

Einführung in die Statistik

Klaus Ritter

SS 2007

Literatur: insbesondere

- A. Irle, Wahrscheinlichkeitstheorie und Statistik, Teubner, 2005.
- U. Krengel, Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg, 2003.
- J. Lehn, H. Wegmann, Einführung in die Statistik, Teubner, 2004.

sowie

- Ch. Hesse, Angewandte Wahrscheinlichkeitstheorie, Vieweg, 2003.
- Ch. H. Snell, J. L. Introduction to Probability, [www](#).
- Meintrup, Schäffler, Stochastik, Springer, 2005.

Kap. I Stochastische Modellierung und optimale Beschickung von Cash Points

Als motivierendes Beispiel. PROJEKTOR

Kap. II Stochastische Modelle

1. Wahrscheinlichkeitsräume
2. Bedingte Wahrscheinlichkeiten und Unabhängigkeit
3. Reellwertige Zufallsvariablen

Mathematisches Modell für zufälliges Phänomen:

- Wahrscheinlichkeitsraum
 - Ergebnisraum
 - Ereignisraum
 - Wahrscheinlichkeitsmaß
- Zufallsvariablen

In den Kapiteln II–VI:

- allgemeine Begriffsbildung und Konstruktionen
- konkrete Modelle
- Analyse und Simulation stochastischer Modelle

1 Wahrscheinlichkeitsräume

Ergebnisraum Ω ist nicht-leere Menge, Elemente $\omega \in \Omega$ sind die möglichen **Ergebnisse** des Zufallsexperimentes.

1. Beispiel

- **Würfeln**, $\Omega := \{1, \dots, 6\}$
- **Anzahl Anrufe in Callcenter**, $\Omega := \mathbb{N}_0$
- **Wartezeit bei Anruf**, $\Omega := \mathbb{R}_+ := [0, \infty[$
- **Verlauf eines Aktienkurses**, $\Omega := C_+([0, 1])$ Menge der nicht-negativen stetigen Funktionen auf $[0, 1]$

Für gewisse Teilmengen $A \subseteq \Omega$, genannt **Ereignisse**, definiert man die Wahrscheinlichkeit ihres Eintretens ($\omega \in A$).

Ereignisraum \mathfrak{A} ist die Menge aller Ereignisse in Ω .

2. Beispiel

- Würfeln, **Ergebnis ist gerade Zahl**, $A := \{2, 4, 6\}$
- Anzahl Anrufe in Callcenter, **Kapazitätsgrenze K nicht überschritten**, $A := \{0, \dots, K\}$
- Wartezeit bei Anruf, **Wartezeit liegt zwischen 1 und 2 (Minuten)**, $A := [1, 2]$

- Verlauf eines Aktienkurses, Kurs weicht von Anfangswert um nicht mehr als 1 (Euro) ab,

$$A := \{\omega \in C_+([0, 1]) : \sup_{0 \leq t \leq 1} |\omega(0) - \omega(t)| \leq 1\}$$

Mengentheoretische Operationen mit Ereignissen.

3. Beispiel

- Ereignis A oder Ereignis B tritt ein, $A \cup B$
- Ereignis A und Ereignis B treten ein, $A \cap B$
- Ereignis A tritt nicht ein, $A^c := \Omega \setminus A$
- (mindestens) eines der Ereignisse A_1, A_2, \dots tritt ein,
$$\bigcup_{i=1}^{\infty} A_i$$
- alle Ereignisse A_1, A_2, \dots treten ein, $\bigcap_{i=1}^{\infty} A_i$

Forderung: obige Operationen liefern wieder Ereignisse.

Dazu Begriff der σ -Algebra.

Bezeichnung $\mathfrak{P}(\Omega)$ **Potenzmenge** von Ω (Menge aller Teilmengen von Ω), $|U|$ **Mächtigkeit** (Anzahl der Elemente) einer endlichen Menge U .

4. Beispiel Münzwurf, $\Omega := \{Z, K\}$,

$$\mathfrak{P}(\Omega) = \{\emptyset, \{Z\}, \{K\}, \{Z, K\}\}.$$

Beachte $Z \notin \mathfrak{P}(\Omega)$, aber $\{Z\} \in \mathfrak{P}(\Omega)$.

5. Bemerkung Falls Ω endlich,

$$|\mathfrak{P}(\Omega)| = 2^{|\Omega|}.$$

Beweis. Durch Induktion über $n := |\Omega|$.

Verankerung: $|\Omega| = 0$, also $\Omega = \emptyset$ und $|\mathfrak{P}(\Omega)| = |\{\emptyset\}| = 1$.

Induktionsschritt: $|\Omega| = n + 1 \geq 1$. Fixiere $\omega^* \in \Omega$. Dann

$$\begin{aligned} |\mathfrak{P}(\Omega)| &= |\{A \subseteq \Omega : \omega^* \in A\}| + |\{A \subseteq \Omega : \omega^* \notin A\}| \\ &= 2^n + 2^n = 2^{n+1}. \end{aligned}$$

□

6. Definition $\mathfrak{A} \subseteq \mathfrak{P}(\Omega)$ σ -Algebra (in Ω), falls:

(i) $\Omega \in \mathfrak{A}$

(ii) $A \in \mathfrak{A} \Rightarrow A^c \in \mathfrak{A}$

(iii) $A_1, A_2, \dots \in \mathfrak{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathfrak{A}$

Vgl. **Topologie**: Menge \mathfrak{D} der **offenen Teilmengen** von Ω , z.B. $\Omega := \mathbb{R}^d$.

Forderung: Ereignisraum ist σ -Algebra.

7. Bemerkung In der Regel betrachtet man $\mathfrak{A} := \mathfrak{P}(\Omega)$, falls Ω abzählbar. Nicht so, falls Ω überabzählbar, siehe Kapitel V.

8. Lemma Für jede σ -Algebra \mathfrak{A} gilt:

(i) $\emptyset \in \mathfrak{A}$

(ii) $A, B \in \mathfrak{A} \Rightarrow A \cup B, A \cap B, A \setminus B \in \mathfrak{A},$

(iii) $A_1, A_2, \dots \in \mathfrak{A} \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathfrak{A}$

Beweis.

(i) $\emptyset = \Omega^c \in \mathfrak{A}$ nach Def. 6.(i), (ii).

(ii) $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \dots$, verwende (i) und Def. 6.(iii).

$A \cap B = (A^c \cup B^c)^c$, verwende Def. 6.(ii) und $A^c \cup B^c \in \mathfrak{A}$.

$A \setminus B = A \cap B^c$, wie zuvor.

(iii) $\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{\infty} A_i^c\right)^c \in \mathfrak{A}$ nach Def. 6.(ii), (iii).

□

Bezeichnung Mengen A_1, A_2, \dots **paarweise disjunkt (p.d.)**,
falls $A_i \cap A_j = \emptyset$ für $i \neq j$.

Im folgenden \mathfrak{A} σ -Algebra in nicht-leerer Menge Ω , z.B.

$\Omega := \mathbb{N}_0$ und $\mathfrak{A} := \mathfrak{P}(\Omega)$.

Zuordnung von Wahrscheinlichkeiten $P(A)$ zu den einzelnen Ereignissen $A \in \mathfrak{A}$. Dabei folgende Vorstellung:

- bei „großer“ Anzahl von „unabhängigen“ Wiederholungen des Zufallsexperimentes liegt die **relative Häufigkeit** des Eintretens von Ereignis A „nahe“ bei $P(A)$.

9. Definition $P : \mathfrak{A} \rightarrow [0, 1]$ **Wahrscheinlichkeitsmaß** oder **Wahrscheinlichkeitsverteilung** (auf \mathfrak{A}), falls:

(i) $P(\Omega) = 1$

(ii) $A_1, A_2, \dots \in \mathfrak{A}$ p.d. $\Rightarrow P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$
(**σ -Additivität**)

Genauer: $\dots \Rightarrow \sum_{i=1}^{\infty} P(A_i)$ (absolut) konvergent und \dots

10. Beispiel Ω endlich, Laplace-Annahme

$$P(A) := |A|/|\Omega|, \quad A \subseteq \Omega.$$

Speziell für jedes $\omega \in \Omega$

$$P(\{\omega\}) = 1/|\Omega|.$$

Beh.: P Wahrscheinlichkeitsmaß auf $\mathfrak{A} := \mathfrak{P}(\Omega)$, genannt **Gleichverteilung** auf Ω .

Beweis. Offensichtlich gilt $0 \leq P(A) \leq 1$ und $P(\Omega) = 1$.

Für $A_1, A_2, \dots \subseteq \Omega$ p.d. (notwendig: $A_i = \emptyset$ bis auf endlich viele i) gilt

$$\left| \bigcup_{i=1}^{\infty} A_i \right| = \sum_{i=1}^{\infty} |A_i|.$$

Dies zeigt die σ -Additivität.

□

11. Definition $(\Omega, \mathfrak{A}, P)$ **Wahrscheinlichkeitsraum**, falls Ω nicht-leere Menge, \mathfrak{A} σ -Algebra in Ω und P W'maß auf \mathfrak{A} .

12. Beispiel **Stochastisches Modell für einmaliges Würfeln:**

W'raum $(\Omega, \mathfrak{A}, P)$ mit:

(i) $\Omega := \{1, \dots, 6\}$

(ii) $\mathfrak{A} := \mathfrak{P}(\Omega)$

(iii) P Gleichverteilung auf Ω

13. Beispiel Stochastisches Modell für Geschlecht eines

Neugeborenen: Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit:

- (i) $\Omega := \{W, M\}$
- (ii) $\mathfrak{A} := \mathfrak{P}(\Omega)$
- (iii) P definiert durch $P(\{W\}) := 0,4863$.

Letzteres empirisch ermittelt als relative Häufigkeit unter den 25 171 123 Lebendgeburten in D in den Jahren 1970–1999. Siehe Hesse (2003, p. 23).

14. Beispiel **Hard core model** der Physik, Gleichverteilung auf sehr großer Menge „unbekannter“ Mächtigkeit.

PROJEKTOR

15. Beispiel (Fragwürdiges) **stochastisches Modell** für **Pfeiltreffer auf Dartscheibe** mit Radius $r > 0$:

(i) $\Omega := \underbrace{\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}}_{=: \omega}$

(ii) \mathfrak{A} eine „ σ -Algebra in Ω , deren Elementen ein ‘**Flächeninhalt**’ $\lambda(A)$ zugeordnet werden kann“, siehe Kapitel V.1.

(iii) $P(A) := \lambda(A) / (\pi r^2)$

Beachte: $P(\{\omega\}) = 0$ für alle $\omega \in \Omega$.

Im folgenden stets $(\Omega, \mathfrak{A}, P)$ W'raum.

16. Satz Rechenregeln für W'maße

Für $A, B \in \mathfrak{A}$ gilt:

$$(i) \quad A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$$

(Additivität)

$$(ii) \quad A \subseteq B \Rightarrow P(B) = P(A) + P(B \setminus A)$$

$$(iii) \quad A \subseteq B \Rightarrow P(A) \leq P(B) \quad (\text{Monotonie})$$

$$(iv) \quad P(A^c) = 1 - P(A)$$

$$(v) \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Beweis. Seien $A, B \subseteq \Omega$.

(i) Gelte $A \cap B = \emptyset$. Dann gilt

$A \cup B = A \cup B \cup \emptyset \cup \emptyset \dots$ mit p.d. Mengen, und somit

$$P(A \cup B) = P(A) + P(B) + \sum_{i=1}^{\infty} P(\emptyset) < \infty.$$

Es folgt $P(\emptyset) = 0$ und die Behauptung.

(ii) Gelte $A \subseteq B$. Dann $B = A \cup (B \setminus A)$. Da A und $B \setminus A$ disjunkt, folgt $P(B) = P(A) + P(B \setminus A)$.

(iii) Verwende (ii) und $P(B \setminus A) \geq 0$.

(iv) Verwende (ii) mit $B = \Omega$, also $P(B) = 1$.

(v) Verwende $A \cup B = A \cup (B \setminus A)$ und $B \setminus A = B \setminus (A \cap B)$.

Somit $A \cap (B \setminus (A \cap B)) = \emptyset$, $A \cap B \subseteq B$, und (i), (ii) liefern

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

□

2 Bedingte Wahrscheinlichkeiten und Unabhängigkeit

Betrachte W'raum $(\Omega, \mathfrak{A}, P)$. Übergang zu neuem W'maß, falls bekannt, daß B eingetreten.

17. Definition Für $A, B \in \mathfrak{A}$ mit $P(B) > 0$ heißt

$$P(A | B) := \frac{P(A \cap B)}{P(B)}$$

bedingte Wahrscheinlichkeit von A gegeben B .

18. Bemerkung $P(\cdot | B)$ ist ein W'maß auf \mathfrak{A} mit $P(B | B) = 1$. „Restriktion“ auf B und Normierung.

19. Beispiel Für Gleichverteilung P auf endlicher Menge Ω ,
 $\emptyset \neq B \subseteq \Omega$ und $A \subseteq \Omega$ gilt

$$P(A | B) = \frac{|A \cap B|}{|B|},$$

d.h. $P(A | B)$ ist **relativer Anteil von Elementen aus A in B** .
„Gleichverteilung“ auf B .

20. Beispiel Einmaliges Würfeln (wie üblich modelliert) und
 $B := \{1, 5, 6\}$. Dann

$$P(\{\omega\} | B) = \begin{cases} \frac{1}{3}, & \text{falls } \omega \in B \\ 0, & \text{falls } \omega \notin B. \end{cases}$$

21. Beispiel 2 weiße (1, 2) und 3 schwarze (3, 4, 5) Kugeln, ziehe 2 Kugeln ohne Zurücklegen. Gesucht:

Wahrscheinlichkeit, daß die 2. Kugel schwarz, falls die 1. Kugel weiß. Modell: Gleichverteilung auf

$$\Omega := \{(\omega_1, \omega_2) \in \{1, \dots, 5\}^2 : \omega_1 \neq \omega_2\}.$$

Für $A := \{(\omega_1, \omega_2) \in \Omega : \omega_2 \geq 3\},$

$$B := \{(\omega_1, \omega_2) \in \Omega : \omega_1 \leq 2\}$$

gilt (wie erwartet)

$$P(A | B) = \frac{|A \cap B|}{|B|} = \frac{6}{8} = \frac{3}{4}.$$

22. Satz Für p.d. Mengen $B_1, \dots, B_n \in \mathfrak{A}$ mit $P(B_i) > 0$ für alle i und $\bigcup_{i=1}^n B_i = \Omega$ gilt für jedes $A \in \mathfrak{A}$ die **Formel von der totalen Wahrscheinlichkeit**,

$$P(A) = \sum_{i=1}^n P(A | B_i) \cdot P(B_i),$$

und, falls $P(A) > 0$, die **Formel von Bayes**,

$$P(B_i | A) = \frac{P(A | B_i) \cdot P(B_i)}{\sum_{j=1}^n P(A | B_j) \cdot P(B_j)}.$$

Analog für abzählbar viele Mengen $B_i, i \in \mathbb{N}$.

Beweis. Totale Wahrscheinlichkeit: Es gilt

$$A = \bigcup_{i=1}^n (A \cap B_i)$$

mit p.d. Mengen $A \cap B_i$. Somit folgt

$$P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A | B_i) \cdot P(B_i).$$

Formel von Bayes: Es gilt für jedes $i \in \{1, \dots, n\}$

$$P(B_i | A) = \frac{P(B_i \cap A)}{P(A)} \cdot \frac{P(B_i)}{P(B_i)} = \frac{P(A | B_i) \cdot P(B_i)}{P(A)}.$$

Die Behauptung folgt nun mit Hilfe der Formel von der totalen Wahrscheinlichkeit. □

23. Beispiel Situation:

- 3 Maschinen, $i = 1, 2, 3$
- Anteil an Tagesproduktion, $r_i = 60\%, 30\%, 10\%$
- Anteil defekter Produkte pro Maschine, $d_i = 1\%, 2\%, 3\%$

Fragen:

- Mit welcher Wahrscheinlichkeit ist ein zufällig gewähltes Produkt defekt?
- Mit welcher Wahrscheinlichkeit stammt ein defektes Produkt von Maschine 1?

Modell:

$$(i) \quad \Omega := \{(1, +), (1, -), (2, +), (2, -), (3, +), (3, -)\}$$

$$(ii) \quad \mathfrak{A} := \mathfrak{P}(\Omega)$$

(iii) P definiert durch

$$P(\{i, -\}) := r_i \cdot d_i, \quad P(\{i, +\}) := r_i \cdot (1 - d_i)$$

Für $D := \{(1, -), (2, -), (3, -)\}$, $M_i := \{(i, +), (i, -)\}$

$$P(M_i) = r_i, \quad P(D | M_i) = \frac{r_i \cdot d_i}{r_i} = d_i.$$

Häufig wie in diesem Beispiel: Modellierung durch **Vorgabe bedingter Wahrscheinlichkeiten**, etwa bei **Markov-Ketten**.

Man erhält

$$P(D) = d_1 r_1 + d_2 r_2 + d_3 r_3 = \frac{3}{200}$$

und

$$P(M_1 | D) = \frac{P(D | M_1) \cdot P(M_1)}{P(D)} = \frac{200}{3} \cdot d_1 r_1 = \frac{2}{5}.$$

24. Definition $A, B \in \mathfrak{A}$ **unabhängig**, falls

$$P(A \cap B) = P(A) \cdot P(B).$$

25. Bemerkung Falls $P(B) > 0$:

$$A, B \text{ unabhängig} \Leftrightarrow P(A | B) = P(A).$$

26. Beispiel Einmaliges Würfeln (wie üblich modelliert)

$$B := \{1, 2, 3, 4\}, \quad A_1 := \{2, 4, 6\}, \quad A_2 := \{1\}.$$

Dann gilt

$$P(A_1 | B) = \frac{1}{2}, \quad P(A_1) = \frac{1}{2},$$

d.h. A_1, B unabhängig. Ferner gilt

$$P(A_2 | B) = \frac{1}{4}, \quad P(A_2) = \frac{1}{6},$$

d.h. A_2, B nicht unabhängig.

27. Beispiel Zweimaliger Wurf einer fairen Münze,

$$\Omega := \{(Z, Z), (Z, K), (K, Z), (K, K)\},$$

$\mathfrak{A} := \mathfrak{P}(\Omega)$ und P Gleichverteilung auf Ω . Betrachte:

$$A_1 := \{(Z, Z), (Z, K)\} \quad \text{1. Wurf Z}$$

$$A_2 := \{(Z, K), (K, K)\} \quad \text{2. Wurf K}$$

$$A_3 := \{(Z, K), (K, Z)\} \quad \text{Würfe verschieden}$$

Es gilt $|A_i| = 2$ und $|A_i \cap A_j| = 1$ für $i \neq j$. Also:

$$A_1, A_2 \text{ unabh.}, \quad A_1, A_3 \text{ unabh.}, \quad A_2, A_3 \text{ unabh.}$$

Im folgenden $I := \{1, \dots, n\}$ oder $I := \mathbb{N}$.

28. Definition Folge $(A_i)_{i \in I}$ von Ereignissen **unabhängig**, falls für jede endliche Menge $\emptyset \neq J \subseteq I$ gilt

$$P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j).$$

Spezialfall $|I| = |J| = 2$ in Definition 24.

29. Bemerkung Falls $(A_i)_{i \in I}$ unabhängig, so folgt die **paarweise Unabhängigkeit**

$$\forall j_1, j_2 \in I, j_1 \neq j_2 : A_{j_1}, A_{j_2} \text{ unabhängig.}$$

Umkehrung ist falsch, wie das folgende Beispiel zeigt.

30. Beispiel Zweimaliger Wurf einer fairen Münze, siehe

Bsp. 27. Ereignisse A_1, A_2, A_3 nicht unabhängig, da

$$|A_1 \cap A_2 \cap A_3| = 1$$

Alternativ: $P(A_3|A_1 \cap A_2) = 1$, aber $P(A_3) = 1/2$, siehe Bem. 31.

31. Bemerkung Gelte $P(A_i) > 0$ für alle $i \in I$. Dann

$(A_i)_{i \in I}$ unabhängig, gdw. für alle $\emptyset \neq J_1, J_2 \subseteq I$ endlich mit $J_1 \cap J_2 = \emptyset$:

$$P\left(\bigcap_{j_1 \in J_1} A_{j_1} \mid \bigcap_{j_2 \in J_2} A_{j_2}\right) = P\left(\bigcap_{j_1 \in J_1} A_{j_1}\right)$$

Beweis. ÜBUNG M:G4, WInf:H6

□

3 Reellwertige Zufallsvariablen

Oft interessiert man sich (nur) für **spezielle Aspekte** eines Zufallsexperimentes. Dazu betrachtet man Abbildungen $\Omega \rightarrow \mathbb{R}$.

Bezeichnung Indikatorfunktion $1_U : V \rightarrow \mathbb{R}$ einer Teilmenge $U \subseteq V$ definiert durch

$$1_U(x) := \begin{cases} 1, & \text{falls } x \in U \\ 0, & \text{falls } x \notin U. \end{cases}$$

32. Beispiel Anzahl Anrufe in Callcenter an Tagen $1, \dots, n$:

$$\Omega := \mathbb{N}_0^n := \underbrace{\{(\omega_1, \dots, \omega_n) : \omega_i \in \mathbb{N}_0 \text{ für } i = 1, \dots, n\}}_{=: \omega}$$

$$\mathfrak{A} := \mathfrak{P}(\Omega)$$

- Anzahl Anrufe an Tag i , $X_i(\omega) := \omega_i$
- Gesamtanzahl der Anrufe, $X(\omega) := \sum_{i=1}^n \omega_i$
- Wurde an Tag i die Kapazitätsgrenze K überschritten?,
 $Y_i(\omega) := 1_{\{K+1, \dots\}}(\omega_i)$
- Anzahl der Tage, an denen Kapazitätsgrenze K überschritten, $Y(\omega) := \sum_{i=1}^n 1_{\{K+1, \dots\}}(\omega_i)$

33. Definition $X : \Omega \rightarrow \mathbb{R}$ (reellwertige) Zufallsvariable (ZV)
(auf W'raum $(\Omega, \mathfrak{A}, P)$), falls

$$\forall x \in \mathbb{R} : \{\omega \in \Omega : X(\omega) \leq x\} \in \mathfrak{A}.$$

Manchmal zugelassen: Funktionswerte $\pm\infty$. Vergleiche [Topologie](#), [stetige Abbildung](#). Siehe auch Lemma 39 und Lemma V.13.

34. Bemerkung

- ZVen sind Abbildungen!
- Für ZV sind die W'keiten $P(\{\omega \in \Omega : X(\omega) \leq x\})$ definiert
- Im Falle $\mathfrak{A} = \mathfrak{P}(\Omega)$ ist jede Abbildung $\Omega \rightarrow \mathbb{R}$ eine ZV

Bezeichnung Kurzschreibweise

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}$$

für Abbildung $X : \Omega \rightarrow \mathbb{R}$ und $B \subseteq \mathbb{R}$ sowie

$$\{X \leq x\} := \{\omega \in \Omega : X(\omega) \leq x\} = \{X \in]-\infty, x]\}$$

für $x \in \mathbb{R}$. Analog mit „ \geq “ usw.

