))^2}{\sigma(P)^2 + \sigma^2(Q)} + \frac{|\sigma^2(P) - \sigma^2(Q)|}{1 + \sigma^2(P) + \sigma^2(Q)} < \varepsilon \right\} \text{ für } \varepsilon > 0,$$

eine Basis des Nachbarschaftsfilters

$$\mathcal{F} := \{V \subseteq \mathcal{X} \times \mathcal{X} : \text{es existiert ein } \varepsilon > 0 \text{ mit } V \supseteq U_\varepsilon\}.$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

*Bemerkung 3.18.* In Satz 3.17 wird ausgenutzt, dass  $\{(\mu(P), \sigma^2(P)) : P \in \text{BP}\}$  nicht dicht in  $[0, \infty[_{\geq}^2 := \{(\mu, \sigma^2) \in [0, \infty[^2 : \mu \geq \sigma^2\}$  liegt. Setzt man

$$\mathcal{X} := \left\{ P \in \text{Prob}(\mathbb{R}) : (\mu(P), \sigma^2(P)) \in [0, \infty[_{\geq}^2 \right\},$$

so ist  $\mathcal{F}$  keine uniforme Struktur, da (2.12) nicht erfüllt ist: Für  $\varepsilon > 0$  und  $\delta > 0$  seien  $P := \delta_0$ ,  $Q := \delta_{\delta/2}$  und  $R := N_{\delta/4, \delta/5}$ . Dann sind  $P, Q, R \in \mathcal{X}$  und  $(P, R), (R, Q) \in U_\delta$ , aber  $(\mu(P) - \mu(Q))^2 / (\sigma^2(P) + \sigma^2(Q)) = \infty > \varepsilon$ , also  $(P, Q) \notin U_\varepsilon$ .

**Satz 3.19.** *Die durch den Totalvariationsabstand auf BP gegebene uniforme Struktur ist gleich der in Satz 3.17 definierten uniformen Struktur.*

*Seien  $(P_k)$  und  $(Q_k)$  Folgen in BP. Dann gilt für  $k \rightarrow \infty$*

$$d_{\text{TV}}(P_k, Q_k) \rightarrow 0 \quad (3.18)$$

*genau dann, wenn*

$$\frac{|\sigma^2(P_k) - \sigma^2(Q_k)|}{1 + \sigma^2(P_k) + \sigma^2(Q_k)} + \frac{(\mu(P_k) - \mu(Q_k))^2}{\sigma^2(P_k) + \sigma^2(Q_k)} \rightarrow 0. \quad (3.19)$$

Dabei folgt (3.18) direkt aus (3.19) mit Hilfe der folgenden Ungleichung:

**Satz 3.20.** *Seien  $P, Q \in \text{BP}$ . Dann ist*

$$d_{\text{TV}}(P, Q) \leq 5 \cdot \sqrt{\frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)}} + 12 \cdot \frac{|\sigma^2(P) - \sigma^2(Q)|}{1 + \sigma^2(P) + \sigma^2(Q)}.$$

Wie oben bereits erwähnt, ist eine untere Abschätzung des Totalvariationsabstandes als Gegenstück zur Ungleichung in Satz 3.20 Gegenstand aktueller Forschung. Das folgende Lemma 3.21 ist ein erster Schritt in diese Richtung.

**Lemma 3.21.** *Für  $j \in \mathbb{N}$  und  $P, Q \in \text{Prob}(\mathbb{R})$  seien  $\mu_j := \int x^j dP(x)$  und  $\lambda_j := \int x^j dQ(x)$ , falls die Integrale existieren.*

*Seien  $k \in \mathbb{N}$  und  $P, Q \in \text{Prob}(\mathbb{R})$  mit  $\mu_{2k}, \lambda_{2k} \in \mathbb{R}$ . Dann ist*

$$\frac{1}{3} \min \left\{ 1, \frac{(\mu_k - \lambda_k)^2}{\mu_{2k} - \mu_k^2 + \lambda_{2k} - \lambda_k^2} \right\} \leq \frac{\frac{(\mu_k - \lambda_k)^2}{2}}{\mu_{2k} - \mu_k^2 + \lambda_{2k} - \lambda_k^2 + \frac{(\mu_k - \lambda_k)^2}{2}} \leq d_{\text{TV}}(P, Q). \quad (3.20)$$

*Bemerkung 3.22.* In der Situation von Lemma 3.21 vereinfacht sich (3.20) für  $k = 1$  mit  $\mu := \mu(P)$ ,  $\sigma^2 := \sigma^2(P)$ ,  $\lambda := \mu(Q)$  und  $\tau^2 := \sigma^2(Q)$  zu

$$\frac{1}{3} \min \left\{ 1, \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2} \right\} \leq \frac{\frac{(\mu - \lambda)^2}{2}}{\sigma^2 + \tau^2 + \frac{(\mu - \lambda)^2}{2}} \leq d_{\text{TV}}(P, Q), \quad (3.21)$$

und in der zweiten Ungleichung in (3.21) gilt für  $P := (1 - \varepsilon)\delta_\varepsilon + \varepsilon\delta_{1+\varepsilon}$  und  $Q := (1 - \varepsilon)\delta_\varepsilon + \varepsilon\delta_{-1+\varepsilon}$  für  $\varepsilon \in [0, 1]$  stets Gleichheit:

$$\frac{\frac{(\mu - \lambda)^2}{2}}{\sigma^2 + \tau^2 + \frac{(\mu - \lambda)^2}{2}} = \frac{2\varepsilon^2}{2\varepsilon(1 - \varepsilon) + 2\varepsilon^2} = \varepsilon = d_{\text{TV}}(P, Q).$$

## 3.6 Offene Fragen

In der vorliegenden Arbeit wird die uniforme Struktur des Totalvariationsabstandes auf der Menge BP der Binomial- und Poisson-Verteilungen mit Hilfe der die Verteilungen eindeutig bestimmenden Erwartungswerte und Varianzen charakterisiert. Insbesondere wird eine obere Schranke für den Totalvariationsabstand zweier Verteilungen durch eine geeignete Funktion der Erwartungswerte und Varianzen angegeben.

Als Vorbereitung der Angabe einer unteren Schranke wurden in einigen Spezialfällen, z. B. (3.7) und (3.16), geeignete untere Schranken bereitgestellt. Neben einer Verbesserung der Schranken durch die Ausarbeitung besserer, wenn möglich optimaler, Konvergenzraten und Konstanten, wäre ein nächster Schritt, die Schranken durch eine Metrik auf  $\{(\mu(P), \sigma^2(P)) : P \in \text{BP}\}$  auszudrücken, und damit eine Quasiisometrie zwischen dieser Menge und der Menge BP charakterisieren zu können.

Außerdem sollten diese Untersuchungen auf weitere Mengen von Verteilungen ausgeweitet werden, z. B. auf die Menge der BP-Faltungen, auf der in Tasto (2011) die Topologie des Totalvariationsabstandes mit Hilfe der beteiligten Parameter charakterisiert wurde, oder sogar auf die Menge der BPBP-Faltungen, auf der in Tasto (2011) eine hinreichende Bedingung an die Konvergenz der beteiligten Parameter für die Konvergenz einer Folge von Verteilungen gegeben wurde.

## 3.7 Beweise der Ergebnisse

### 3.7.1 Beweise der Abschnitte 3.1 – 3.3

*Beweis von Satz 3.1.* Es ist

$$\begin{aligned}
 d_{\text{H}}(\text{B}_{n,p}, \text{B}_{n,q}) &= \left( \frac{1}{2} \sum_{k=0}^n \left( \sqrt{\text{B}_{n,p}(\{k\})} - \sqrt{\text{B}_{n,q}(\{k\})} \right)^2 \right)^{1/2} \\
 &= \left( 1 - \sum_{k=0}^n \sqrt{\text{B}_{n,p}(\{k\}) \cdot \text{B}_{n,q}(\{k\})} \right)^{1/2} \\
 &= \left( 1 - \left( \sqrt{pq} + \sqrt{(1-p)(1-q)} \right)^n \right)^{1/2} \\
 &= \left( 1 - \left( 1 - \frac{1}{2} \left( (\sqrt{p} - \sqrt{q})^2 + (\sqrt{1-p} - \sqrt{1-q})^2 \right) \right)^n \right)^{1/2} \quad (3.22) \\
 &= \left( 1 - \left( 1 - d_{\text{H}}^2(\text{B}_p, \text{B}_q) \right)^n \right)^{1/2},
 \end{aligned}$$

wobei die Gleichheit des ersten und letzten Ausdrucks schon bekannt ist (vgl. z. B. Verknüpfung von Liese und Miescke (2008, Theorem 4.45, dort  $D := \sqrt{2}d_{\text{H}}$ ) und Strasser (1985, Lemma 2.16)). Insbesondere ist nach (3.22) unter Anwendung der Bernoulli-

### 3 Proximität von Binomial- und Poisson-Verteilungen

Ungleichung  $((1+x)^n \geq 1+nx$  für  $x \geq -1, n \in \mathbb{N}$ ) in der ersten Ungleichung

$$d_{\mathbb{H}}^2(\mathbb{B}_{n,p}, \mathbb{B}_{n,q}) = 1 - \left(1 - \frac{1}{2} \left( (\sqrt{p} - \sqrt{q})^2 + (\sqrt{1-p} - \sqrt{1-q})^2 \right)\right)^n \quad (3.23)$$

$$\leq \frac{1}{2} \left( (\sqrt{np} - \sqrt{nq})^2 + (\sqrt{n(1-p)} - \sqrt{n(1-q)})^2 \right) \quad (3.24)$$

$$\begin{aligned} &= n \left(1 - \sqrt{pq} - \sqrt{(1-p)(1-q)}\right) \cdot \frac{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}}{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}} \\ &= n \cdot \frac{(\sqrt{p} - \sqrt{q})^2}{1 - \sqrt{pq} + \sqrt{(1-p)(1-q)}} \cdot \frac{(\sqrt{p} + \sqrt{q})^2}{(\sqrt{p} + \sqrt{q})^2} \\ &= \frac{n(p-q)^2}{p(1-p) + q(1-q) + \sqrt{pq}(2 - (p+q)) + (p-q)^2 + \sqrt{(1-p)(1-q)}(\sqrt{p} + \sqrt{q})^2} \\ &\leq \frac{(np - nq)^2}{np(1-p) + nq(1-q)}. \end{aligned} \quad (3.25)$$

Die Behauptung folgt damit nach (2.2).  $\square$

*Beweis von Lemma 3.2.* Nach der Konvention auf Seite 9 gilt (3.2), wenn  $n(p-q) = 0$ . Sei also o. E.  $n(p-q) \neq 0$ , und damit nach (3.22)

$$n \cdot d_{\mathbb{H}}^2(\mathbb{B}_p, \mathbb{B}_q) = \frac{1}{2} \left( (\sqrt{np} - \sqrt{nq})^2 + (\sqrt{n(1-p)} - \sqrt{n(1-q)})^2 \right) > 0.$$

Für  $0 < x \leq 1$  ist nach Anwendung der Bernoulli-Ungleichung in der ersten Ungleichung

$$1 + nx \leq (1+x)^n \leq \frac{1}{(1-x)^n}. \quad (3.26)$$

Anwendung von (3.26) in der ersten Ungleichung und einer Exponentialungleichung ( $\exp(-y) \leq 1/(1+y)$  für  $y \geq -1$ ) in der zweiten liefert

$$\exp\left(-\frac{1}{nx}\right) = \exp\left(\frac{1}{1-(1+nx)}\right) \leq \exp\left(\frac{(1-x)^n}{(1-x)^n - 1}\right) \leq 1 - (1-x)^n,$$

und damit ist (mit  $x := d_{\mathbb{H}}^2(\mathbb{B}_p, \mathbb{B}_q)$ )

$$\begin{aligned} \exp\left(-\frac{2}{(\sqrt{np} - \sqrt{nq})^2 + (\sqrt{n(1-p)} - \sqrt{n(1-q)})^2}\right) &\leq 1 - \left(1 - d_{\mathbb{H}}^2(\mathbb{B}_p, \mathbb{B}_q)\right)^n \\ &= d_{\mathbb{H}}^2(\mathbb{B}_{n,p}, \mathbb{B}_{n,q}). \end{aligned} \quad \square$$

*Beweis von Satz 3.3.* Die Folgerung von (3.3) zu (3.4) folgt direkt aus (3.25). Die Äquivalenz zwischen (3.4) und (3.5) gilt nach (3.2) und (3.24).

Aus (3.5) folgt (3.6), denn es ist

$$\begin{aligned} \left| \sqrt{p(1-p)} - \sqrt{q(1-q)} \right| &\leq |(\sqrt{p} - \sqrt{q})\sqrt{1-q}| + |(\sqrt{1-p} - \sqrt{1-q})\sqrt{p}| \\ &\leq |\sqrt{p} - \sqrt{q}| + |\sqrt{1-p} - \sqrt{1-q}|. \end{aligned} \quad (3.27)$$

Die Funktionen  $[0, \sqrt{\frac{2}{3}}] \ni x \mapsto \sqrt{1-x^2}$  und  $[\sqrt{\frac{1}{3}}, 1] \ni x \mapsto \frac{1}{x}$  sind Lipschitz-stetig, also existiert ein  $C > 0$  mit

$$\left| \sqrt{1-x} - \sqrt{1-y} \right| \leq C \left| \sqrt{x} - \sqrt{y} \right|, \quad \text{falls } x, y \in [0, \frac{2}{3}],$$

und

$$\begin{aligned} \left| \sqrt{1-x} - \sqrt{1-y} \right| &\leq \sqrt{3} \left| \sqrt{x(1-x)} - \sqrt{y(1-y)} \right| + \left| \frac{1}{\sqrt{x}} - \frac{1}{\sqrt{y}} \right| \\ &\leq \sqrt{3} \left| \sqrt{x(1-x)} - \sqrt{y(1-y)} \right| + C \left| \sqrt{x} - \sqrt{y} \right|, \quad \text{falls } x, y \in [\frac{1}{3}, 1], \end{aligned}$$

wobei in der vorletzten Ungleichung ein Analogon zu (3.27) verwendet wurde. Also folgt aus (3.6), da (o. E.  $n \neq 0$ )  $|p - q| \rightarrow 0$ , dass schließlich

$$\left| \sqrt{n(1-p)} - \sqrt{n(1-q)} \right| \leq \sqrt{3} \left| \sqrt{np(1-p)} - \sqrt{nq(1-q)} \right| + C \left| \sqrt{np} - \sqrt{nq} \right|,$$

also (3.5). □

*Beweis von Satz 3.4.* Anwendung von Roos und Pfeifer (2003a, Theorem 1, dort  $N := n, M := m$ ) in der ersten Ungleichung liefert

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) &\leq \min \left\{ p |n - m|, \sqrt{\frac{p}{1-p}} |\sqrt{n} - \sqrt{m}| \right\} \\ &= |np - mp| \cdot \min \left\{ 1, \frac{1}{\sqrt{np(1-p)} + \sqrt{mp(1-p)}} \right\} \\ &\leq \frac{|np - mp|}{\max \left\{ 1, \sqrt{np(1-p)} + \sqrt{mp(1-p)} \right\}} \\ &\leq \sqrt{\frac{(np - mp)^2}{np(1-p) + mp(1-p)}}. \end{aligned} \quad \square$$

*Beweis von Satz 3.6.* Es sei o. E.  $np(1-p) \neq 0$ . Die Ungleichung  $\Sigma^2 \geq \sigma^2$  folgt z. B. aus

### 3 Proximität von Binomial- und Poisson-Verteilungen

Ehm (1991, (3)). Weiter existiert eine Konstante  $C \geq 1/124$  mit

$$\begin{aligned}
C \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2} &\leq C \left( |\Sigma^2 - \sigma^2| \wedge \left( 1 - \frac{\sigma^2}{\Sigma^2} \right) \right) \\
&\leq d_{\text{TV}}(\mathbb{B}_{n,p}, \bigstar_{j=1}^n \mathbb{B}_{p_j}) \\
&\leq \frac{n}{n+1} \left( 1 - p^{n+1} - (1-p)^{n+1} \right) \left( 1 - \frac{\sigma^2}{\Sigma^2} \right) \\
&\leq \left( \Sigma^2 \wedge \frac{n}{n+1} \right) \left( 1 - \frac{\sigma^2}{\Sigma^2} \right) \\
&\leq \begin{cases} \frac{\Sigma^2 - \sigma^2}{\Sigma^2} \leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}, & \text{falls } \Sigma^2 \geq 1, \\ |\Sigma^2 - \sigma^2| \leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}, & \text{falls } \Sigma^2 \leq 1, \end{cases} \\
&\leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2},
\end{aligned}$$

wobei aus Ehm (1991) verwendet wurde: (3), in der zweiten und dritten Ungleichung Theorem 1, und in der vierten Ungleichung Lemma 2.

Hieraus folgt sofort die Abschätzung für zwei Binomialverteilungen: o. E.  $m \leq n$ ,  $p_1 := \dots := p_m := q$ ,  $p_{m+1} := \dots := p_n := 0$ .  $\square$

*Beweis von Satz 3.7.* Ohne Einschränkung sei  $m \leq n$ . Die zweite Ungleichung in Example 1 von Roos und Pfeifer (2003b, dort  $N \sim \mathbb{B}_{n, \frac{m}{n}}$ ) liefert

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \leq m \left( 1 - \frac{m}{n} \right) q^2 \min \left\{ \frac{1}{\sqrt{2}q(1-q)(m+1/2)}, 1 \right\},$$

mit  $\Sigma^2 := np(1-p) \geq mq(1-q) =: \sigma^2$  also

$$\begin{aligned}
d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) &\leq \min \left\{ 1, \Sigma^2 - \sigma^2, \frac{1}{\sqrt{2} \left( 1 + \frac{1}{2m} \right)} \cdot \frac{\Sigma^2 - \sigma^2}{\sigma^2} \right\} \\
&\leq \min \left\{ 1, \Sigma^2 - \sigma^2, \frac{1}{\sqrt{2}} \cdot \frac{\Sigma^2 - \sigma^2}{\sigma^2} \right\} \\
&= \begin{cases} 1 & \Leftrightarrow \sigma^2(\sqrt{2}+1) \leq \Sigma^2, \quad 1 + \sigma^2 \leq \Sigma^2 \\ \Sigma^2 - \sigma^2 & \Leftrightarrow \Sigma^2 \leq 1 + \sigma^2 \leq 1 + \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \cdot \frac{\Sigma^2 - \sigma^2}{\sigma^2} & \Leftrightarrow \frac{1}{\sqrt{2}} \leq \sigma^2 \leq \Sigma^2 \leq \sigma^2(\sqrt{2}+1) \end{cases} \\
&\leq \sqrt{2}(\sqrt{2}+1) \cdot \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}. \quad \square
\end{aligned}$$

*Beweis von Satz 3.8.* Unter Anwendung von (3.1) in der ersten Ungleichung ist

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{n,q}) \leq \sqrt{2} \sqrt{\frac{(np - nq)^2}{np(1-p) + nq(1-q)}} \leq \sqrt{2} \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}}. \quad (3.28)$$

Wenn  $q(1 - q) \leq p(1 - p)$ , ist  $q < 1/2$ , da  $q \leq p$ . Dann ist also  $n \geq 2mq$ , und so  $x(n - 2mq) + mq(1 - 2q) \geq 0$  für alle  $x \in [q, 1]$ . Damit ist  $f: [q, 1] \rightarrow [0, \infty[$ ,  $x \mapsto \sqrt{\frac{(nx - mq)^2}{nx(1 - x) + mq(1 - q)}}$ , wachsend, denn für  $x \in ]q, 1[$  ist

$$f'(x) = n \cdot \frac{nx(1 - x) + mq(1 - q) - \frac{1}{2}(nx - mq)(1 - 2x)}{(nx(1 - x) + mq(1 - q))^{3/2}} \geq 0.$$

In der nächsten Ungleichung gilt die erste Abschätzung nach Satz 3.4. Die zweite Ungleichung gilt, da entweder  $q(1 - q) \leq p(1 - p)$  ist, und damit das obige Argument greift, oder  $p \geq q$  und  $q(1 - q) \geq p(1 - p)$  ist:

$$d_{\text{TV}}(\mathbb{B}_{n,q}, \mathbb{B}_{m,q}) \leq \sqrt{\frac{(nq - mq)^2}{nq(1 - q) + mq(1 - q)}} \leq \sqrt{\frac{(np - mq)^2}{np(1 - p) + mq(1 - q)}}. \quad (3.29)$$

Damit folgt die Behauptung nach der Dreiecksungleichung aus (3.28) und (3.29).  $\square$

*Beweis von Bemerkung 3.9.* Es ist

$$np - mq \geq np(1 - p) - mq(1 - p) \geq np(1 - p) - mq(1 - q),$$

und wenn  $np(1 - p) < mq(1 - q)$ , ist  $m(1 - q) \geq n(1 - p)\frac{p}{q} \geq n(1 - p)$ , und damit

$$np - mq \geq m(1 - q) - n(1 - p) \geq mq(1 - q) - nq(1 - p) \geq mq(1 - q) - np(1 - p),$$

also ist für  $np(1 - p) \neq mq(1 - q)$

$$\frac{\sqrt{np(1 - p) + mq(1 - q)}}{1 + np(1 - p) + mq(1 - q)} \leq 1 \leq \frac{|np - mq|}{|np(1 - p) - mq(1 - q)|}. \quad \square$$

*Beweis von Satz 3.10.* Es ist

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,r_1}) + d_{\text{TV}}(\mathbb{B}_{m,r_1}, \mathbb{B}_{m,q}).$$

Dabei ist unter Anwendung von (3.1) in der ersten und  $r_1 < \min\{q, 1/2\} \leq 1/2$  in der zweiten Abschätzung

$$d_{\text{TV}}(\mathbb{B}_{m,r_1}, \mathbb{B}_{m,q}) \leq \sqrt{2} \sqrt{\frac{(np - mq)^2}{np(1 - r_1) + mq(1 - q)}} < 2 \sqrt{\frac{(np - mq)^2}{np(1 - p) + mq(1 - q)}}.$$

Sei  $f := g + h$  mit  $g, h: [r_1, 1] \rightarrow [0, 1]$ ,

$$g(x) := \frac{|np(1 - p) - mx(1 - x)|}{1 + np(1 - p) + mx(1 - x)}, \quad h(x) := \sqrt{\frac{(mx - np)^2}{np(1 - p) + mx(1 - x)}}.$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

Für  $x \in ]r_1, 1[$  ist  $mx \geq np$  und  $p + x \leq \frac{np}{m} + x < \frac{3}{2}$ , also

$$h'(x) = \frac{m(mx + np(3 - 2(p + x)))}{2(np(1 - p) + mx(1 - x))^{3/2}} \geq 0.$$

Da  $mr_1(1 - r_1) = np(1 - r_1) \leq np(1 - p)$ , ist

$$r_1 \leq y := \inf\{x \in [r_1, 1]: mx(1 - x) = np(1 - p)\}$$

und  $y = 1$  oder  $y \leq 1/2$ .