Also **Urbilder** von Mengen: $\{X \in B\} = X^{-1}(B)$.

35. Beispiel Callcenter

$$\Omega := \mathbb{N}_0^n, \quad \mathfrak{A} := \mathfrak{P}(\Omega), \quad X_i(\omega) := \omega_i.$$

Es gilt (**Verknüpfung von ZVen**)

$$X = \sum_{i=1}^n X_i, \quad Y_i = 1_{\{K+1, \dots\}} \circ X_i, \quad Y = \sum_{i=1}^n Y_i.$$

Gängige Schreibweise $1_A(X_i)$ für $1_A \circ X_i$.

Spezielle Ereignisse: **Gesamtanzahl der Anrufe liegt zwischen 1000 und 2000**, $\{1000 \leq X \leq 2000\}$, **Kapazitätsgrenze K wurde nie überschritten**,

$$\bigcap_{i=1}^n \{X_i \leq K\} = \bigcap_{i=1}^n \{Y_i = 0\} = \{Y = 0\}.$$

36. Definition **Verteilungsfunktion** $F_X : \mathbb{R} \rightarrow [0, 1]$ einer ZV X auf $(\Omega, \mathfrak{A}, P)$ definiert durch

$$F_X(x) := P(\{X \leq x\}).$$

Gleichheit von ZVen in folgendem schwachen Sinn.

37. Definition ZVen X auf $(\Omega, \mathfrak{A}, P)$ und X' auf $(\Omega', \mathfrak{A}', P')$ **identisch verteilt**, falls $F_X = F_{X'}$.

38. Beispiel P Gleichverteilung auf $\Omega := \{1, \dots, n\}$ und $X(\omega) := \omega$. Dann:

$$P(\{X = x\}) = \begin{cases} 1/n, & \text{falls } x \in \{1, \dots, n\} \\ 0, & \text{sonst} \end{cases}$$

Betrachte den W'raum $(\Omega', \mathfrak{A}', P')$ zur Modellierung von **Pfeiltreffer auf Dartscheibe**, siehe Bsp. 15. Definiere

$A'_0 = \{(0, 0)\}$ und für $x = 1, \dots, n$ Sektoren:

$$A'_x := \{\rho(\cos \alpha, \sin \alpha) : \rho \in]0, r], \\ \alpha \in](x-1)/n \cdot 2\pi, x/n \cdot 2\pi]\}$$

Sei $X'(\omega')$ der getroffene Sektor, d.h. $X'(\omega') := x$, falls $\omega' \in A'_x$. Dann für $x = 1, \dots, n$

$$P'(\{X' = x\}) = P'(A'_x) = \lambda(A'_x)/\lambda(\Omega') = 1/n,$$

sowie $P'(\{X' = x\}) = 0$ für $x \in \mathbb{R} \setminus \{1, \dots, n\}$.

Also: $\forall x \in \mathbb{R} : P(\{X = x\}) = P'(\{X' = x\})$.

Beh.: X und X' sind identisch verteilt.

Beweis. Für $x \in \mathbb{R}$, $M :=]-\infty, x]$ und $D := \{1, \dots, n\}$

$$\begin{aligned} P(\{X \leq x\}) &= P(\{X \in M\} \cap \{X \in D\}) \\ &= \sum_{y \in D} P(\{X \in M\} \cap \{X = y\}) \\ &= \sum_{y \in M \cap D} \underbrace{P(\{X = y\})}_{=P'(\{X' = y\})} \\ &= P'(\{X' \leq x\}). \end{aligned}$$

Siehe auch Lemma III.23. Warnung!

□

Bezeichnung $\mathfrak{M} := \{M \subseteq \mathbb{R} : M \text{ oder } M^c \text{ Intervall}\}$

39. Lemma Für jede Zufallsvariable X auf $(\Omega, \mathfrak{A}, P)$ gilt

$$\forall M \in \mathfrak{M} : \{X \in M\} \in \mathfrak{A}.$$

Beweis. Seien $a, b \in \mathbb{R}$ mit $a \leq b$.

0. Für $M = \mathbb{R}$ gilt $\{X \in M\} = \Omega \in \mathfrak{A}$.

1. Für $M =]-\infty, b]$ gilt $\{X \in M\}$ nach Def. einer ZV.

2. Für $M =]a, \infty[$ gilt $M = \mathbb{R} \setminus]-\infty, a]$, also:

$$\{X \in M\} = \{X \in \mathbb{R}\} \setminus \{X \in]-\infty, a]\} \in \mathfrak{A}$$

3. Für $M =]-\infty, b[$ gilt $M = \bigcup_{n=1}^{\infty}]-\infty, b - \frac{1}{n}]$, also:

$$\{X \in M\} = \bigcup_{i=1}^{\infty} \{X \in]-\infty, b - \frac{1}{n}]\} \in \mathfrak{A}.$$

4. Für $M = [a, \infty[$ gilt $M = \mathbb{R} \setminus]-\infty, a[$, also:

$$\{X \in M\} = \{X \in \mathbb{R}\} \setminus \{X \in]-\infty, a[\} \in \mathfrak{A}.$$

5. Beschränkte Intervalle sind Durchschnitte der unter 1.–4. betrachteten Intervalle, und es gilt

$$\{X \in I_1 \cap I_2\} = \{X \in I_1\} \cap \{X \in I_2\}.$$

Demnach gilt die Aussage für alle Intervalle, unter Benutzung der Definition der σ -Algebra auch für deren Komplemente. □

40. Satz ZVen X auf $(\Omega, \mathfrak{A}, P)$ und X' auf $(\Omega', \mathfrak{A}', P')$
genau dann identisch verteilt, wenn

$$\forall M \in \mathfrak{M} : P(\{X \in M\}) = P'(\{X' \in M\}).$$

Beweis. Zu zeigen ist nur „ \Rightarrow “. Seien $a, b \in \mathbb{R}$ mit $a < b$.

1. Für $M :=]a, b]$ gilt

$$\begin{aligned} P(\{X \in M\}) &= P(\{X \leq b\} \setminus \{X \leq a\}) \\ &= P(\{X \leq b\}) - P(\{X \leq a\}) \\ &= P'(\{X' \leq b\}) - P'(\{X' \leq a\}) = P'(\{X' \in M\}). \end{aligned}$$

2. Für $M :=]a, b[$ gilt $M = \bigcup_{n=1}^{\infty}]a, b - 1/n]$. Also nach ÜBUNG M:G2 und 1.)

$$\begin{aligned} P(\{X \in M\}) &= P\left(\bigcup_{n=1}^{\infty} \{X \in]a, b - 1/n]\}\right) \\ &= \lim_{n \rightarrow \infty} P(\{X \in]a, b - 1/n]\}) \\ &= \lim_{n \rightarrow \infty} P'(\{X' \in]a, b - 1/n]\}) = P'(\{X' \in M\}). \end{aligned}$$

3. Für $M := \{a\}$ gilt

$$M = \bigcap_{n=1}^{\infty}]a - 1/n, a + 1/n[.$$

Also nach ÜBUNG M:G2 und 2.)

$$\begin{aligned} P(\{X \in M\}) &= \lim_{n \rightarrow \infty} P(\{X \in]a - 1/n, a + 1/n[\}) \\ &= \lim_{n \rightarrow \infty} P'(\{X' \in]a - 1/n, a + 1/n[\}) \\ &= P'(\{X' \in M\}). \end{aligned}$$

Für alle weiteren Typen von Mengen $M \in \mathfrak{M}$ nutze man die σ -Additivität, Additivität und die Rechenregel für Komplemente. □

41. Bemerkung Gemäß Satz 40 bestimmt die Verteilungsfunktion F_X die Wahrscheinlichkeiten $P(\{X \in M\})$ für $M \in \mathfrak{M}$ eindeutig. Siehe auch Satz V.55.

42. Definition Familie $(X_i)_{i \in I}$ von ZVen **identisch verteilt**, falls für alle $i, j \in I$ die ZVen X_i und X_j identisch verteilt.

Im folgenden:

- $I := \{1, \dots, n\}$ oder $I := \mathbb{N}$
- $(X_i)_{i \in I}$ Folge von ZVen auf $(\Omega, \mathfrak{A}, P)$

43. Definition $(X_i)_{i \in I}$ **unabhängig**, falls für jede Folge $(x_i)_{i \in I}$ in \mathbb{R} gilt: $(\{X_i \leq x_i\})_{i \in I}$ unabhängig.

44. Bemerkung Nach Definition äquivalent: für alle endlichen Mengen $J \subseteq I$ mit $|J| \geq 2$ und Folgen $(x_j)_{j \in J}$ in \mathbb{R} gilt

$$P\left(\bigcap_{j \in J} \{X_j \leq x_j\}\right) = \prod_{j \in J} P(\{X_j \leq x_j\}).$$

45. Beispiel **Zweimaliger Münzwurf**, siehe Bsp. 27. Betrachte für $i = 1, 2$ die ZVen $X_i(\omega) := 1_{\{z\}}(\omega_i)$.

Beh.: X_1 und X_2 sind unabhängig und identisch verteilt.

Beweis. Es gilt:

$$\{X_i \leq x\} = \begin{cases} \emptyset & \text{für } x < 0 \\ \{\omega \in \Omega : \omega_i = K\} & \text{für } 0 \leq x < 1 \\ \Omega & \text{für } x \geq 1 \end{cases}$$

Also:

$$P(\{X_i \leq x\}) = \begin{cases} 0 & , \text{ falls } x < 0 \\ \frac{1}{2} & , \text{ falls } 0 \leq x < 1 \\ 1 & , \text{ falls } x \geq 1 \end{cases}$$

Insbesondere sind X_1 und X_2 identisch verteilt.

□

Für die Unabhängigkeit ist noch zu zeigen:

$$P(\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}) = P(\{X_1 \leq x_1\}) \cdot P(\{X_2 \leq x_2\}) \quad (1)$$

Klar: (1) gilt, falls $x_1 < 0$, $x_2 < 0$, $x_1 \geq 1$ oder $x_2 \geq 1$. Für $x_i \in [0, 1[$ gilt:

$$\begin{aligned} P(\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}) &= P(\{(K, K)\}) = \frac{1}{4} \\ &= P(\{X_1 \leq x_1\}) \cdot P(\{X_2 \leq x_2\}) \end{aligned}$$

□

46. Satz $(X_i)_{i \in \mathbb{N}}$ genau dann unabhängig, wenn

$\forall n \in \mathbb{N} \forall M_1, \dots, M_n \in \mathfrak{M} :$

$$P\left(\bigcap_{j=1}^n \{X_j \in M_j\}\right) = \prod_{j=1}^n P(\{X_j \in M_j\}).$$

Analog für $I := \{1, \dots, n\}$. Siehe auch Satz V.21.

Beweis. TUTORIUM T1:2

□

Stochastische Modelle beruhen sehr häufig auf einer unabhängigen Folge von identisch verteilten ZVen. Abkürzung: **iid** für **independent and identically distributed**.

Kap. III Diskrete Wahrscheinlichkeitsräume und diskrete Zufallsvariablen

1. Wahrscheinlichkeitsfunktionen
2. Elementare Kombinatorik
3. Produkträume
4. Diskrete Zufallsvariablen

In diesem Kapitel

stochastische Modelle für Zufallsexperimente, bei denen die Menge der möglichen Ausgänge abzählbar ist.

1 Wahrscheinlichkeitsfunktionen

1. Definition $(\Omega, \mathfrak{A}, P)$ **diskreter Wahrscheinlichkeitsraum**, falls Ω abzählbar und $\mathfrak{A} = \mathfrak{P}(\Omega)$.

In diesem Abschnitt Ω und \mathfrak{A} wie oben. Erste Frage: Konstruktion von Wahrscheinlichkeitsmaßen auf \mathfrak{A} ?

2. Definition $f : \Omega \rightarrow \mathbb{R}_+$ **Wahrscheinlichkeitsfunktion** (auf Ω), falls

$$\sum_{\omega \in \Omega} f(\omega) = 1.$$

Interpretation: **Punktmassen** $f(\omega)$. Graphische Darstellung: **Stabdiagramm**.

3. Satz W 'maße und W 'funktionen

(i) Jede Wahrscheinlichkeitsfunktion $f : \Omega \rightarrow \mathbb{R}_+$ definiert durch

$$P(A) := \sum_{\omega \in A} f(\omega), \quad A \subseteq \Omega, \quad (1)$$

ein Wahrscheinlichkeitsmaß auf \mathfrak{A} , und es gilt

$$f(\omega) = P(\{\omega\}), \quad \omega \in \Omega. \quad (2)$$

(ii) Für jedes Wahrscheinlichkeitsmaß P auf \mathfrak{A} definiert (2) eine Wahrscheinlichkeitsfunktion f auf Ω , und es gilt (1) .

Beweis. Ad (i): Offenbar gilt $P(A) \in [0, 1]$ und $P(\Omega) = 1$.

Ferner gilt für $A := \bigcup_{i=1}^{\infty} A_i$ mit p.d. Mengen $A_i \subseteq \Omega$

$$P(A) = \sum_{\omega \in A} f(\omega) = \sum_{i=1}^{\infty} \sum_{\omega \in A_i} f(\omega) = \sum_{i=1}^{\infty} P(A_i),$$

ggf. aufgrund der absoluten Konvergenz der Reihen.

Ad (ii): klar.

□

Modellierung durch Wahl von f . Siehe bereits Bsp. II.13 und II.23. Dazu:

- **kombinatorische Methoden**, oft ausgehend von Gleichverteilungsannahmen
- **statistische Schätzung**, siehe bereits Bsp. II.13

4. Beispiel Gleichverteilung auf endlicher Menge Ω entspricht der Wahrscheinlichkeitsfunktion

$$f(\omega) := \frac{1}{|\Omega|}, \quad \omega \in \Omega.$$

2 Elementare Kombinatorik

Abzählung von endlichen Mengen zur Berechnung von Wahrscheinlichkeiten (unter Gleichverteilungsannahme).

Erinnerung $n! := 1 \cdot \dots \cdot n$ für $n \in \mathbb{N}$ und $0! := 1$, sowie

- **Binomialkoeffizienten** für $n \in \mathbb{N}_0$ und $k \in \{0, \dots, n\}$

$$\binom{n}{k} := \frac{n!}{k!(n-k)!}$$

und Rekursionsformel für $k \neq 0$

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$

- **Binomischer Lehrsatz**: für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}_0$

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} \cdot a^k \cdot b^{n-k}$$

Im folgenden N, N_1, \dots, N_k endliche nicht-leere Mengen
und $n := |N|$.

5. Satz

$$|N_1 \times \dots \times N_k| = |N_1| \cdot \dots \cdot |N_k|$$

6. Bemerkung Obiger Satz mit $N = N_1 = \dots = N_k$:

Anzahl der **Stichproben** vom Umfang k aus der Menge N
unter Berücksichtigung der Reihenfolge mit Zurücklegen
beträgt n^k .

Ebenso: Anzahl der Abbildungen $\{1, \dots, k\} \rightarrow N$.

Beweis von Satz 5. Induktion über k .

Für $x \in N_{k+1}$ sei

$$A_x := \{(\omega_1, \omega_2, \dots, \omega_{k+1}) \in N_1 \times \dots \times N_{k+1} : x = \omega_{k+1}\}.$$

Dann $A_x \cap A_y = \emptyset$ für $x \neq y$ sowie

$$N_1 \times \dots \times N_{k+1} = \bigcup_{x \in N_{k+1}} A_x.$$

Ferner, unter Verwendung der Induktionsannahme,

$$|A_x| = |N_1 \times \dots \times N_k| = |N_1| \cdot \dots \cdot |N_k|.$$

Fazit

$$|N_1 \times \dots \times N_{k+1}| = \sum_{x \in N_{k+1}} |N_1 \times \dots \times N_k| = |N_{k+1}| \cdot |N_1| \cdot \dots \cdot |N_k|.$$



7. Satz

$$\begin{aligned} & |\{(\omega_1, \dots, \omega_k) \in N^k : \omega_1, \dots, \omega_k \text{ paarw. verschieden}\}| \\ &= \frac{n!}{(n-k)!} = n \cdot (n-1) \cdots (n-k+1). \end{aligned}$$

Für $k = n$: Anzahl der Permutationen von N ist $n!$

8. Bemerkung Obiger Satz: Anzahl der **Stichproben** vom Umfang k aus der Menge N **unter Berücksichtigung der Reihenfolge ohne Zurücklegen** beträgt $n!/(n-k)!$

Ebenso: Anzahl der Injektionen $\{1, \dots, k\} \rightarrow N$

Beweis von Satz 7. Seien $2 \leq k + 1 \leq n$ und $x \in N$. Setze

$$Q := \{(\omega_1, \dots, \omega_{k+1}) \in N^{k+1} : \omega_1, \dots, \omega_{k+1} \text{ paarw. verschieden}\}$$

und

$$A_x := \{\omega \in Q : \omega_{k+1} = x\}.$$

Dann

$$\begin{aligned} |A_x| &= |\{(\omega_1, \dots, \omega_k) \in (N \setminus \{x\})^k : \omega_1, \dots, \omega_k \text{ paarw. verschieden}\}| \\ &= (n - 1) \cdot \dots \cdot (n - 1 - k + 1) = (n - 1) \cdot \dots \cdot (n - k). \end{aligned}$$

Also

$$\begin{aligned} |Q| &= (n - 1) \cdot \dots \cdot (n - k) \cdot n \\ &= \frac{n!}{(n - (k + 1))!}. \end{aligned}$$

□

9. Beispiel $\Omega := N^k$ mit Gleichverteilung P . Gesucht:
Wahrscheinlichkeit, daß in Stichprobe $\omega \in \Omega$ **mindestens**
2 Komponenten übereinstimmen. Also

$$A := \{(\omega_1, \dots, \omega_k) \in \Omega : \exists i \neq j : \omega_i = \omega_j\}.$$

Es gilt $P(A) = 1 - P(A^c)$ und

$$\begin{aligned} P(A^c) &= \frac{|A^c|}{|\Omega|} = \frac{n!}{(n-k)! \cdot n^k} = \frac{n \cdot \dots \cdot (n - (k-1))}{n \cdot \dots \cdot n} \\ &= 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{n}\right). \end{aligned}$$

Für $n = 365$ (Geburtstagszwillinge) ergibt sich
näherungsweise

k	4	16	22	23	40	64
$P(A)$	0,016	0,284	0,476	0,507	0,891	0,997

Implizit angenommen: Geburtstage unabhängig und jeweils gleichverteilt
auf $N = \{1, \dots, 365\}$, siehe Bsp. 16 und Bem. 19.

10. Satz Für $k \in \{0, \dots, n\}$ gilt

$$|\{K \subseteq N : |K| = k\}| = \binom{n}{k}.$$

11. Bemerkung Obiger Satz: Anzahl der **Stichproben** vom Umfang k aus Menge N **ohne Berücksichtigung der Reihenfolge ohne Zurücklegen**.

Vergleich der Sätze 7 und 10. Es gilt

$$\binom{n}{k} \cdot k! = \frac{n!}{(n-k)!}.$$

Interpretation: Teilmenge auswählen und anordnen.

Beweis von Satz 10. Induktion nach $n = |N|$.

Sei $|N'| = n + 1 \in \mathbb{N}$. Die Behauptung gilt offenbar für $k = 0$ und $k = n + 1$, also fortan $k \in \{1, \dots, n\}$.

Fixiere $x \in N'$, setze $N = N' \setminus \{x\}$. Dann folgt

$$\begin{aligned} |\{K' \subseteq N' : |K'| = k\}| &= |\{K' \subseteq N' : |K'| = k \wedge x \in K'\}| \\ &\quad + |\{K' \subseteq N' : |K'| = k \wedge x \notin K'\}| \\ &= |\{K \subseteq N : |K| = k - 1\}| \\ &\quad + |\{K \subseteq N : |K| = k\}| \\ &= \binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k} \end{aligned}$$

□

12. Beispiel **Lotto**: Gleichverteilung auf

$$\Omega := \{K \subseteq \{1, \dots, 49\} : |K| = 6\}.$$

Also für jedes $\omega \in \Omega$

$$P(\{\omega\}) = 1 / \binom{49}{6} = 7,15 \dots \cdot 10^{-8}.$$

3 Produkträume

Gegeben:

diskrete W'räume $(\Omega_1, \mathfrak{A}_1, P_1), \dots, (\Omega_n, \mathfrak{A}_n, P_n)$

Gesucht:

Modell für „unabh. Hintereinanderausführung“ der
Einzelexperimente ([Produktexperiment](#))

13. Beispiel n -maliges Würfeln, n Geburten, . . .

Fragwürdig bei Callcenter an n Tagen.

Definiere

$$\Omega := \Omega_1 \times \cdots \times \Omega_n,$$

$$\mathfrak{A} := \mathfrak{P}(\Omega),$$

$$f(\omega) := f_1(\omega_1) \cdot \dots \cdot f_n(\omega_n), \quad \omega \in \Omega,$$

wobei f_i die zu P_i gehörige W'funktion auf Ω_i .

14. Lemma f ist Wahrscheinlichkeitsfunktion auf Ω .

Beweis. Klar: $f \geq 0$. Ferner, ggf. aufgrund der absoluten Konvergenz,

$$\begin{aligned} \sum_{\omega \in \Omega} f(\omega) &= \sum_{\omega_1 \in \Omega_1} \cdots \sum_{\omega_n \in \Omega_n} f_1(\omega_1) \cdot \dots \cdot f_n(\omega_n) \\ &= \sum_{\omega_1 \in \Omega_1} f_1(\omega_1) \cdot \dots \cdot \sum_{\omega_n \in \Omega_n} f_n(\omega_n) = 1. \quad \square \end{aligned}$$

15. Definition Sei P das durch f definierte W'maß auf \mathfrak{A} .

Dann heißt $(\Omega, \mathfrak{A}, P)$ das **Produkt der W'räume** $(\Omega_i, \mathfrak{A}_i, P_i)$

und P das **Produkt der W'maße** P_i .

16. Beispiel Für endliche Mengen Ω_i und Gleichverteilungen

P_i ist das Produktmaß P die **Gleichverteilung** auf Ω , siehe

Satz 5.

So etwa für n -maliges Würfeln:

$$\Omega = \{1, \dots, 6\}^n$$

und

$$P(\{\omega\}) = 1/6^n$$

für $\omega \in \Omega$.

17. Beispiel Geschlecht von n Neugeborenen,

$$\Omega_i := \{W, M\}, f_i(W) := p, f_i(M) := 1 - p.$$

Also

$$f(\omega) = p^{k(\omega)} \cdot (1 - p)^{n-k(\omega)}$$

mit

$$k(\omega) := |\{i \in \{1, \dots, n\} : \omega_i = W\}|.$$

Im folgenden: $(\Omega, \mathfrak{A}, P)$ Produkt der W'räume $(\Omega_i, \mathfrak{A}_i, P_i)$.

Zurück zu den Einzelexperimenten durch die **Projektionen**

$X_i : \Omega \rightarrow \Omega_i$, d.h.

$$X_i(\omega) := \omega_i.$$

Siehe etwa Bsp. II.35.