Für  $x \in ]r_1, 1[ \setminus [y, 1 - y]$  ist  $mx(1 - x) < np(1 - p)$ , also

$$\begin{aligned} f'(x) &= -\frac{m(2np(1 - p) + 1)(1 - 2x)}{(1 + np(1 - p) + mx(1 - x))^2} + \frac{m(mx + np(3 - 2(p + x)))}{2(np(1 - p) + mx(1 - x))^{3/2}} \\ &\geq 0 \end{aligned}$$

genau dann, wenn

$$\frac{mx + np(3 - 2(p + x))}{2(np(1 - p) + mx(1 - x))^{3/2}} \geq \frac{(2np(1 - p) + 1)(1 - 2x)}{(1 + np(1 - p) + mx(1 - x))^2}, \quad (3.30)$$

und dies gilt, denn wenn  $x \geq \frac{1}{2}$ , ist  $p + x < \frac{3}{2}$ , und wenn  $x < \frac{1}{2}$ , ist

$$\begin{aligned} \frac{mx + np(3 - 2(p + x))}{(2np(1 - p) + 1)(1 - 2x)} &= \frac{2np(1 - p) + np(1 - 2x) + mx}{(2np(1 - p) + 1)(1 - 2x)} \\ &\geq \frac{2np(1 - p) + 2np(1 - x)}{(2np(1 - p) + 1)(1 - 2x)} \\ &\geq \frac{2np(1 - p) + np}{(2np(1 - p) + 1)(1 - 2x)} \\ &\geq \frac{3np(1 - p)}{2np(1 - p) + 1} \\ &\geq \frac{2(2np(1 - p))^{3/2}}{(1 + \frac{3}{2}np(1 - p))^2} \\ &\geq \frac{2(np(1 - p) + mx(1 - x))^{3/2}}{(1 + np(1 - p) + mx(1 - x))^2}, \end{aligned}$$

wobei im letzten Schritt genutzt wurde, dass  $1/2 > x \in ]r_1, 1[ \setminus [y, 1 - y]$ , und damit  $\frac{1}{2}np(1 - p) \leq \frac{1}{2}np < mx(1 - x) < np(1 - p)$ .

Falls  $y = 1$ , ist also  $f(r_1) \leq f(x)$  für alle  $x \in [r_1, 1]$ , da  $f' \geq 0$  auf  $]r_1, 1[$ .

Falls  $y \leq \frac{1}{2}$ , ist, da  $h' \geq 0$  auf  $]r_1, 1[$ ,

$$f(y) = g(y) + h(y) = h(y) \leq \begin{cases} g(x) + h(x) = f(x) \text{ für } x \in [y, 1 - y], \\ h(1 - y) = f(1 - y), \end{cases}$$

also ist  $f(r_1) \leq f(x)$  für alle  $x \in [r_1, 1]$ , denn

$$f(r_1) \leq \begin{cases} f(x) \leq f(y), & \text{falls } x \in [r_1, y], \\ f(y) \leq f(x), & \text{falls } x \in [y, 1-y], \\ f(y) \leq f(1-y) \leq f(x), & \text{falls } x \in [1-y, 1]. \end{cases}$$

Also ist nach (3.7) in der ersten Ungleichung

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,r_1}) &\leq 3 \frac{|np(1-p) - mr_1(1-r_1)|}{1 + np(1-p) + mr_1(1-r_1)} \\ &= 3f(r_1) \\ &\leq 3f(q) \\ &= 3 \sqrt{\frac{(np-mq)^2}{np(1-p) + mq(1-q)}} + 3 \left( \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} \right). \end{aligned}$$

Damit folgt die Behauptung nach der Dreiecksungleichung.  $\square$

*Beweis von Satz 3.11.* Es ist  $mq(1-q) \leq n \frac{m}{n} q \left(1 - \frac{m}{n} q\right) \leq np(1-p)$ . Also ist unter Anwendung von (3.1) und (3.7) in der zweiten Ungleichung

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) &\leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{n, \frac{m}{n}q}) + d_{\text{TV}}(\mathbb{B}_{n, \frac{m}{n}q}, \mathbb{B}_{m,q}) \\ &\leq \sqrt{2} \sqrt{\frac{(np-mq)^2}{np(1-p) + mq(1 - \frac{m}{n}q)}} + 3 \frac{|n \frac{m}{n} q(1 - \frac{m}{n} q) - mq(1-q)|}{1 + n \frac{m}{n} q(1 - \frac{m}{n} q) + mq(1-q)} \\ &\leq \sqrt{2} \sqrt{\frac{(np-mq)^2}{np(1-p) + mq(1-q)}} + 3 \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}. \quad \square \end{aligned}$$

*Beweis von Lemma 3.12.* Es ist  $q \geq a := k(p-q) + q \geq 0$ , also  $(k-1)q + a = kp$  und  $(k-1)q(1-q) \leq (k-1)q(1-q) + a(1-a) \leq kp(1-p)$ . Also ist nach Satz 3.6

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{k,p}, \mathbb{B}_{k-1,q} * \mathbb{B}_a) &\leq 3 \frac{|kp(1-p) - ((k-1)q(1-q) + a(1-a))|}{1 + kp(1-p) + ((k-1)q(1-q) + a(1-a))} \\ &\leq 3 \frac{|kp(1-p) - (k-1)q(1-q)|}{1 + kp(1-p) + (k-1)q(1-q)}. \end{aligned}$$

Weiter ist  $p < q$  und  $kp \geq (k-1)q$ , also  $1-p \leq 1 - \frac{k-1}{k}q$ , und damit  $kp(1-p) \leq kq(1-q) + q^2$ . Wenn  $\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)} \geq 1$ , ist also

$$\begin{aligned} \frac{1}{\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)}} &\leq \frac{2}{\sqrt{1+q^2}} \frac{1}{\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)}} \\ &\leq 2 \frac{1}{\sqrt{(k-1)q(1-q) + kq(1-q) + q^2}} \\ &\leq \frac{2}{\sqrt{(k-1)q(1-q) + kp(1-p)}}, \end{aligned}$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

und wenn  $\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)} \leq 1$ , ist

$$\begin{aligned} 1 &\leq \frac{\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)} + q}{\sqrt{(k-1)q(1-q)} + kp(1-p)} \\ &\leq \frac{2}{\sqrt{(k-1)q(1-q)} + kp(1-p)}, \end{aligned}$$

also ist in beiden Fällen

$$\min \left\{ 1, \frac{1}{\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)}} \right\} \leq \frac{2}{\sqrt{(k-1)q(1-q)} + kp(1-p)}.$$

Unter Ausnutzung von (2.1) ist nach Roos (2000, Lemma 4)

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{k-1,q} * \mathbb{B}_a, \mathbb{B}_{k-1,q}) &\leq \frac{a}{2} \frac{1}{\sqrt{kq(1-q)}} \\ &\leq a \frac{1}{\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)}}. \end{aligned}$$

Damit ist, da  $d_{\text{TV}}(\mathbb{B}_{k-1,q} * \mathbb{B}_a, \mathbb{B}_{k-1,q}) \leq a$ ,

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{k-1,q} * \mathbb{B}_a, \mathbb{B}_{k-1,q}) &\leq a \cdot \min \left\{ 1, \frac{1}{\sqrt{(k-1)q(1-q)} + \sqrt{kq(1-q)}} \right\} \\ &\leq a \cdot \frac{2}{\sqrt{(k-1)q(1-q)} + kp(1-p)} \\ &= 2 \sqrt{\frac{(kp - (k-1)q)^2}{kp(1-p) + (k-1)q(1-q)}}. \end{aligned}$$

Es folgt die Behauptung nach der Dreiecksungleichung.  $\square$

*Beweis von Satz 3.13.* Nach Voraussetzung ist  $m < n$  und  $1/2 < p$ , und o. E.  $m > 0$ .

Mit  $k := \max\{k \in \{m, \dots, n-1\} : \text{es existiert ein } r \in [p, q] : kr = mq\}$  ist  $r := \frac{mq}{k} > 0$ . Dann ist  $kr \leq (k+1)p$ , denn wenn  $kr > (k+1)p$ , ist  $(k+1)p < kr < (k+1)r$ , also existiert ein  $s \in [p, r] \subseteq [p, q]$  mit  $(k+1)s = kr = mq$ , und dies steht im Widerspruch zur Definition von  $k$ .

Also ist  $\frac{1}{2} < p \leq r \leq q$  und  $(k+1)(1-p) \leq n(1-p) < mq = kr \leq (k+1)p \leq np$ , und damit

$$mq(1-q) \leq kr(1-r) \leq (k+1)p(1-p) \leq np(1-p). \quad (3.31)$$

Für  $f: ]mq, np[ \rightarrow ]0, \infty[$ ,  $f(x) := \frac{x-mq}{\sqrt{mq(1-q)+x(1-p)}}$  ist  $f' \geq 0$ .

Also gilt, nach Anwendung von Lemma 3.12 in der ersten Ungleichung, (3.31) in der zweiten, und  $f' \geq 0$  in der dritten,

$$\begin{aligned} d_{\text{TV}}(B_{k,r}, B_{k+1,p}) &\leq 2\sqrt{\frac{(kr - (k+1)p)^2}{kr(1-r) + (k+1)p(1-p)}} + 3\frac{|kr(1-r) - (k+1)p(1-p)|}{1 + kr(1-r) + (k+1)p(1-p)} \\ &\leq 2\sqrt{\frac{(mq - (k+1)p)^2}{mq(1-q) + (k+1)p(1-p)}} + 3\frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} \\ &\leq 2\sqrt{\frac{(mq - np)^2}{mq(1-q) + np(1-p)}} + 3\frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}. \end{aligned}$$

Nach Satz 3.6 ist weiter, unter Anwendung von (3.31) in der zweiten Ungleichung,

$$d_{\text{TV}}(B_{m,q}, B_{k,r}) \leq 3\frac{|mq(1-q) - kr(1-r)|}{1 + mq(1-q) + kr(1-r)} \leq 3\frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

Weiter ist nach Satz 3.4

$$d_{\text{TV}}(B_{k+1,p}, B_{n,p}) \leq \sqrt{\frac{(np - (k+1)p)^2}{np(1-p) + (k+1)p(1-p)}} \leq \sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}}.$$

Die Behauptung folgt nach der Dreiecksungleichung.  $\square$

*Beweis von Satz 3.14.* Es ist  $r_2 := \frac{mq}{n} \leq q$ , also  $mq(1-q) \leq \min\{np(1-p), nr_2(1-r_2)\}$ . Nach (3.1) ist also

$$d_{\text{TV}}(B_{n,p}, B_{n,r_2}) \leq \sqrt{2}\sqrt{\frac{(np - nr_2)^2}{np(1-p) + nr_2(1-r_2)}} \leq \sqrt{2}\sqrt{\frac{(np - mq)^2}{np(1-p) + mq(1-q)}}.$$

Wenn  $nr_2(1-r_2) \leq np(1-p)$ , ist nach (3.7)

$$d_{\text{TV}}(B_{n,r_2}, B_{m,q}) \leq 3\frac{|nr_2(1-r_2) - mq(1-q)|}{1 + nr_2(1-r_2) + mq(1-q)} \leq 3\frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

Wenn  $nr_2(1-r_2) > np(1-p)$ , ist  $\frac{m}{2n} \leq p < \min\{r_2, 1-r_2\} \leq \frac{1}{2}$ , also  $np(1-p) - mq(1-q) \geq \frac{m}{4}\left(1 - \frac{m}{n}\right) > 0$ , also

$$\frac{nr_2(1-r_2) - mq(1-q)}{np(1-p) - mq(1-q)} \leq \frac{mq^2\left(1 - \frac{m}{n}\right)}{\frac{m}{4}\left(1 - \frac{m}{n}\right)} \leq 4 < 4\frac{1 + nr_2(1-r_2) + mq(1-q)}{1 + np(1-p) + mq(1-q)},$$

und nach (3.7)

$$d_{\text{TV}}(B_{n,r_2}, B_{m,q}) \leq 3\frac{|nr_2(1-r_2) - mq(1-q)|}{1 + nr_2(1-r_2) + mq(1-q)} \leq 12\frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)}.$$

Damit folgt die Behauptung nach der Dreiecksungleichung.  $\square$

### 3 Proximität von Binomial- und Poisson-Verteilungen

*Beweis von Satz 3.15.* Falls  $n = 0$ , ist alles klar. Falls  $n \in \mathbb{N}$ , gilt nach Barbour und Hall (1984, Theorem 1 und 2)

$$\frac{1}{32} \left(1 \wedge \frac{1}{\lambda}\right) \sum_{j=1}^n p_j^2 \leq d_{\text{TV}}(\ast \mathbf{B}_{p_j}, \mathbf{P}_\lambda) \leq \frac{1 - \exp(-\lambda)}{\lambda} \sum_{j=1}^n p_j^2 \leq \left(1 \wedge \frac{1}{\lambda}\right) \sum_{j=1}^n p_j^2, \quad (3.32)$$

wobei in der letzten Ungleichung  $1 - \exp(-\lambda) \leq \min\{1, \lambda\}$  benutzt wurde. Also gilt (3.14), da

$$\begin{aligned} \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2} &\leq \min \left\{ |\Sigma^2 - \sigma^2|, \frac{|\Sigma^2 - \sigma^2|}{\Sigma^2} \right\} = \left(1 \wedge \frac{1}{\lambda}\right) \sum_{j=1}^n p_j^2 \\ &= \begin{cases} \frac{|\Sigma^2 - \sigma^2|}{\Sigma^2} \leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}, & \text{falls } \Sigma^2 \geq 1, \\ |\Sigma^2 - \sigma^2| \leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}, & \text{falls } \Sigma^2 \leq 1, \end{cases} \\ &\leq 3 \frac{|\Sigma^2 - \sigma^2|}{1 + \Sigma^2 + \sigma^2}. \end{aligned}$$

Die Aussagen im Binomialfall ergeben sich durch Spezialisierung auf  $p = p_1 = \dots = p_n$ , da

$$p \cdot \min\{1, np\} = \min \left\{ |np(1-p) - np|, \frac{|np(1-p) - np|}{np} \right\}.$$

□

*Beweis von Satz 3.16.* Nach (3.13) ist

$$d_{\text{TV}}(\mathbf{P}_{np}, \mathbf{P}_\lambda) \leq \sqrt{\frac{2}{e}} \sqrt{\frac{(np - \lambda)^2}{np(1-p) + \lambda}}. \quad (3.33)$$

Es ist  $f := g + h$ , mit  $g, h: [0, \infty[ \rightarrow \mathbb{R}$ ,

$$g(x) := \frac{|np(1-p) - x|}{1 + np(1-p) + x}, \quad h(x) := \sqrt{\frac{(np - x)^2}{np(1-p) + x}},$$

antiton auf  $[0, np(1-p)]$ , da  $g$  und  $h$  hier antiton sind. Außerdem ist  $f$  antiton auf  $]np(1-p), np]$ , da für  $x \in ]np(1-p), np[$

$$\begin{aligned} f'(x) &= \frac{1 + 2np(1-p)}{(1 + np(1-p) + x)^2} - \frac{np(1-p) + x + \frac{1}{2}(np - x)}{(np(1-p) + x)^{3/2}} \\ &\leq \frac{1}{1 + np(1-p) + x} - \frac{1}{\sqrt{np(1-p) + x}} \\ &\leq 0, \end{aligned}$$

und isoton auf  $]np, \infty[$ , da für  $x \in ]np, \infty[$

$$\begin{aligned} f'(x) &= \frac{1 + 2np(1-p)}{(1 + np(1-p) + x)^2} + \frac{np(1-p) + \frac{1}{2}(np+x)}{(np(1-p) + x)^{3/2}} \\ &\geq 0. \end{aligned}$$

Also ist  $f$  minimal in  $np$ , und nach (3.16) ist

$$\begin{aligned} d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{P}_{np}) &\leq 3 \frac{|np(1-p) - np|}{1 + np(1-p) + np} = 3f(np) \\ &\leq 3f(\lambda) \\ &= 3 \frac{|np(1-p) - \lambda|}{1 + np(1-p) + \lambda} + 3 \sqrt{\frac{(np - \lambda)^2}{np(1-p) + \lambda}}. \end{aligned}$$

Damit folgt die Behauptung nach der Dreiecksungleichung.  $\square$

### 3.7.2 Beweise des Abschnitts 3.5

*Beweis von Satz 3.17.* Es muss gezeigt werden, dass  $\mathcal{F}$  eine uniforme Struktur auf BP ist. Da (2.7), (2.8), (2.10) und (2.11) direkt nach der Definition von  $\mathcal{F}$  gelten, muss nur (2.12) für  $\mathcal{F}$  nachgewiesen werden:

Sei  $V \in \mathcal{F}$ . Dann existiert ein  $\varepsilon > 0$  mit  $V \supseteq U_\varepsilon$ . Also reicht zu zeigen: Für alle  $\varepsilon > 0$  existiert ein  $\delta > 0$  mit  $U_\delta \circ U_\delta \subseteq U_\varepsilon$ . Seien  $\varepsilon > 0$ ,

$$\delta := \min \left\{ \frac{1}{51}, \left( \frac{\sqrt{1+2\varepsilon} - 1}{2} \right)^2 \right\}$$

und  $(P, Q) \in U_\delta \circ U_\delta$  so, dass ein  $R \in \text{BP}$  existiert mit  $(P, R), (R, Q) \in U_\delta$ . Zu zeigen ist, dass  $(P, Q) \in U_\varepsilon$  gilt.

Nach der Dreiecksungleichung gilt, da  $(P, R), (R, Q) \in U_\delta$ ,

$$\frac{|\sigma^2(P) - \sigma^2(Q)|}{1 + \sigma^2(P) + \sigma^2(Q)} \leq \frac{|\sigma^2(P) - \sigma^2(R)|}{1 + \sigma^2(P) + \sigma^2(R)} + \frac{|\sigma^2(R) - \sigma^2(Q)|}{1 + \sigma^2(R) + \sigma^2(Q)} < 2\delta,$$

und es reicht, zu zeigen, dass

$$\frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)} < \varepsilon - 2\delta. \quad (3.34)$$

1. Fall:  $\sigma^2(R) \leq \max\{\sigma^2(P), \sigma^2(Q)\}$ .

Dann ist  $\frac{1}{2}(\max\{\sigma^2(P), \sigma^2(Q)\} + \min\{\sigma^2(P), \sigma^2(Q)\} + \sigma^2(R)) \leq \sigma^2(Q) + \sigma^2(P)$ , und damit

$$\frac{1}{\sigma^2(P) + \sigma^2(Q)} \leq \min \left\{ \frac{2}{\sigma^2(P) + \sigma^2(R)}, \frac{2}{\sigma^2(Q) + \sigma^2(R)} \right\},$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

also

$$\begin{aligned}
 \frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)} &\leq 2 \left( \frac{(\mu(P) - \mu(R))^2}{\sigma^2(P) + \sigma^2(Q)} + \frac{(\mu(R) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)} \right) \\
 &\leq 4 \left( \frac{(\mu(P) - \mu(R))^2}{\sigma^2(P) + \sigma^2(R)} + \frac{(\mu(R) - \mu(Q))^2}{\sigma^2(R) + \sigma^2(Q)} \right) \\
 &< 8\delta \\
 &\leq 2\sqrt{\delta} \\
 &\leq \varepsilon - 2\delta,
 \end{aligned} \tag{3.35}$$

und es gilt (3.34).

2. Fall:  $\sigma^2(R) > \max\{\sigma^2(P), \sigma^2(Q)\}$  und  $\min\{\sigma^2(P) + \sigma^2(R), \sigma^2(Q) + \sigma^2(R)\} \geq \frac{1}{8}$ .  
Dann ist  $\sigma^2(R) \geq 1/16$  und, da  $(P, R) \in U_\delta$ ,

$$\frac{\sigma^2(R) - \sigma^2(P)}{1 + \sigma^2(P) + \sigma^2(R)} = \frac{|\sigma^2(P) - \sigma^2(R)|}{1 + \sigma^2(P) + \sigma^2(R)} < \delta < \frac{1}{50},$$

und damit

$$\frac{\delta}{1 - 3\delta} < \frac{1}{47} < \frac{1 - 17\delta}{16(1 + \delta)} \leq \frac{\sigma^2(R)(1 - \delta) - \delta}{1 + \delta} < \sigma^2(P),$$

also

$$\sigma^2(R) + \sigma^2(P) < \frac{\delta + 2\sigma^2(P)}{1 - \delta} < 3\sigma^2(P).$$

Analog zeigt man, dass  $\sigma^2(R) + \sigma^2(Q) < 3\sigma^2(Q)$ , und zusammen ergibt sich

$$\frac{1}{\sigma^2(P) + \sigma^2(Q)} < \frac{3}{\max\{3\sigma^2(P), 3\sigma^2(Q)\}} < \min \left\{ \frac{3}{\sigma^2(P) + \sigma^2(R)}, \frac{3}{\sigma^2(Q) + \sigma^2(R)} \right\},$$

und analog zu (3.35) ergibt sich

$$\frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)} < 12\delta \leq 2\sqrt{\delta} \leq \varepsilon - 2\delta.$$

3. Fall:  $\sigma^2(R) > \max\{\sigma^2(P), \sigma^2(Q)\}$  und  $\min\{\sigma^2(P) + \sigma^2(R), \sigma^2(Q) + \sigma^2(R)\} < \frac{1}{8}$ .  
Dann ist

$$\sigma^2(R) = \max\{\sigma^2(P), \sigma^2(Q), \sigma^2(R)\} < \frac{1}{8} \tag{3.36}$$

und

$$\max\{\sigma^2(P) + \sigma^2(Q), \sigma^2(P) + \sigma^2(R), \sigma^2(Q) + \sigma^2(R)\} < \frac{1}{4}. \tag{3.37}$$

Da  $(P, R), (R, Q) \in U_\delta$ , ist  $\max\{(\mu(P) - \mu(R))^2, (\mu(Q) - \mu(R))^2\} < \delta/4$ , nach der Dreiecksungleichung also

$$\max\{(\mu(P) - \mu(R))^2, (\mu(Q) - \mu(R))^2, (\mu(P) - \mu(Q))^2\} < \delta. \quad (3.38)$$

Wir nehmen an, es sei

$$l := \max\{\mu(P), \mu(Q), \mu(R)\} \geq \frac{1}{2} \quad \text{und} \quad u := \min\{\mu(P), \mu(Q), \mu(R)\} < \frac{1}{2}.$$

Nach (3.38) ist  $(u - l)^2 < \delta$ , also

$$\frac{1}{2} - \sqrt{\delta} < l < \frac{1}{2} \leq u < \frac{1}{2} + \sqrt{\delta}.$$

Da  $P, Q, R \in \text{BP}$ , ist damit  $\min\{\sigma^2(P), \sigma^2(Q), \sigma^2(R)\} > (1/2 - \sqrt{\delta})(1/2 + \sqrt{\delta}) = 1/4 - \delta$ , und dies steht im Widerspruch zu (3.37).

Also ist entweder

$$\max\{\mu(P), \mu(Q), \mu(R)\} < \frac{1}{2} \quad (3.39)$$

oder

$$\min\{\mu(P), \mu(Q), \mu(R)\} \geq \frac{1}{2}. \quad (3.40)$$

Falls (3.39) gilt, existieren  $p, q \in [0, 1/2]$  mit

$$\frac{\mu(Q)}{\sigma^2(Q)}, \frac{\mu(P)}{\sigma^2(P)} \in \left\{0, 1, \frac{1}{1-p}, \frac{1}{1-q}\right\},$$

also ist

$$\frac{|\mu(P) - \mu(Q)|}{\sigma^2(P) + \sigma^2(Q)} \leq \frac{\mu(P) + \mu(Q)}{\sigma^2(P) + \sigma^2(Q)} \leq \frac{\mu(P)}{\sigma^2(P)} + \frac{\mu(Q)}{\sigma^2(Q)} \leq 2,$$

und damit, unter Anwendung von (3.38),

$$\begin{aligned} \frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)} &= \frac{|\mu(P) - \mu(Q)|}{\sigma^2(P) + \sigma^2(Q)} \cdot |\mu(P) - \mu(Q)| \\ &< 2\sqrt{\delta} \\ &\leq \varepsilon - 2\delta. \end{aligned}$$

Falls (3.40) gilt, ist wegen (3.36)  $\min\left\{\frac{\mu(P)}{\sigma^2(P)}, \frac{\mu(Q)}{\sigma^2(Q)}, \frac{\mu(R)}{\sigma^2(R)}\right\} > 2$ , und es sind  $P, Q, R \in \{\mathbb{B}_{n,p} : n \in \mathbb{N}, p \in ]1/2, 1]\}$ .

Wenn  $m > n \in \mathbb{N}$  und  $p, q \in ]1/2, 1]$  mit  $np(1-p) + mq(1-q) < 1/4$  und  $\min\{np, mq\} \geq 1/2$ , ist  $q(1-q) < \frac{1}{4m}$ , also  $q > \frac{1}{2} + \frac{1}{2}\sqrt{\frac{m-1}{m}}$ , also  $mq - np \geq mq - n \geq \frac{1}{2}$ , und damit  $(np - mq)^2 / (np(1-p) + mq(1-q)) \geq 1 \geq \delta$ .

### 3 Proximität von Binomial- und Poisson-Verteilungen

Da  $(P, R), (R, Q) \in U_\delta$ , existiert also ein  $n \in \mathbb{N}$  mit  $P, Q, R \in \{B_{n,p} : p \in ]1/2, 1]\}$ .  
Damit ist, mit  $n \in \mathbb{N}$  und  $p, q \in ]1/2, 1]$  so, dass  $P = B_{n,p}$  und  $Q = B_{n,q}$ , nach (3.38)

$$\begin{aligned} \frac{(\mu(P) - \mu(Q))^2}{\sigma^2(P) + \sigma^2(Q)} &= \frac{|p - q|}{p(1 - p) + q(1 - q)} |\mu(P) - \mu(Q)| \\ &\leq 2 |\mu(P) - \mu(Q)| \\ &< 2\sqrt{\delta} \\ &\leq \varepsilon - 2\delta. \end{aligned} \quad \square$$

Als Vorbereitung des Beweises von Satz 3.19 beweisen wir einige Ungleichungen:

*Bemerkung 3.23.* Für  $n \in \mathbb{N}_0$ ,  $p \in [0, 1]$  und  $\lambda \in [0, \infty[$  gilt

$$\frac{(np - \lambda)^2}{np(1 - p) + \lambda} \leq 6 \frac{(\sqrt{np} - \sqrt{\lambda})^2}{1 - p \min\{1, np\}}, \quad (3.41)$$

$$\frac{|np(1 - p) - \lambda|}{1 + np(1 - p) + \lambda} \leq 2(\sqrt{np} - \sqrt{\lambda})^2 + 4|\sqrt{np} - \sqrt{\lambda}| + p \min\{1, np\} \quad (3.42)$$

und, unter Anwendung von (3.15) in der ersten und (3.12) in der dritten Ungleichung

$$\begin{aligned} \frac{p \min\{1, np\}}{32} &\leq d_{\text{TV}}(P_{np}, B_{n,p}) \\ &\leq d_{\text{TV}}(P_{np}, P_\lambda) + d_{\text{TV}}(P_\lambda, B_{n,p}) \\ &\leq \sqrt{\frac{2}{e}} \cdot |\sqrt{np} - \sqrt{\lambda}| + d_{\text{TV}}(P_\lambda, B_{n,p}). \end{aligned} \quad (3.43)$$

Weiter ist

$$(\sqrt{np} - \sqrt{\lambda})^2 = \frac{(np - \lambda)^2}{(\sqrt{np} + \sqrt{\lambda})^2} \leq \frac{(np - \lambda)^2}{np + \lambda} \leq \frac{(np - \lambda)^2}{np(1 - p) + \lambda}. \quad (3.44)$$

*Beweis.* Zu (3.41): Wenn  $np = 0$ , ist

$$\frac{(np - \lambda)^2}{np(1 - p) + \lambda} = \lambda = \frac{(\sqrt{np} - \sqrt{\lambda})^2}{1 - p \min\{1, np\}}.$$

Sei also o. E.  $np \neq 0$ , also

$$\frac{1}{1 - p} = \frac{1 + p}{1 - p^2} \leq \frac{2}{1 - p^2} \leq \frac{2}{1 - p \min\{1, np\}}. \quad (3.45)$$

Wenn  $\lambda = 0$ , ist

$$\frac{(np - \lambda)^2}{np(1 - p) + \lambda} = \frac{np}{1 - p} \leq 2 \frac{(\sqrt{np} - \sqrt{\lambda})^2}{1 - p \min\{1, np\}}.$$

Wenn  $\lambda \neq 0$  und  $p = 1$ , ist

$$\frac{(np - \lambda)^2}{np(1 - p) + \lambda} = \frac{(n - \lambda)^2}{\lambda} \leq \infty \cdot \mathbb{1}_{\{n \neq \lambda\}} = 6 \frac{(\sqrt{n} - \sqrt{\lambda})^2}{0} = 6 \frac{(\sqrt{np} - \sqrt{\lambda})^2}{1 - p \min\{1, np\}}.$$

Wenn  $\lambda \neq 0$  und  $p < 1$ , ist, unter Verwendung von (3.45) im dritten Schritt,

$$\begin{aligned} \frac{(np - \lambda)^2}{np(1 - p) + \lambda} &\leq (\sqrt{np} - \sqrt{\lambda})^2 \frac{1}{1 - p} \left( \frac{(\sqrt{np} + \sqrt{\lambda})^2}{np + \lambda} \right) \\ &\leq (\sqrt{np} - \sqrt{\lambda})^2 \frac{1}{1 - p} \cdot 2 \\ &\leq 4 \frac{(\sqrt{np} - \sqrt{\lambda})^2}{1 - p \min\{1, np\}} \\ &\leq 6 \frac{(\sqrt{np} - \sqrt{\lambda})^2}{1 - p \min\{1, np\}}. \end{aligned}$$

Zu (3.42): Es ist

$$p \cdot np \leq p \min\{1, np\} \max\{1, \lambda\} + (\sqrt{np} - \sqrt{\lambda})^2 + 2\sqrt{\lambda}|\sqrt{np} - \sqrt{\lambda}|,$$

denn für  $np \geq 1$  ist

$$\begin{aligned} p \cdot np &= p(np - \lambda + \lambda) \\ &\leq |\sqrt{np} - \sqrt{\lambda}| \cdot |\sqrt{np} - \sqrt{\lambda} + 2\sqrt{\lambda}| + p\lambda \\ &\leq (\sqrt{np} - \sqrt{\lambda})^2 + 2\sqrt{\lambda}|\sqrt{np} - \sqrt{\lambda}| + p \min\{1, np\} \lambda. \end{aligned}$$

Also ist

$$\begin{aligned} \frac{|np(1 - p) - \lambda|}{1 + np(1 - p) + \lambda} &\leq \frac{|np - \lambda| + p \cdot np}{1 + np(1 - p) + \lambda} \\ &\leq \frac{(\sqrt{np} - \sqrt{\lambda})^2 + 2\sqrt{\lambda}|\sqrt{np} - \sqrt{\lambda}| + p \cdot np}{1 + np(1 - p) + \lambda} \\ &\leq 2(\sqrt{np} - \sqrt{\lambda})^2 + 4|\sqrt{np} - \sqrt{\lambda}| + p \min\{1, np\}. \quad \square \end{aligned}$$

Wir kommen nun zum Beweis von Satz 3.19. Da die Formulierung einer geeigneten unteren Schranke für den Totalvariationsabstand der betrachteten Verteilungen Gegenstand aktueller Forschung ist, kann es eventuell zu Inhomogenitäten in der Beweisführung kommen. Wenn z. B. an einigen Stellen davon die Rede ist, dass  $a \leq b$  für zwei Folgen  $(a) := (a_n)_n$  und  $(b) := (b_n)_n$  gilt, so ist dies natürlich punktweise zu verstehen. Der Gebrauch dieser oder anderer salopper Schreibweisen dürfte aber keinen Einfluss auf die Korrektheit der Ergebnisse haben.

*Beweis von Satz 3.19.* Für  $\varepsilon > 0$  sei  $V_\varepsilon := \{(P, Q) \in \text{BP} : d_{\text{TV}}(P, Q) < \varepsilon\}$ , also  $\{V_\varepsilon : \varepsilon > 0\}$  die Basis der uniformen Struktur des Totalvariationsabstandes auf BP, und sei

### 3 Proximität von Binomial- und Poisson-Verteilungen

$\{U_\varepsilon: \varepsilon > 0\}$  die Basis der uniformen Struktur aus Satz 3.17. Wir zeigen unten die Äquivalenz von (3.18) und (3.19). Daraus folgt, dass es für jedes  $\varepsilon > 0$  ein  $\delta > 0$  mit  $V_\varepsilon \subseteq U_\delta$  und  $V_\delta \subseteq U_\varepsilon$  gibt. Also sind diese Basen die Basen der gleichen uniformen Struktur.

Wir kommen jetzt zur Äquivalenz von (3.18) und (3.19) und schreiben hier abkürzend  $B_{\mathbb{N}_0, [0,1]} := \{B_{n,p}: n \in \mathbb{N}_0, p \in [0,1]\}$  und  $P_{[0,\infty[} := \{P_\lambda: \lambda \in [0,\infty[ \}$ .

Es gelte (3.18). Wir nutzen ein Teiltonfolgenargument und nehmen o. E.  $((P_k), (Q_k)) \in \left( (B_{\mathbb{N}_0, [0,1]})^{\mathbb{N}} \cup (P_{[0,\infty[})^{\mathbb{N}} \right)^2$  an und zeigen, dass (3.19) in diesen Spezialfällen folgt. Damit folgt (3.19) allgemein.

1. Fall: Es ist  $((P_k), (Q_k)) \in \left( (P_{[0,\infty[})^{\mathbb{N}} \right)^2$ . Dann gilt nach (2.2) und (3.11) für alle  $k \in \mathbb{N}$

$$1 - \exp \left( - \left( \frac{1}{12} \frac{(\mu(P_k) - \mu(Q_k))^2}{\sigma^2(P_k) + \sigma^2(Q_k)} + \frac{1}{4} \left( \frac{|\sigma^2(P_k) - \sigma^2(Q_k)|}{1 + \sigma^2(P_k) + \sigma^2(Q_k)} \right)^2 \right) \right) \leq d_{\text{TV}}(P_k, Q_k),$$

und in diesem Fall gilt (3.19).

2. Fall: Es ist  $((P_k), (Q_k)) \in \left( (P_{[0,\infty[})^{\mathbb{N}} \times (B_{\mathbb{N}_0, [0,1]})^{\mathbb{N}} \right) \cup \left( (B_{\mathbb{N}_0, [0,1]})^{\mathbb{N}} \times (P_{[0,\infty[})^{\mathbb{N}} \right)$ , o. E.  $(P_k) = (P_{\lambda_k})$  und  $(Q_k) = (B_{n_k, p_k})$  mit  $(\lambda) := (\lambda_k) \in [0, \infty[^{\mathbb{N}}$ ,  $(n) := (n_k) \in \mathbb{N}_0^{\mathbb{N}}$  und  $(p) := (p_k) \in [0, 1]^{\mathbb{N}}$ . Wir nutzen ein Teiltonfolgenargument, um zu zeigen, dass  $|\sqrt{np} - \sqrt{\lambda}| \rightarrow 0$ : Sei eine Teilfolge von  $(k)$  gegeben.

( $\alpha$ ): Die Teilfolge  $(np(1-p))$  ist beschränkt. Dann ist die Teilfolge  $(B_{n,p} * \delta_{-\lfloor np \rfloor})$   $L^1$ -beschränkt und zentriert  $L^2$ -beschränkt, also  $L^2$ -beschränkt und nach Lemma 3.25 unten straff. Nach dem Satz von Prochorov existieren zu jeder Teilfolge hiervon eine Teiltonfolge und ein  $P \in \text{Prob}(\mathbb{R})$  mit  $B_{n,p} * \delta_{-\lfloor np \rfloor} \rightsquigarrow P$ . Da schwache Konvergenz äquivalent ist zur Konvergenz in  $d_P$  (vgl. Billingsley (1999, S.72)), die auf  $\text{Prob}(\mathbb{Z})$  nach (2.6) äquivalent zur Konvergenz in  $d_{\text{TV}}$  ist, hat nach (3.18) also jede Teilfolge eine Teiltonfolge, entlang der  $P_\lambda * \delta_{-\lfloor np \rfloor} \rightsquigarrow P$  und  $d_{\text{TV}}(P_\lambda * \delta_{-\lfloor np \rfloor}, P) \rightarrow 0$  gilt, und damit nach Barbour und Jensen (1989, Lemma 1)  $(P_\lambda * \delta_{-\lfloor np \rfloor})$  zentriert  $L^2$ -beschränkt ist.

Nach dem Satz von Prochorov ist  $(P_\lambda * \delta_{-\lfloor np \rfloor})$  straff, und aufgrund der zentrierten  $L^2$ -Beschränktheit nach Lemma 3.26 unten  $L^1$ -beschränkt. Wegen der zentrierten  $L^2$ -Beschränktheit und der  $L^1$ -Beschränktheit ist  $(P_\lambda * \delta_{-\lfloor np \rfloor})$   $L^2$ -beschränkt.

Aus der jeweiligen  $L^2$ -Beschränktheit folgt, dass  $(B_{n,p} * \delta_{-\lfloor np \rfloor})$  und  $(P_\lambda * \delta_{-\lfloor np \rfloor})$  gleichgradig integrierbar sind (vgl. Gut (2005, Chapter 5, Theorem 4.2.)). Nach obigem hat also jede Teilfolge eine Teiltonfolge, entlang der  $\mu(B_{n,p}) - \lfloor np \rfloor \rightarrow \mu(P)$  und  $\mu(P_\lambda) - \lfloor np \rfloor \rightarrow \mu(P)$  gilt. Also gilt  $|\sqrt{np} - \sqrt{\lambda}| \rightarrow 0$ .

( $\beta$ ): Es existiert eine Teiltonfolge mit  $np(1-p) \rightarrow \infty$ , nach Barbour und Jensen (1989, Lemma 1) also  $\sup\{B_{n,p}(\{z\}): z \in \mathbb{N}_0\} \rightarrow 0$ . Dann gilt wegen (3.18) nach Lemma 1.5 in Mattner und Schulz (2014), welches in einer schärferen Version auf Paul Lévy zurückgeht, auch  $\lambda \rightarrow \infty$ . In der folgenden Abschätzung gilt die erste Ungleichung nach (3.44), die zweite Ungleichung nach (3.17), und nach der Berry-Esseen-Ungleichung existiert ein

$C \in ]0, \infty[$ , so dass

$$\begin{aligned}
 \frac{1}{2} \left( 1 - \exp \left( -\frac{1}{4} (\sqrt{np} - \sqrt{\lambda})^2 \right) \right) &\leq \frac{1}{2} \left( 1 - \exp \left( -\frac{1}{4} \frac{(np - \lambda)^2}{np(1-p) + \lambda} \right) \right) \\
 &\leq d_K(N_{np, np(1-p)}, N_{\lambda, \lambda}) \\
 &\leq d_K(N_{np, np(1-p)}, B_{n,p}) + d_K(B_{n,p}, P_\lambda) + d_K(P_\lambda, N_{\lambda, \lambda}) \\
 &\leq C \left( \frac{1}{\sqrt{np(1-p)}} + \frac{1}{\sqrt{\lambda}} \right) + d_K(B_{n,p}, P_\lambda) \\
 &\rightarrow 0,
 \end{aligned}$$

also gilt hier  $|\sqrt{np} - \sqrt{\lambda}| \rightarrow 0$ .

Nach Betrachtung der Fälle  $(\alpha)$  und  $(\beta)$  und dem Teiltailfolgenargument konvergiert also im 2. Fall  $|\sqrt{np} - \sqrt{\lambda}| \rightarrow 0$ , also nach (3.43) auch  $p \cdot \min\{1, np\} \rightarrow 0$ , und nach (3.41) und (3.42) gilt damit im 2. Fall (3.19).

**3. Fall:** Es ist  $((P_k), (Q_k)) \in \left( (B_{\mathbb{N}_0, [0,1]})^{\mathbb{N}} \right)^2$ . Seien  $(n) := (n_k), (m) := (m_k) \in \mathbb{N}_0^{\mathbb{N}}$  und  $(p) := (p_k), (q) := (q_k) \in [0, 1]^{\mathbb{N}}$  so, dass  $(P_k) = (B_{n,p})$  und  $(Q_k) = (B_{m,q})$ . Wir nutzen wiederum ein Teiltailfolgenargument und zeigen in den folgenden Fällen  $(\alpha)$  bis  $(\gamma)$ , dass (3.19) gilt. Sei eine Teilfolge gegeben. Da  $d_{TV}(B_{n,p}, B_{m,q}) \rightarrow 0$ , nehmen wir an, dass  $d_{TV}(B_{n,p}, B_{m,q}) < 1/3$ .

$(\alpha)$ : Es existieren ein  $0 \leq \Delta' < \infty$  und eine Teilfolge, entlang der  $0 \leq np \leq \Delta'$ . Nach (3.21) und der Annahme  $d_{TV}(B_{n,p}, B_{m,q}) < 1/3$  ist

$$\frac{(np - mq)^2}{np(1-p) + mq(1-q)} \leq 3d_{TV}(B_{n,p}, B_{m,q}) < 1.$$

Da  $np(1-p) \leq np \leq \Delta'$ , ist  $\sqrt{mq} (\sqrt{mq} - \sqrt{1-q}) \leq \Delta' + \sqrt{\Delta'}$ , also

$$mq(1-q) \leq mq \leq \sup \{2, \Delta' + \sqrt{\Delta'}\} =: \Delta.$$

Seien  $\mu_j := \int x^j dB_{n,p}(x)$  und  $\lambda_j := \int x^j dB_{m,q}(x)$  für  $j \in \mathbb{N}$ . Nach zweifacher Anwendung von (3.20) in der dritten Ungleichung, und (2.13) in der vierten Ungleichung ist

$$\begin{aligned}
 \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} &\leq \left| \mu_2 - \mu_1^2 - \lambda_2 + \lambda_1^2 \right| \\
 &\leq |\mu_2 - \lambda_2| + |\mu_1 - \lambda_1|(\mu_1 + \lambda_1) \\
 &\leq \sqrt{3d_{TV}(B_{n,p}, B_{m,q}) (\mu_4 - \mu_2^2 + \lambda_4 - \lambda_2^2)} \\
 &\quad + \sqrt{3d_{TV}(B_{n,p}, B_{m,q}) (\mu_2 - \mu_1^2 + \lambda_2 - \lambda_1^2)} (\mu_1 + \lambda_1) \\
 &\leq \sqrt{3d_{TV}(B_{n,p}, B_{m,q}) 12\Delta(1 + \Delta)^3} \\
 &\quad + \sqrt{3d_{TV}(B_{n,p}, B_{m,q}) 2\Delta(1 + \Delta) 2\Delta} \\
 &= \sqrt{d_{TV}(B_{n,p}, B_{m,q})} \sqrt{\Delta(1 + \Delta)} (\sqrt{6} (\sqrt{6} + 2) \Delta + 6).
 \end{aligned}$$

### 3 Proximität von Binomial- und Poisson-Verteilungen

Also ist, wenn  $d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) < 1/3$ , mit (3.21)

$$\frac{1}{4} \left( \frac{\left( \frac{|np(1-p) - mq(1-q)|}{1 + np(1-p) + mq(1-q)} \right)^2}{\Delta(1 + \Delta) \left( \sqrt{6} (\sqrt{6} + 2) \Delta + 6 \right)^2} + \frac{(np - mq)^2}{np(1-p) + mq(1-q)} \right) \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}).$$

( $\beta$ ): Es existieren ein  $0 \leq \Delta < \infty$  und eine Teilteifolge, entlang der  $np(1-p) \leq \Delta$  und  $np \rightarrow \infty$ . Dann konvergiert

$$p = 1 - \frac{np(1-p)}{np} \geq 1 - \frac{\Delta}{np} \rightarrow 1,$$

also  $p \rightarrow 1$ , und es ist

$$\limsup n(1-p) \leq \limsup \frac{\Delta}{p} = \Delta.$$

Da  $\mathbb{B}_{n,1-p} = \check{\mathbb{B}}_{n,p} * \delta_n$  ist, erhalten wir, unter Anwendung von (3.15) in der dritten Ungleichung,

$$\begin{aligned} \left| \check{\mathbb{B}}_{m,q} * \delta_n(\{0\}) - \mathbb{P}_{n(1-p)}(\{0\}) \right| &\leq d_{\text{TV}}(\check{\mathbb{B}}_{m,q} * \delta_n, \mathbb{P}_{n(1-p)}) \\ &\leq d_{\text{TV}}(\check{\mathbb{B}}_{m,q} * \delta_n, \check{\mathbb{B}}_{n,p} * \delta_n) + d_{\text{TV}}(\mathbb{B}_{n,1-p}, \mathbb{P}_{n(1-p)}) \\ &\leq d_{\text{TV}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + (1-p) \cdot \min\{1, n(1-p)\} \\ &\rightarrow 0. \end{aligned}$$

Da  $\limsup n(1-p) \leq \Delta$ , ist also

$$\liminf \check{\mathbb{B}}_{m,q} * \delta_n(\{0\}) \geq \liminf \mathbb{P}_{n(1-p)}(\{0\}) = \liminf \exp(-n(1-p)) \geq \exp(-\Delta).$$

Nach Baillon et al. (2013) existiert eine Konstante  $\eta \approx 0,4688$  mit

$$\limsup mq(1-q) = \limsup \sigma^2(\check{\mathbb{B}}_{m,q} * \delta_n) \leq \left( \frac{\eta}{\exp(-\Delta)} \right)^2 = \eta^2 \exp(2\Delta).$$

Also ist auch  $(mq(1-q))$  beschränkt.