18. Satz Für $A_1 \subseteq \Omega_1, \dots, A_n \subseteq \Omega_n$ gilt

$$P\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) = \prod_{i=1}^n P(\{X_i \in A_i\}) = \prod_{i=1}^n P_i(A_i).$$

Beweis. Es gilt (vgl. Beweis Lemma 14)

$$\begin{aligned} P\left(\bigcap_{i=1}^n \{X_i \in A_i\}\right) &= P(A_1 \times \dots \times A_n) \\ &= \sum_{\omega \in A_1 \times \dots \times A_n} f(\omega) \\ &= \sum_{\omega_1 \in A_1} \dots \sum_{\omega_n \in A_n} f_1(\omega_1) \cdot \dots \cdot f_n(\omega_n) \\ &= P_1(A_1) \cdot \dots \cdot P_n(A_n). \end{aligned}$$

Die Wahl von $A_j = \Omega_j$ für $j \neq i$ zeigt

$$P(\{X_i \in A_i\}) = P_i(A_i).$$

□

19. Bemerkung Ziel erreicht. Genauer

- (i) Produktraum-Modell beinhaltet Modelle der Einzelexperimente, da $P(\{X_i \in A_i\}) = P_i(A_i)$ für $i \in \{1, \dots, n\}$ und $A_i \subseteq \Omega_i$.
- (ii) Falls $\Omega_1, \dots, \Omega_n \subseteq \mathbb{R}$ (oder bei Verwendung eines allgemeineren Begriffs von ZVen und deren Unabhängigkeit), so sind X_1, \dots, X_n unabhängige Zufallsvariablen.
Beweis: Wähle $A_i =]-\infty, x_i] \cap \Omega_i$ bzw. Ω_i in Satz 18.
Spezialfall $P_1 = \dots = P_n$ liefert iid-Folge X_1, \dots, X_n .

4 Diskrete Zufallsvariablen

Im folgenden X, X_1, \dots Zufallsvariablen auf $(\Omega, \mathfrak{A}, P)$.

20. Definition X **diskrete Zufallsvariable**, falls

$P(\{X \in D\}) = 1$ für eine abzählbare Menge $D \subset \mathbb{R}$.

21. Bemerkung $(\Omega, \mathfrak{A}, P)$ diskret $\Rightarrow X(\Omega)$ abzählbar
 $\Rightarrow X$ diskret.

22. Beispiel **Pfeiltreffer auf Dartscheibe**, X Nummer des getroffenen Sektors, siehe Beispiele II.15 und II.38.

23. Lemma Diskrete ZVen X, X' genau dann identisch verteilt, wenn

$$\forall x \in \mathbb{R} : P(\{X = x\}) = P'(\{X' = x\}).$$

Beweis. „ \Rightarrow “: Wende Satz II.40 an.

„ \Leftarrow “: Betrachte $D := X(\Omega) \cup X'(\Omega')$ im Beweis in Bsp. II.38 □

24. Definition X **Bernoulli-verteilt** mit Parameter $p \in [0, 1]$, falls $P(\{X = 1\}) = p$ und $P(\{X = 0\}) = 1 - p$.

Bez.: $X \sim \mathbf{B}(1, p)$.

25. Beispiel n gleichartige Produkte, voneinander unabhängig

- mit Wahrscheinlichkeit p funktionstüchtig
- mit Wahrscheinlichkeit $1 - p$ defekt.

Hierbei $p \in [0, 1]$, z.B. empirisch bestimmt als relative Häufigkeit.

Gesucht: Wahrscheinlichkeit, daß **genau k Produkte funktionstüchtig** sind.

Daraus durch Summation: W'keit, daß mindestens k Produkte funktionstüchtig sind.

Konkretes Modell: Produktexperiment mit $\Omega_i := \{0, 1\}$ und

$$f_i(\omega_i) := \begin{cases} p, & \text{falls } \omega_i = 1 \\ 1 - p, & \text{falls } \omega_i = 0. \end{cases}$$

Also $\Omega := \{0, 1\}^n$ Menge der Produktionsergebnisse und für $\omega \in \Omega$

$$f(\omega) := f_1(\omega_1) \cdot \dots \cdot f_n(\omega_n).$$

Berechne bzgl. des Produktmaßes P

$$P(\{\omega \in \Omega : \sum_{i=1}^n \omega_i = k\}).$$

Abstraktes Modell: X_1, \dots, X_n iid mit $X_1 \sim \mathbf{B}(1, p)$.

Im konkreten Modell: $X_i(\omega) = \omega_i$.

Anzahl funktionstüchtiger Produkte

$$X := \sum_{i=1}^n X_i.$$

Berechne

$$P(\{X = k\}).$$

Modellierung analog bei n -fachem Münzwurf oder n Geburten, X Anzahl der geworfenen K bzw. Anzahl der weiblichen Neugeborenen.

26. Satz Seien X_1, \dots, X_n iid mit $X_1 \sim \mathbf{B}(1, p)$. Ferner sei

$$X := \sum_{i=1}^n X_i.$$

Dann gilt für $k \in \{0, \dots, n\}$

$$P(\{X = k\}) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}. \quad (3)$$

27. Definition X **binomialverteilt** mit Parametern $n \in \mathbb{N}$ und $p \in [0, 1]$, falls (3) für alle $k \in \{0, \dots, n\}$ gilt.

Bez.: $X \sim \mathbf{B}(n, p)$.

Beweis von Satz 26. Es gilt

$$\begin{aligned} P(\{(X_1, \dots, X_n) \in \{0, 1\}^n\}) &= P\left(\bigcap_{i=1}^n \{X_i \in \{0, 1\}\}\right) \\ &= \prod_{i=1}^n P(\{X_i \in \{0, 1\}\}) = 1. \end{aligned}$$

Setze

$$A_k := \{x \in \{0, 1\}^n : \sum_{i=1}^n x_i = k\}.$$

Gemäß Satz 10

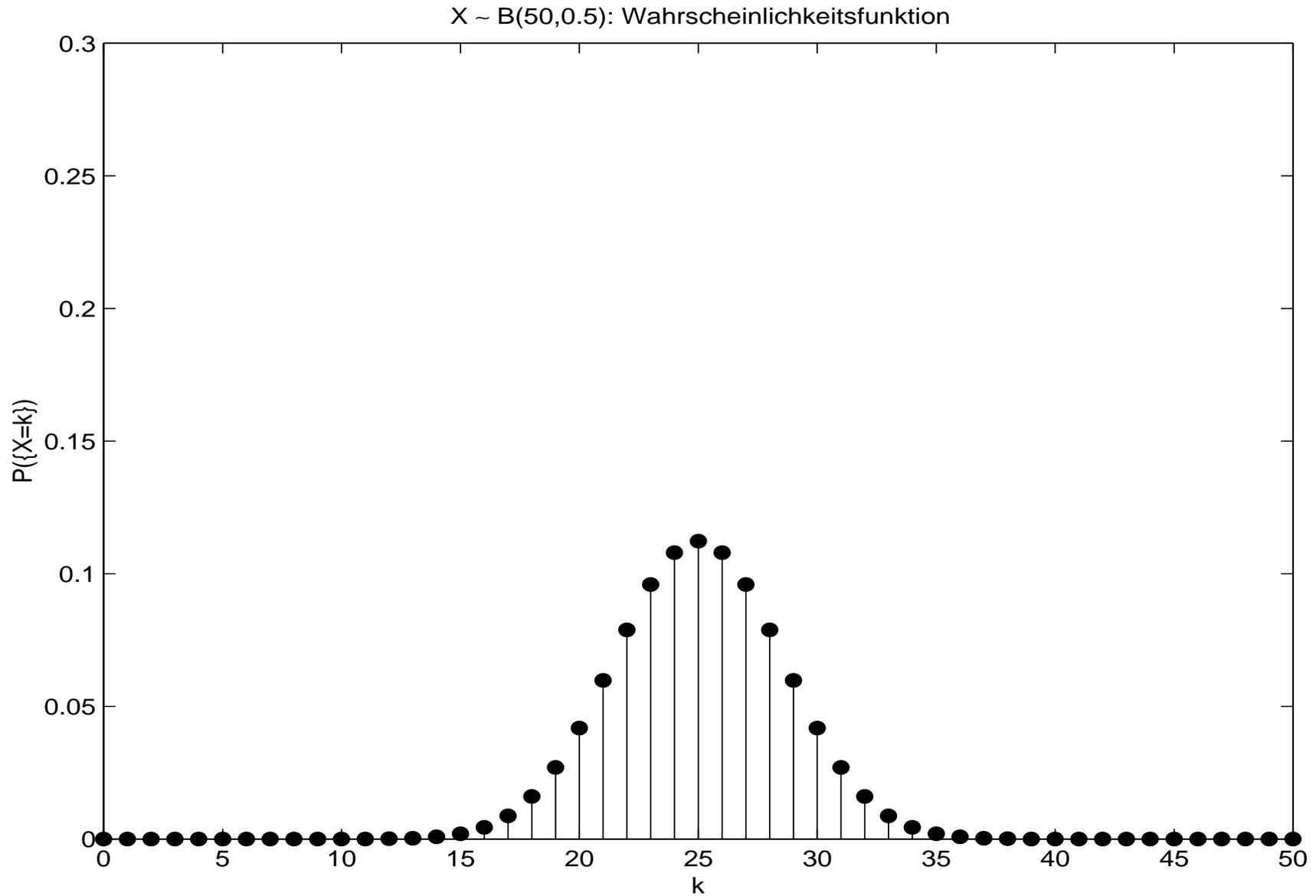
$$|A_k| = \binom{n}{k}.$$

Damit folgt für $k \in \{0, \dots, n\}$

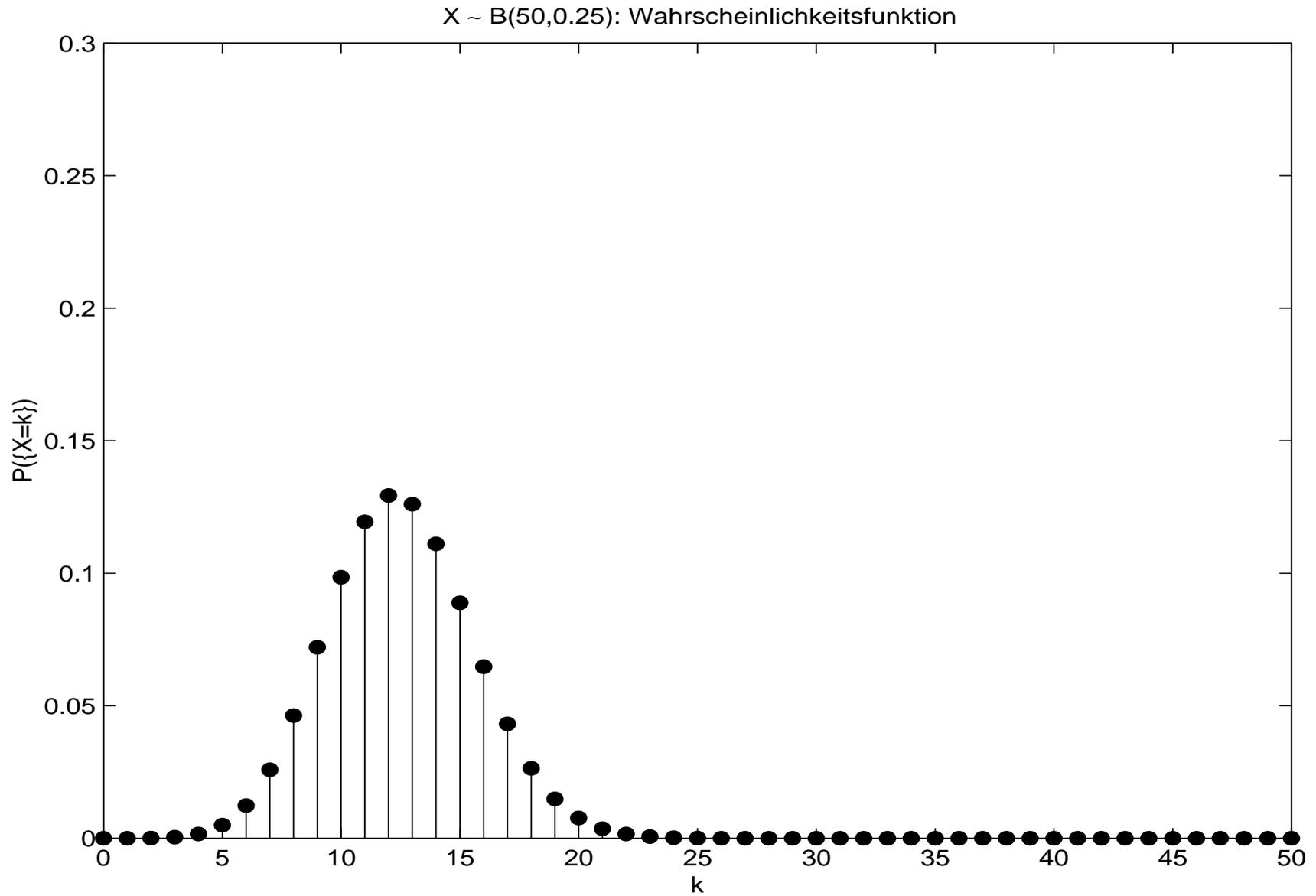
$$\begin{aligned} P(\{X = k\}) &= P\left(\{X = k\} \cap \bigcup_{x \in \{0,1\}^n} \{(X_1, \dots, X_n) = x\}\right) \\ &= \sum_{x \in \{0,1\}^n} P(\{X = k\} \cap \{(X_1, \dots, X_n) = x\}) \\ &= \sum_{x \in A_k} \prod_{i=1}^n P(\{X_i = x_i\}) \\ &= |A_k| \cdot p^k \cdot (1 - p)^{n-k} \\ &= \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}. \end{aligned}$$

□

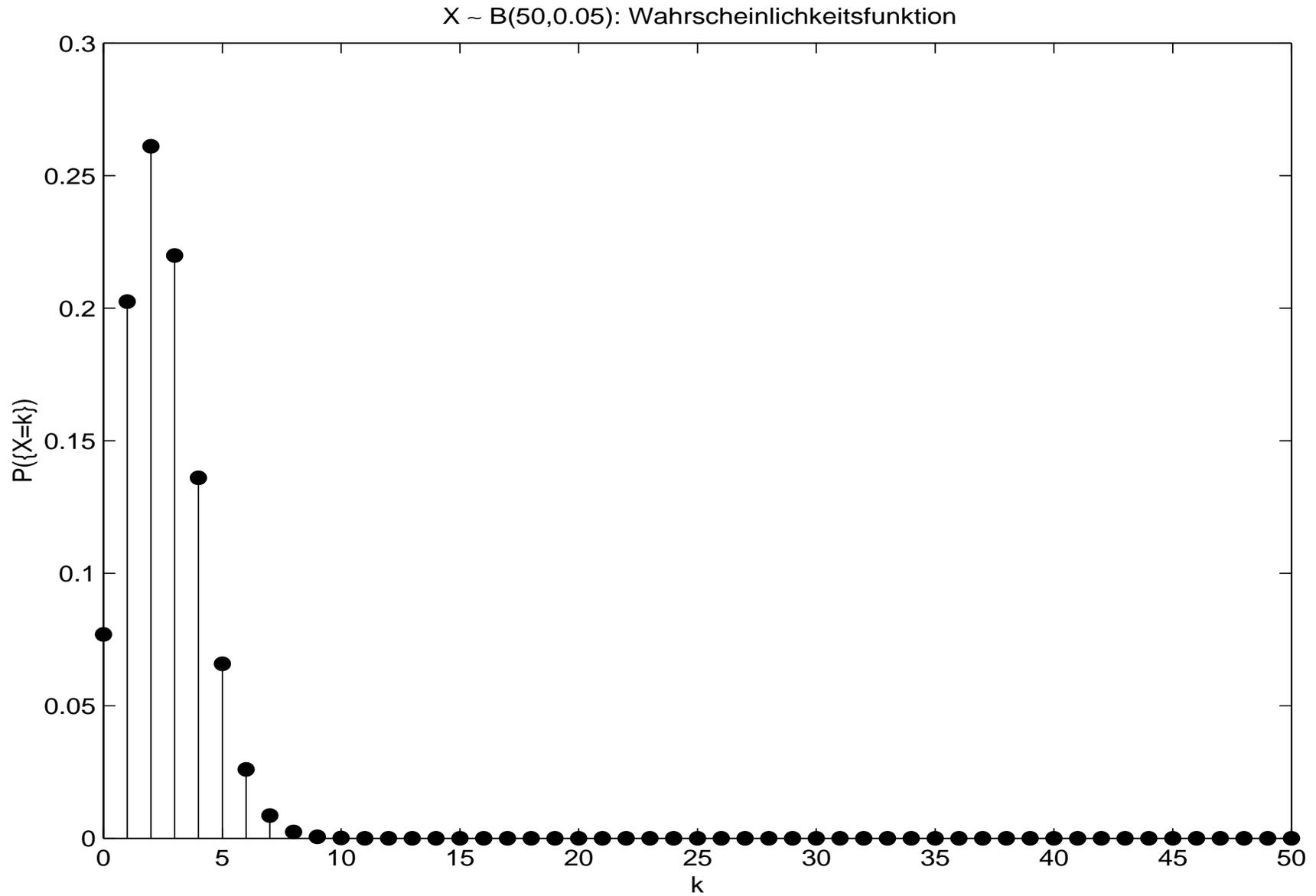
28. Beispiel $X \sim B(50, 0.5)$



$$X \sim \mathbf{B}(50, 0.25)$$



$$X \sim \mathbf{B}(50, 0.05)$$



29. Beispiel n Produkte, davon n_0 defekt. Gesucht:

Wahrscheinlichkeit, daß bei Auswahl von k Produkten genau ℓ Produkte defekt.

Modell: Gleichverteilung P auf

$$\Omega := \{K \subseteq N : |K| = k\}.$$

Berechne $P(A_\ell)$ für

$$A_\ell := \{K \in \Omega : |K \cap N_0| = \ell\},$$

wobei $N_0 \subseteq N$ fest gewählt mit $|N_0| = n_0$, d.h. bestimme $|\Omega|$ und $|A_\ell|$.

Es gilt: $|\Omega| = \binom{n}{k}$ und für

$$\ell \in \{\max(0, k - (n - n_0)), \dots, \min(n_0, k)\},$$

daß

$$\begin{aligned} |A_\ell| &= |\{(K_0, K_1) : K_0 \subseteq N_0, |K_0| = \ell, \\ &\quad K_1 \subseteq N \setminus N_0, |K_1| = k - \ell\}| \\ &= \binom{n_0}{\ell} \cdot \binom{n - n_0}{k - \ell}. \end{aligned}$$

Also

$$P(A_\ell) = \binom{n_0}{\ell} \cdot \binom{n - n_0}{k - \ell} / \binom{n}{k}.$$

Hiermit auch die Wahrscheinlichkeit, beim Skat genau 3 Asse zu erhalten:

$$\binom{4}{3} \cdot \binom{28}{7} / \binom{32}{10} = 66/899 = 0,0734\dots$$

Ausblick auf **statistische Fragestellungen**. Bekannt

- Gesamtanzahl n der Produkte,
- Stichprobengröße k ,
- Anzahl ℓ defekter Produkte in Stichprobe.

Unbekannt

- Gesamtanzahl n_0 defekter Produkte.

Aufgaben:

- (i) Schätze n_0
- (ii) Entscheide, ob $n_0/n \leq 0.02$

30. Definition X **hypergeometrisch verteilt** mit Parametern

$n \in \mathbb{N}$, $n_0 \in \{0, \dots, n\}$, $k \in \{1, \dots, n\}$, falls

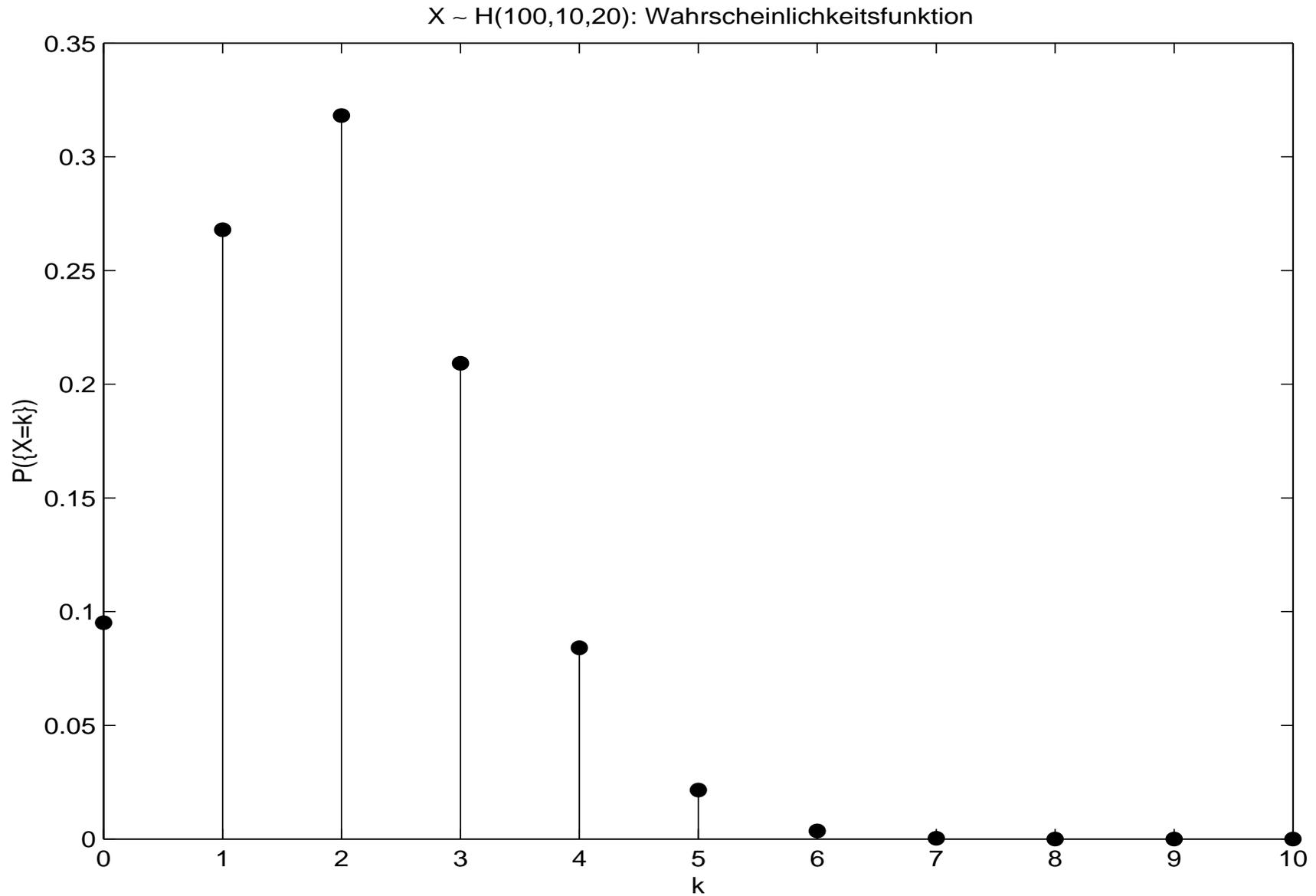
$$P(\{X = \ell\}) = \binom{n_0}{\ell} \cdot \binom{n - n_0}{k - \ell} / \binom{n}{k}$$