Wir zeigen, dass schließlich  $m = n$  gilt. Angenommen, es gibt eine Teilfolge, entlang der  $m < n$  gilt. Dies führt zum Widerspruch

$$0 = \lim d_{\text{TV}}(\check{\mathbb{B}}_{m,q} * \delta_n, \mathbb{P}_{n(1-p)}) \geq \liminf \mathbb{P}_{n(1-p)}(\{0\}) > 0. \quad (3.46)$$

Angenommen, es gibt eine Teilfolge, entlang der  $m > n$  gilt. Nach (3.21) ist

$$\frac{1}{3} \min \left\{ 1, \frac{(np - mq)^2}{np(1-p) + mq(1-q)} \right\} \leq d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) \rightarrow 0.$$

Da  $np(1-p)$  und  $mq(1-q)$  beschränkt sind, und da  $np \rightarrow \infty$ , gilt auch  $mq \rightarrow \infty$ . Also kann man das gleiche Argument wie oben, mit vertauschten Rollen von  $(n, p)$  und  $(m, q)$  anwenden, und erhält den gleichen Widerspruch wie in (3.46).

Also ist schließlich  $m = n$  und damit

$$d_{\text{TV}}(\mathbb{B}_{n,p}, \mathbb{B}_{m,q}) = d_{\text{TV}}(\mathbb{B}_{n,1-p}, \mathbb{B}_{m,1-q})$$

mit  $\limsup n(1-p) \leq \Delta$ , und man verfährt weiter wie im Fall  $(\alpha)$ .

$(\gamma)$ : Es gibt eine Teilteilstolge, entlang der  $np(1-p) \rightarrow \infty$ . Seien  $\mu := np, \sigma^2 := np(1-p), \lambda := mq, \tau^2 := mq(1-q)$ . Die erste Ungleichung in der folgenden Abschätzung ist eine erneute Anwendung von Lemma 1.5 in Mattner und Schulz (2014); die letzte Ungleichung gilt nach (2.4) und der Berry-Esseen-Ungleichung, nach der ein  $C \in ]0, \infty[$  existiert, mit

$$\begin{aligned} \frac{1}{\sqrt{1+12\tau^2}} &\leq \max\{\mathbb{B}_{m,q}(\{k\}) : k \in \mathbb{N}_0\} \\ &\leq 2d_{\mathbb{K}}(\mathbb{B}_{m,q}, \mathbb{N}_{np, np(1-p)}) \\ &\leq 2d_{\mathbb{K}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + 2d_{\mathbb{K}}(\mathbb{B}_{n,p}, \mathbb{N}_{np, np(1-p)}) \\ &\leq 2d_{\text{TV}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + 2C/\sigma. \end{aligned}$$

Also ist

$$\frac{1}{\tau} \leq 4\sqrt{3} \cdot \frac{d_{\text{TV}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + C/\sigma}{\sqrt{1 - (2(d_{\text{TV}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + C/\sigma))^2}},$$

und damit

$$\frac{1}{\sigma} + \frac{1}{\tau} \leq \frac{4\sqrt{3}(d_{\text{TV}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + C/\sigma) + 1/\sigma}{\sqrt{1 - (2(d_{\text{TV}}(\mathbb{B}_{m,q}, \mathbb{B}_{n,p}) + C/\sigma))^2}}.$$

Weiter ist

$$\begin{aligned} \left( \frac{|\sigma^2 - \tau^2|}{1 + \sigma^2 + \tau^2} \right)^2 &\leq \frac{(\sigma + \tau)^2(\sigma - \tau)^2}{(\sigma^2 + \tau^2)^2} \\ &= \left( 1 + \frac{2\sigma\tau}{\sigma^2 + \tau^2} \right) \left( 1 + \sqrt{\frac{2\sigma\tau}{\sigma^2 + \tau^2}} \right) \left( 1 - \sqrt{\frac{2\sigma\tau}{\sigma^2 + \tau^2}} \right) \\ &\leq 2(1 + \sqrt{2}) \left( 1 - \sqrt{\frac{2\sigma\tau}{\sigma^2 + \tau^2}} \right). \end{aligned}$$

Also existiert, unter Anwendung von (3.17) in der zweiten Ungleichung, nach der Berry-

### 3 Proximität von Binomial- und Poisson-Verteilungen

Esseen-Ungleichung ein  $C \in ]0, \infty[$  mit

$$\begin{aligned}
& \frac{1}{2} \left( 1 - \left( 1 - \frac{1}{2(1+\sqrt{2})} \left( \frac{|\sigma^2 - \tau^2|}{1 + \sigma^2 + \tau^2} \right)^2 \right) \exp \left( -\frac{1}{4} \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2} \right) \right) \\
& \leq \frac{1}{2} \left( 1 - \sqrt{\frac{2\sigma\tau}{\sigma^2 + \tau^2}} \exp \left( -\frac{1}{4} \frac{(\mu - \lambda)^2}{\sigma^2 + \tau^2} \right) \right) \\
& \leq d_K(N_{\mu, \sigma^2}, N_{\lambda, \tau^2}) \\
& \leq d_K(N_{\mu, \sigma^2}, B_{n,p}) + d_K(B_{n,p}, B_{m,q}) + d_K(B_{m,q}, N_{\lambda, \tau^2}) \\
& \leq C \left( \frac{1}{\sigma} + \frac{1}{\tau} \right) + d_K(B_{n,p}, B_{m,q}) \\
& \leq C \left( \frac{4\sqrt{3}(d_{TV}(B_{m,q}, B_{n,p}) + C/\sigma) + 1/\sigma}{\sqrt{1 - (2(d_{TV}(B_{m,q}, B_{n,p}) + C/\sigma))^2}} \right) + d_{TV}(B_{n,p}, B_{m,q}).
\end{aligned}$$

Also gilt (3.19).

Damit ist im 3. Fall nach dem Teilleitfolgenargument gezeigt, dass (3.19) gilt. Also folgt in allen Fällen aus (3.18) (3.19).

Es gelte (3.19). Dann folgt (3.18) aus der Ungleichung in Satz 3.20, für dessen Beweis Satz 3.19 nicht benutzt wird.  $\square$

Der folgende Beweis der Ungleichung in Satz 3.20 kombiniert die Ergebnisse der Abschnitte 3.1 – 3.3:

*Beweis von Satz 3.20. 1. Fall:*  $P$  und  $Q$  sind Poisson-Verteilungen. Dann folgt die Behauptung aus (3.13).

*2. Fall:*  $P$  und  $Q$  sind Binomialverteilungen. Dann existieren  $m, n \in \mathbb{N}_0$ ,  $p, q \in [0, 1]$  mit  $P = B_{n,p}$  und  $Q = B_{m,q}$ . Sei o. E.  $m \leq n$ . Wenn  $q \leq p$ , folgt die Behauptung nach Satz 3.8. Sei also o. E.  $p < q$ . Wenn  $p < \min \left\{ \frac{m}{n}q, 1 - \frac{m}{n}q \right\}$ , ist  $p < \min \left\{ \frac{m}{n}q, \frac{m}{n} \frac{1}{2} \right\}$  oder  $p \in \left[ \frac{m}{n} \frac{1}{2}, \frac{m}{n}q \right]$ , und die Behauptung folgt nach Satz 3.10 bzw. 3.14. Wenn  $p > \max \left\{ \frac{m}{n}q, 1 - \frac{m}{n}q \right\}$ , ist  $1 - \frac{m}{n}q < p < q < \frac{n}{m}p$ , und die Behauptung folgt nach Satz 3.13. Sonst ist  $p \in \left[ \frac{m}{n}q, 1 - \frac{m}{n}q \right] \cup \left[ 1 - \frac{m}{n}q, \frac{m}{n}q \right]$ , und die Behauptung folgt nach Satz 3.11.

*3. Fall:*  $P$  und  $Q$  sind eine Poisson-Verteilung und eine Binomialverteilung. Dann folgt die Behauptung aus Satz 3.16.  $\square$

*Beweis von Lemma 3.21.* Die erste Ungleichung verifiziert man durch eine Unterscheidung der Fälle  $(\mu_k - \lambda_k)^2 \leq \mu_{2k} - \mu_k^2 + \lambda_{2k} - \lambda_k^2$  und  $(\mu_k - \lambda_k)^2 \geq \mu_{2k} - \mu_k^2 + \lambda_{2k} - \lambda_k^2$ .

Zur zweiten Ungleichung: Mit  $\nu := P + Q$ ,  $f := \frac{dP}{d\nu}$  und  $g := \frac{dQ}{d\nu}$  ist

$$\mu_k - \lambda_k = \int \left( x^k - \frac{\mu_k + \lambda_k}{2} \right) (f(x) - g(x)) d\nu(x),$$

und damit, nach Jensen-Ungleichung in der ersten, und Cauchy-Schwarz-Ungleichung in der zweiten Ungleichung, und unter Ausnutzung von (2.1) in der zweiten Gleichheit,

$$\begin{aligned}
 |\mu_k - \lambda_k| &\leq \int \left| x^k - \frac{\mu_k + \lambda_k}{2} \right| |f(x) - g(x)| d\nu(x) \\
 &= \int \left| x^k - \frac{\mu_k + \lambda_k}{2} \right| d(|f - g|\nu)(x) \\
 &\leq \left( \int 1^2 |f(x) - g(x)| d\nu(x) \cdot \int \left( x^k - \frac{\mu_k + \lambda_k}{2} \right)^2 d(P + Q)(x) \right)^{1/2} \\
 &= \left( 2d_{\text{TV}}(P, Q) \cdot \left( \mu_{2k} - \mu_k^2 + \lambda_{2k} - \lambda_k^2 + \frac{(\mu_k - \lambda_k)^2}{2} \right) \right)^{1/2}. \quad \square
 \end{aligned}$$

## 3.8 Ergebnisse für BPBP-Faltungen

Dieser Abschnitt enthält Ergebnisse, die schon aus Tasto (2011) bekannt sind oder nur durch kleine Änderungen aus Ergebnissen dort hervorgehen. Der Teilabschnitt 3.8.1 enthält Ergebnisse, die in der vorliegenden Arbeit verwendet werden, und deren Beweise. Der Teilabschnitt 3.8.2 enthält Ergebnisse, die im vorliegenden Themenkreis interessant sind, aber hier nicht verwendet werden, und deren Beweis daher hier nicht erneut gegeben wird.

### 3.8.1 Verwendete Ergebnisse aus Tasto (2011)

**Definition 3.24.** Für  $(p_j)_j \in [0, 1]^{\mathbb{N}}$  mit  $\sum_{j=1}^{\infty} p_j < \infty$  ist  $\bigstar_{j=1}^{\infty} B_{p_j}$  der Grenzwert  $\lim_{n \rightarrow \infty} \bigstar_{j=1}^n B_{p_j}$  in der schwachen Topologie auf  $\text{Prob}(\mathbb{R})$ .

Sind  $(p_j)_j, (q_j)_j \in [0, 1]^{\mathbb{N}}$  mit  $\sum_{j=1}^{\infty} p_j < \infty$  und  $\sum_{j=1}^{\infty} q_j < \infty$ , und  $\mu, \lambda \in [0, \infty[$ , so heißen  $\bigstar_{j=1}^{\infty} B_{p_j} * P_{\lambda}$  BP-Faltung und  $\bigstar_{j=1}^{\infty} B_{p_j} * P_{\lambda} * \bigstar_{j=1}^{\infty} \check{B}_{q_j} * \check{P}_{\mu}$  BPBP-Faltung.

**Lemma 3.25.** Sei  $(P_n) \in \text{Prob}(\mathbb{R})^{\mathbb{N}}$  mit  $\sup \{ |\mu(P_n)|, \sigma^2(P_n) : n \in \mathbb{N} \} < \infty$ . Dann ist  $(P_n)$  straff.

*Beweis.* Seien  $\varepsilon > 0$  und  $c > \sup \{ |\mu(P_n)| \} + \sqrt{\sup \{ \sigma^2(P_n) \} / \varepsilon}$ . Dann ist  $[-c, c]$  kompakt, und nach der Chebyshev-Ungleichung gilt für alle  $n \in \mathbb{N}$

$$P_n(\mathbb{R} \setminus [-c, c]) \leq P_n(\{x : |x - \mu(P_n)| \geq c - |\mu(P_n)|\}) \leq \frac{\sigma^2(P_n)}{(c - |\mu(P_n)|)^2} < \varepsilon. \quad \square$$

**Lemma 3.26.** Eine Folge von reellwertigen Zufallsgrößen ist genau dann straff und zentriert  $L^1$ -beschränkt, wenn sie  $L^1$ -beschränkt ist.

### 3 Proximität von Binomial- und Poisson-Verteilungen

*Beweis.* Seien  $t > 0$ ,  $(X_n) \in \text{Prob}(\mathbb{R})^{\mathbb{N}}$  und  $\mu_n := \mathbb{E}(X_n)$  für alle  $n \in \mathbb{N}$ .

Wenn  $(X_n)$  straff und zentriert  $L^1$ -beschränkt ist, gilt  $\sup_{n \in \mathbb{N}} \mathbb{P}(|X_n| \geq t) \rightarrow 0$ ,  $t \rightarrow \infty$ , und nach der Markov-Ungleichung

$$\sup_{n \in \mathbb{N}} \mathbb{P}(|X_n - \mu_n| \geq t) \leq \frac{1}{t} \sup_{n \in \mathbb{N}} \mathbb{E}|X_n - \mu_n| \rightarrow 0, \quad t \rightarrow \infty,$$

und zusammen

$$\sup_{n \in \mathbb{N}} \mathbb{P}(|\mu_n| \geq 2t) \leq \sup_{n \in \mathbb{N}} \mathbb{P}(|X_n| \geq t) + \sup_{n \in \mathbb{N}} \mathbb{P}(|X_n - \mu_n| \geq t) \rightarrow 0, \quad t \rightarrow \infty.$$

Da  $\mathbb{P}(|\mu_n| \geq 2t) \in \{0, 1\}$  für  $t \in \mathbb{R}$ , ist  $\sup_{n \in \mathbb{N}} |\mu_n| < \infty$ , und damit

$$\sup_{n \in \mathbb{N}} \mathbb{E}|X_n| \leq \sup_{n \in \mathbb{N}} \mathbb{E}|X_n - \mu_n| + \sup_{n \in \mathbb{N}} |\mu_n| < \infty,$$

also ist  $(X_n)$   $L^1$ -beschränkt.

Wenn  $(X_n)$   $L^1$ -beschränkt ist, gilt

$$\sup_{n \in \mathbb{N}} \mathbb{E}|X_n - \mu_n| \leq \sup_{n \in \mathbb{N}} |\mu_n| + \sup_{n \in \mathbb{N}} \mathbb{E}|X_n| \leq 2 \sup_{n \in \mathbb{N}} \mathbb{E}|X_n| < \infty,$$

also ist  $(X_n)$  zentriert  $L^1$ -beschränkt, und damit gilt nach der Markov-Ungleichung

$$\sup_{n \in \mathbb{N}} \mathbb{P}(|X_n| \geq t) \leq \frac{1}{t} \sup_{n \in \mathbb{N}} \mathbb{E}|X_n| \rightarrow 0, \quad t \rightarrow \infty,$$

also ist  $(X_n)$  straff. □

### 3.8.2 Nicht verwendete Ergebnisse aus Tasto (2011)

**Lemma 3.27.** Sei  $(P_n)_{n \in \mathbb{N}}$  eine Folge von BPBP-Faltungen mit  $P_n \rightsquigarrow P$ ,  $n \rightarrow \infty$ . Dann konvergieren für  $n \rightarrow \infty$  für alle  $r \in \mathbb{N}_0$  die Kumulanten  $\kappa_r(P_n) \rightarrow \kappa_r(P)$  und alle Momente  $\mu_r(P_n) \rightarrow \mu_r(P)$ .

**Satz 3.28.** Seien  $(P_n) := \left( \bigstar_{j=1}^{\infty} B_{p_{j_n}} * P_{\lambda_n} \right)$  eine Folge von BP-Faltungen und  $P \in \text{Prob}(\mathbb{N}_0)$ . Dann sind äquivalent:

(1)  $P_n \rightsquigarrow P$ ,  $n \rightarrow \infty$ .

(2)  $P = \bigstar_{j=1}^{\infty} B_{p_j} * P_{\lambda}$  ist eine BP-Faltung,

$$\lim_{n \rightarrow \infty} \left( \lambda_n + \sum_{j=1}^{\infty} p_{j_n} - \sum_{j=1}^{\infty} p_j \right) = \lambda \text{ und}$$

für alle  $j \in \mathbb{N}$  ist  $\lim_{n \rightarrow \infty} p_{j_n} = p_j$ .

Da die Parameter einer BPBP-Faltung im Allgemeinen nicht identifizierbar sind, gilt hier ein Analogon zu Satz 3.28 nur eingeschränkt:

**Satz 3.29.** Seien  $(P_n) := \left( \bigstar_{j=1}^{\infty} B_{p_{j_n}} * P_{\lambda_n} * \bigstar_{i=1}^{\infty} \check{B}_{q_{i_n}} * \check{P}_{\mu_n} \right)$  eine Folge von BPBP-Faltungen und  $P \in \text{Prob}(\mathbb{Z})$ . Für die Aussagen

$$(1) P_n \rightsquigarrow P, n \rightarrow \infty,$$

$$(2) P = \bigstar_{j=1}^{\infty} B_{p_j} * P_{\lambda} * \bigstar_{i=1}^{\infty} \check{B}_{q_i} * \check{P}_{\mu} \text{ ist eine BPBP-Faltung,}$$

$$\lim_{n \rightarrow \infty} \left( \lambda_n + \sum_{j=1}^{\infty} p_{j_n} - \sum_{j=1}^{\infty} p_j \right) = \lambda, \quad \lim_{n \rightarrow \infty} \left( \mu_n + \sum_{i=1}^{\infty} q_{i_n} - \sum_{i=1}^{\infty} q_i \right) = \mu,$$

$$\text{für alle } j, i \in \mathbb{N} \text{ ist } \lim_{n \rightarrow \infty} p_{j_n} = p_j, \quad \lim_{n \rightarrow \infty} q_{i_n} = q_i,$$

gilt dann

$$(1) \Rightarrow (2), \text{ falls } p_{j_n}, q_{i_n} \leq 1/2 \text{ für alle } i, j, n \in \mathbb{N},$$

und  $(2) \Rightarrow (1)$ .



# 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

Das Thema der Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten wird in Abschnitt 4.2 beispielhaft anhand der Betrachtung von MRSA-Fällen in einem Krankenhaus eingeführt. Natürlich bleiben aber alle mathematischen Aussagen auch außerhalb dieses Kontextes gültig. Um Artefakte zu vermeiden, beschränken wir uns bei allen statistischen Untersuchungen auf die Betrachtung von Niveaus  $\beta \in ]0, 1[$ .

Vorbereitend werden im folgenden Abschnitt 4.1 einige Definitionen und Eigenschaften von Konfidenzregionen und der Binomialverteilung bereitgestellt.

Die Untersuchung von Konfidenzintervallen für die durchschnittliche Erfolgswahrscheinlichkeit im Bernoulli-Faltungs-Modell gipfelt in der Angabe eines optimalen isotonen einseitigen Konfidenzintervalls in Satz 4.7.

In Abschnitt 4.3 wird die Untersuchung von Bernoulli-Faltungen auf hypergeometrische Verteilungen spezialisiert. In Satz 4.22 wird für die Schätzung der Erfolgswahrscheinlichkeit bei unbekannter Urnengröße ein optimales isotones einseitiges Konfidenzintervall angegeben.

An die Betrachtung offener Forschungsfragen in Abschnitt 4.4 schließen sich in Abschnitt 4.5 die Beweise der Aussagen dieses Kapitels an.

## 4.1 Konfidenzregionen und Binomialverteilungen

**Definition und Bemerkung 4.1.** Sei  $n \in \mathbb{N}_0$ . Eine Menge  $J \subseteq [0, 1]$  heißt *Upray in*  $[0, 1]$ , wenn aus  $x \in J$ ,  $y \in [0, 1]$  und  $x \leq y$  zusammen  $y \in J$  folgt. Dies ist äquivalent dazu, dass  $J$  von der Form  $[a, 1]$  oder  $]a, 1]$  für ein  $a \in [0, 1]$  ist.

Eine Funktion  $K: \{0, \dots, n\} \rightarrow 2^{[0,1]}$  heißt *Upray*, wenn jeder Funktionswert  $K(x)$  ein Upray in  $[0, 1]$  ist.

Analog definieren wir *Downrays*, so dass jeder Funktionswert von der Form  $[0, b]$  oder  $[0, b[$  für ein  $b \in [0, 1]$  ist.

Ein Upray  $K: \{0, \dots, n\} \rightarrow 2^{[0,1]}$  heißt *isoton*, wenn es isoton ist bezüglich der gewöhnlichen Ordnung auf  $\{0, \dots, n\}$  und der der Mengeneinklusison entgegengesetzten Ordnung auf  $2^{[0,1]}$ , wenn also die Folgerung

$$x, y \in \{0, \dots, n\}, x < y \Rightarrow K(x) \supseteq K(y)$$

gilt, und *streng isoton*, wenn oben „ $\supseteq$ “ zu „ $\supsetneq$ “ verschärft werden kann. Analog heißt ein Downray  $\Lambda$  *isoton*, wenn aus  $x < y$  folgt, dass  $\Lambda(x) \subseteq \Lambda(y)$ , und *streng isoton*, wenn „ $\subseteq$ “ zu „ $\subsetneq$ “ verschärft werden kann. Beispielsweise ist jedes der  $K_{CP,n}$  und  $\Lambda_{CP,n}$  aus Definition und Bemerkung 4.5 unten streng isoton.

Seien  $\mathcal{P} := (P_\vartheta: \vartheta \in \Theta)$  ein statistisches Experiment auf  $(\{0, \dots, n\}, 2^{\{0, \dots, n\}})$  und  $\kappa: \Theta$  ein interessierender Parameter. Ein isotones  $\beta$ -Konfidenz-Upray  $K$  für das Schätzproblem  $(\mathcal{P}, \kappa)$  heißt (*Buehler-*)*optimal* (vgl. Buehler (1957) und, für eine neuere Diskussion, Lloyd und Kabaila (2010), die durch Wiedererentdeckungen in Wang (2010) hervorgerufen wurde) (siehe auch Bemerkung 4.2), wenn jedes andere isotone  $\beta$ -Konfidenz-Upray  $K^*$  für  $(\mathcal{P}, \kappa)$  die Bedingung

$$K(x) \subseteq K^*(x) \quad \text{für jedes } x \in \{0, \dots, n\} \quad (4.1)$$

erfüllt. Schließlich heißt ein nicht notwendig isotones  $\beta$ -Konfidenz-Upray *zulässig* in der Menge aller Konfidenz-Uprays für  $(\mathcal{P}, \kappa)$  wenn für jedes andere  $\beta$ -Konfidenz-Upray  $K^*$  für  $(\mathcal{P}, \kappa)$  mit  $K^*(x) \subseteq K(x)$  für jedes  $x \in \{0, \dots, n\}$  schon gilt, dass  $K^* = K$ .