für

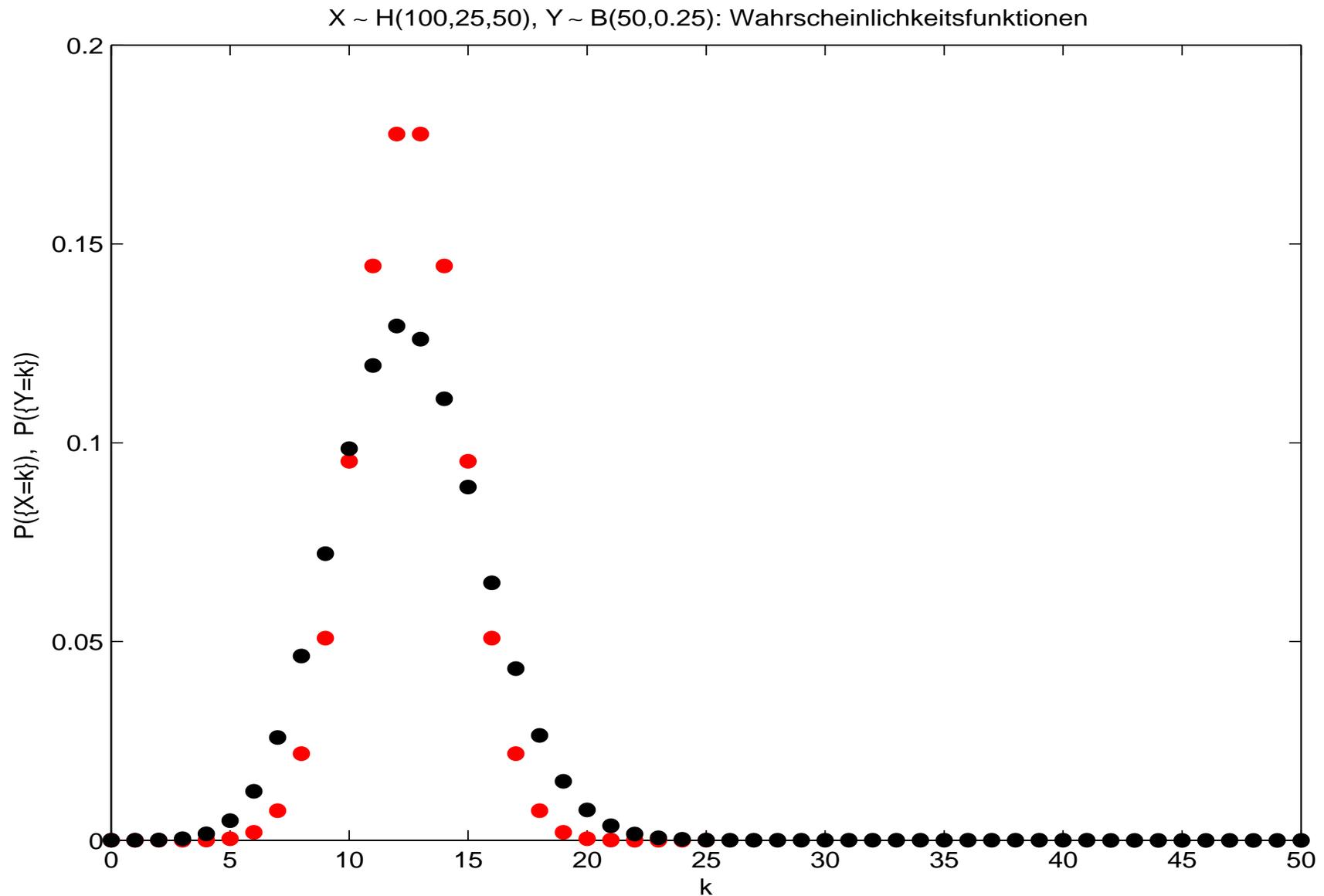
$$\ell \in \{\max(0, k - (n - n_0)), \dots, \min(n_0, k)\}.$$

Bez.: $X \sim \mathbf{H}(n, n_0, k)$.

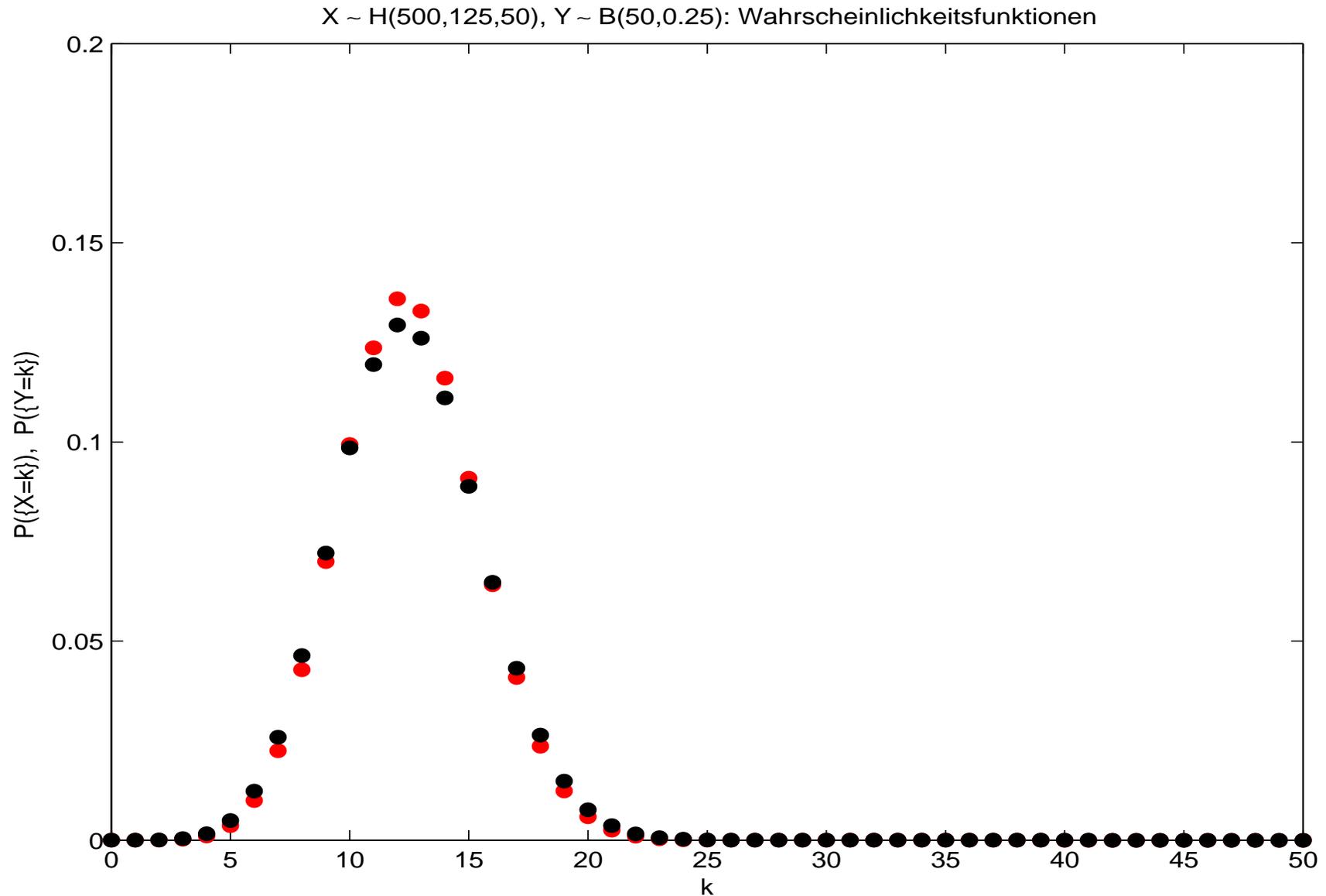
31. Beispiel $X \sim \mathbf{H}(100, 10, 20)$



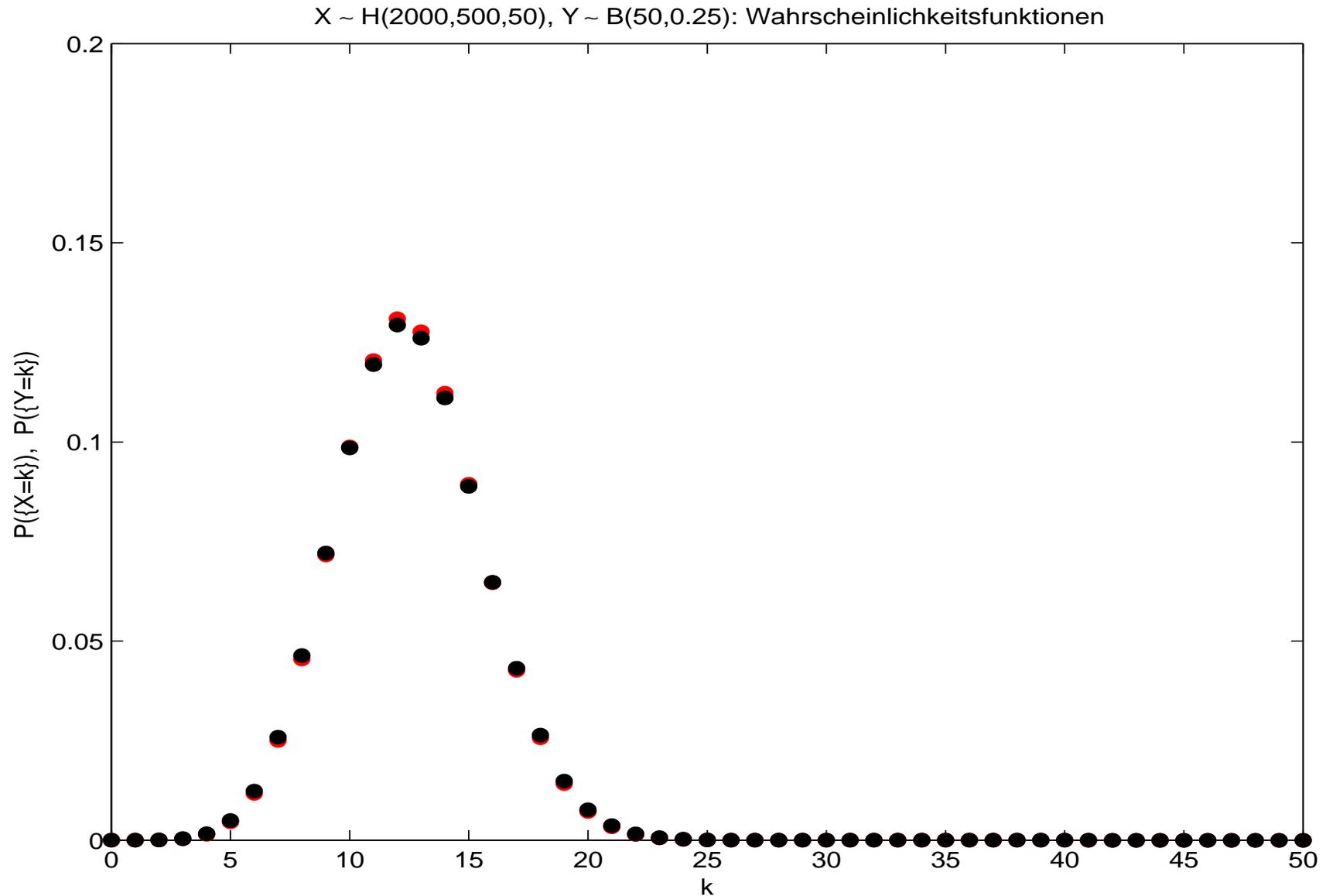
$X \sim \mathbf{H}(100, 25, 50)$ (rot), $Y \sim \mathbf{B}(50, 0.25)$ (schwarz)



$X \sim \mathbf{H}(500, 125, 50)$ (rot), $Y \sim \mathbf{B}(50, 0.25)$ (schwarz)



$X \sim \mathbf{H}(2000, 500, 50)$ (rot), $Y \sim \mathbf{B}(50, 0.25)$ (schwarz)



Vermutung: Konvergenz. Bestätigung: ÜBUNG M:H9, WInf:H9.

32. Beispiel n unabhängige Würfe auf Dartscheibe. Gesucht: Wahrscheinlichkeit, daß im k -ten Wurf erstmals oberes rechtes Viertel getroffen wird.

Abstraktes Modell: X_1, \dots, X_n iid mit $X_1 \sim \mathbf{B}(1, p)$, wobei $p := 1/4$.

Zeitpunkt des ersten Treffers im oberen rechten Viertel:

$$T_n(\omega) := \begin{cases} \inf\{i \in \{1, \dots, n\} : X_i(\omega) = 1\}, & \text{falls } \{\dots\} \neq \emptyset \\ 0, & \text{falls } \forall i : X_i(\omega) = 0 \end{cases}$$

Für $k \in \{1, \dots, n\}$ gilt

$$\{T_n = k\} = \bigcap_{i=1}^{k-1} \{X_i = 0\} \cap \{X_k = 1\},$$

also, unabhängig von n ,

$$\begin{aligned} P(\{T_n = k\}) &= \prod_{i=1}^{k-1} P(\{X_i = 0\}) \cdot P(\{X_k = 1\}) \\ &= (1 - p)^{k-1} \cdot p. \end{aligned}$$

Ebenso $P(\{T_n = 0\}) = (1 - p)^n$, so daß

$$\lim_{n \rightarrow \infty} P(\{T_n = 0\}) = 0,$$

33. Definition X **geometrisch verteilt** mit Parameter $p \in]0, 1]$, falls

$$\forall k \in \mathbb{N} : P(\{X = k\}) = p \cdot (1 - p)^{k-1}.$$

Bez.: $X \sim \mathbf{G}(p)$.

34. Bemerkung Sei $p \in]0, 1]$. Auf $\Omega := \mathbb{N}$ definiert

$$f(\omega) := p \cdot (1 - p)^{\omega-1}$$

eine Wahrscheinlichkeitsfunktion. Auf dem zugehörigen W'raum $(\Omega, \mathfrak{P}(\Omega), P)$ gilt $X \sim \mathbf{G}(p)$ für

$$X(\omega) := \omega.$$

35. Bemerkung Sei $p \in]0, 1]$. Für iid ZVen X_1, X_2, \dots mit $X_1 \sim \mathbf{B}(1, p)$ sei

$$T_\infty(\omega) := \begin{cases} \inf\{i \in \mathbb{N} : X_i(\omega) = 1\}, & \text{falls } \{\dots\} \neq \emptyset \\ 0, & \text{falls } \forall i : X_i(\omega) = 0 \end{cases}$$

Die Rechnung aus Bsp. 32 zeigt für $k \in \mathbb{N}$

$$P(\{T_\infty = k\}) = p \cdot (1 - p)^{k-1},$$

so daß $T_\infty \sim \mathbf{G}(p)$.

Beachte, daß $P(\{T_\infty = 0\}) = 0$, da $\sum_{k=1}^{\infty} P(\{T_\infty = k\}) = 1$.

36. Bemerkung Frage: Gibt es einen W'raum $(\Omega, \mathfrak{A}, P)$ und darauf eine unendliche Folge X_1, X_2, \dots von iid ZVen mit $X_1 \sim \mathbf{B}(1, p)$?

Antwort: Ja.

Beweis ist nicht-trivial.

Verwendet Maßtheorie, z.B. Existenz des Lebesgue-Maßes auf $[0, 1]$ für den Fall $p = 1/2$.

Siehe Billingsley (1995), Probability and Measure, Seiten 1–4, oder Vorlesung „Probability Theory“.

37. Satz Sei $X_n \sim \mathbf{B}(n, p_n)$ mit $p_n \in]0, 1[$, und gelte $\lim_{n \rightarrow \infty} n \cdot p_n = \lambda$ für $\lambda > 0$. Dann

$$\forall k \in \mathbb{N}_0 : \lim_{n \rightarrow \infty} P(\{X_n = k\}) = \exp(-\lambda) \cdot \frac{\lambda^k}{k!}.$$

Beweis. Für $n \geq k$

$$\begin{aligned} & \binom{n}{k} \cdot p_n^k \cdot (1 - p_n)^{n-k} \\ &= \frac{\lambda^k}{k!} \cdot \underbrace{\left(\frac{n \cdot p_n}{\lambda}\right)^k}_{\rightarrow 1} \cdot \underbrace{\frac{\left(1 - \frac{n \cdot p_n}{n}\right)^n}{(1 - p_n)^k}}_{\rightarrow \exp(-\lambda)} \cdot \underbrace{\prod_{i=0}^{k-1} \frac{n-i}{n}}_{\rightarrow 1}. \end{aligned}$$

□

38. Definition X **Poisson-verteilt** mit Parameter $\lambda > 0$, falls

$$\forall k \in \mathbb{N}_0 : P(\{X = k\}) = \exp(-\lambda) \cdot \frac{\lambda^k}{k!}.$$

Bez.: $X \sim \mathbf{P}(\lambda)$.

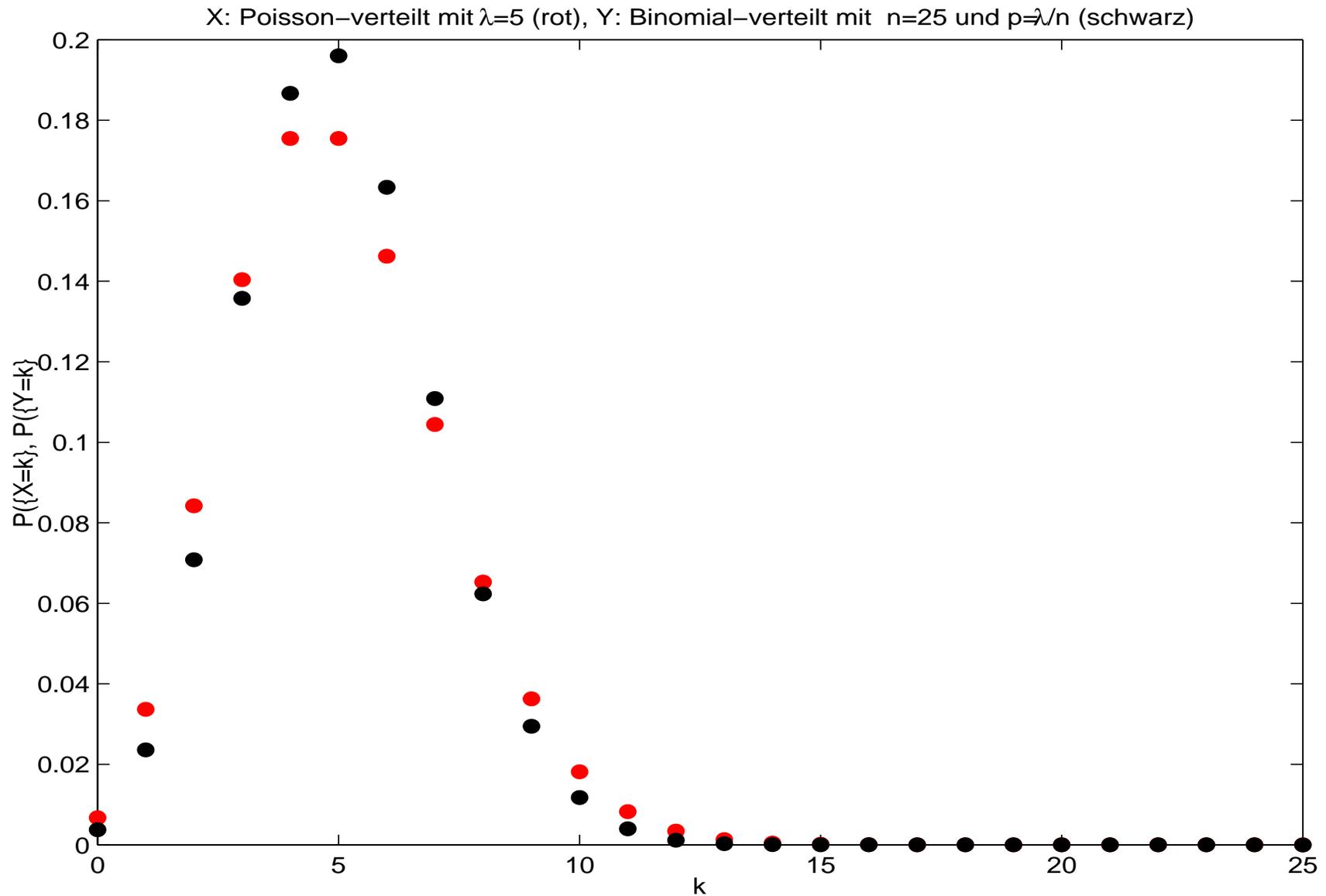
39. Bemerkung In Satz 37: Approximation von $\mathbf{B}(n, p_n)$ durch $\mathbf{P}(\lambda)$, falls n „groß“ und p_n „klein“. Siehe Bsp. VI.6 zur Bedeutung von $n \cdot p_n$.

Siehe Hesse (2003, p. 190–192) zur Approximationsgüte.

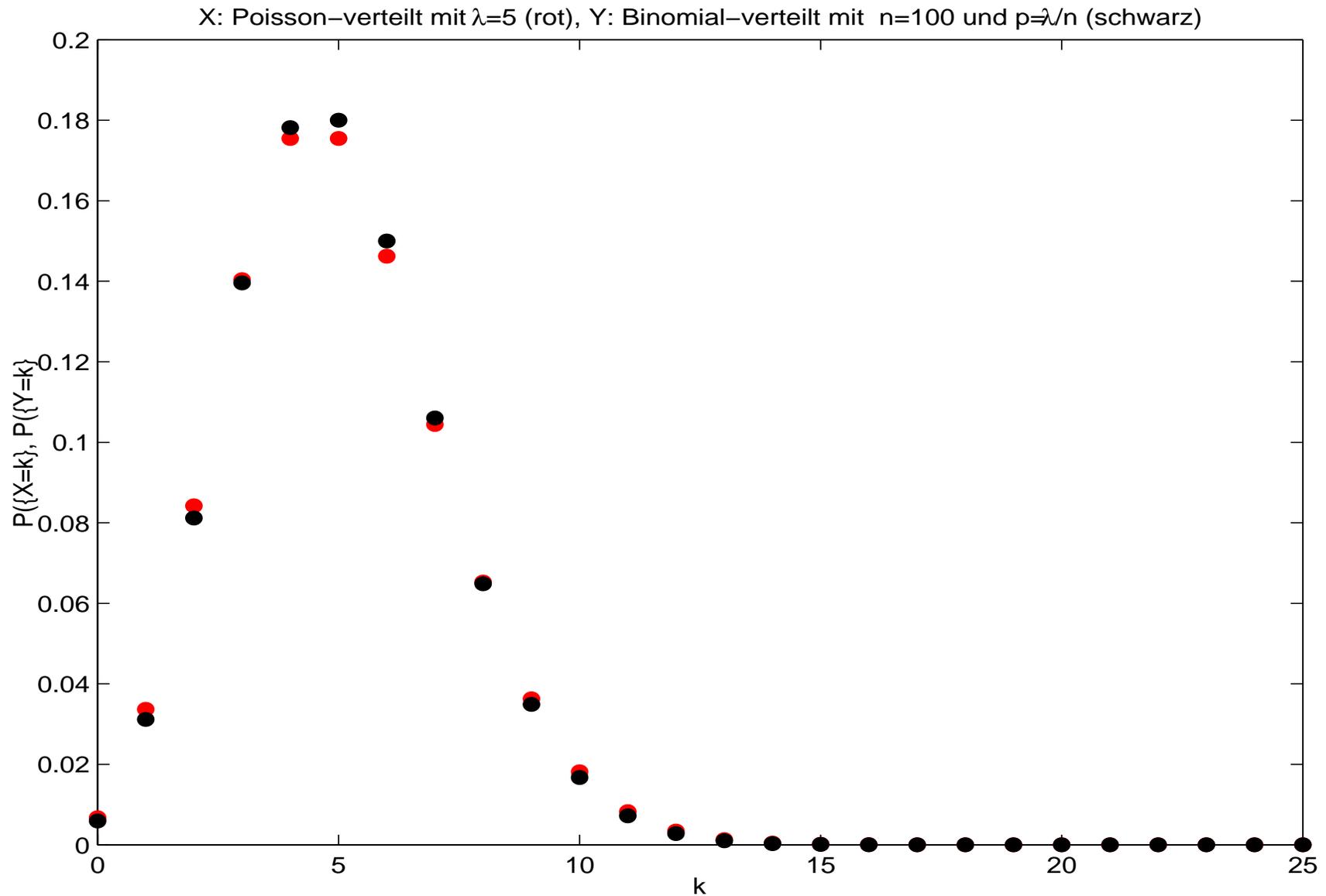
Anwendung: Modellierung der Anzahl von

- Druckfehlern in Manuskript,
- Anrufen in Call-Center pro Tag,
- radioaktiven Zerfällen pro Zeiteinheit.

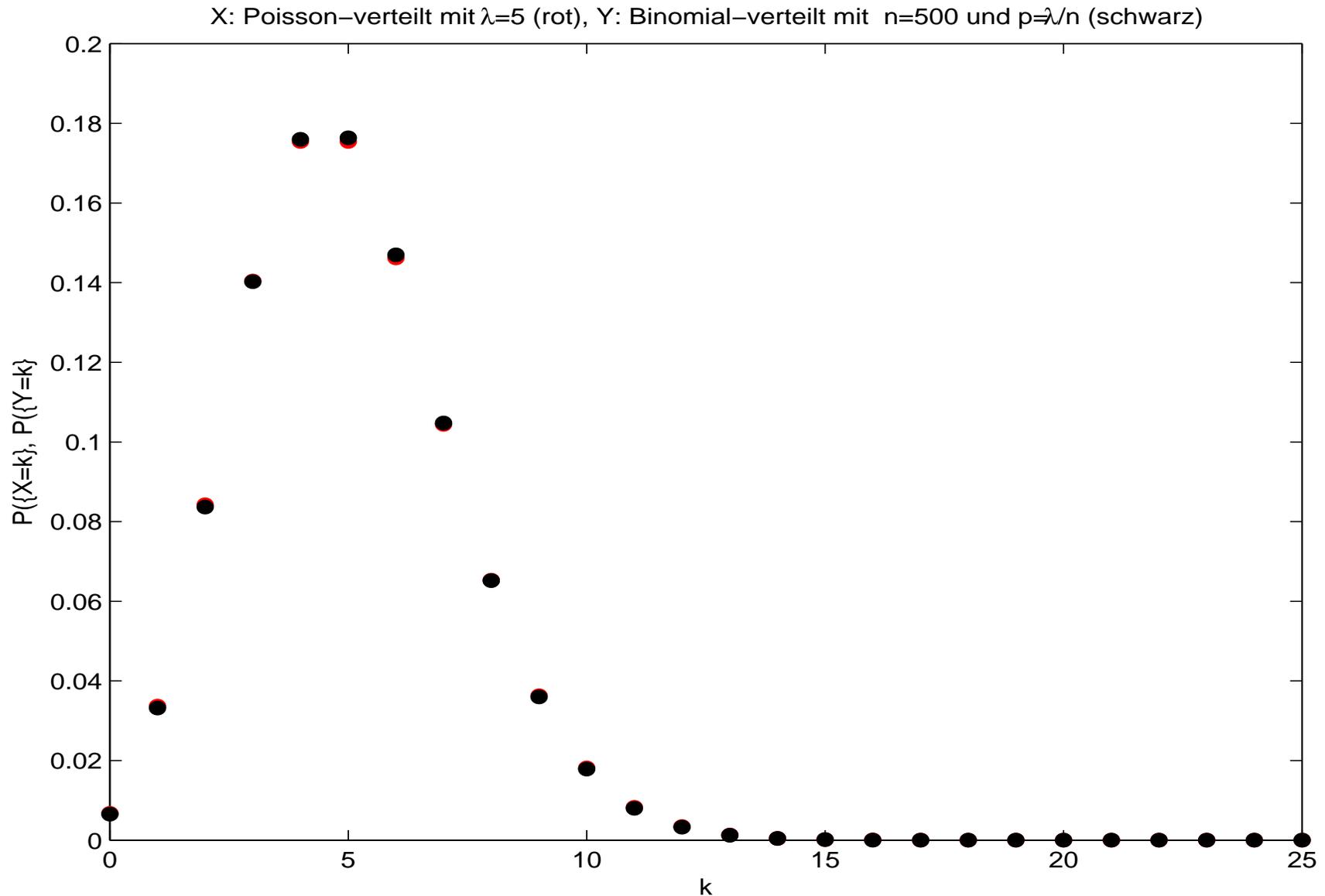
$X \sim \mathbf{P}(5)$ (rot), $Y \sim \mathbf{B}(n, 5/n)$ mit $n = 25$ (schwarz)



$X \sim \mathbf{P}(5)$ (rot), $Y \sim \mathbf{B}(n, 5/n)$ mit $n = 100$ (schwarz)



$X \sim \mathbf{P}(5)$ (rot), $Y \sim \mathbf{B}(n, 5/n)$ mit $n = 500$ (schwarz)



Kap. IV Stochastische Simulation und der Hauptsatz der Mathematischen Statistik

1. Problemstellungen und Methoden
2. Borels starkes Gesetz der großen Zahlen
3. Der Satz von Glivenko-Cantelli
4. Zufallszahlen
5. Anwendung: Simulation einer Irrfahrt

1 Problemstellungen und Methoden

Problem A: Gegeben: ZV X' auf $(\Omega', \mathfrak{A}', P')$ und $M \in \mathfrak{M}$.

Berechne $P'(\{X' \in M\})$.

Problem B: Gegeben: Daten $x_1, \dots, x_n \in \mathbb{R}$ als Resultate „unabhängiger“ Wiederholungen eines Zufallsexperimentes.

Gesucht: Stochastisches Modell für Einzelexperiment.

1. Beispiel Problem A. Erfolgswahrscheinlichkeit einer Strategie beim Patience-Spiel:

- P' Gleichverteilung auf der Menge Ω' aller Permutationen von $\{1, \dots, 52\}$ (Kartenverteilungen)
- $X' := 1_A$, wobei A die Menge aller Kartenverteilungen, bei denen die Strategie gewinnt, und $M := \{1\}$

Somit

$$P'(\{X' \in M\}) = P'(A) = |A|/|\Omega'|.$$

Hier: $|\Omega'| = 52! = 8,06 \dots \cdot 10^{67}$, also Ω' sehr groß.

Ulam (1946), siehe Los Alamos Science, Special Issue.

2. Beispiel Problem A. **Durchgang von Neutronen durch Materie**, siehe PROJEKTOR . Schließlich tritt einer dieser Fälle ein:

- Neutron wird von Abschirmung absorbiert, $X'(\omega) := 0$
- Neutron wird von Abschirmung reflektiert, $X'(\omega) := 1$
- Neutron passiert Abschirmung, $X'(\omega) := 2$

Gesucht ist $P'(\{X' = 2\})$.

Hier: $(\Omega', \mathfrak{A}', P')$ und X' „kompliziert“.

3. Beispiel Problem B. **Geschlecht eines Neugeborenen.**

Anzahl der Daten $n = 25\ 171\ 123$ in Bsp. II.13.

Modell: $\mathbf{B}(1, p)$ -verteilte ZV mit unbekanntem $p \in]0, 1[$.

Gesucht ist p .

4. Beispiel Problem B. **Callcenter.**

Daten: x_1, \dots, x_n Anzahl Anrufe an Tagen $i = 1, \dots, n$.

Modell: $\mathbf{P}(\lambda)$ -verteilte ZV mit unbekanntem $\lambda > 0$.

Gesucht ist λ .

Stochastische Simulation zur Lösung von A:

- Konstruiere iid-ZVen X_1, \dots, X_n auf $(\Omega, \mathfrak{A}, P)$, wobei X' und X_1 identisch verteilt
- „Erzeuge“ eine **Realisierung** x_1, \dots, x_n der ZVen X_1, \dots, X_n , d.h. für ein $\omega \in \Omega$

$$x_1 = X_1(\omega), \dots, x_n = X_n(\omega)$$

- Approximiere $P'(\{X' \in M\}) = P(\{X_1 \in M\})$ durch die **relative Häufigkeit**

$$1/n \cdot \sum_{i=1}^n 1_M(x_i)$$

Formale Beschreibung von Problem B:

- X_1, \dots, X_n iid-ZVen auf $(\Omega, \mathfrak{A}, P)$, wobei P unbekannt.
Jede dieser ZVen modelliert ein Einzelexperiment.
- Annahme: Daten x_1, \dots, x_n sind eine **Realisierung** der ZVen X_1, \dots, X_n , d.h. für ein $\omega \in \Omega$

$$x_1 = X_1(\omega), \dots, x_n = X_n(\omega)$$

Methode zur Lösung von Problem B:

- Schätze $P(\{X_1 \in M\})$ für $M \in \mathfrak{M}$ durch die **relative Häufigkeit**

$$1/n \cdot \sum_{i=1}^n 1_M(x_i)$$

5. Beispiel Problem B. [Geschlecht eines Neugeborenen](#),
siehe Hesse (2003, p. 24). PROJEKTOR

6. Beispiel Problem B. [Münzwurf](#), siehe Hesse (2003, p. 241).
PROJEKTOR

Befund: In beiden Beispielen scheinen die Folgen der
relativen Häufigkeiten zu „konvergieren“.

2 Borels starkes Gesetz der großen Zahlen

7. Satz Sei $(Z_i)_{i \in \mathbb{N}}$ iid mit $Z_1 \sim \mathbf{B}(1, p)$. Dann

$$P\left(\left\{\omega \in \Omega : \lim_{n \rightarrow \infty} 1/n \cdot \sum_{i=1}^n Z_i(\omega) = p\right\}\right) = 1.$$

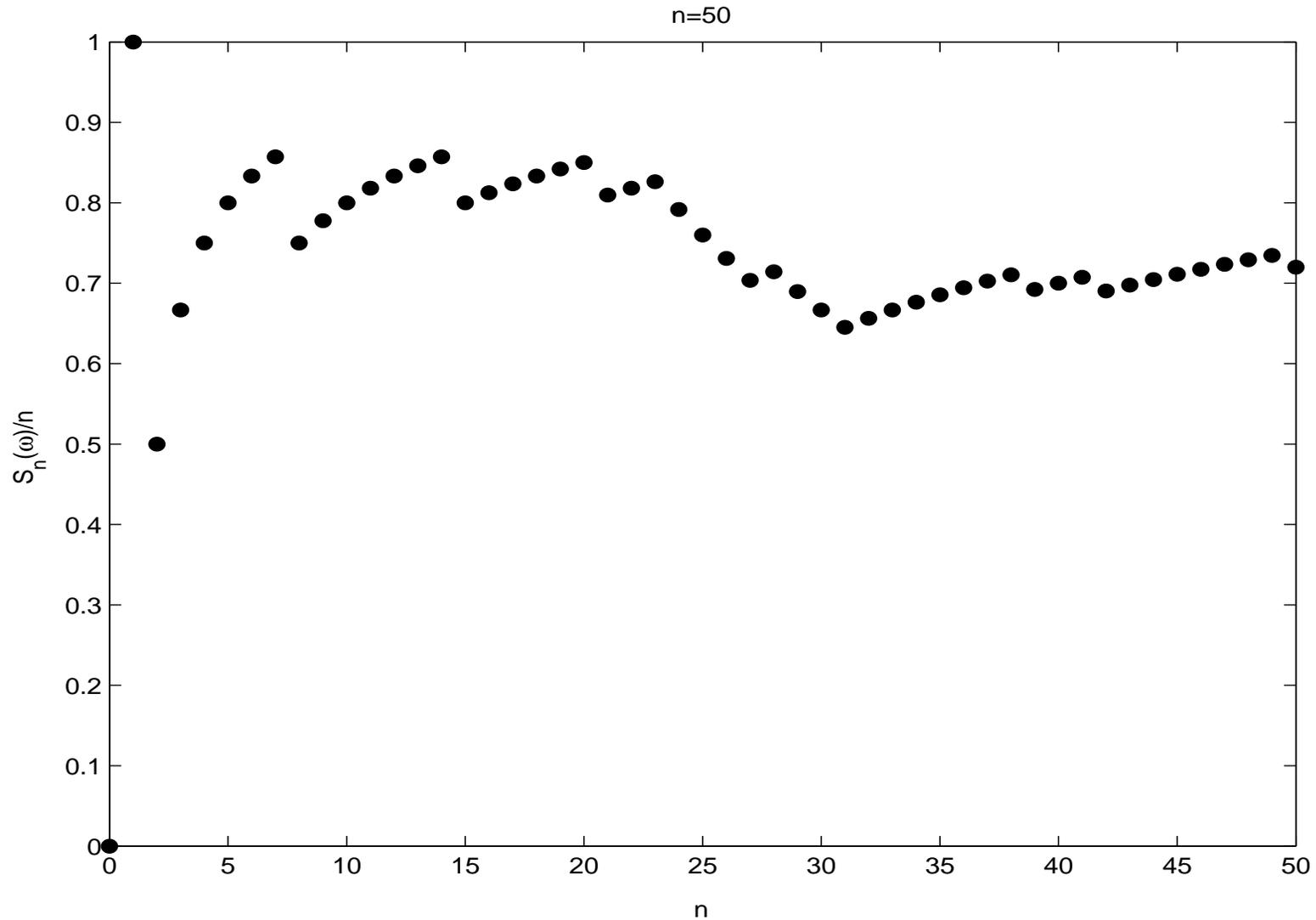
Beweis. Skizze PROJEKTOR, Details TUTORIUM T4. □

Siehe auch Satz VI.35. **Sprechweise:** eine Eigenschaft gilt für **fast alle** ω bzw. **fast sicher**, falls sie für alle ω aus einer Menge $A \in \mathfrak{A}$ mit $P(A) = 1$ gilt.

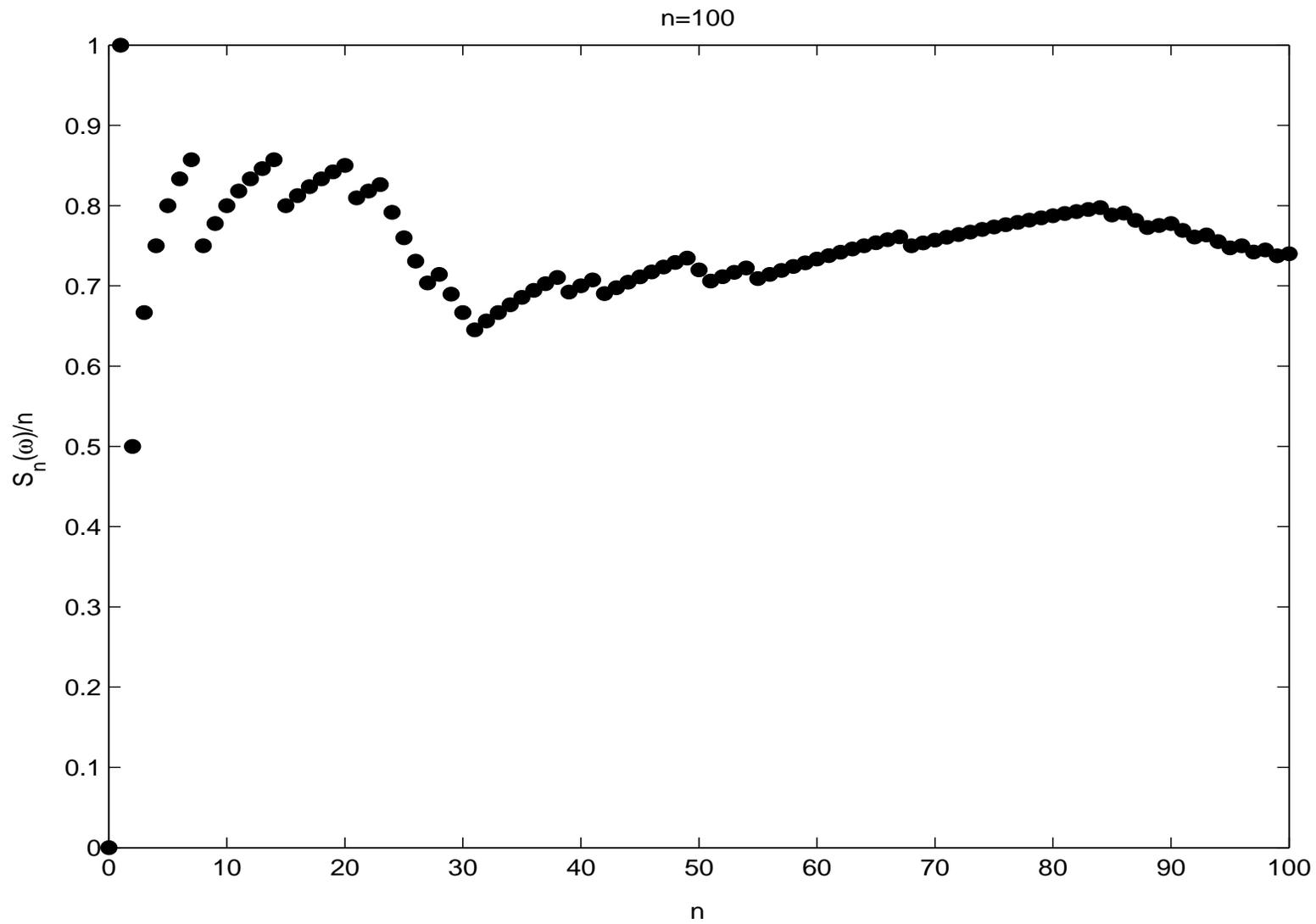
8. Beispiel Unabhängige, unendliche Folge von Münzwürfen.

9. Beispiel Simulationsbeispiele

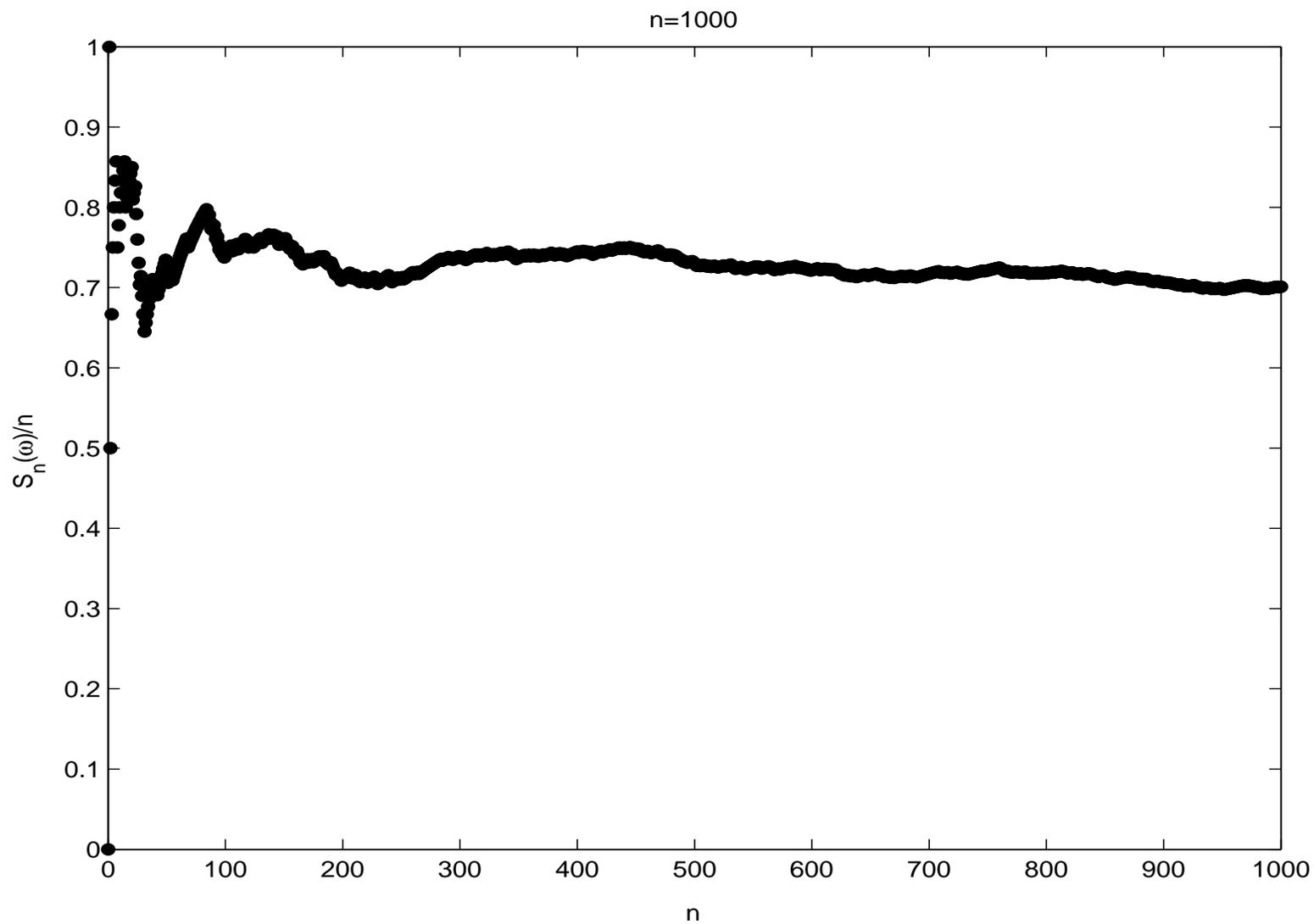
$$Z_1 \sim \mathbf{B}(1, 0.7), n = 50.$$



$$Z_1 \sim \mathbf{B}(1, 0.7), n = 100.$$



$$Z_1 \sim \mathbf{B}(1, 0.7), n = 1000.$$



10. Korollar Sei $(X_i)_{i \in \mathbb{N}}$ iid und $M \in \mathfrak{M}$. Dann gilt für fast alle ω

$$\lim_{n \rightarrow \infty} 1/n \cdot \sum_{i=1}^n 1_M(X_i(\omega)) = P(\{X_1 \in M\})$$

Beweis. Für

$$Z_i := 1_M \circ X_i$$

gilt $(Z_i)_{i \in \mathbb{N}}$ iid mit $Z_1 \sim \mathbf{B}(1, p)$ und $p = P(\{X_1 \in M\})$,
siehe ÜBUNG M:H8, WInf:G6. Wende Satz 7 an. \square

11. Bemerkung Damit

- **bei Problem A:** stochastische Simulation liefert für fast alle ω für „große“ Anzahl n von Wiederholungen „gute“ Näherung. Genauer gilt für eine Menge $A \in \mathfrak{A}$ mit $P(A) = 1$

$$\forall \omega \in A \forall \varepsilon > 0 \exists n_0 = n_0(\omega, \varepsilon) \dots$$

- **bei Problem B:** theoretisches Gegenstück zur beobachteten Konvergenz.