*Bemerkung 4.2.* Optimalität, wie sie durch (4.1) definiert wird, ist nicht exakt Optimalität im Sinne von Buehler (1957), erstere impliziert aber letztere:

Für  $n \in \mathbb{N}_0$ ,  $\beta \in ]0, 1[$ ,  $\mathcal{X} := \{0, \dots, n\}$ , ein Experiment  $\mathcal{P} = (P_\vartheta: \vartheta \in \Theta)$  auf  $(\mathcal{X}, 2^{\mathcal{X}})$  und einen interessierenden Parameter  $\kappa: \Theta$  betrachten wir das Schätzproblem  $(\mathcal{P}, \kappa)$ . Dann ist die im Sinne von Buehler (1957) definierte optimale isotone untere  $\beta$ -Konfidenzschranke  $l_{B,\beta}$  definiert durch

$$l_{B,\beta}(x) := \inf \{ \kappa(\vartheta) : P_\vartheta(\{x, \dots, n\}) > 1 - \beta, \vartheta \in \Theta \} \quad \text{für } x \in \{0, \dots, n\}.$$

Ein isotones  $\beta$ -Konfidenz-Upray  $K$  ist optimal im Sinne von (4.1), wenn für jedes isotone  $\beta$ -Konfidenz-Upray  $K^*$  gilt, dass  $K(x) \subseteq K^*(x)$  für alle  $x \in \{0, \dots, n\}$ , also insbesondere  $K(x) \in ]l_{B,\beta}(x), 1], [l_{B,\beta}(x), 1]$  für alle  $x \in \{0, \dots, n\}$ .

**Lemma 4.3.** *Seien  $n \in \mathbb{N}_0$ ,  $\beta \in ]0, 1[$ ,  $\mathcal{X} := \{0, \dots, n\}$ ,  $\mathcal{P} = (P_\vartheta: \vartheta \in \Theta)$  ein Experiment auf  $(\mathcal{X}, 2^{\mathcal{X}})$  und  $\kappa: \Theta$  ein interessierender Parameter, und sei  $K$  das optimale isotone  $\beta$ -Konfidenz-Upray für  $(\mathcal{P}, \kappa)$  und streng isoton. Dann ist  $K$  zulässig in der Menge aller  $\beta$ -Konfidenz-Uprays für  $(\mathcal{P}, \kappa)$ .*

*Bemerkung 4.4.* Für  $n \in \mathbb{N}_0$  und  $x \in \{0, \dots, n\}$  ist die Abbildung  $[0, 1] \ni p \mapsto B_{n,p}(\{0, \dots, x\})$  antiton, denn für  $n \in \mathbb{N}$  und  $x \in \{0, \dots, n-1\}$  ist die Abbildung  $[0, 1] \ni p \mapsto B_{n,p}(\{0, \dots, x\})$  streng antiton, denn für  $p \in ]0, 1[$  ist

$$\frac{d}{dp} B_{n,p}(\{0, \dots, x\}) = -n B_{n-1,p}(\{x\}) < 0. \quad (4.2)$$

**Definition und Bemerkung 4.5.** Für  $\beta \in ]0, 1[$ , und mit

$$g_n(x) := g_{n,\beta}(x) := \text{das } p \in [0, 1] \text{ mit } B_{n,p}(\{x, \dots, n\}) = 1 - \beta$$

für  $n \in \mathbb{N}$  und  $x \in \{1, \dots, n\}$ , welches nach (4.2) wohldefiniert ist und insbesondere

$$g_n(1) = 1 - \beta^{1/n} \quad \text{und} \quad g_n(n) = (1 - \beta)^{1/n} \quad (4.3)$$

und

$$g_{n,\beta}(x) \quad \text{ist streng} \quad \left\{ \begin{array}{l} \text{isoton} \\ \text{antiton} \end{array} \right\} \quad \text{in} \quad \left\{ \begin{array}{l} x \\ \beta \end{array} \right\} \quad (4.4)$$

liefert, sind die *Clopper-Pearson- $\beta$ -Konfidenz-Uprays*  $K_{\text{CP},n}: \{0, \dots, n\} \rightarrow 2^{[0,1]}$  definiert durch

$$K_{\text{CP},n}(x) := K_{\text{CP},n,\beta}(x) := \begin{cases} [0, 1], & \text{falls } x = 0, \\ ]g_n(x), 1], & \text{falls } x \in \{1, \dots, n\}, \end{cases} \quad (4.5)$$

für  $n \in \mathbb{N}_0$ . Insbesondere ist

$$K_{\text{CP},n}(1) = ]1 - \beta^{1/n}, 1] \quad \text{und} \quad K_{\text{CP},n}(n) = ](1 - \beta)^{1/n}, 1]$$

für  $n \in \mathbb{N}$ .

Für  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$  ist  $K_{\text{CP},n}$  streng isoton, das optimale isotone  $\beta$ -Konfidenz-Upray für das Schätzproblem der Erfolgswahrscheinlichkeit in Binomialmodell (4.7) unten und zulässig in der Menge aller Konfidenz-Uprays für (4.7).

Für  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$  sind die *Clopper-Pearson- $\beta$ -Konfidenz-Downrays*  $\Lambda_{\text{CP},n} := \Lambda_{\text{CP},n,\beta}$  definiert durch

$$\Lambda_{\text{CP},n,\beta}(x) := 1 - K_{\text{CP},n}(n - x) \quad \text{für } x \in \{0, \dots, n\}$$

und die *zweiseitigen Clopper-Pearson- $\beta$ -Konfidenzintervalle*  $M_{\text{CP},n} := M_{\text{CP},n,\beta}$  durch

$$M_{\text{CP},n,\beta}(x) := K_{\text{CP},n,\frac{1+\beta}{2}}(x) \cap \Lambda_{\text{CP},n,\frac{1+\beta}{2}}(x) \quad \text{für } x \in \{0, \dots, n\}. \quad (4.6)$$

## 4.2 Der Bernoulli-Faltungs-Fall

Der Methicillin-resistente Staphylococcus aureus (MRSA) ist ein „Staphylococcus aureus, der aufgrund eines veränderten Penicillin-bindenden Proteins resistent gegenüber allen Betalaktam-Antibiotika ist (Oxacillinresistenz); das Auftreten von MRSA-Stämmen im Krankenhaus erfordert gezielte antiepidemische Maßnahmen mit Isolierung des Patienten“ (Pschyrembel (2002, S. 1087)). Wie man der Berichterstattung in den Medien entnehmen kann, ist MRSA heute ein weit verbreitetes und großes Problem in Krankenhäusern, weshalb sein Auftreten weitreichender Untersuchungen bedarf. Unter anderem ist es sinnvoll, die Verteilung von MRSA-Fällen zu untersuchen:

Nehmen wir an, in einem Krankenhaus werden in einem Jahr  $n$  Patienten behandelt, die sich alle unabhängig voneinander mit der gleichen Wahrscheinlichkeit  $p$  mit MRSA infizieren, so ist die Gesamtzahl der MRSA-Fälle in diesem Krankenhaus  $B_{n,p}$ -verteilt.

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

Möchte man Aussagen über die Größe des unbekanntes Wertes von  $p$  machen, so gibt man typischerweise Konfidenzbereiche an:

Bei gegebener Stichprobengröße  $n \in \mathbb{N}_0$  und gegebenem Niveau  $\beta \in ]0, 1[$  ist für das Schätzproblem

$$((B_{n,p}: p \in [0, 1]), p \mapsto p) \quad (4.7)$$

$K: \{0, \dots, n\} \rightarrow 2^{[0,1]}$  eine  $\beta$ -Konfidenzregion, wenn  $B_{n,p}(K \ni p) \geq \beta$  für alle  $p \in [0, 1]$ .

Zu den bekanntesten und gebräuchlichsten Konfidenzbereichen für das Schätzproblem (4.7) zählen die Konfidenz-Uprays und -Downrays (einseitige Konfidenzintervalle) und die zweiseitigen Konfidenzintervalle von Clopper und Pearson (1934), vgl. Definition 4.5.

Im Jahr 2011 wurden in dem betrachteten Krankenhaus  $n = 27976$  Patienten behandelt, von denen sich  $x = 28$  mit MRSA infizierten. Die Realisierung des Konfidenz-Downrays  $\Lambda_{CP,n}$  zum Niveau  $\beta = 0,95$  für die Infektionswahrscheinlichkeit ist also

$$\Lambda_{CP,n}(x) = 1 - K_{CP,n}(n - x) = [0, 1 - g_{27976}(27976 - 28)[ = [0, 0,00137 \dots].$$

Schaut man sich die Verteilung der MRSA-Fälle über das Jahr genauer an, so sieht man aber, dass das Verhältnis von Fällen zu Patienten über das Jahr nicht konstant ist:

Monat	Jan	Feb	Mär	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez
MRSA-Fälle	3	2	0	3	3	2	1	1	1	3	4	5
Patienten	2326	2289	2511	2295	2432	2240	2254	2320	2315	2347	2298	2349

Dies gibt Grund zu der Annahme, dass die Wahrscheinlichkeit der Infektion nicht für jeden Patienten gleich ist. Nehmen wir an, dass jeder Patient  $j \in \{1, \dots, n\}$  eine Infektionswahrscheinlichkeit  $p_j \in [0, 1]$  hat, und ist der interessierende Parameter die durchschnittliche Infektionswahrscheinlichkeit aller Patienten  $(p_1, \dots, p_n) = p \mapsto \bar{p} := \sum_{j=1}^n p_j$ , so betrachten wir das Schätzproblem

$$((BC_p: p \in [0, 1]^n), p \mapsto \bar{p}). \quad (4.8)$$

Dieses Schätzproblem ist allgemeiner als (4.7) im folgenden Sinne: Jede Binomialverteilung  $B_{n,p}$  ist auch eine Bernoulli-Faltung  $BC_{(p, \dots, p)}$ , und der Wert des interessierenden Parameters in (4.8) ist für jede Binomialverteilung  $B_{n,p}$  gleich  $p$ , so dass jede  $\beta$ -Konfidenzregion für (4.8) auch  $\beta$ -Konfidenzregion für (4.7) ist.

Für das allgemeinere Schätzproblem (4.8) ist  $K_{CP,n}$  aber kein gültiger  $\beta$ -Konfidenzbereich, was schon von Agnew (1974) festgestellt wurde und in Bemerkung 4.9 nochmal gezeigt wird.

Gültige  $\beta$ -Konfidenzbereiche werden im folgenden Satz angegeben, der auf klassischen Ergebnissen von Chebyshev und Hoeffding basiert.

**Satz 4.6.** *Seien  $n \in \mathbb{N}$  und sei  $\beta \in ]0, 1[$ . Für jedes  $m \in \{0, \dots, n\}$  sei  $K'_m$  ein  $\beta$ -Konfidenzbereich für  $((B_{m,p}: p \in [0, 1]), \text{id}_{[0,1]})$ . Dann ist ein  $\beta$ -Konfidenzbereich  $K$  für (4.8) gegeben durch*

$$K(x) := \bigcup_{\substack{l \in \{0, \dots, x\}, \\ m \in \{x-l, \dots, n-l\}}} \left( \frac{m}{n} K'_m(x-l) + \frac{l}{n} \right) \supseteq K'_n(x) \quad \text{für } x \in \{0, \dots, n\}.$$

Wenn man in Satz 4.6 als  $K'_m$  die Konfidenz-Uprays von Clopper und Pearson (1934) wie in Definition 4.5 nimmt, ist das resultierende  $K$  Buehler-optimal (vgl. Seite 58) und, wenn  $\beta$  nicht ungewöhnlich klein ist, vereinfacht sich die Formel für  $K$  stark, wie man in Satz 4.7 unten sieht. Dazu definieren wir

$$\beta_n := B_{n, \frac{1}{n}}(\{0, 1\}) \quad \text{für } n \in \mathbb{N},$$

so dass  $\beta_1 = 1$ ,  $\beta_2 = \frac{3}{4}$ ,  $\beta_3 = \frac{20}{27}$  und  $\beta_n \downarrow \frac{2}{e} = 0,735\dots$ , wobei die strenge Antitonie von  $(\beta_n)_n$  aus Jogdeo und Samuels (1968, Theorem 2.1 mit  $m_n := n, p_n := \frac{1}{n}, r := 0$ ) folgt, so dass insbesondere

$$\beta_n \leq \frac{3}{4} \quad \text{für } n \geq 2$$

ist.

**Satz 4.7.** *Seien  $n \in \mathbb{N}$ ,  $\beta \in ]0, 1[$  und  $K$  wie in Satz 4.6 mit  $K'_m := K_{\text{CP}, m}$ , wie in (4.5) definiert. Dann ist  $K$  das optimale isotone  $\beta$ -Konfidenz-Upray für (4.8), zulässig in der Menge aller  $\beta$ -Konfidenz-Uprays, strikt isoton, und hat effektives Niveau  $\inf_{p \in [0, 1]^n} \text{BC}_p(K \ni \bar{p}) = \beta$ . Es ist*

$$K(x) = \begin{cases} [0, 1], & \text{falls } x = 0, \\ \left] \frac{1-\beta}{n}, 1 \right], & \text{falls } x = 1, \\ [g_n(x), 1], & \text{falls } x \in \{2, \dots, n\} \text{ und } \beta \geq \beta_n. \end{cases} \quad (4.9)$$

*Bemerkung 4.8. Schachtelung* (engl.: *nestedness*) wird durch die Konstruktion in Satz 4.6 erhalten: Nehmen wir an, wir wenden Satz 4.6 für verschiedene  $\beta \in ]0, 1[$  an und schreiben dementsprechend  $K'_{m, \beta}$  und  $K_\beta$  statt  $K'_m$  und  $K$ . Sind nun  $\beta, \tilde{\beta} \in ]0, 1[$  so, dass  $\beta < \tilde{\beta}$  und  $K'_{m, \beta}(x) \subseteq K'_{m, \tilde{\beta}}(x)$  für alle  $m \in \{0, \dots, n\}$  und  $x \in \{0, \dots, m\}$ , so gilt nach Satz 4.6  $K_\beta(x) \subseteq K_{\tilde{\beta}}(x)$  für alle  $x \in \{0, \dots, n\}$ . Nach der zweiten Zeile in (4.4) und nach (4.5) sind die Clopper-Pearson-Uprays geschachtelt, und damit sind es auch die Uprays aus Satz 4.7. Analoge Bemerkungen gelten für die Konfidenz-Downrays aus Bemerkung 4.13 und die zweiseitigen Konfidenzintervalle aus Satz 4.17.

*Bemerkung 4.9.* Seien  $n \geq 2$  und  $\beta \in ]0, 1[$ . Wie schon in Agnew (1974) bemerkt, aber von späteren Autoren ignoriert (vgl. Bemerkung 4.15) wird, ist  $K_{\text{CP}, n}$  kein  $\beta$ -Konfidenzbereich für (4.8). Dies folgt sofort aus Satz 4.7 und der Tatsache, dass  $K_{\text{CP}, n}(1) \subsetneq K(1)$ , indem man entweder die Optimalität von  $K$  und die Isotonie von  $K_{\text{CP}, n}$  ausnutzt, oder die Zulässigkeit von  $K$  und, dass  $K_{\text{CP}, n}(x) \subseteq K(x)$  für alle  $x$ . Wenn  $\beta \geq \beta_n$ , folgt aus Satz 4.7 außerdem, dass das effektive Niveau von  $K_{\text{CP}, n}$  als Konfidenzbereich für (4.8)

$$\gamma_n := 1 - n \left(1 - \beta^{1/n}\right) \in ]1 + \log(\beta), \beta[$$

ist, denn für  $p \in [0, 1]^n$  mit  $\bar{p} \notin \left] \frac{1-\beta}{n}, g_n(1) \right]$  liefert (4.9), dass  $\text{BC}_p(K_{\text{CP}, n} \ni \bar{p}) = \text{BC}_p(K \ni \bar{p}) \geq \beta$ , und man erhält

$$\inf_{\bar{p} \in \left] \frac{1-\beta}{n}, g_n(1) \right]} \text{BC}_p(K_{\text{CP}, n} \ni \bar{p}) = \inf_{\bar{p} \in \left] \frac{1-\beta}{n}, g_n(1) \right]} \prod_{j=1}^n (1 - p_j) = 1 - n g_n(1) = \gamma_n,$$

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

wenn man im zweiten Schritt  $p$  mit  $p_1 = ng_n(1) = n(1 - \beta^{1/n}) \leq 1$  und  $p_2 = \dots = p_n = 0$  betrachtet. Da  $\gamma_n \downarrow 1 + \log(\beta) < \beta$  für  $n \rightarrow \infty$ , folgt für  $\beta > \frac{2}{e}$ , dass die  $K_{CP,n}$  noch nicht einmal asymptotische  $\beta$ -Konfidenzbereiche für (4.8) sind.

*Bemerkung 4.10.* Das einzige uns bisher bekannte  $\beta$ -Konfidenz-Upray für (4.8) wurde in Agnew (1974, section 3) bereitgestellt als  $K_A(x) := [g_A(x), 1]$ , mit  $g_A(0) := 0$  und  $g_A(x) := g_n(x) \wedge \frac{x-1}{n}$  für  $x \in \{1, \dots, n\}$ . Dieses  $K_A$  ist aber echt schlechter als das optimale isotone  $K$  aus Satz 4.7, denn  $K_A$  ist auch isoton mit  $K_A(1) = [0, 1] \not\supseteq K(1)$ . Andererseits zeigt Lemma 4.28 unten, dass faktisch  $g_A(x) = g_n(x)$  für  $\beta \geq \beta_n$  und  $x \in \{2, \dots, n\}$ , was eine präzisere Version einer unbewiesenen Behauptung in Agnew (1974) ist.

*Bemerkung 4.11.* Die Bedingung  $\beta \geq \beta_n$  in (4.9) kann nicht weggelassen werden: Für  $n \in \mathbb{N}$  sei  $A_n := \{\beta \in ]0, 1[ : \text{Wenn } K \text{ wie in Satz 4.7 ist, dann ist } K(x) = ]g_n(x), 1] \text{ für } x \in \{2, \dots, n\}\}$ . Dann ist nach Satz 4.7  $[\beta_n, 1[ \subseteq A_n$ . Numerische Untersuchungen haben ergeben, dass auch  $\beta_n - 0.001 \in A_n$  für  $2 \leq n \leq 123$ , aber  $K(2) \not\supseteq ]g_n(2), 1]$  für  $\beta = \beta_n - 0.001$  und  $124 \leq n \leq 3000$ .

*Bemerkung 4.12.* Das  $K$  aus Satz 4.7 ist als  $\beta$ -Konfidenzintervall für (4.8) ebenso wenig zulässig in der Menge *aller*  $\beta$ -Konfidenzintervalle wie  $K_{CP,n}$  als  $\beta$ -Konfidenzintervall für (4.7), denn mit  $c := (\inf K(n)) \vee (1 - (1 - \beta)^{1/n})$  und

$$K^*(x) := \begin{cases} [0, c] \subsetneq K(0), & \text{falls } x = 0, \\ K(x), & \text{falls } x \in \{1, \dots, n\}, \end{cases}$$

ergibt sich  $BC_p(K^* \ni \bar{p}) = BC_p(K \ni \bar{p}) \geq \beta$ , wenn  $\bar{p} \leq c$ , und  $BC_p(K^* \ni \bar{p}) = BC_p(\{1, \dots, n\}) = 1 - \prod_{j=1}^n (1 - p_j) \geq 1 - (1 - \bar{p})^n > 1 - (1 - c)^n \geq \beta$ , wenn  $\bar{p} > c$ .

*Bemerkung 4.13.* Da für  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$  genau dann  $K$  ein  $\beta$ -Konfidenzbereich für (4.8) ist, wenn  $\{0, \dots, n\} \ni x \mapsto 1 - K(n - x)$  eines ist, gelten analoge Aussagen zu Satz 4.7 und den Bemerkungen 4.8 – 4.12 für Downrays: Satz 4.7 bleibt gültig, wenn man  $K_{CP,m}$  durch  $\Lambda_{CP,m}$ , „Upray“ durch „Downray“ und (4.9) durch

$$K(x) = \begin{cases} [0, 1 - g_n(n - x)[, & \text{falls } x \in \{0, \dots, n - 2\} \text{ und } \beta \geq \beta_n, \\ [0, 1 - \frac{1-\beta}{n}[, & \text{falls } x = n - 1, \\ [0, 1], & \text{falls } x = n, \end{cases} \quad (4.10)$$

ersetzt.

*Bemerkung 4.14.* Für  $n \in \mathbb{N}$ ,  $x \in \{0, \dots, n\}$  und  $\text{beta} \geq 3/4$  ist nach Satz 4.7 ein R-Code zum Berechnen der unteren **beta**-Konfidenzschranke

```
(x==1)*((1-beta)/n)+
(x!=1)*binom.test(x,n,alt="g",conf.level=beta)$conf.int[1]
```

und nach Bemerkung 4.13 ist der entsprechende Code für die obere **beta**-Konfidenzschranke

```
(x==n-1)*(1-(1-beta)/n)+
(x!=n-1)*binom.test(x,n,alt="1",conf.level=beta)$conf.int[2]
```

Beispielsweise ist in Byers et al. (1979, S. 249, Zeilen 13-21)  $n = 7$  und  $x = 6$ , wodurch für  $\beta = 0,99, 0,98$  und  $0,95$  die unteren Konfidenzschranken  $0,356 \dots, 0,404 \dots$  bzw.  $0,479 \dots$  sind, so dass die in Byers et al. (1979) angegebenen Schranken tatsächlich gültig sind, aber erst jetzt in Satz 4.7 als gültig nachgewiesen, vgl. auch Bemerkung 4.15.

*Bemerkung 4.15.* Zu den Artikeln, die fälschlicherweise behaupten, dass die Clopper-Pearson-Upper- oder -Downrays  $\beta$ -Konfidenzbereiche für (4.8) sind, gehören Kappauf und Bohrer (1974, S. 652, Zeilen 3–5), Byers et al. (1979, S. 249, erste Spalte, Zeilen 15–18), und Cheng et al. (2010, S. 7, Zeilen 10–8 von unten). Die entsprechende Behauptung von Ollero und Ramos (1995, S. 247, Zeilen 9–12) für eine gewisse Teilfamilie von  $(BC_p: p \in [0, 1]^n)$ , die die hypergeometrischen Verteilungen mit Stichprobengrößenparameter  $n$  enthält, wird in Bemerkung 4.21 unten widerlegt. Die gemeinsame Fehlerquelle in diesen Artikeln scheint eine missverständliche Bemerkung in Hoeffding (1956, S. 720, erster Absatz von Kapitel 5) darüber zu sein, dass, nach Hoeffding (1956, Theorem 4) oder David (1960), gewisse Tests für  $\Theta_0 \subseteq [0, 1]$  im Binomialmodell  $(B_{n,p}: p \in [0, 1])$  ihr Niveau als Tests für  $\tilde{\Theta}_0 := \{p \in [0, 1]^n: \bar{p} \in \Theta_0\}$  im Modell  $(BC_p: p \in [0, 1]^n)$  halten (vgl. Bemerkung 4.16). Es sei außerdem angemerkt, dass das Teilresultat in Ollero und Ramos (1995) bezüglich der Darstellbarkeit von hypergeometrischen Verteilungen als Bernoulli-Faltungen schon aus Vatutin und Mikhailov (1983) bekannt war.

*Bemerkung 4.16.* Der Kern der missverständlichen Bemerkung in Hoeffding (1956), die in Bemerkung 4.15 erwähnt wird, ist: „that the usual (one-sided and two-sided) tests for the constant probability of ‘success’ in  $n$  independent (Bernoulli) trials can be used as tests for the average probability of success when the probability of success varies from trial to trial.“ Wir spezifizieren und verallgemeinern diese Aussage wie folgt: Seien  $n \in \mathbb{N}$ ,  $p_1 \leq p_2 \in [0, 1]$ ,  $\gamma_-, \gamma_+ \in [0, 1]$ ,  $c_- \leq \lfloor np_1 \rfloor - 1$  und  $c_+ \geq \lceil np_2 \rceil + 1$ . Dann hält die kritische Funktion

$$\psi := \mathbb{1}_{\{0, \dots, c_- - 1\}} + \gamma_- \mathbb{1}_{\{c_-\}} + \gamma_+ \mathbb{1}_{\{c_+\}} + \mathbb{1}_{\{c_+ + 1, \dots, n\}}$$

für die Hypothese  $[p_1, p_2]$  im Binomialmodell  $(B_{n,p}: p \in [0, 1])$  ihr Niveau als kritische Funktion für  $\{p \in [0, 1]^n: \bar{p} \in [p_1, p_2]\}$  im Modell  $(BC_p: p \in [0, 1]^n)$ , denn für jedes  $p$  mit  $\bar{p} \in [p_1, p_2]$  folgt aus Hoeffding (1956, Theorem 4)

$$\begin{aligned} BC_p \psi &= \gamma_- BC_p(\{0, \dots, c_- \}) + (1 - \gamma_-) BC_p(\{0, \dots, c_- - 1\}) \\ &\quad + \gamma_+ BC_p(\{c_+, \dots, n\}) + (1 - \gamma_+) BC_p(\{c_+ + 1, \dots, n\}) \\ &\leq B_{n, \bar{p}} \psi. \end{aligned}$$

Diese Aussage ist aber nicht immer auf die einseitigen, auf den Clopper-Pearson-Upper- basierenden Tests anwendbar:

Seien  $n = 2$  und  $\beta \in ]0, 1[$ . Seien  $r \in [0, 1]$ ,  $H := [0, r]$  und  $\psi := \mathbb{1}_{\{K_{CP,n} \cap H = \emptyset\}}$ , so dass  $\sup_{p \in H} B_{n,p} \psi \leq 1 - \beta$ . Wenn aber z. B.  $r = 1 - \sqrt{\beta}$  ist, vereinfacht sich der Test zu  $\psi = \mathbb{1}_{\{1,2\}}$ , und für  $p := (r - \varepsilon, r + \varepsilon)$  für ein genügend kleines  $\varepsilon > 0$  ist  $\bar{p} \in H$  und  $BC_p \psi = 1 - BC_p(\{0\}) = 1 - \beta + \varepsilon^2 > 1 - \beta$ .

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

Im Gegensatz zu Bemerkung 4.9 haben wir insbesondere für  $\beta \geq 1/2$  das folgende positive Resultat für die zweiseitigen Clopper-Pearson- $\beta$ -Konfidenzintervalle  $M_{\text{CP},n}$ , die in (4.6) definiert werden, als gültige Intervalle für (4.8).

**Satz 4.17.** *Seien  $n \in \mathbb{N}$  und  $\beta \geq 2\beta_n - 1$  oder  $n = 1$ . Dann ist  $M_{\text{CP},n}$  aus (4.6) ein  $\beta$ -Konfidenzintervall für (4.8).*

*Bemerkung 4.18.* In der Situation von Satz 4.17 ist  $M_{\text{CP},n}$  eine Verbesserung des zweiseitigen Intervalls für (4.8) von Agnew (1974), das analog zu (4.6) aus seinen einseitigen Intervallen gebildet wird.

*Bemerkung 4.19.* Im Gegensatz zu Bemerkung 4.11 wissen wir nicht, ob die Bedingung „ $\beta \geq 2\beta_n - 1$  oder  $n = 1$ “ in Satz 4.17 weggelassen werden kann.

*Bemerkung 4.20.* Die Robustheit der zweiseitigen Clopper-Pearson-Intervalle aus Satz 4.17 beim Übergang zu (4.8) gilt nicht für alle zweiseitigen Intervalle für (4.7), z. B. nicht für das auf Sterne (1954) basierende  $\beta$ -Konfidenzintervall  $K_{\text{S},n}$  für (4.7) von Dümbgen (2004, S. 5):

Für  $n \in \mathbb{N}$  und  $\beta \in ]0, 1[$  sei die  $\beta$ -Konfidenzregion  $C_n^{\text{S}}$  gegeben durch

$$C_n^{\text{S}}(x) := C_{n,\beta}^{\text{S}}(x) := \{p \in [0, 1] : B_{n,p}(\{k : B_{n,p}(\{k\}) \leq B_{n,p}(\{x\})\}) > 1 - \beta\},$$

welche nicht immer ein Intervall bildet. Letzteres wird durch eine numerische Überprüfung der Beobachtung aus Dümbgen (2004, Figure 2) verifiziert: Es ist  $\{0,4, 0,425\} \subseteq C_{20,0,8269}^{\text{S}}(5) \not\subseteq \{0,424\}$ . Es ist aber durch

$$K_{\text{S},n}(x) := [\inf(C_n^{\text{S}}(x)), \sup(C_n^{\text{S}}(x))] \quad \text{für alle } x \in \{0, \dots, n\}$$

das  $\beta$ -Konfidenzintervall  $K_{\text{S},n}$  für (4.7) definiert. Wenn z. B.  $n = 2$  und  $\beta > 3/4$ , ist speziell

$$K_{\text{S},2}(0) = [0, 1 - g_2(2)[, \quad K_{\text{S},2}(1) = ]g_2(1), 1 - g_2(1)[ \quad \text{und} \quad K_{\text{S},2}(2) = ]g_2(2), 1],$$

und  $K_{\text{S},2}$  ist für (4.8) nicht gültig, da für  $p \in [0, 1]^2$  mit  $\bar{p} = g_2(1)$  und  $p_1 \neq p_2$  gilt, dass

$$\text{BC}_p(K_{\text{S},2} \ni \bar{p}) = \text{BC}_p(\{0\}) = \prod_{j=1}^2 (1 - p_j) < (1 - \bar{p})^2 = (1 - g_2(1))^2 = \beta.$$

Für  $n = 2$  und  $\beta > 3/4$  erhalten wir mit Satz 4.6 ein  $\beta$ -Konfidenzintervall für (4.8), das wir  $\tilde{K}$  nennen, indem wir dort  $K'_m := K_{\text{S},m}$  für  $m \in \{0, 1, 2\}$  setzen, und es ist dann

$$\tilde{K}(0) = [0, 1 - (1 - \beta)^{1/2}[, \quad \tilde{K}(1) = ]\frac{1-\beta}{2}, \frac{1+\beta}{2}[ \quad \text{und} \quad \tilde{K}(2) = ](1 - \beta)^{1/2}, 1].$$

Man sieht, dass  $\tilde{K}(x) \subsetneq M_{\text{CP},2}(x)$  für  $x \in \{0, 1, 2\}$ , wobei  $M_{\text{CP},2}$  wie in (4.6) definiert ist. Wir wissen nicht, ob diese Inklusion für alle üblichen  $n$  und  $\beta$  gilt, aber es scheint nicht einmal im Allgemeinen  $K_{\text{S},n}(x) \subseteq M_{\text{CP},n}(x)$  zu gelten (vgl. Dümbgen (2004, Figure 3,  $x = 2$ )).

Im Fall  $n = 2$  und  $\beta > 3/4$  ist  $\tilde{K}$  auch echt besser als  $K$  aus Satz 4.7, denn  $\tilde{K}(x) \subsetneq K(x)$  für  $x \in \{0, 1\}$  und  $\tilde{K}(2) = K(2)$ . Dies widerspricht aber nicht der Optimalität von  $K$ , denn  $\tilde{K}$  ist kein Upray.

### 4.3 Spezialisierung auf hypergeometrische Verteilungen

In Vatutin und Mikhailov (1983, Corollary 5) wird gezeigt, dass jede hypergeometrische Verteilung eine Bernoulli-Faltung ist, also können wir die obigen Ergebnisse auf hypergeometrische Verteilungen anwenden und spezialisieren.

Man könnte vermuten, dass die im Binomialmodell gültigen Clopper-Pearson-Uprays, wenn schon nicht im Bernoulli-Faltungs-Schätzproblem (4.8), so doch in den hypergeometrischen Schätzproblemen (4.11) oder (4.12) unten für die Erfolgswahrscheinlichkeit bei bekannter bzw. unbekannter Urnengröße gültig sind, was aber in der folgenden Bemerkung widerlegt wird, deren Beweis man in Abschnitt 4.5 findet.

*Bemerkung 4.21.* Für  $N \in \mathbb{N}_0$ ,  $n \in \{0, \dots, N\}$  und  $\beta \in ]0, 1[$  ist  $K_{\text{CP},n}$  im Allgemeinen kein  $\beta$ -Konfidenzbereich für das Schätzproblem

$$\left( \left( H_{n,p,N} : p \in \left\{ \frac{j}{N} : j \in \{0, \dots, N\} \right\} \right), p \mapsto p \right). \quad (4.11)$$

Für  $\beta \in ]0, 1[$ ,  $n \in \mathbb{N}_0$  und unbekanntes  $\mathbb{N}_0 \ni N \geq n$  ist im Allgemeinen  $K_{\text{CP},n}$  auch nicht gültig für das allgemeinere Schätzproblem

$$\left( \left( H_{n,p,N} : n \leq N \in \mathbb{N}_0, p \in \left\{ \frac{j}{N} : j \in \{0, \dots, N\} \right\} \right), (N, p) \mapsto p \right). \quad (4.12)$$

Satz 4.22 unten liefert aber das folgende positive Resultat: Wenn z. B.  $\beta \geq 3/4$ , ist  $K^*$  aus (4.14) das optimale isotone Konfidenz-Upray für das Schätzproblem (4.12). Falls  $n \geq 3$  und  $\beta \geq \beta_n$ , gilt für  $K^*$  die Darstellung in (4.15):

$$K^*(x) = \begin{cases} K_{\text{CP},n}(x), & \text{falls } x \in \{0, \dots, n\} \setminus \{1\}, \\ \left[ \frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1}, 1 \right], & \text{falls } x = 1. \end{cases}$$

Um dies zu beweisen, definieren wir zunächst die Hilfsfunktion  $f$ , deren Wohldefiniertheit und weitere Eigenschaften in Lemma 4.29 in Abschnitt 4.5.2 unten gezeigt werden. Die Funktion  $f$  weist die folgende Analogie zu den Clopper-Pearson-Schranken auf: wenn  $n \in \mathbb{N}$  und  $f(m) = \frac{m}{N}$  für ein  $N \geq n$ , ist  $H_{n,f(m),N}(\{0\}) = \beta$ .

Für  $n \in \mathbb{N}$  und  $\beta \in ]0, 1[$  sei

$$f: \mathbb{N} \rightarrow ]0, 1[, \quad f(m) := \text{das } p \in \left] 0, \frac{m}{m+n-1} \right[ \text{ mit } \prod_{j=0}^{n-1} \frac{m(1-p) - pj}{m - pj} = \beta. \quad (4.13)$$

Für  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$  definieren wir damit  $K^*$  durch

$$K^*(x) := \begin{cases} K(x), & \text{falls } x \in \{0, \dots, n\} \setminus \{1\}, \\ \bigcup_{m=1}^{\infty} \left[ \frac{m}{\lceil \frac{m}{f(m)} \rceil - 1}, 1 \right], & \text{falls } x = 1, \end{cases} \quad (4.14)$$

wobei  $K$  wie in Satz 4.7 gegeben ist, und formulieren das folgende Ergebnis:

**Satz 4.22.** Seien  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$ . Für das Schätzproblem (4.12) ist  $K$  aus Satz 4.6 ein  $\beta$ -Konfidenzbereich.

Weiter ist  $K^*$  aus (4.14) ein  $\beta$ -Konfidenz-Upray für (4.12). Falls  $n = 1$  oder  $\beta \geq \beta_n$ , ist  $K^*$  streng isoton, das optimale isotone  $\beta$ -Konfidenz-Upray für (4.12) und zulässig in der Menge aller  $\beta$ -Konfidenz-Uprays für (4.12).

Falls  $n \geq 3$  und  $\beta \geq \beta_n$ , gilt für  $K^*$

$$K^*(x) = \begin{cases} K_{\text{CP},n}(x), & \text{falls } x \in \{0, \dots, n\} \setminus \{1\}, \\ \left[ \frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1}, 1 \right], & \text{falls } x = 1. \end{cases} \quad (4.15)$$

*Bemerkung 4.23.* In der Situation von Satz 4.22 gelten die folgenden Abschätzungen für  $K^*$  aus (4.14) und  $K$  aus Satz 4.7: Es ist

$$K^*(x) \subseteq K(x) \quad \text{für alle } x \in \{0, \dots, n\}, \quad (4.16)$$

und, falls  $n \leq 1$ , ist  $K^* = K$ , falls  $n = 2$ , ist

$$K^*(1) \subseteq \left[ \frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1}, 1 \right] \cup \left[ \frac{\frac{1-\beta}{n} + 2 - \sqrt{(\frac{1-\beta}{n})^2 + 4\beta}}{n+1}, 1 \right] \subsetneq \left] \frac{1-\beta}{n}, 1 \right] = K(1), \quad (4.17)$$

und falls  $n \geq 3$ , ist

$$K^*(1) = \left[ \frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1}, 1 \right] \subsetneq \left] \frac{1-\beta}{n}, 1 \right] = K(1). \quad (4.18)$$

*Bemerkung 4.24.* Seien  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$ . In Analogie zu den Clopper-Pearson-Uprays  $K_{\text{CP},n}$  für (4.7) werden die Uprays  $K_{\text{H},n,N}$  für das hypergeometrische Schätzproblem (4.11) bei bekanntem  $\mathbb{N}_0 \ni N \geq n$  definiert:

Für  $x \in \{0, \dots, n\}$  sei  $K_{\text{H},n,N}(x) := [g_{\text{H},n,N}(x), 1]$ , wobei

$$g_{\text{H},n,N}(x) := \min \left\{ \frac{j}{N} : j \in \{0, \dots, N\}, H_{n, \frac{j}{N}, N}(\{x, \dots, n\}) > 1 - \beta \right\}. \quad (4.19)$$

Diese finden sich z. B. in Konijn (1973). In Buonaccorsi (1987) wird aber bemerkt, dass sie schon in Katz (1953) erwähnt werden. Nach Wang (2014, Lemma 8) ist  $K_{\text{H},n,N}$  das optimale isotone  $\beta$ -Konfidenz-Upray für (4.11).

Für  $N \geq n$  besteht zwischen  $K^*$  aus (4.14) und  $K_{\text{H},n,N}$  der folgende Zusammenhang: Da  $K^*$  das optimale isotone  $\beta$ -Konfidenz-Upray für (4.12) ist, ist es auch ein isotones  $\beta$ -Konfidenz-Upray für (4.11). Da  $K_{\text{H},n,N}$  das optimale isotone  $\beta$ -Konfidenz-Upray für dieses Schätzproblem ist, muss also

$$K_{\text{H},n,N}(x) \subseteq K^*(x) \quad \text{für alle } x \in \{0, \dots, n\} \quad (4.20)$$

gelten. Da  $K^*$  optimal isoton für (4.12) ist, kann aber  $K_{\text{H},n,N}$  nicht echt besser als  $K^*$  für (4.12) sein.

### 4.3 Spezialisierung auf hypergeometrische Verteilungen

*Bemerkung 4.25.* In Anwendungen kann es vorkommen, dass die wahre Populationsgröße  $N$  nicht bekannt ist, dass aber eine obere Grenze  $N_0$  für  $N$  bekannt ist. Diese Information kann man nutzen, und den Parameterraum in (4.12) verkleinern:

Seien  $n \in \mathbb{N}_0$ ,  $\beta \geq \beta_n$  und  $\mathbb{N}_0 \ni N_0 \geq n$ . Dann sind für das Schätzproblem

$$\left( \left( \mathbb{H}_{n,p,N} : n \leq N \leq N_0, p \in \left\{ \frac{j}{N} : j \in \{0, \dots, N\} \right\} \right) \right), (N, p) \mapsto p \quad (4.21)$$

sowohl  $K^*$  als auch  $\Lambda_{N_0}$ , definiert durch

$$\Lambda_{N_0}(x) := \bigcup_{N=n}^{N_0} K_{\mathbb{H},n,N}(x) \quad \text{für } x \in \{0, \dots, n\},$$

$\beta$ -Konfidenz-Uprays. Man sieht, dass für  $n \leq N'_0 \leq N''_0$

$$\Lambda_n(x) = \left[ \frac{x}{n}, 1 \right] \subseteq \Lambda_{N'_0}(x) \subseteq \Lambda_{N''_0}(x) \subseteq K^*(x) \quad \text{für alle } x \in \{0, \dots, n\}, \quad (4.22)$$

wobei die letzte Inklusion aus (4.20) folgt. Bei bekanntem  $N_0$  ist also  $\Lambda_{N_0}$  besser als  $K^*$  für (4.21). Zur Beantwortung der Frage, wie groß hier in der Anwendung der vermeidbare Informationsverlust durch Benutzung des in der Situation von (4.15) leichter zu implementierenden Uprays  $K^*$  ist, wurden numerische Untersuchungen zu (4.22) durchgeführt, deren Ergebnisse für die folgenden beiden Beispiele grafisch dargestellt sind:

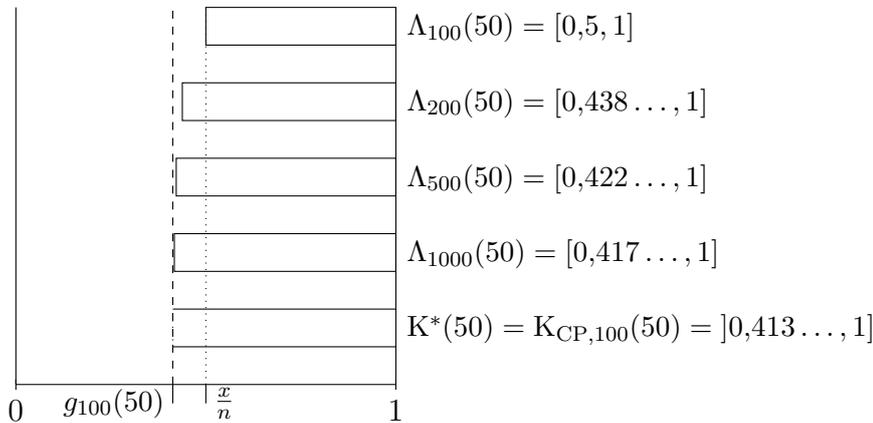


Abbildung 4.1: Vergleich von  $\Lambda_{N_0}$  und  $K^*$  für  $\beta = 0,95, n = 100, x = 50$  und  $N_0 \in \{100, 200, 500, 1000\}$ .

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

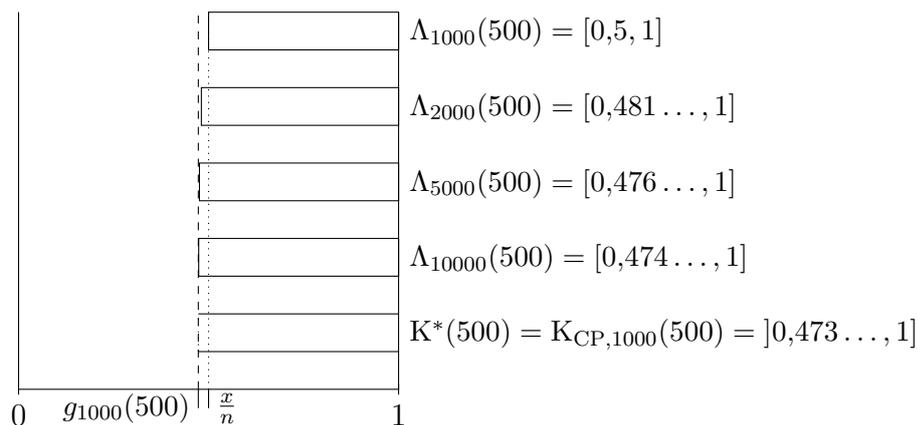


Abbildung 4.2: Vergleich von  $\Lambda_{N_0}$  und  $K^*$  für  $\beta = 0,95$ ,  $n = 1000$ ,  $x = 500$  und  $N_0 \in \{1000, 2000, 5000, 10000\}$ .

Die Abbildungen 4.1 und 4.2 skizzieren für  $\beta = 0,95$  und  $(n, x) = (100, 50)$  bzw.  $(n, x) = (1000, 500)$  die Größe der Intervalle  $K^*(x)$  und  $\Lambda_{N_0}(x)$  für verschiedene  $N_0$ . In beiden Fällen ist  $x/n = 1/2$ .

In Abbildung 4.1 sieht man, dass die Größen der Intervalle relativ stark variieren, so dass die Ausnutzung einer oberen Grenze  $N_0$  einen relativ starken Effekt hat. Im Gegensatz dazu sieht man in Abbildung 4.2, dass die Größen der Intervalle weniger stark variieren, so dass die Ausnutzung einer oberen Grenze  $N_0$  hier einen weniger starken Effekt hat. In beiden Fällen hat eine Ausnutzung der oberen Schranke  $N_0$  nur einen starken Effekt, wenn  $N_0$  nicht bedeutend größer als  $n$  ist.

Diese Beobachtungen sollten in einer folgenden Untersuchung konkretisiert werden, indem man Schranken für den Fehler angibt, den man begeht, wenn man  $K^*$  als Konfidenz-Upperay für (4.21) nutzt.

*Bemerkung 4.26.* Da  $K^*$  genau dann ein  $\beta$ -Konfidenzbereich für (4.12) ist, wenn  $\{0, \dots, n\} \ni x \mapsto 1 - K^*(n - x)$  eines ist, gelten auch in diesem Abschnitt 4.3 analoge Aussagen für Downrays.

## 4.4 Offene Fragen

Für das Schätzproblem (4.8) der Schätzung des arithmetischen Mittels der Erfolgswahrscheinlichkeiten im Bernoulli-Faltungs-Modell sind die optimalen isotonen Up- und Downrays in Satz 4.7 bzw. Bemerkung 4.13 gegeben.

Neben der Frage, wie geeignete Up- oder Downrays für die Schätzung anderer interessierender Parameter aussehen könnten, stellt sich die Frage, welche zweiseitigen Konfidenzintervalle durch Anwendung von Satz 4.6 geeignete Konfidenzregionen für (4.8) ergeben. Beispielfhaft wurden in Satz 4.17 und Bemerkung 4.20 die Verallgemeinerungen der Clopper-Pearson- und Sterne-Intervalle betrachtet, und im Spezialfall  $n = 2$ ,  $\beta > 3/4$  gezeigt, dass die Sterne-Verallgemeinerung echt besser als das Clopper-Pearson-Intervall

für (4.8) ist. Eine Verallgemeinerung der isotonen Optimalität auf zweiseitige Konfidenzintervalle ist nicht sinnvoll, aber man könnte zumindest systematisch untersuchen, ob gewisse Intervalle in bestimmten Fällen besser sind als andere.