3 Der Satz von Glivenko-Cantelli

Nun: gleichzeitige Betrachtung der Mengen $M :=]-\infty, x]$,
siehe Korollar 10.

Vermutung: Konvergenz gegen Verteilungsfunktion F_{X_1} in
geeignetem Sinn.

12. Bemerkung Falls X diskret mit $P(\{X \in D\}) = 1$ für
 $D \subset \mathbb{R}$ abzählbar,

$$F_X(x) = \sum_{y \in]-\infty, x] \cap D} P(\{X = y\}).$$

Siehe Bsp. II.38.

13. Beispiel P Gleichverteilung auf $\Omega := \{1, \dots, n\}$ und $X(\omega) := \omega$. Dann:

$$F_X(x) = |\{X \leq x\}|/|\Omega| = |[1, x] \cap \Omega|/n$$
$$= \begin{cases} 0, & \text{falls } x < 1 \\ \lfloor x \rfloor / n, & \text{falls } 1 \leq x < n \\ 1, & \text{falls } x \geq n \end{cases}$$

14. Satz Für jede Verteilungsfunktion F_X gilt:

- (i) F_X ist monoton wachsend
- (ii) F_X ist rechtsseitig stetig
- (iii) $\lim_{x \rightarrow \infty} F_X(x) = 1$ und $\lim_{x \rightarrow -\infty} F_X(x) = 0$

Ferner gilt für alle $x \in \mathbb{R}$

$$P(\{X = x\}) = F_X(x) - \lim_{y \rightarrow x^-} F_X(y).$$

und

$$P(\{X = x\}) = 0 \quad \Leftrightarrow \quad F_X \text{ stetig in } x.$$

Beweis. ÜBUNG M:G9.

□

15. Definition Für $q \in]0, 1[$ heißt

$$\inf\{v \in \mathbb{R} : F_X(v) \geq q\}$$

q -Quantil der Verteilungsfunktion F_X .

Speziell für $q = 1/2$: z **Median**.

16. Beispiel PROJEKTOR

17. Lemma z genau dann q -Quantil von F_X , wenn

$$F_X(z) \geq q \quad \text{und} \quad \forall y \in \mathbb{R} : y < z \Rightarrow F_X(y) < q.$$

Beweis. Da F_X rechtsseitig stetig und monoton wachsend, ex. $\tilde{z} \in \mathbb{R}$ mit

$$\{v \in \mathbb{R} : F_X(v) \geq q\} = [\tilde{z}, \infty[.$$



Approximation (Problem A) bzw. Schätzung (Problem B) der zugrundeliegenden Vert'funktion.

18. Definition **Empirische Verteilungsfunktion**

$F_n(\cdot; x_1, \dots, x_n) : \mathbb{R} \rightarrow [0, 1]$ zu $x_1, \dots, x_n \in \mathbb{R}$ definiert durch

$$F_n(x; x_1, \dots, x_n) := 1/n \cdot \sum_{i=1}^n 1_{]-\infty, x]}(x_i).$$

19. Beispiel PROJEKTOR

Im folgenden $(X_i)_{i \in \mathbb{N}}$ iid. Sei $F := F_{X_1}$ die Vert'funktion jeder der ZVen X_i . Bei Problem A: F „schwer“ zu berechnen. Bei Problem B: F unbekannt.

20. Bemerkung Korollar 10 zeigt für alle $x \in \mathbb{R}$, daß für fast alle ω gilt

$$\lim_{n \rightarrow \infty} F_n(x; X_1(\omega), \dots, X_n(\omega)) = F(x).$$

Verschärfung im folgenden **Satz von Glivenko-Cantelli (Hauptsatz der Mathematischen Statistik)**: fast sicher gleichmäßige Konvergenz.

21. Satz Für fast alle ω gilt

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n(x; X_1(\omega), \dots, X_n(\omega)) - F(x)| = 0.$$

Beweis. PROJEKTOR: vorab Spezialfall F stetig, streng mon. wachsend.

Bezeichnungen:

$$G^-(x) := \lim_{y \rightarrow x^-} G(y),$$

$$F_n(x, \omega) := F_n(x; X_1(\omega), \dots, X_n(\omega)),$$

$$\mathfrak{E} := \{A \in \mathfrak{A} : P(A) = 1\}.$$

Fixiere $k \in \mathbb{N}$. Betrachte für $\ell = 1, \dots, k - 1$ die ℓ/k -Quantile $z_{\ell,k}$ von F , setze ferner $z_{0,k} := -\infty$ und $z_{k,k} := \infty$.

Korollar 10 und Satz 14 zeigen

$$\forall \ell \in \{1, \dots, k - 1\} \exists A \in \mathfrak{E} \forall \omega \in A :$$

$$\lim_{n \rightarrow \infty} F_n^{(-)}(z_{\ell,k}, \omega) = F^{(-)}(z_{\ell,k}).$$

Für

$$\Delta_{n,k}(\omega) := \max_{\ell \in \{0, \dots, k\}} \max \left(|F_n(z_{\ell,k}, \omega) - F(z_{\ell,k})|, \right. \\ \left. |F_n^-(z_{\ell,k}, \omega) - F^-(z_{\ell,k})| \right)$$

folgt

$$\exists A_k \in \mathfrak{E} \forall \omega \in A_k : \lim_{n \rightarrow \infty} \Delta_{n,k}(\omega) = 0. \quad (1)$$

Lemma 17 zeigt

$$F^-(z_{\ell,k}) - F(z_{\ell-1,k}) \leq \ell/k - (\ell - 1)/k = 1/k.$$

Hiermit folgt für $x \in]z_{\ell-1,k}, z_{\ell,k}[$ und $\omega \in \Omega$

$$\begin{aligned} F_n(x, \omega) &\leq F_n^-(z_{\ell,k}, \omega) \leq F^-(z_{\ell,k}) + \Delta_{n,k}(\omega) \\ &\leq F(z_{\ell-1,k}) + 1/k + \Delta_{n,k}(\omega) \\ &\leq F(x) + 1/k + \Delta_{n,k}(\omega), \end{aligned}$$

$$\begin{aligned} F_n(x, \omega) &\geq F_n(z_{\ell-1,k}, \omega) \geq F(z_{\ell-1,k}) - \Delta_{n,k}(\omega) \\ &\geq F^-(z_{\ell,k}) - 1/k - \Delta_{n,k}(\omega) \\ &\geq F(x) - 1/k - \Delta_{n,k}(\omega). \end{aligned}$$

Für $\omega \in A_k$ ergibt sich gemäß (1)

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n(x, \omega) - F(x)| \leq 1/k.$$

Fazit: für

$$A := \bigcap_{k \in \mathbb{N}} A_k \in \mathfrak{A}$$

gilt

$$P(A) = 1$$

und

$$\forall \omega \in A : \lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_n(x, \omega) - F(x)| = 0.$$

□

22. Beispiel Empirische Verteilungsfunktionen schwarz.

(Zugrundeliegende Verteilung gegeben durch

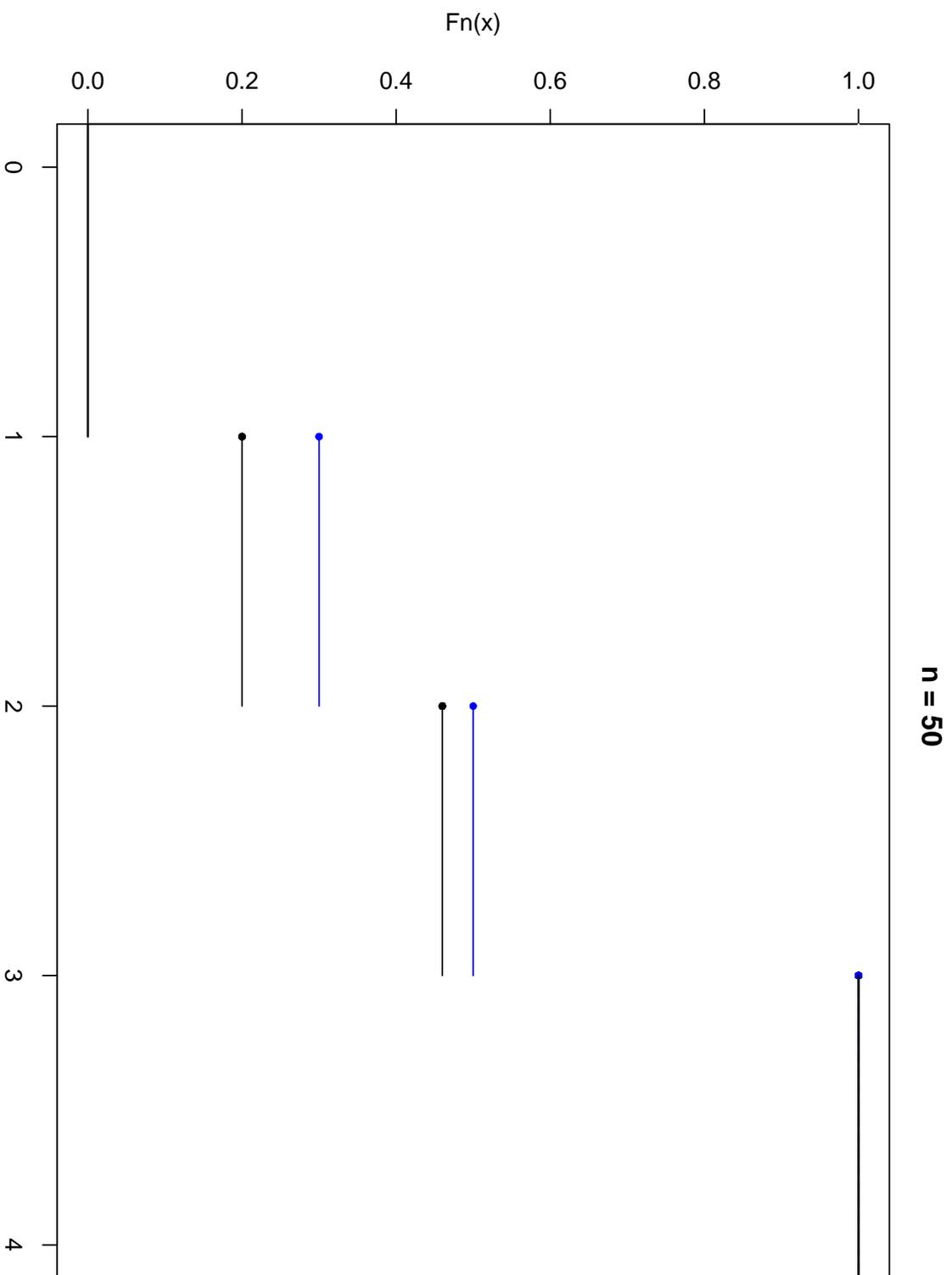
$$P(\{X_1 = 1\}) = 0,3$$

$$P(\{X_1 = 2\}) = 0,2$$

$$P(\{X_1 = 3\}) = 0,5$$

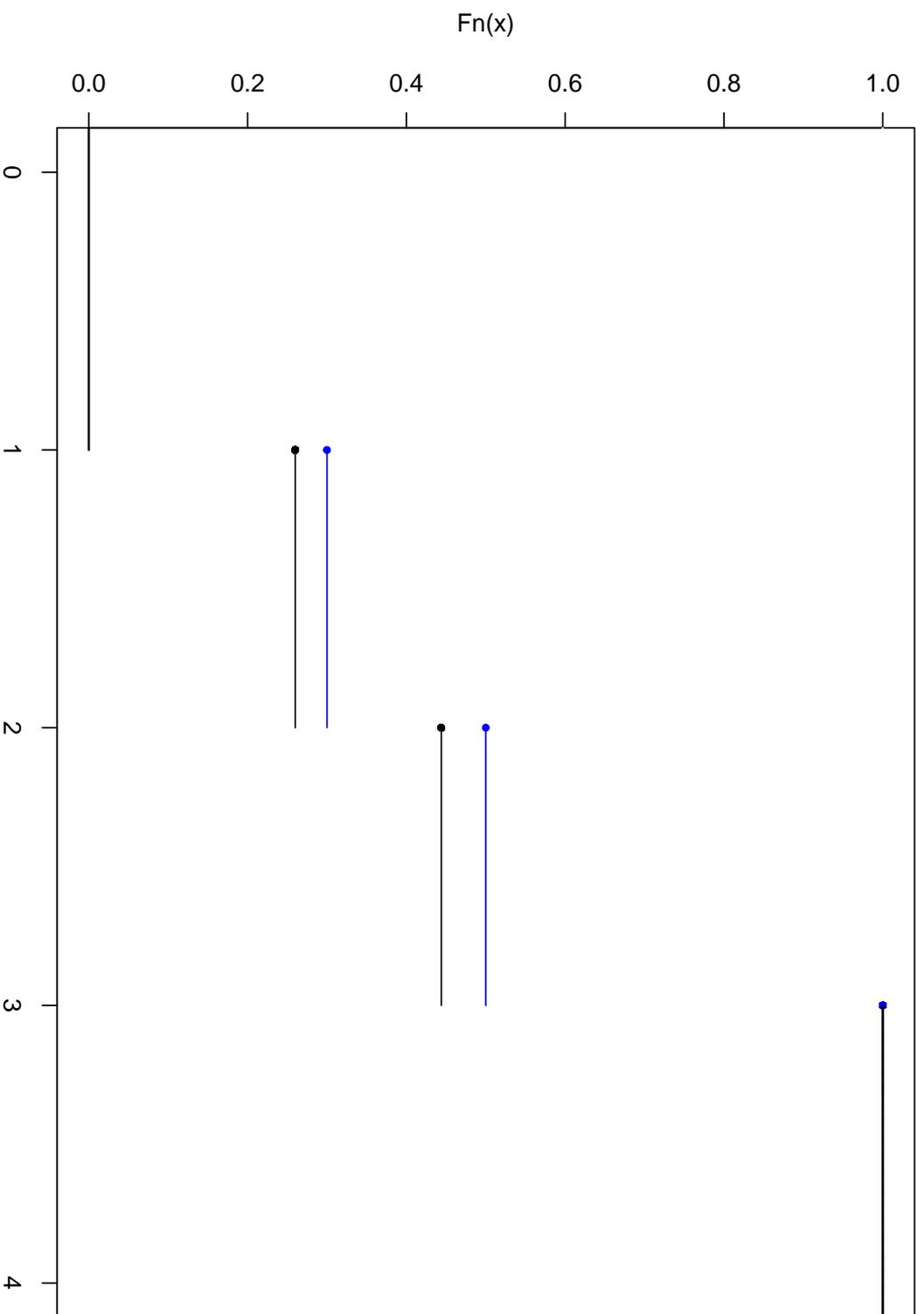
Verteilungsfunktion **blau**.)