In Bemerkung 4.25 sieht man, dass  $K_{\text{CP},1000}(500) \subsetneq K_{\text{CP},100}(50)$  für  $\beta = 0,95$ , wobei in beiden Fällen  $x/n = 1/2$ . Erste numerische Untersuchungen geben Grund zu der Annahme, dass allgemeiner für  $\beta \in ]0, 1[$ ,  $n \in \mathbb{N}$ ,  $x \in \{1, \dots, n\}$  und  $j \in ]1, \infty[$  mit  $jn \in \mathbb{N}$  und  $jx \in \{x + 1, \dots, jn\}$  gilt, dass  $K_{\text{CP},jn}(jx) \subsetneq K_{\text{CP},n}(x)$ . Diese Beobachtung sollte analytisch verifiziert werden.

Für die hypergeometrischen Schätzmodelle (4.11) und (4.12) sind die für eine normale Wahl von  $n$  und  $\beta$  optimalen Uprays in Bemerkung 4.24 bzw. Satz 4.22 gegeben. Es sollte eine analytische Untersuchung der in Bemerkung 4.25 erwähnten Fehlerschranken bei Nutzung der unnötig großen Konfidenzintervalle im Modell (4.21) erfolgen.

## 4.5 Beweise der Ergebnisse

### 4.5.1 Beweise des Abschnitts 4.2

*Beweis von Lemma 4.3.* Wir nehmen an, es gibt ein  $\beta$ -Konfidenz-Upray  $K^*$  für  $(\mathcal{P}, \kappa)$  mit  $K^*(x) \subseteq K(x)$  für jedes  $x \in \{0, \dots, n\}$  und  $K^*(x') \subsetneq K(x')$  für ein  $x'$ . Dann definiert, da  $K$  strikt isoton ist,

$$K^{**}(x) := \left\{ \begin{array}{ll} K(x), & \text{falls } x \neq x' \\ K^*(x') \cup K(x' + 1), & \text{falls } x = x' < n, \\ K^*(x'), & \text{falls } x = x' = n, \end{array} \right\} \supseteq K^*(x)$$

ein isotones  $\beta$ -Konfidenz-Upray für  $(\mathcal{P}, \kappa)$  mit  $K^{**}(x') \subsetneq K(x')$ , was der Optimalität von  $K$  widerspricht.  $\square$

*Beweis von Definition und Bemerkung 4.5.* Es sind nur die Optimalität und Zulässigkeit von  $K_{\text{CP},n}$  zu zeigen. Seien  $n \in \mathbb{N}_0$  und  $\beta \in ]0, 1[$ . Um die Optimalität von  $K_{\text{CP},n}$  zu zeigen, nehmen wir an, dass  $\tilde{K}: \{0, \dots, n\} \rightarrow 2^{[0,1]}$  ein weiteres isotones Upray ist und dass es ein  $x' \in \{0, \dots, n\}$  gibt mit

$$\tilde{K}(x') \subsetneq K_{\text{CP},n}(x'). \quad (4.23)$$

Wir müssen zeigen, dass  $\inf_{p \in [0,1]} B_{n,p}(\tilde{K} \ni \bar{p}) < \beta$ . Wenn  $x' = 0$ , ist  $K_{\text{CP},n}(x') = [0, 1]$  und, da  $\tilde{K}(0)$  ein Upray in  $[0, 1]$  ist, folgt aus (4.23), dass  $0 \notin \tilde{K}(0)$ . Also ist dann

$$\inf_{p \in [0,1]} B_{n,p}(\tilde{K} \ni p) \leq \delta_0(\tilde{K} \ni 0) = 0 < \beta.$$

Wenn  $x' \in \{1, \dots, n\}$ , existiert nach (4.23) und wegen der Isotonie von  $\tilde{K}$  ein  $a \in ]g_n(x'), 1]$  mit  $a \notin \tilde{K}(y)$  für alle  $y \in \{x', \dots, n\}$ . Damit ist

$$\begin{aligned} \inf_{p \in [0,1]} B_{n,p}(\tilde{K} \ni p) &\leq B_{n,a}(\{0, \dots, x' - 1\}) \\ &< B_{n,g_n(x')}(\{0, \dots, x' - 1\}) \\ &= \beta, \end{aligned}$$

wobei in der zweiten Ungleichung (4.2) verwendet wurde. Die Zulässigkeit folgt nach Lemma 4.3.  $\square$

*Beweis von Satz 4.6.* Man kann leicht zeigen, dass  $K(x) \subseteq [0, 1]$  und, wenn man  $l = 0$  und  $m = n$  berücksichtigt,  $K(x) \supseteq K'_n(x)$  für alle  $x \in \{0, \dots, n\}$ . Wenn  $\varphi: \{0, \dots, n\} \rightarrow \mathbb{R}$  eine Funktion und  $\pi \in [0, 1]$  ist, wird nach Hoeffdings (1956, Corollary 2.1) Verallgemeinerung von Tchebichef (1846, zweites Théorème) das Minimum des Erwartungswertes  $BC_p \varphi$  als Funktion von  $p \in [0, 1]^n$  unter der Nebenbedingung, dass  $\bar{p} = \pi$  ist, an einem Punkt  $p$  angenommen, dessen Koordinaten höchstens drei verschiedene Werte annehmen, von denen höchstens einer verschieden von 0 und 1 ist. Für gegebenes  $p \in [0, 1]^n$  liefert der vorangehende Satz, angewendet auf  $\pi := \bar{p}$  und den Indikator  $\varphi$  von  $\{K \ni \pi\}$ , dass  $r, s \in \{0, \dots, n\}$  mit  $r + s \leq n$  und  $a \in [0, 1]$  so existieren, dass  $r + sa = n\pi$  und

$$\begin{aligned} BC_p(K \ni \bar{p}) &\geq (\delta_r * B_{s,a}) (\{x \in \{r, \dots, r+s\} : K(x) \ni \pi\}) \\ &\geq (\delta_r * B_{s,a}) \left( \{x \in \{r, \dots, r+s\} : \frac{s}{n} K'_s(x-r) + \frac{r}{n} \ni \pi \right) \\ &= B_{s,a} (K'_s \ni a) \\ &\geq \beta, \end{aligned}$$

wobei im zweiten Schritt die  $K(x)$  definierende Vereinigung auf die Menge mit dem Index  $(l, m) = (r, s)$  eingeschränkt wurde.  $\square$

Um Satz 4.7 beweisen zu können, beweisen wir Lemma 4.28 und vorbereitend Lemma 4.27. Es bezeichnen  $F_{n,p}$  und  $f_{n,p}$  die Verteilungsfunktion bzw. Dichte der Binomialverteilung  $B_{n,p}$ .

**Lemma 4.27.** *Sei  $n \in \mathbb{N}$ . Dann ist*

$$F_{n, \frac{x}{n}}(x) < F_{n, \frac{1}{n}}(1) \quad \text{für alle } x \in \{2, \dots, n-1\}. \quad (4.24)$$

*Beweis.* Wenn  $x \in \mathbb{N}$  und  $x \leq \frac{n-1}{2}$ , dann gilt für  $p \in ]\frac{x}{n}, \frac{x+1}{n}[$ , dass  $y := x+1 - np > 0$ , und damit

$$\frac{f_{n-1,p}(x)}{f_{n, \frac{x+1}{n}}(x+1)} = \frac{f_{n-1,p}(x)}{f_{n-1, \frac{x+1}{n}}(x)} = \frac{\left(1 + \frac{y}{n-x-1}\right)^{n-x-1}}{\left(1 + \frac{y}{np}\right)^x} > \frac{\left(1 + \frac{y}{n-x-1}\right)^{n-x-1}}{\left(1 + \frac{y}{x}\right)^x} \geq 1,$$

wobei im letzten Schritt die Isotonie von  $]0, \infty[ \ni t \mapsto \left(1 + \frac{y}{t}\right)^t$  ausgenutzt wird, und wir erhalten

$$F_{n, \frac{x}{n}}(x) - F_{n, \frac{x+1}{n}}(x+1) = n \int_{\frac{x}{n}}^{\frac{x+1}{n}} f_{n-1,p}(x) dp - f_{n, \frac{x+1}{n}}(x+1) > 0;$$

also gilt (4.24) unter der Einschränkung  $x \leq \frac{n+1}{2}$ . Wenn nun  $x \in \mathbb{N}$  und  $\frac{n+1}{2} \leq x \leq n-1$ , dann ist  $1 \leq k := n-x < \frac{n}{2}$ , und eine Ungleichung, die Jogdeo und Samuels (1968, Corollary 4.2) Simmons zuschreiben, liefert  $F_{n, \frac{k}{n}}(k-1) > 1 - F_{n, \frac{k}{n}}(k)$ , so dass

$$F_{n, \frac{x}{n}}(x) = 1 - F_{n, \frac{k}{n}}(k-1) < F_{n, \frac{k}{n}}(k) \leq F_{n, \frac{1}{n}}(1),$$

wobei im letzten Schritt (4.24) in dem Fall benutzt wurde, der im vorangegangenen Satz bewiesen wurde.  $\square$

**Lemma 4.28.** *Seien  $n \in \mathbb{N}$ ,  $\beta \in [\beta_n, 1[$  und  $x \in \{2, \dots, n\}$ . Dann ist  $g_n(x) \leq \frac{x-1}{n}$ .*

*Beweis.* Mit Lemma 4.27 erhalten wir  $F_{n, \frac{x-1}{n}}(x-1) \leq F_{n, \frac{1}{n}}(1) = \beta_n \leq \beta = F_{n, g_n(x)}(x-1)$ , und damit die Behauptung.  $\square$

Damit können wir Satz 4.7 beweisen:

*Beweis von Satz 4.7.* Um die definierende Darstellung von  $K$  zu vereinfachen, setzen wir

$$g(x) := \min_{\substack{l \in \{0, \dots, x-1\}, \\ m \in \{x-l, \dots, n-l\}}} \left( \frac{m}{n} g_m(x-l) + \frac{l}{n} \right) \quad \text{für } x \in \{1, \dots, n\}. \quad (4.25)$$

Für  $x \in \{0, \dots, n\}$  erhalten wir mit (4.5)

$$K(x) \supseteq \frac{n-x}{n} K_{\text{CP}, n-x}(x-x) + \frac{x}{n} = \left[ \frac{x}{n}, 1 \right],$$

also insbesondere  $K(0) = [0, 1]$ . Für  $x \in \{1, \dots, n\}$  ist, mit einem Paar  $(l, m)$ , in dem das Minimum in (4.25) angenommen wird,

$$K(x) \supseteq \frac{m}{n} K_{\text{CP}, m}(x-l) + \frac{l}{n} = \left] g(x), \frac{l+m}{n} \right] \supseteq \left] g(x), \frac{x}{n} \right]$$

und, unter Ausnutzung von  $g_n(x) < 1$  im dritten Schritt unten,

$$\begin{aligned} K(x) \setminus \left] g(x), 1 \right] &\subseteq \bigcup_{m \in \{0, \dots, n-x\}} \left( \frac{m}{n} K_{\text{CP}, m}(x-x) + \frac{x}{n} \right) \subseteq \left[ \frac{x}{n}, 1 \right] \\ &\subseteq \left[ \frac{x}{n} g_x(x-0) + \frac{0}{n}, 1 \right] \subseteq \left] g(x), 1 \right]. \end{aligned}$$

Diese Inklusionen zusammen ergeben

$$K(x) = \begin{cases} [0, 1], & \text{falls } x = 0, \\ \left] g(x), 1 \right], & \text{falls } x \in \{1, \dots, n\}, \end{cases} \quad (4.26)$$

also ist  $K$  tatsächlich ein Upray und die Behauptung (4.9) gilt in ihrem trivialen ersten Fall. Mit (4.3) und der Isotonie von  $t \mapsto (\beta^t - 1)/t$ , die aus der Konvexität von  $t \mapsto \beta^t$  folgt, erhalten wir

$$g(1) = \min_{m=1, \dots, n} \frac{m}{n} g_m(1) = \frac{1}{n} \min_{m=1, \dots, n} m \left( 1 - \beta^{1/m} \right) = \frac{1-\beta}{n}$$

und so den zweiten Fall von (4.9). Der letzte Fall wird am Ende dieses Beweises nachgewiesen.

Um die strikte Isotonie von  $K$  zu zeigen, benutzen wir, dass  $g_m(x-1) < g_m(x)$  für alle  $2 \leq x \leq m \leq n$ , und erhalten für alle  $x \in \{2, \dots, n\}$

$$\begin{aligned} g(x) &= \min_{m \in \{x, \dots, n\}} \frac{m}{n} g_m(x) \wedge \min_{\substack{l \in \{1, \dots, x-1\}, \\ m \in \{x-(l-1)-1, \dots, n-(l-1)-1\}}} \left( \frac{m}{n} g_m(x-1-(l-1)) + \frac{l-1}{n} + \frac{1}{n} \right) \\ &> \min_{m \in \{x-1, \dots, n\}} \frac{m}{n} g_m(x-1) \wedge \min_{\substack{l \in \{0, \dots, x-1-1\}, \\ m \in \{x-1-l, \dots, n-1-l\}}} \left( \frac{m}{n} g_m(x-1-l) + \frac{l}{n} \right) \\ &\geq g(x-1). \end{aligned}$$

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

Wenn man unten im ersten Schritt  $p = (1 - \beta, 0, \dots, 0) \in [0, 1]^n$  betrachtet und im zweiten Schritt  $K(1) = ]\frac{1-\beta}{n}, 1] \not\supseteq \frac{1-\beta}{n}$  und die Isotonie von  $K$  ausnutzt, erhält man

$$\inf_{p \in [0, 1]^n} \text{BC}_p(K \ni \bar{p}) \leq B_{1-\beta} \left( K \ni \frac{1-\beta}{n} \right) = B_{1-\beta}(\{0\}) = \beta$$

und so mit Satz 4.6, dass  $\inf_{p \in [0, 1]^n} \text{BC}_p(K \ni \bar{p}) = \beta$ .

Um die Optimalität von  $K$  zu zeigen, nehmen wir an, dass  $\tilde{K}: \{0, \dots, n\} \rightarrow 2^{[0, 1]}$  ein weiteres isotones Upray ist und dass es ein  $x' \in \{0, \dots, n\}$  gibt mit

$$\tilde{K}(x') \subsetneq K(x'). \quad (4.27)$$

Wir müssen zeigen, dass  $\inf_{p \in [0, 1]^n} \text{BC}_p(\tilde{K} \ni \bar{p}) < \beta$ . Wenn  $x' = 0$ , ist  $K(x') = [0, 1]$  und, da  $\tilde{K}(0)$  ein Upray in  $[0, 1]$  ist, folgt aus (4.27), dass  $0 \notin \tilde{K}(0)$ , und so

$$\inf_{p \in [0, 1]^n} \text{BC}_p(\tilde{K} \ni \bar{p}) \leq \delta_0(\tilde{K} \ni 0) = 0 < \beta.$$

Wenn  $x' \in \{1, \dots, n\}$ , erhalten wir mit (4.26) und (4.25), dass  $K(x') = ]\frac{m}{n}g_m(x'-l) + \frac{l}{n}, 1]$  für ein Paar  $(l, m)$  mit  $l \in \{0, \dots, x'-1\}$  und  $m \in \{x'-l, \dots, n-l\}$ , und weil  $g_m(x'-l) < 1$ , finden wir ein  $a \in ]g_m(x'-l), 1]$  mit  $\frac{m}{n}a + \frac{l}{n} \notin \tilde{K}(x')$ . Aufgrund der Isotonie von  $\tilde{K}$  ist  $\frac{m}{n}a + \frac{l}{n} \notin \tilde{K}(y)$  für alle  $y \in \{x', \dots, n\}$  und damit

$$\begin{aligned} \inf_{p \in [0, 1]^n} \text{BC}_p(\tilde{K} \ni \bar{p}) &\leq B_{m,a} \left( \left\{ x \in \{0, \dots, n\} : \tilde{K}(x+l) \ni \frac{l+ma}{n} \right\} \right) \\ &\leq B_{m,a}(\{0, \dots, x'-l-1\}) \\ &< B_{m, g_m(x'-l)}(\{0, \dots, x'-l-1\}) \\ &= \beta. \end{aligned}$$

Die Zulässigkeit folgt dann nach Lemma 4.3.

Um schließlich den letzten Fall von (4.9) zu beweisen, seien  $n \geq 2$  und  $\beta \geq \beta_n$  und sei nun  $\tilde{K}: \{0, \dots, n\} \rightarrow 2^{[0, 1]}$  durch die rechte Seite von (4.9) definiert. Wenn  $p \in [0, 1]^n$  und  $\bar{p} \in [0, \frac{1-\beta}{n}]$ , ist

$$\text{BC}_p(\tilde{K} \ni \bar{p}) \geq \text{BC}_p(\{0\}) = \prod_{j=1}^n (1 - p_j) \geq 1 - \sum_{j=1}^n p_j = 1 - n\bar{p} \geq \beta.$$

Wenn  $p \in [0, 1]^n$  und  $\bar{p} \in ]\frac{1-\beta}{n}, 1]$ , existiert, mit  $g_n(n+1) := 1$ , entweder ein  $c \in \{2, \dots, n\}$  mit  $\bar{p} \in ]g_n(c), g_n(c+1)]$  oder es ist  $\bar{p} \in ]\frac{1-\beta}{n}, g_n(2)]$  und wir setzen  $c := 1$ . In beiden Fällen ist nach Lemma 4.28  $n\bar{p} \leq ng_n(c+1) \leq c \leq n$  und eine Anwendung von Hoeffding (1956, Theorem 4, (26)) im zweiten Schritt unten liefert

$$\text{BC}_p(\tilde{K} \ni \bar{p}) = \text{BC}_p(\{0, \dots, c\}) \geq F_{n, \bar{p}}(c) \geq F_{n, g_n(c+1)}(c) \geq \beta.$$

Also ist  $\tilde{K}$  ein  $\beta$ -Konfidenz-Upray für (4.8) und erfüllt  $\tilde{K}(x) \subseteq K(x)$  für jedes  $x$ . Aufgrund der Zulässigkeit von  $K$  ist  $\tilde{K}(x) = K(x)$ , also gilt (4.9).  $\square$

*Beweis von Satz 4.17.* Sei  $\gamma := \frac{1+\beta}{2}$ ,  $K_\gamma$  das  $\gamma$ -Konfidenz-Upray aus Satz 4.7 und  $\Lambda_\gamma$  das analoge  $\gamma$ -Konfidenz-Downray aus Bemerkung 4.13. Aufgrund der Subadditivität definiert dann  $M_\beta(x) := K_\gamma(x) \cap \Lambda_\gamma(x)$  für  $x \in \{0, \dots, n\}$  ein  $\beta$ -Konfidenzintervall für (4.8).

Wenn  $n = 1$ , ist  $M_{\text{CP},n} = M_\beta$ , und es gilt die Behauptung. Sei also  $\beta \geq 2\beta_n - 1$ , also  $\gamma \geq \beta_n$ . Dann folgt mit (4.9) und (4.10), mit  $\gamma$  statt  $\beta$ ,  $M_{\text{CP},n}(x) = M_\beta(x)$  für  $x \notin \{1, n-1\}$ . Also ist  $\text{BC}_p(M_{\text{CP},n} \ni \bar{p}) = \text{BC}_p(M_\beta \ni \bar{p}) \geq \beta$ , falls  $\bar{p} \notin (M_{\text{CP},n}(1) \setminus M_\beta(1)) \cup (M_{\text{CP},n}(n-1) \setminus M_\beta(n-1))$ . Sonst ist  $\bar{p} \in ]\frac{1-\gamma}{n}, g_{n,\gamma}(1)]$  oder  $\bar{p} \in [g_{n,\gamma}(n-1), 1 - \frac{1-\gamma}{n}[$ . Im ersten Fall ist  $\bar{p} \in ]\frac{1-\gamma}{n}, g_{n,\gamma}(1)] = ]\frac{1-\gamma}{n}, 1 - \gamma^{1/n}] \subseteq [0, 1 - (1-\gamma)^{1/n}] = M_{\text{CP},n}(0)$ , und aus  $\bar{p} \in M_{\text{CP},n}(0)$  und  $\bar{p} \leq 1 - \gamma^{1/n}$  folgt

$$\text{BC}_p(M_{\text{CP},n} \ni \bar{p}) \geq \text{BC}_p(\{0\}) = \prod_{j=1}^n (1 - p_j) \geq 1 - n\bar{p} \geq 1 - n(1 - \gamma^{1/n}) \geq \gamma > \beta.$$

Im zweiten, analogen Fall ist  $\bar{p} \in [g_{n,\gamma}(n-1), 1 - \frac{1-\gamma}{n}[ = [\gamma^{1/n}, 1 - \frac{1-\gamma}{n}[ \subseteq M_{\text{CP},n}(n)$ , und aus  $\bar{p} \in M_{\text{CP},n}(n)$  und  $\bar{p} \geq \gamma^{1/n}$  folgt  $\text{BC}_p(M_{\text{CP},n} \ni \bar{p}) \geq \text{BC}_p(\{n\}) = \prod_{j=1}^n p_j \geq \bar{p}^n \geq \gamma > \beta$ .  $\square$

## 4.5.2 Beweise des Abschnitts 4.3

*Beweis von Bemerkung 4.21.* Wenn  $n \geq 2$  und  $\beta = (1 - 1/N)^n$  ist, gilt für  $p = g_n(1)$ , dass  $p = 1 - \beta^{1/n} = 1/N$ , und damit

$$H_{n,p,N}(K_{\text{CP},n} \ni p) = H_{n,p,N}(\{0\}) = \frac{\binom{N(1-p)}{n}}{\binom{N}{n}} = \prod_{j=0}^{n-1} \frac{N(1-p)-j}{N-j} < (1-p)^n = \beta.$$

Damit ist  $K_{\text{CP},n}$  nicht gültig für (4.11) und damit auch nicht gültig für das allgemeinere Schätzproblem (4.12).  $\square$

Um Satz 4.22 zu beweisen, untersuchen wir zunächst die Hilfsfunktion  $f$  aus (4.13).

**Lemma 4.29.** *Seien  $n \in \mathbb{N}$ ,  $\beta \in ]0, 1[$  und  $M := \{(m, p) : m \in \mathbb{N}, p \in ]0, \frac{m}{m+n-1}[$ . Sei  $k : M \rightarrow ]0, 1[$  definiert durch*

$$k(m, p) := \prod_{j=0}^{n-1} \frac{m(1-p) - pj}{m - pj} = \prod_{j=0}^{m-1} \frac{m - p(j+n)}{m - pj}, \quad (4.28)$$

so dass

$$f : \mathbb{N} \rightarrow ]0, 1[, \quad f(m) := \text{das } p \in ]0, \frac{m}{m+n-1}[ \text{ mit } k(m, p) = \beta$$

wohldefiniert, isoton, mit  $\lim_{m \rightarrow \infty} f(m) = 1 - \sqrt[n]{\beta} = g_n(1)$ , und, falls  $n \geq 2$ , streng isoton

ist. Speziell ist  $f(1) = \frac{1-\beta}{n}$  und  $f(2) = \frac{\frac{1-\beta}{n} + 2 - \sqrt{(\frac{1-\beta}{n})^2 + 4\beta}}{n+1}$ .