$n = 50$



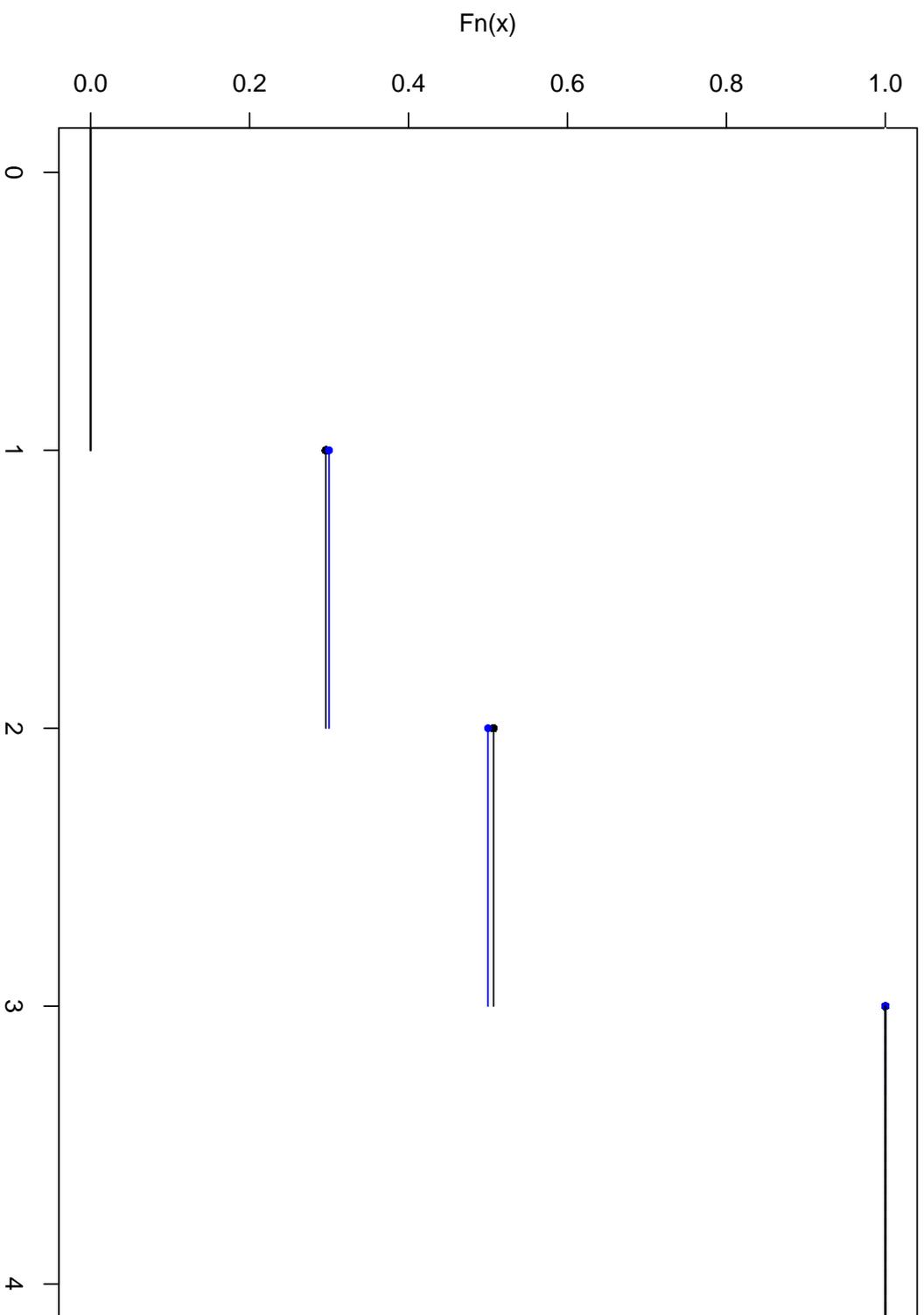
$n = 250$

$n = 250$



$n = 1000$

$n = 1000$

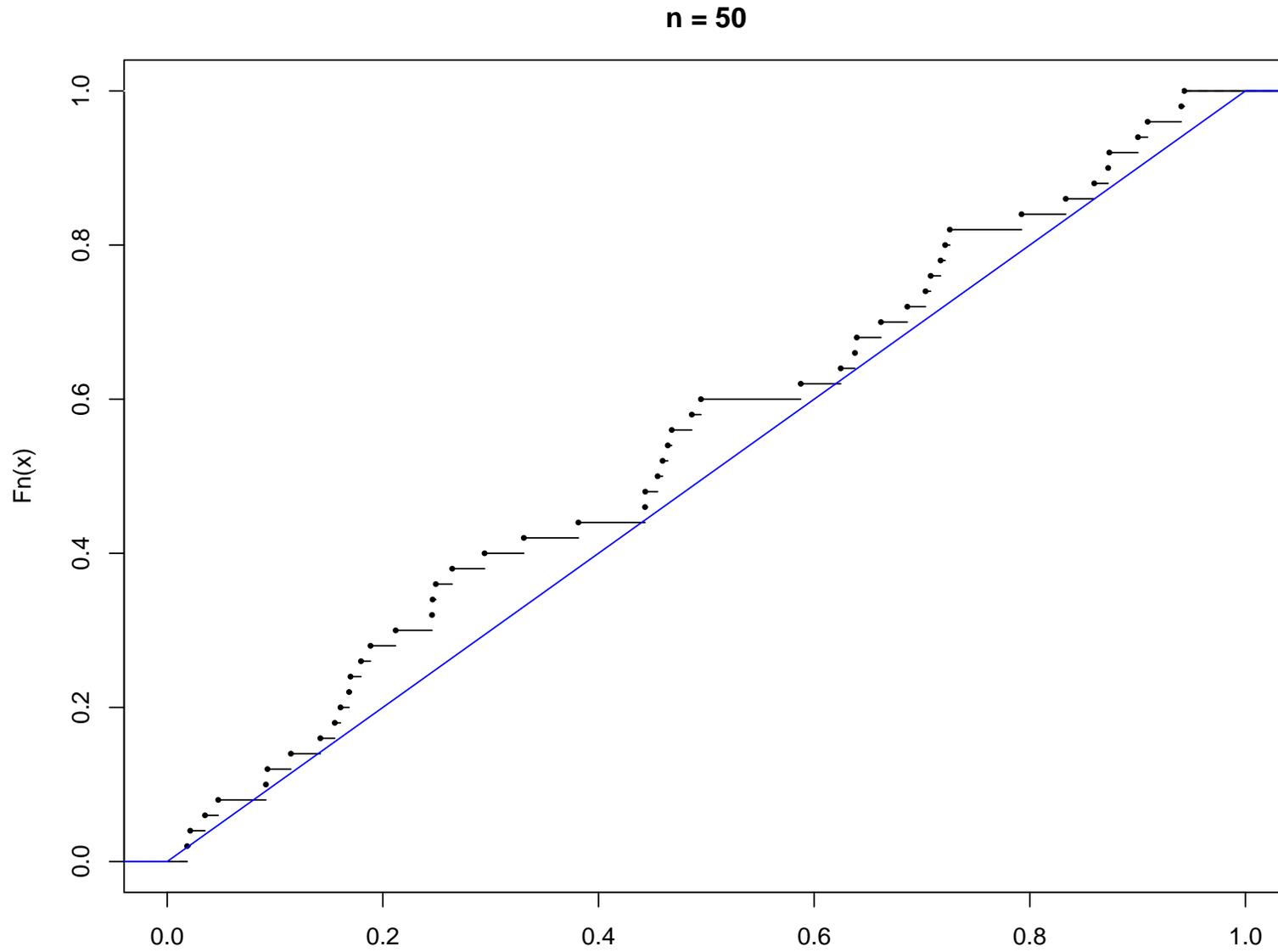


23. Beispiel Empirische Verteilungsfunktionen schwarz.

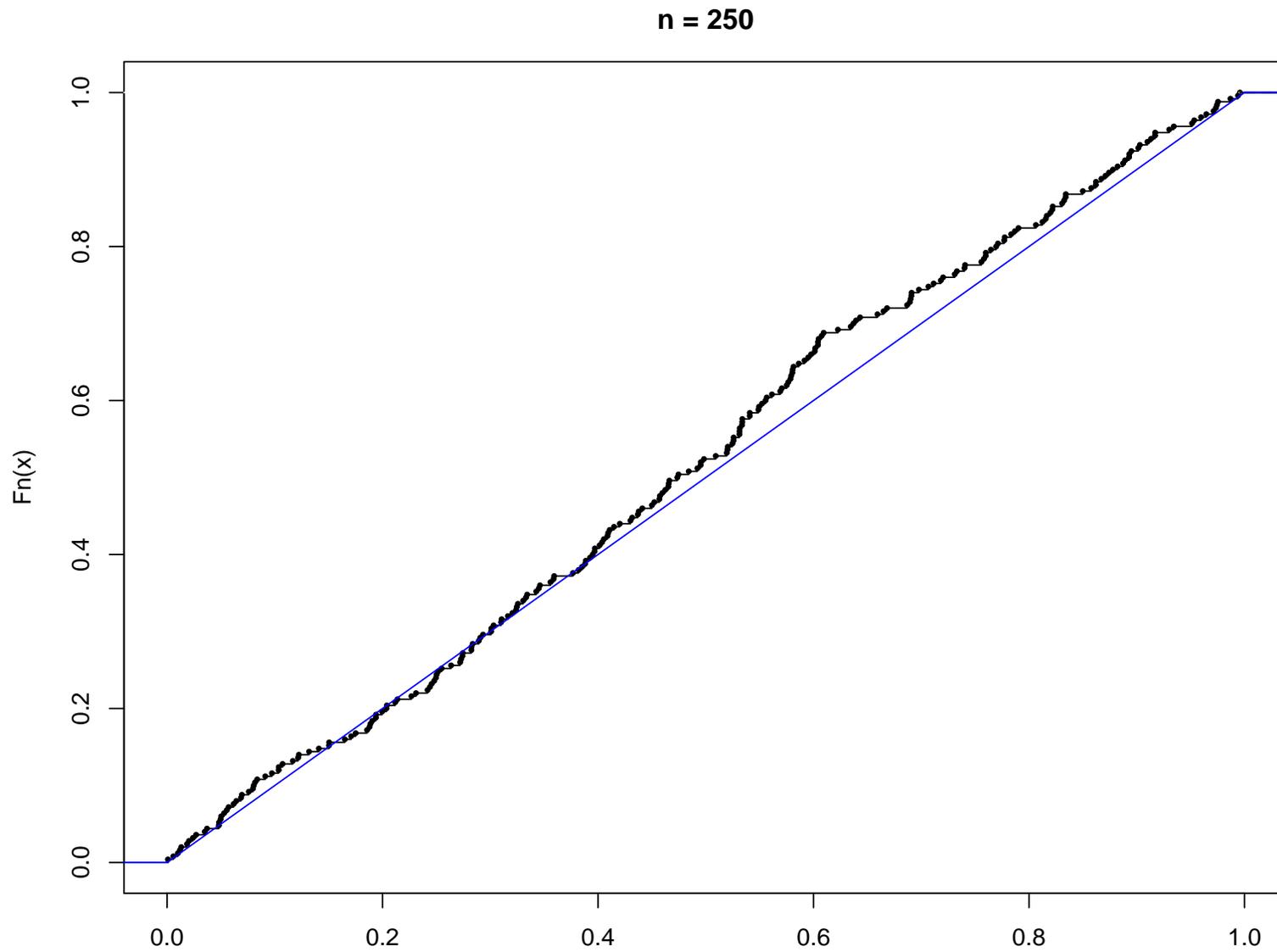
(Zugrundeliegende Verteilung gegeben durch

$X_1 \sim \mathbf{U}([0, 1])$, siehe Def. 25, Verteilungsfunktion **blau**.)

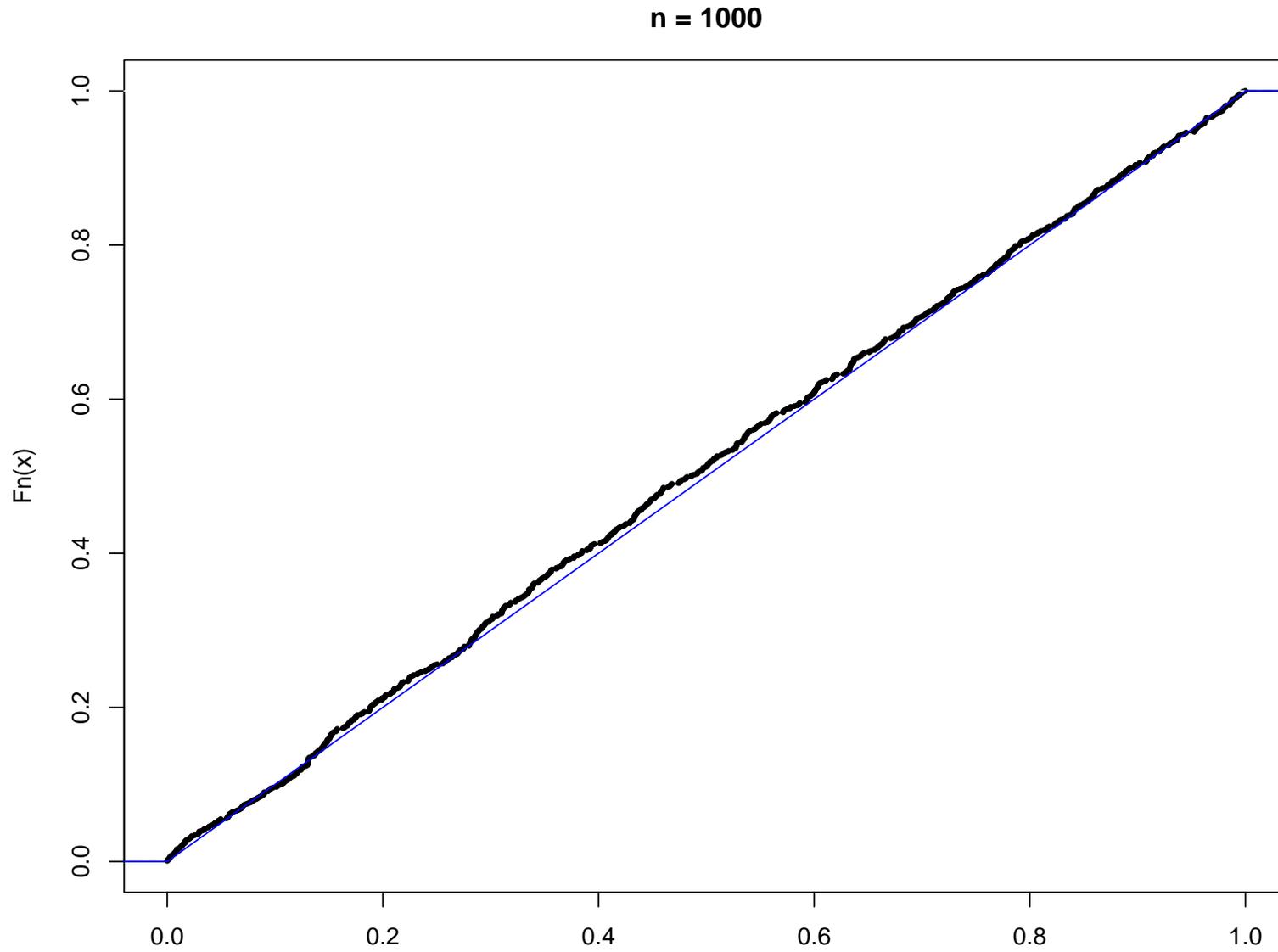
$n = 50$



$n = 250$



$n = 1000$

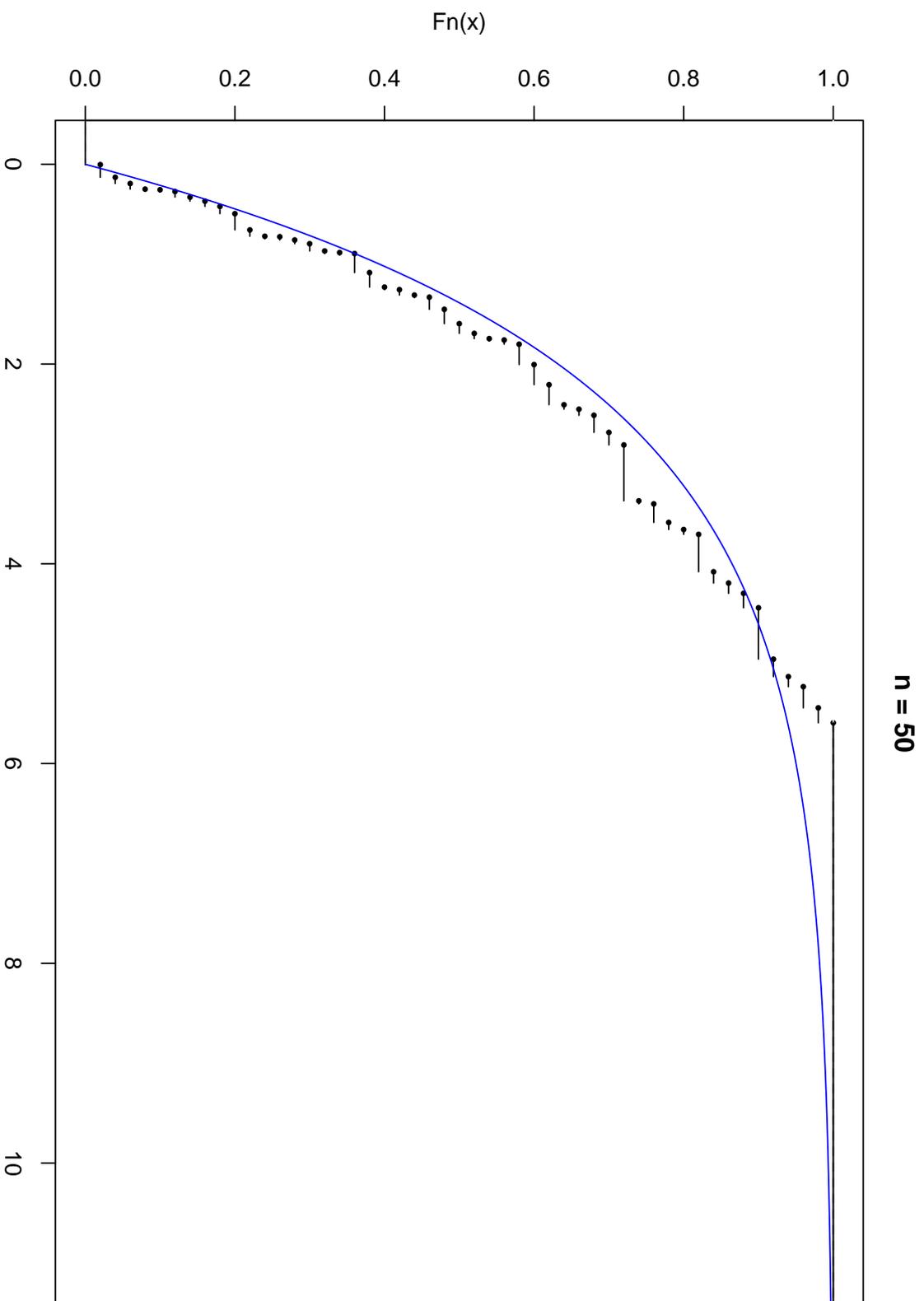


24. Beispiel Empirische Verteilungsfunktionen schwarz.

(Zugrundeliegende Verteilung gegeben durch

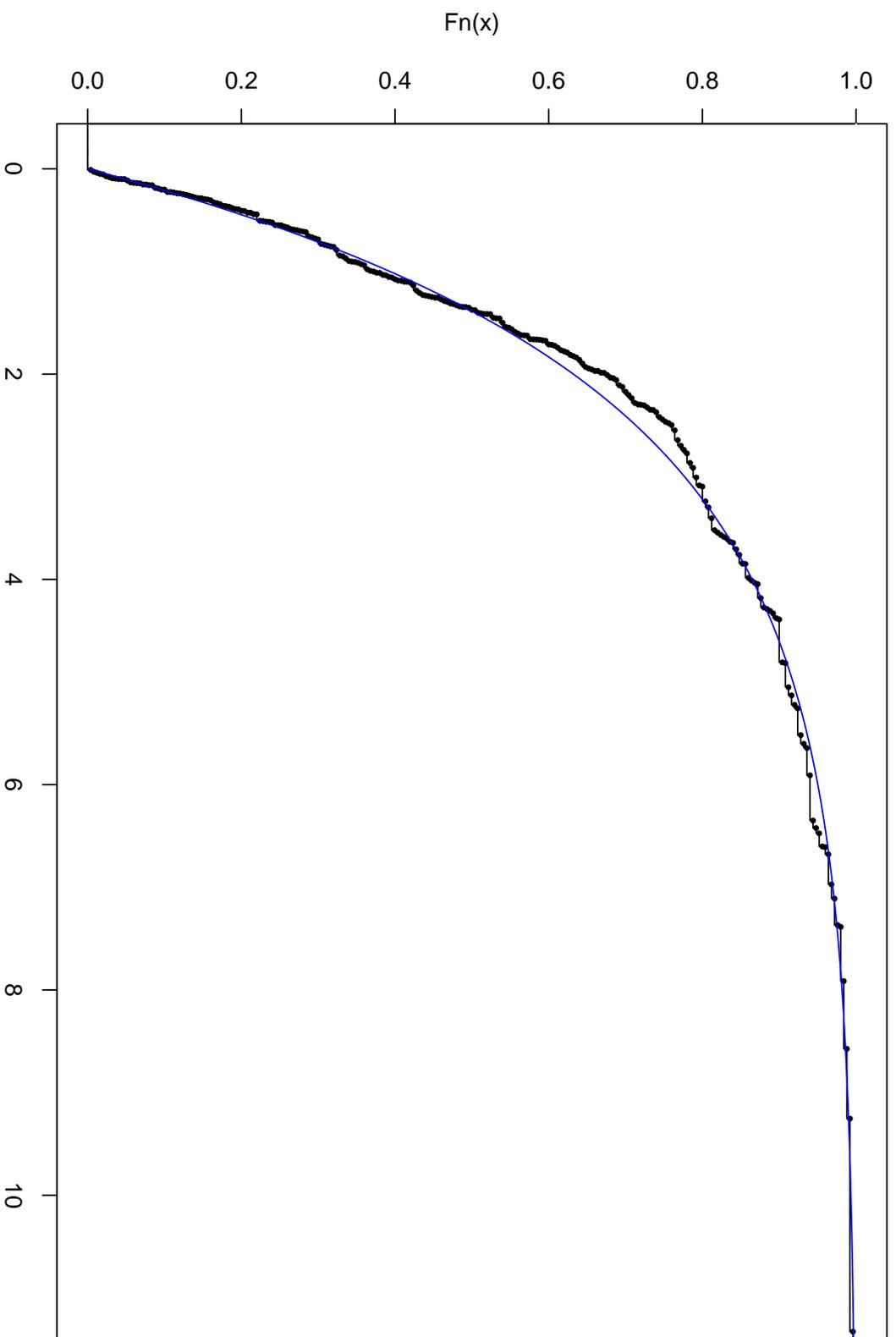
$X_1 \sim \mathbf{Exp}(1/2)$, siehe Def. V.32, Verteilungsfunktion **blau**.)

$n = 50$



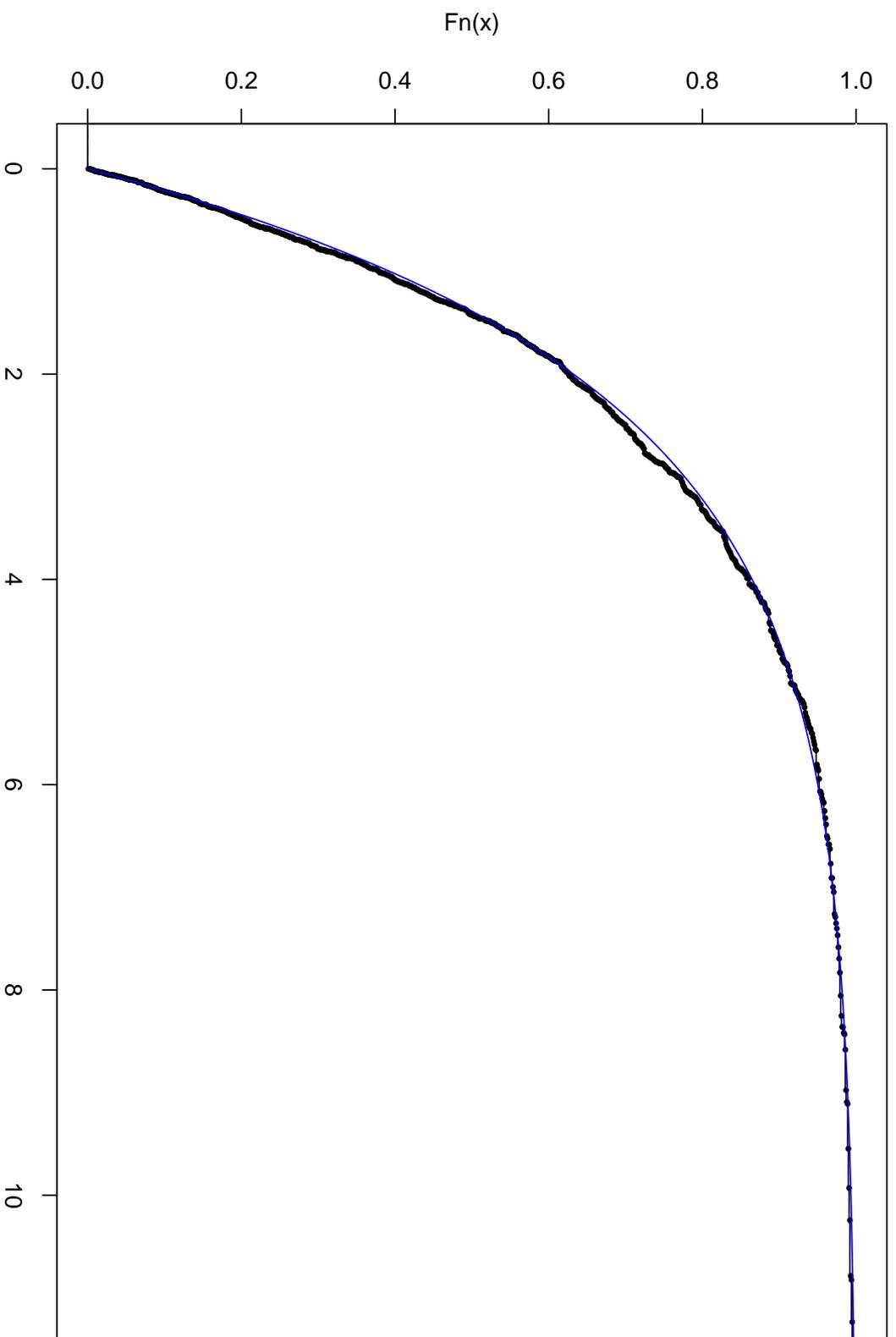
$n = 250$

$n = 250$



$n = 1000$

$n = 1000$



4 Zufallszahlen

Praktische Durchführung der stochastischen Simulation beruht auf **Zufallszahlengeneratoren**, d.h. auf geeigneten

Abbildungen $f : K \rightarrow K$, wobei $K := \{0, \dots, k - 1\}$ und k sehr groß, etwa $k = 2^{128}$ oder $k = 2^{19937}$.

Man erhält endliche Folgen $v_1, v_2, \dots \in K$ durch

- Initialisierung $v_0 \in K$,
- Iteration $v_\ell := f(v_{\ell-1})$.

„Kleine“ Abschnitte dieser Folgen „verhalten sich annähernd“ wie Realisierungen von iid-ZVen, die jeweils gleichverteilt auf K sind.

25. Definition ZV U **gleichverteilt** auf $[0, 1]$, falls

$$F_U(x) = \begin{cases} 0, & \text{falls } x < 0 \\ x, & \text{falls } 0 \leq x \leq 1 \\ 1, & \text{falls } x > 1 \end{cases}$$

Bez.: $U \sim \mathbf{U}([0, 1])$.

26. Bemerkung Für $U \sim \mathbf{U}([0, 1])$, $x \in \mathbb{R}$ und Intervall M mit Endpunkten $0 \leq a < b \leq 1$:

$$P(\{U = x\}) = 0$$

$$P(\{U \in M\}) = F_U(b) - F_U(a) = b - a$$

Insbesondere ist $U \sim \mathbf{U}([0, 1])$ keine diskrete ZV.

In Bsp. 23: zugrundeliegende Verteilungsfunktion F_U .

Siehe Kapitel V.2 zur Frage der Existenz.

Zurück zu den Zufallszahlen $v_1, v_2, \dots \in K$.

Setze $u_\ell := v_\ell/k$. „Kleine Abschnitte“ der Folgen

$u_1, u_2, \dots \in [0, 1[$ „verhalten sich annähernd“ wie

Realisierungen von iid-ZVen, die jeweils gleichverteilt auf

$[0, 1]$ sind.

Kurz: u_1, u_2, \dots **gleichverteilte Zufallszahlen** in $[0, 1]$.

In dieser Vorlesung: Zufallszahlengenerator als black-box.

Verwendung etwa in der Form

- Initialisierung `init(u0)`,
- Iteration `u := rand()`.

Stochastische Simulation (Monte-Carlo-Algorithmen) und
Zufallszahlengeneratoren Gegenstand eigener Vorlesungen.

Siehe auch Hesse (2003, Kap. 10).

27. Bemerkung Die Entwicklung des Computers und militärische Anwendungen setzten zur Mitte des letzten Jahrhunderts eine rapide Entwicklung der stochastischen Simulation in Gang. Den Namen 'The Monte Carlo Method' trägt ein Tagungsband aus dem Jahr 1951, aus dem wir hier zitieren. John von Neumann (1951) war der folgenden Ansicht:

We see then that we could build a physical instrument to feed random digits directly into a high-speed computing machine and could have the control call for these numbers as needed.

Von Neumann fährt fort mit einem Einwand gegen diesen „echten“ Zufall, der uns auch heute noch einleuchtet:

The real objection to this procedure is the practical need for checking computations. If we suspect that a calculation is wrong, almost any reasonable check involves repeating something done before.

Zum Einsatz von auf deterministischen Methoden basierender
Generatoren sagt von Neumann in derselben Arbeit:

*Any one who considers arithmetical methods of
producing random digits is, of course, in a state of sin.*

Frage: Simulation „komplizierter“ Verteilungen.

Eine „universelle“ Möglichkeit: **Inversionsmethode**.