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

*Beweis.* Die behauptete Gleichheit in (4.28) gilt, da für  $(m, p) \in M$

$$\prod_{j=0}^{n-1} \frac{m(1-p) - pj}{m - pj} = \frac{\prod_{j=m-n}^{m-1} m - p(j+n)}{\prod_{j=0}^{n-1} m - pj} = \prod_{j=0}^{m-1} \frac{m - p(j+n)}{m - pj}.$$

Für  $(m, p) \in M$  und  $j \in \{0, \dots, m-1\}$  ist  $m = \frac{m(m-1+n)}{m+n-1} > p(j+n) > pj$ , also ist die Funktion  $h: ]0, \frac{m}{m+n-1}[ \rightarrow ]0, 1[, p \mapsto \frac{m-p(j+n)}{m-pj}$  streng antiton mit  $\lim_{p \rightarrow 0} h(p) = 1$  und  $\lim_{p \rightarrow \frac{m}{m+n-1}} h(p) = \frac{m-1-j}{m+n-1-j} \geq 0$ . Damit ist für  $m \in \mathbb{N}$

$$\left] 0, \frac{m}{m+n-1} \right[ \ni p \mapsto k(m, p) \quad \text{streng antiton} \quad (4.29)$$

mit  $\lim_{p \rightarrow 0} k(m, p) = 1$  und  $\lim_{p \rightarrow \frac{m}{m+n-1}} k(m, p) = 0$ , und es existiert die eindeutige Lösung  $p = f(m)$  von  $k(m, p) = \beta$ . Wenn  $n = 1$ , ist  $f: m \mapsto 1 - \beta$  isoton, und wenn  $n \geq 2$ , ist für jedes  $(m, p) \in M$  und  $j \in \{0, \dots, n-1\}$

$$\frac{(m+1)(1-p) - pj}{(m+1) - pj} \bigg/ \frac{m(1-p) - pj}{m - pj} = \frac{1 + \frac{1}{m - \frac{pj}{1-p}}}{1 + \frac{1}{m - pj}} > 1,$$

also  $k(m+1, p) > k(m, p)$ , und da  $k$  für alle  $m$  streng antiton in  $p$  ist, ist  $f$  somit streng isoton. Außerdem ist  $\lim_{m \rightarrow \infty} k(m, p) = (1-p)^n$ , also  $\lim_{m \rightarrow \infty} f(m) = 1 - \sqrt[n]{\beta}$ .

Die expliziten Formeln für  $f(1)$  und  $f(2)$  kann man direkt ausrechnen.  $\square$

**Lemma 4.30.** Sei  $n \geq 3$  und  $\beta \in ]0, 1[$ . Dann gilt für  $f$  aus (4.13)

$$\frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1} = \frac{1}{\lceil \frac{1}{f(1)} \rceil - 1} = \min \left\{ \frac{m}{\lceil \frac{m}{f(m)} \rceil - 1} : m \in \mathbb{N} \right\}. \quad (4.30)$$

*Beweis.* Die erste Gleichheit gilt, da  $f(1) = (1-\beta)/n$ . Für  $j \in \{0, \dots, n-1\}$  und  $r \in \mathbb{N}$  sei

$$p(r, j) := \frac{rn - (1-\beta)(2r-1+j)}{rn - (1-\beta)(r-1+j)}.$$

Man rechnet nach, dass  $p(m, j) \geq p(1, j)$  genau dann, wenn  $j \geq 1$ . Weiter ist  $p(m, 2) \cdot p(m, 0) \geq p(1, 2) \cdot p(1, 0)$  genau dann, wenn  $\frac{1-\beta}{n} \leq \frac{2}{3}$ , also wahr. Damit ist, da  $n \geq 3$ ,

$$\prod_{j=0}^{n-1} p(m, j) \geq \prod_{j=0}^{n-1} p(1, j).$$

Sei  $m \in \mathbb{N}$  und  $k(m, \cdot)$  aus (4.29). Da  $n > 2$ , ist  $\tilde{f}(m) := \frac{m(1-\beta)}{mn - (m-1)(1-\beta)} < \frac{m}{m+n-1}$ , und damit

$$k(m, \tilde{f}(m)) = \prod_{j=0}^{n-1} p(m, j) \geq \prod_{j=0}^{n-1} p(1, j) = k(1, \tilde{f}(1)) = \prod_{j=0}^{n-1} p(1, j) = \beta.$$

Weil  $k(m, f(m)) = \beta$  und  $k(m, \cdot)$  streng antiton ist, folgt

$$f(m) \geq \tilde{f}(m) = \frac{m(1-\beta)}{mn - (m-1)(1-\beta)} = \frac{mf(1)}{m(1-f(1)) + f(1)}. \quad (4.31)$$

Also ist  $m/f(m) \leq m\lceil 1/f(1) \rceil - (m-1)$ , und damit folgt die Behauptung.  $\square$

*Beweis von Satz 4.22.* Wenn  $n = 0$ , ist alles klar, also können wir im Folgenden annehmen, dass  $n \in \mathbb{N}$ .

Nach Vatutin und Mikhailov (1983) ist jede hypergeometrische Verteilung  $H_{n,p,N}$  eine Bernoulli-Faltung  $BC_q$  für ein  $q \in [0, 1]^n$ . Damit ist  $K$ , wie in Satz 4.6 definiert, nach jenem Satz ein  $\beta$ -Konfidenzbereich für (4.12), denn

$$H_{n,p,N}(K \ni p) = H_{n,p,N}\left(K \ni \frac{\mu(H_{n,p,N})}{n}\right) = BC_q\left(K \ni \frac{\mu(BC_q)}{n}\right) = BC_q(K \ni \bar{q}) \geq \beta.$$

Wir zeigen, dass  $K^*$  ein Upray ist. Wenn  $n = 1$ , ist  $f(m) = 1 - \beta$ . Wenn  $n \geq 2$ , ist  $f(m) < \frac{m}{m+n-1} \leq \frac{m}{m+1}$ . Also gilt in beiden Fällen  $0 < m \leq \lceil \frac{m}{f(m)} \rceil - 1$  für alle  $m \in \mathbb{N}$ , und damit ist  $K^*$  ein Upray, da  $K$  aus Satz 4.7 ein Upray ist,

Wir zeigen nun, dass  $\inf_{N,p} H_{n,p,N}(K^* \ni p) \geq \beta$ . Für alle  $p \notin K(1) \setminus K^*(1)$  ist

$$H_{n,p,N}(K^* \ni p) = H_{n,p,N}(K \ni p) \geq \beta.$$

Wenn  $p = \frac{i}{N} \in K(1) \setminus K^*(1)$  für ein  $N \geq n$  und  $i \in \{0, \dots, N\}$ , ist, da  $0 \notin K(1)$ ,  $0 < p < \frac{i}{\lceil \frac{i}{f(i)} \rceil - 1}$ , und somit  $\frac{i}{p} > \lceil \frac{i}{f(i)} \rceil - 1$ , also  $\frac{i}{p} = N \geq \frac{i}{f(i)}$ , also  $0 < p \leq f(i) = f(Np)$ , und damit

$$H_{n,p,N}(K^* \ni p) = H_{n,p,N}(\{0\}) = \prod_{j=0}^{n-1} \frac{Np(1-p) - pj}{Np - pj} = k(Np, p) \geq k(Np, f(Np)) = \beta,$$

wobei  $k(Np, \cdot)$  die antitone Funktion aus (4.29) ist.

Kommen wir nun zur strengen Isotonie, Optimalität und Zulässigkeit. Wenn  $n = 1$ , werden die folgenden Beweisschritte analog durchgeführt. Daher nehmen wir bis zum Ende des Beweises an, dass  $\beta \geq \beta_n$ . Dann ist  $K$  wie in (4.9) gegeben.

Wir zeigen die strenge Isotonie von  $K^*$ . Da  $K$  streng isoton ist, reicht es, zu zeigen, dass  $K^*(0) = [0, 1] \subsetneq K^*(1) \subsetneq ]g_n(2), 1] = K^*(2)$ . Da  $\lim_{m \rightarrow \infty} f(m) = g_n(1)$ , ist auch

$$\lim_{m \rightarrow \infty} \frac{m}{\lceil \frac{m}{f(m)} \rceil - 1} = g_n(1). \quad (4.32)$$

Da  $f$  isoton ist, und  $\lceil \frac{m}{f(m)} \rceil - 1 \leq \frac{m}{f(m)}$  für alle  $m \in \mathbb{N}$  gilt, ist

$$0 < f(1) = \inf\{f(m) : m \in \mathbb{N}\} \leq \inf\left\{\frac{m}{\lceil \frac{m}{f(m)} \rceil - 1} : m \in \mathbb{N}\right\} \leq g_n(1) < g_n(2).$$

#### 4 Konfidenzintervalle für durchschnittliche Erfolgswahrscheinlichkeiten

Damit ist die strenge Isotonie nachgewiesen.

Um die Optimalität von  $K^*$  nachzuweisen, nehmen wir an, dass  $\tilde{K}: \{0, \dots, n\} \rightarrow 2^{[0,1]}$  ein weiteres isotones Upray ist, mit  $\tilde{K}(x') \subsetneq K^*(x')$  für ein  $x' \in \{0, \dots, n\}$ , und zeigen, dass dann  $\inf_{N,p} H_{n,p,N}(\tilde{K} \ni p) < \beta$  gilt.

Wenn  $x' = 0$ , ist  $0 \notin \tilde{K}(0)$ , und damit  $\inf_{N,p} H_{n,p,N}(\tilde{K} \ni p) \leq \delta_0(\tilde{K} \ni 0) = 0 < \beta$ ,

Wenn  $x' = 1$ , ist wegen (4.32) entweder

$$K^*(1) = ]1 - \sqrt[n]{\beta}, 1] = K_{\text{CP},n}(1) \supsetneq \tilde{K}(1) \quad (4.33)$$

oder

$$\text{es existiert ein } m \in \mathbb{N} \text{ mit } \frac{m}{\lceil \frac{m}{f(m)} \rceil - 1} =: p \notin \tilde{K}(1). \quad (4.34)$$

Im Fall von (4.33) folgert man weiter wie unten im Fall  $x' \in \{2, \dots, n\}$ . Im Fall von (4.34) ist  $f(m) < p$ . Falls  $N := \lceil \frac{m}{f(m)} \rceil - 1 < n$ , ist  $0 < \frac{m}{n} < \frac{m}{N} = p$ , also auch  $\frac{m}{n} \notin \tilde{K}(1)$ , und damit  $H_{n, \frac{m}{n}, n}(\tilde{K} \ni \frac{m}{n}) \leq H_{n, \frac{m}{n}, n}(\{0\}) = 0$ . Falls  $N \geq n$ , ist  $H_{n,p,N}(\tilde{K} \ni p) \leq H_{n,p,N}(\{0\}) = k(Np, p) = k(m, p) < k(m, f(m))$ , wobei  $k(Np, \cdot)$  die antitone Funktion aus (4.29) ist.

Wenn  $x' \in \{2, \dots, n\}$  oder (4.33) gilt, sei  $a \in K(x') \setminus \tilde{K}(x')$  und  $b := \frac{g_n(x') + a}{2}$ , also  $g_n(x') < b < a$ . Es gilt  $\frac{\lceil Nb \rceil}{N} = \frac{\lceil Nb \rceil}{Nb} b \rightarrow b$  für  $N \rightarrow \infty$  und es existiert ein  $N_0$ , so dass  $\frac{\lceil Nb \rceil}{N} < a$  für alle  $N \geq N_0$ , und damit auch  $\frac{\lceil Nb \rceil}{N} \notin \tilde{K}(y)$  für alle  $y \in \{x', \dots, n\}$ . Damit gilt, wobei in der dritten Ungleichung ausgenutzt wird, dass  $H_{n, \lceil Nb \rceil / N, N} \rightsquigarrow B_{n,b}$  für  $N \rightarrow \infty$ ,

$$\begin{aligned} \inf_{N,p} H_{n,p,N}(\tilde{K} \ni p) &\leq \inf_N H_{n, \lceil Nb \rceil / N, N}(\tilde{K} \ni \frac{\lceil Nb \rceil}{N}) \\ &\leq \inf_{N \geq N_0} H_{n, \lceil Nb \rceil / N, N}(\{0, \dots, x' - 1\}) \\ &\leq B_{n,b}(\{0, \dots, x' - 1\}) \\ &< B_{n, g_n(x')}(\{0, \dots, x' - 1\}) \\ &= \beta. \end{aligned}$$

Die Zulässigkeit folgt mit Lemma 4.3 aus der Optimalität und der strengen Isotonie.

Zu (4.15): Die Gleichheit für  $x = 1$  gilt nach (4.30). Die Gleichheit für  $x \in \{0, \dots, n\} \setminus \{1\}$  gilt nach (4.9).  $\square$

*Beweis von Bemerkung 4.23.* Falls  $n = 0$ , ist  $K^*(0) = [0, 1] = K(0)$ . Falls  $n = 1$ , ist  $K^*(0) = [0, 1] = K(0)$  und  $K^*(1) = ]1 - \beta, 1] = K(1)$ , also  $K^* = K$ , und damit auch (4.16). Falls  $n = 2$ , gilt (4.17), da

$$K^*(1) = \left[ \frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1}, 1 \right] \cup \bigcup_{m=2}^{\infty} \left[ \frac{m}{\lceil \frac{m}{f(m)} \rceil - 1}, 1 \right]$$

und

$$\frac{1-\beta}{n} < \begin{cases} \frac{1}{\lceil \frac{n}{1-\beta} \rceil - 1}, \\ \frac{\frac{1-\beta}{n} + 2 - \sqrt{(\frac{1-\beta}{n})^2 + 4\beta}}{n+1} = f(2) \leq f(m) < \frac{m}{\lceil \frac{m}{f(m)} \rceil - 1} \text{ für alle } m \geq 2, \end{cases}$$

und daraus folgt (4.16). Falls  $n \geq 3$ , gilt (4.18) nach Lemma 4.30 und damit (4.16).  $\square$



# Literaturverzeichnis

- ADELL, J.A., JODRÁ, P. (2006) Exact Kolmogorov and total variation distances between some familiar discrete distributions. *J. Inequal. Appl.* 1–8.
- ADELL, J.A., LEKUONA, A. (2005) Sharp estimates in signed Poisson approximation of Poisson mixtures. *Bernoulli* **11**, 47–65.
- AGNEW, R.A. (1974). Confidence sets for binary response models. *J. Amer. Statist. Assoc.* **69**, 522–524.
- BAILLON, J.-B., COMINETTI, R., VAISMAN, J. (2013). A sharp uniform bound for the distribution of sums of Bernoulli trials. Preprint, <http://arxiv.org/abs/0806.2350>
- BARBOUR, A.D., CHEN, L.H.Y. An introduction to Stein’s method. Vol. 4. *World Scientific* (2005).
- BARBOUR, A.D., HALL, P. (1984). On the rate of Poisson convergence. *Math. Proc. Cambridge Philos. Soc.* **95**, 473–480.
- BARBOUR, A.D., JENSEN, J.L. (1989). Local and Tail Approximations near the Poisson Limit. *Scand. J. Statist.* **16**, 75–87.
- BILLINGSLEY, P. Convergence of probability measures. *John Wiley & Sons* (2009).
- BOURBAKI, N. Elements of Mathematics. General Topology, Chapters 1–4. *Springer* (1989).
- BUEHLER, R.J. (1957). Confidence intervals for the product of two binomial parameters. *J. Amer. Statist. Assoc.* **52**, 482–493.
- BUONACCORSI, J.P. (1987). A Note on Confidence Intervals for Proportions in Finite Populations. *Amer. Statist.* **41**, 215–218.
- BYERS, V.S., LECAM, L., LEVIN, A.S., JOHNSTON, J.O., HACKETT, A.J. (1979). Immunotherapy of osteogenic sarcoma with transfer factor. Long-term follow-up. *Cancer Immunol., Immunother.* **6**, 243–253.
- CHEN, L.H.Y. (1975). Poisson approximation for dependent trials. *Ann. Probab.* **3**, 534–545.
- CHENG, S.L., MICHEALS, R.J., LU, J. (2010). Comparison of confidence intervals for large operational biometric data by parametric and non-parametric methods. *NIST Interagency/Internal Report (NISTIR) - 7740*.

- CLOPPER, C.J., PEARSON, E.S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404–413.
- D'ARISTOTILE, A., DIACONIS, P., FREEDMAN, D. (1988) On Merging of Probabilities. *Sankhyā* **50**, 363–380.
- DAVID, H.A. (1960). A conservative property of binomial tests. *Ann. Math. Statist.* **31**, 1205–1207.
- DAVYDOV, Y., ROTAR, V. (2009). On Asymptotic Proximity of Distributions. *J. Theoret. Probab.* **22**, 82–98.
- DE MOIVRE, A. The Doctrine of Chances. *Pearson* (1718).
- DUDLEY, R.M. Real Analysis and Probability. *Cambridge University Press* (2002).
- DÜMBGEN, L. (2004). Exact confidence bounds in discrete models - Algorithmic aspects of Sterne's method. Preprint, [www.imsv.unibe.ch/unibe/phlnat/imsv/content/e6030/e7196/e7932/e8042/e8109/e8111/Sterne\\_eng.pdf](http://www.imsv.unibe.ch/unibe/phlnat/imsv/content/e6030/e7196/e7932/e8042/e8109/e8111/Sterne_eng.pdf), abgerufen am 24. Februar 2015.
- EHM, W. (1991). Binomial approximation to the Poisson binomial distribution. *Statist. Probab. Lett.* **11**, 7–16.
- ELSTRODT, J. Maß- und Integrationstheorie. *Springer* (2009).
- HOEFFDING, W. (1956). On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* **27**, 713–721.
- GUT, A. Probability: A Graduate Course. *Springer, New York* (2005).
- JOGDEO, K., SAMUELS, S.M. (1968). Monotone convergence of binomial probabilities and a generalization of Ramanujan's equation. *Ann. Math. Statist.* **39**, 1191–1195.
- KAPPAUF, W.E., BOHRER, R. (1974). Observations on mixed binomials. *The American Journal of Psychology* **87**, 643–665.
- KATZ, L. (1953). Confidence intervals for the number showing a certain characteristic in a population when sampling is without replacement. *J. Amer. Statist. Assoc.* **48**, 256–261.
- KELLEY, J.L. General Topology. *Springer* (1975).
- KONIJN, H.S. Statistical theory of sample survey design and analysis. *North-Holland, Amsterdam* (1973).
- KLENKE, A., MATTNER, L. (2010). Stochastic ordering of classical discrete distributions. *Adv. Appl. Prob.* **42**, 392–410.

- KOOPMAN, B.O. (1950). Necessary and sufficient conditions for Poisson's distribution. *Proc. Amer. Math. Soc.* **1**, 813–823.
- LIESE, F., MIESCKE, K.-J. Statistical Decision Theory: Estimation, Testing, and Selection. *Springer* (2008).
- LLOYD, C., KABAILA, P. (2010). Letter to the editor: Some comments on: On construction of the smallest one-sided confidence interval for the difference of two proportions [*Ann. Statist.* **38** (2010), 1227–1243]. *Ann. Statist.* **38**, 3840–3841.
- MATTNER, L., SCHULZ, J. (2014). On normal approximations to symmetric hypergeometric laws. <http://arxiv.org/abs/1404.7657>
- MATTNER, L., TASTO, C. (2014). Confidence Intervals for Average Success Probabilities. Erscheint in *Probab. Math. Statist.* <http://arxiv.org/abs/1403.0229>
- OLLERO, J., RAMOS, H.M. (1995). Description of a subfamily of the discrete Pearson system as generalized-binomial distributions. *J. Ital. Statist. Soc.* **4**, 235–249.
- PSCHYREMBEL, W. Klinisches Wörterbuch. 259. Auflage. *Walter de Gruyter* (2002).
- QUERENBURG, B.V. Mengentheoretische Topologie. *Springer, Berlin Heidelberg* (2001).
- RIORDAN, J. (1937). Moment recurrence relations for binomial, Poisson and hypergeometric frequency distributions. *Ann. Math. Statist.* **8**, 103–111.
- ROOS, B. (2000). Binomial approximation to the Poisson binomial distribution: The Krawtchouk expansion. *Theory Probab. Appl.* **45**, 258–272.
- ROOS, B. (2003). Improvements in the Poisson approximation of mixed Poisson distributions. *J. Statist. Planning Infer.* **113**, 467–483.
- ROOS, B., PFEIFER, D. (2003a). On the distance between the distribution of random sums. *J. Appl. Prob.* **40**, 87–106.
- ROOS, B., PFEIFER, D. (2003b). On Error Bounds for the Approximation of Random Sums. Preprint, [www.math.uni-trier.de/~roos/preprint\\_2003\\_04.pdf](http://www.math.uni-trier.de/~roos/preprint_2003_04.pdf), abgerufen am 13. Januar 2015.
- RUZANKIN, P.S. (2004). On the rate of Poisson process approximation to a Bernoulli process. *J. Appl. Prob.* **41**, 271–276.
- SCHUBERT, H. Topologie. *BG Teubner Verlagsgesellschaft* (1964).
- STEIN, C. (1970). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2**, 583–602.

- STERNE, T.E. (1954). Some remarks on confidence or fiducial limits. *Biometrika* **41**, 275–278.
- STRASSER, H. Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory. Vol. 7. *Walter de Gruyter* (1985).
- TASTO, C. (2011). Grenzwertsätze für Bernoulli-Faltungen. Diplomarbeit, Universität Trier.
- TCHEBICHEF, P. (1846). Démonstration élémentaire d’une proposition générale de la théorie des probabilités. *J. Reine Angew. Math.* **33**, 259–267.
- VATUTIN, V.A., MIKHAILOV, V.G. (1983). Limit theorems for the number of empty cells in an equiprobable scheme for group allocation of particles. *Theory Probab. Appl.* **27**, 734–743. Russisches Original in *Teor. Veroyatnost. i Primenen.* **27**, 684–692 (1982).
- WANG, W. (2010). On construction of the smallest one-sided confidence interval for the difference of two proportions. *Ann. Statist.* **38**, 1227–1243.
- WANG, W. (2014). Exact Optimal Confidence Intervals for Hypergeometric Parameters. Erscheint in *J. Amer. Statist. Assoc.* DOI: 10.1080/01621459.2014.966191.
- WANG, Y.H. (1993). On the number of successes in independent trials. *Statist. Sinica* **3**, 295–312.
- ZACHAROVAS, V., HWANG, H.-K. (2010). A Charlier-Parseval Approach to Poisson Approximation and its Applications. *Lith. Math. J.* **50** 88–119.