Gegeben:

- $F : \mathbb{R} \rightarrow [0, 1]$ mit Eigenschaften (i)–(iii) aus Satz 14,
- ZV $U \sim \mathbf{U}([0, 1])$

Definiere $T_F : \mathbb{R} \rightarrow \mathbb{R}$ durch

$$T_F(u) := \begin{cases} \inf\{v \in \mathbb{R} : F(v) \geq u\}, & \text{falls } u \in]0, 1[\\ 0, & \text{sonst} \end{cases}$$

28. Satz $X := T_F \circ U$ ist ZV mit $F_X = F$.

Beweis. Für $x \in \mathbb{R}$ und $u \in]0, 1[$:

$$\begin{aligned} T_F(u) \leq x &\Leftrightarrow \inf\{v \in \mathbb{R} : F(v) \geq u\} \leq x \\ &\Leftrightarrow F(x) \geq u. \end{aligned}$$

Also

$$\{X \leq x\} = \{U \in]0, 1[\} \cap \{F(x) \geq U\} \cup \underbrace{\{U \notin]0, 1[\} \cap \{0 \leq x\}}_{\in \{\emptyset, \Omega\}},$$

so daß insbesondere $\{X \leq x\} \in \mathfrak{A}$. Weiter

$$P(\{X \leq x\}) = P(\{F(x) \geq U\}) = F(x).$$

□

29. Bemerkung Die Sätze 14 und 28 zeigen:

$F : \mathbb{R} \rightarrow [0, 1]$ genau dann Verteilungsfunktion einer ZV, wenn F die Eigenschaften (i)–(iii) erfüllt.

Also Modellierung eines beliebigen Zufallsexperimentes mit reellwertigen Ergebnissen durch Vorgabe der Verteilungsfunktion F .

30. Bemerkung **Inversionsmethode**: Simulation einer ZV mit Verteilungsfunktion $F := F_X$ mittels Transformation

$$T_F(u_1), T_F(u_2), \dots$$

gleichverteilter Zufallszahlen in $[0, 1]$.

5 Anwendung: Simulation einer Irrfahrt

Im folgenden Y'_1, \dots, Y'_T iid mit $Y'_1 \sim \mathbf{SB}$, d.h.

$$P'(\{Y'_1 = 1\}) = P'(\{Y'_1 = -1\}) = 1/2.$$

Definiere **symmetrische Bernoulli-Irrfahrt** S'_0, \dots, S'_T durch

$$S'_t := \sum_{j=1}^t Y'_j = S'_{t-1} + Y'_t.$$

Per def. $S'_0 = 0$.

Ausblick: **Irrfahrten auf Gruppen**. **Finanzmathematik**: $]0, \infty[$ mit Multiplikation, **Kartenspielen**: symmetrische Gruppe mit Komposition.

31. Beispiel Faires Spiel zweier Spieler I und II, S'_t Stand aus Sicht von I nach t Runden.

Physik: S'_t eindimensionale Position eines Teilchen nach t Kollisionen.

Problem A: gesucht sind die Vert'funktionen der ZV

- (i) **Spielstand nach letzter Runde**, $X' := S'_T$
- (ii) **maximaler Spielstand**, $X' := \max_{t=0, \dots, T} S'_t$
- (iii) **Anzahl der Runden, nach denen Spieler I führt**,

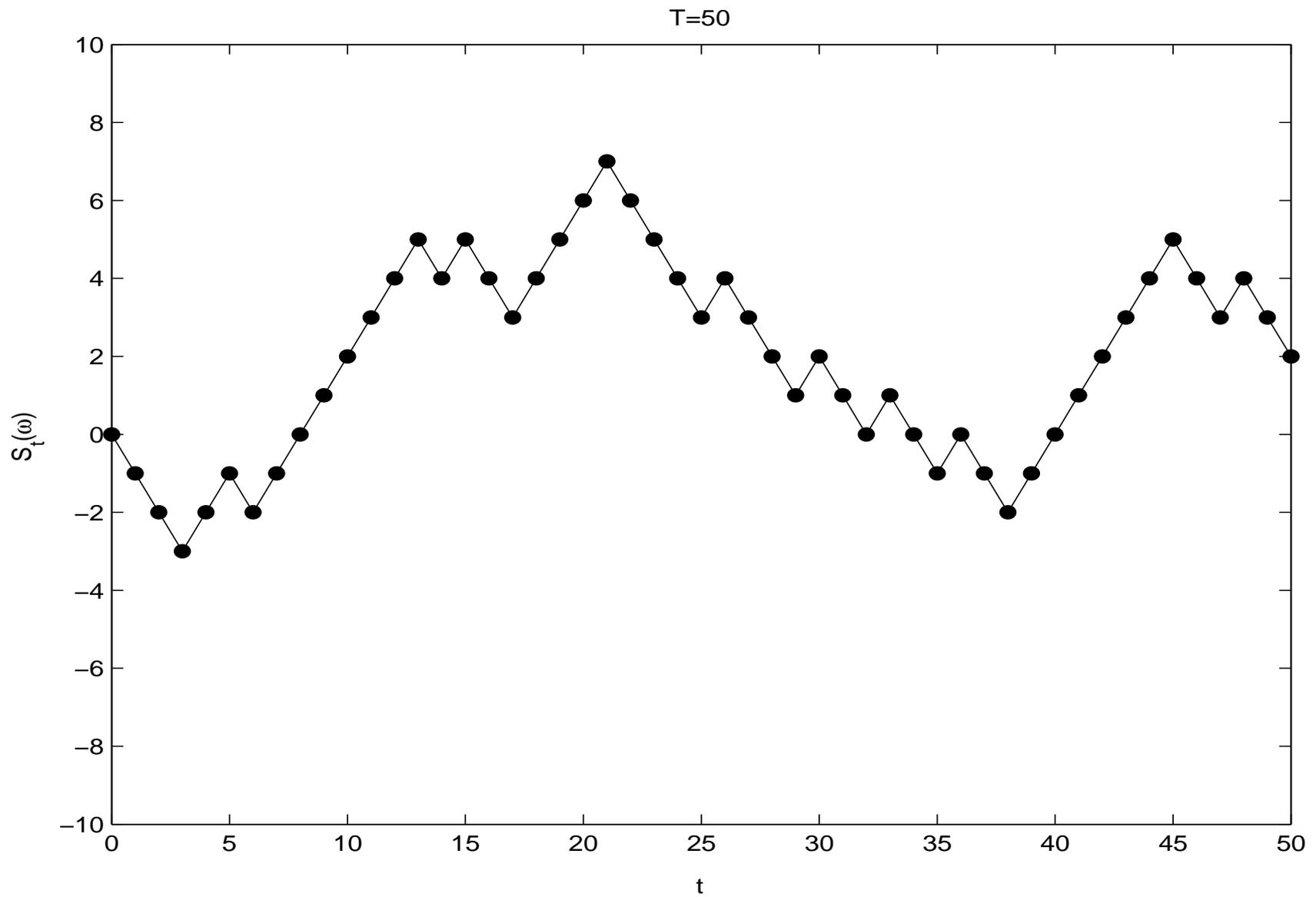
$$X' := |\{t \in \{1, \dots, T\} : S'_t \geq 0 \text{ und } S'_{t-1} \geq 0\}|$$

Beachte $\{S'_t \geq 0\} \cap \{S'_{t-1} \geq 0\}$

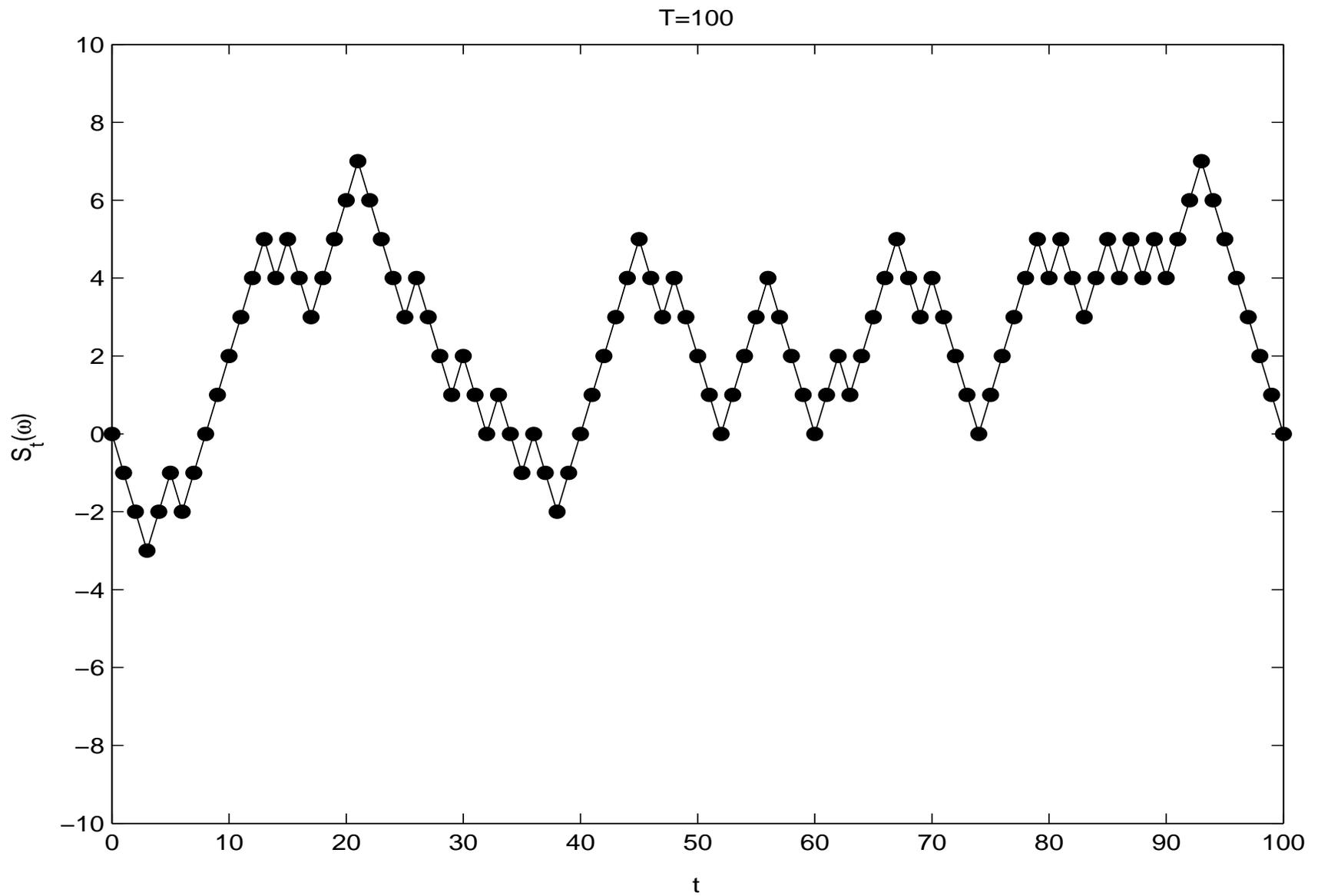
$$= \{S'_t > 0\} \cup (\{S'_t = 0\} \cap \{S'_{t-1} > 0\}).$$

32. Beispiel Je eine Simulation der Irrfahrt für
 $T = 50, 100, 1000$.

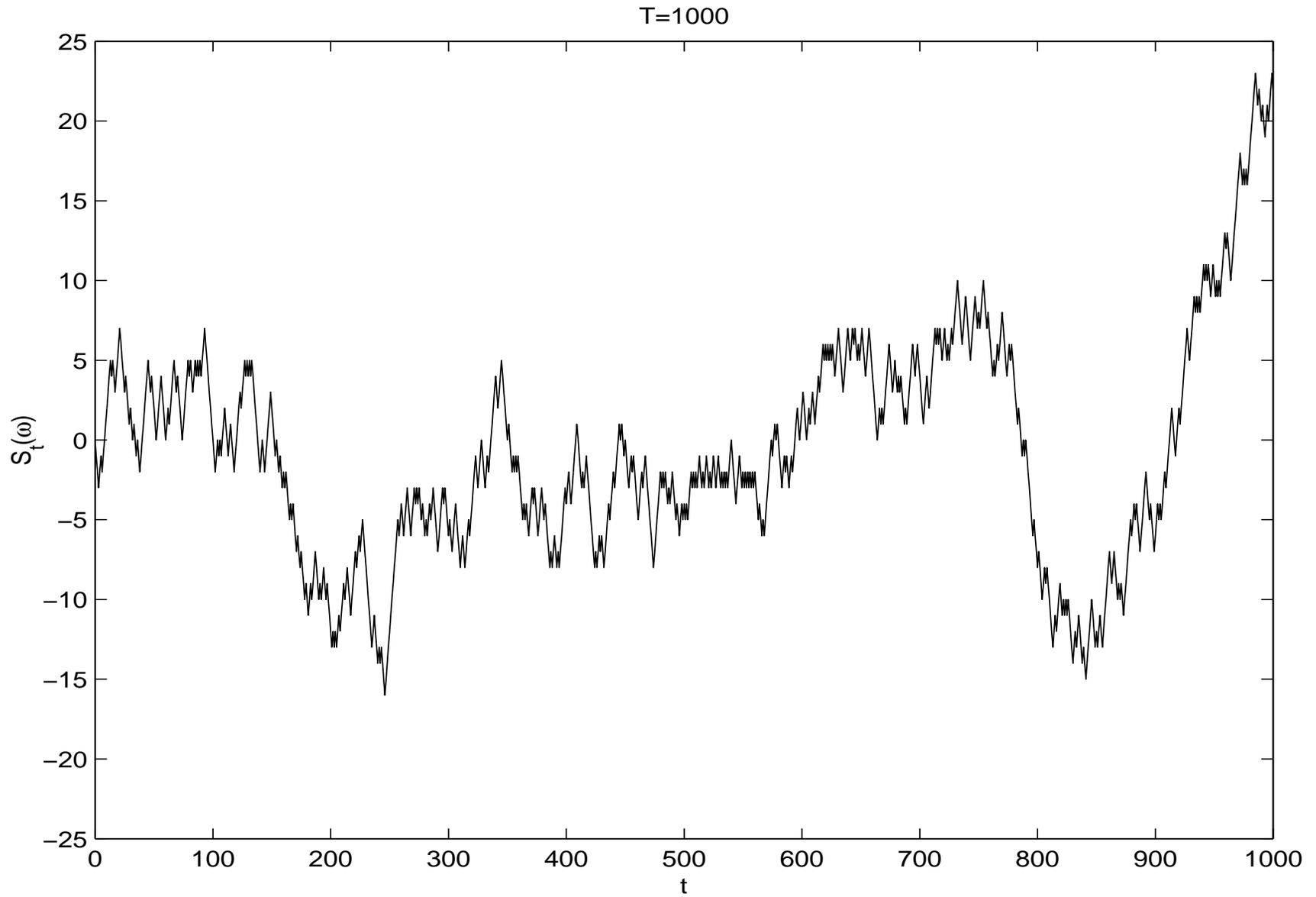
$T = 50$: end=2, max=7, fueh=36.



$T = 100$: end=0, max=7, fueh=86.



$T = 1000$: end=22, max=23, fueh=420.



33. Bemerkung Setze $s_0 := 0$ und

$$D := \{(s_0, \dots, s_T) \in \mathbb{Z}^{T+1} : \forall 1 \leq t \leq T : |s_t - s_{t-1}| = 1\}.$$

Dann gilt:

- $|D| = 2^T$
- $P'(\{(S'_0, \dots, S'_T) = s\}) = 1/|D|$ für alle $s \in D$

Fazit: Irrfahrt entspricht **Gleichverteilung** auf der Menge D ihrer **Pfade**.

Irrfahrten als einfaches Beispiel eines **stochastischen Prozesses** zur Modellierung zeitabhängiger zufälliger Phänomene.

Definiere zugehörige Abbildung $f : D \rightarrow \mathbb{R}$

- bei (i): $f(s_0, \dots, s_T) = s_T$
- bei (ii): $f(s_0, \dots, s_T) = \max_{t=0, \dots, T} s_t$
- bei (iii): $f(s_0, \dots, s_T) = |\{t \in \{1, \dots, T\} : s_t \geq 0 \text{ und } s_{t-1} \geq 0\}|$

Somit

$$X' = f \circ (S'_0, \dots, S'_T).$$

Gesucht ist

$$P'(\{X' = x\})$$

für $x \in \mathbb{Z}$.

Simulation von n unabh. Wiederholungen der Irrfahrt:

Betrachte iid-Folge $Y_1, \dots, Y_{n \cdot T}$ mit $Y_1 \sim \text{SB}$.

Definiere für $t = 0, \dots, T$ und $i = 1, \dots, n$

$$S_t^i := \sum_{j=1}^t Y_{(i-1) \cdot T + j}.$$

Also

$$\begin{array}{ccc} \underbrace{Y_1, \dots, Y_T}, \dots, \underbrace{Y_{(n-1) \cdot T + 1}, \dots, Y_{n \cdot T}} & & \\ \downarrow & & \downarrow \\ \underbrace{S_0^1, \dots, S_T^1}, \dots, \underbrace{S_0^n, \dots, S_T^n} & & \\ f \downarrow & & f \downarrow \\ X^1, \dots, X^n & & \end{array}$$

Es gilt bei (i), (ii) und (iii):

X^1, \dots, X^n iid, X^1 und X' identisch verteilt.

Beweis: Siehe Irle (2001, p. 163, 169) oder Vorlesung „Probability Theory“.

Gem. Kor. 10 existiert $A \in \mathfrak{A}$ mit $P(A) = 1$, so daß für alle

$\omega \in A$ und alle $x \in \mathbb{Z}$

$$\begin{aligned} & \lim_{n \rightarrow \infty} 1/n \cdot \sum_{i=1}^n 1_{\{x\}} \left(\underbrace{f(S_0^i(\omega), \dots, S_T^i(\omega))}_{=X^i(\omega)} \right) \\ & = P'(\{X' = x\}). \end{aligned}$$

34. Bemerkung Für $U \sim \mathbf{U}([0, 1])$ und

$$Y(\omega) := \begin{cases} 1, & \text{falls } U(\omega) > 1/2 \\ -1, & \text{falls } U(\omega) \leq 1/2 \end{cases}$$

gilt $Y \sim \mathbf{SB}$. Vgl. Inversionsmethode.

Zur Simulation von Irrfahrten entsprechende Transformation

$$y_\ell := \begin{cases} 1, & \text{falls } u_\ell > 1/2 \\ -1, & \text{falls } u_\ell \leq 1/2 \end{cases}$$

von in $[0, 1]$ gleichverteilten Zufallszahlen u_1, u_2, \dots

Setze

$$s_t^i := \sum_{j=1}^t y_{(i-1) \cdot T + j}$$

Zusammenfassung: Approximation der gesuchten

Wahrscheinlichkeit $P'(\{X' = x\})$ durch relative Häufigkeit

$$1/n \cdot \sum_{i=1}^n 1_{\{x\}}(f(s_0^i, \dots, s_T^i))$$

der Simulationen, bei denen f den Wert x liefert.

Im folgenden: Simulationsergebnisse für $T = 50$.

Beachte: Für die Menge D der Pfade gilt dann

$$|D| = 2^{50} = 1,1 \dots \cdot 10^{15}.$$

Dargestellt wird

$$x \mapsto 1/n \cdot \sum_{i=1}^n 1_{\{x\}}(f(s_0^i, \dots, s_T^i)).$$

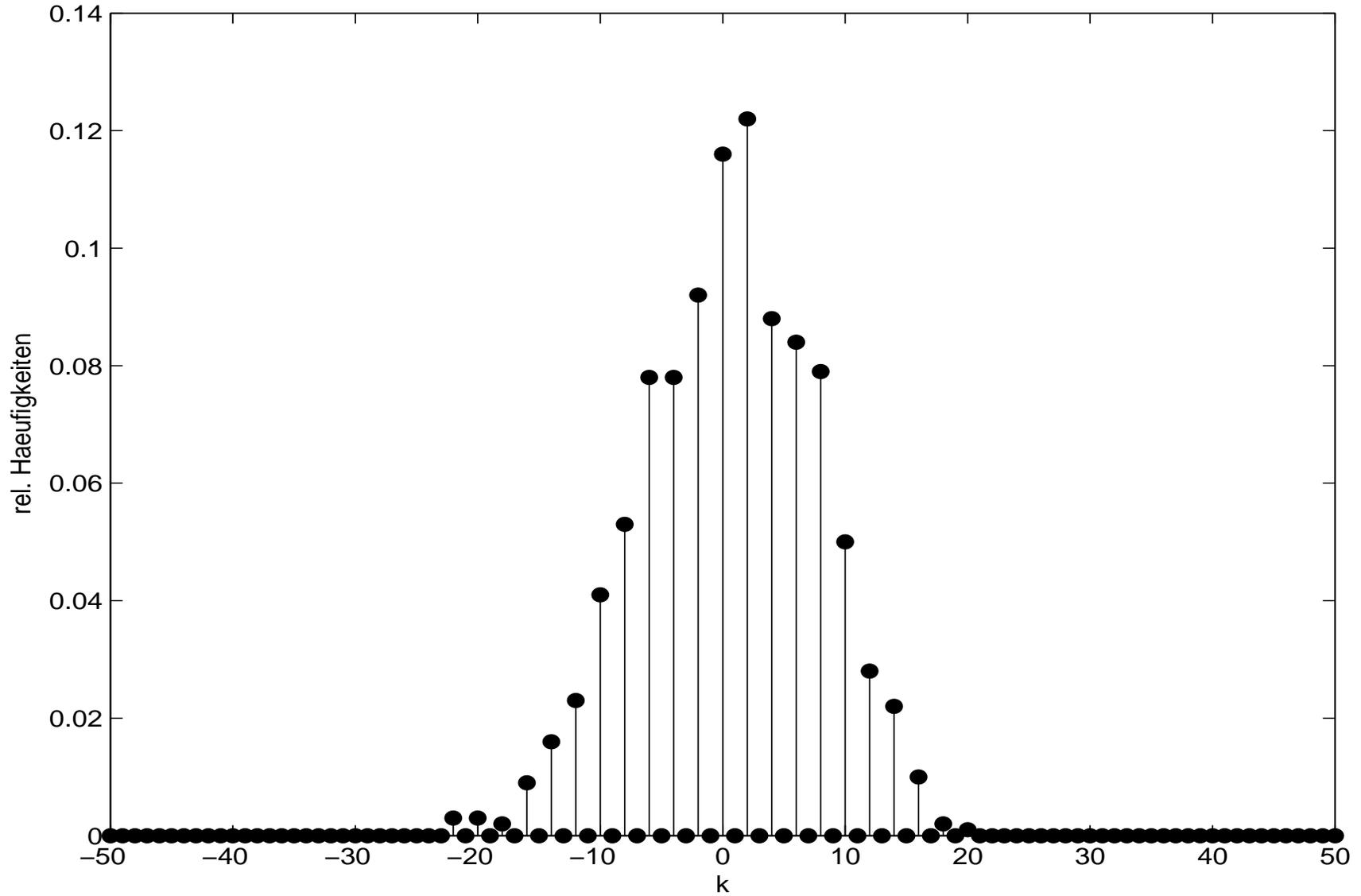
35. Beispiel Approximative Berechnung von

$$P'(\{S'_T = x\})$$

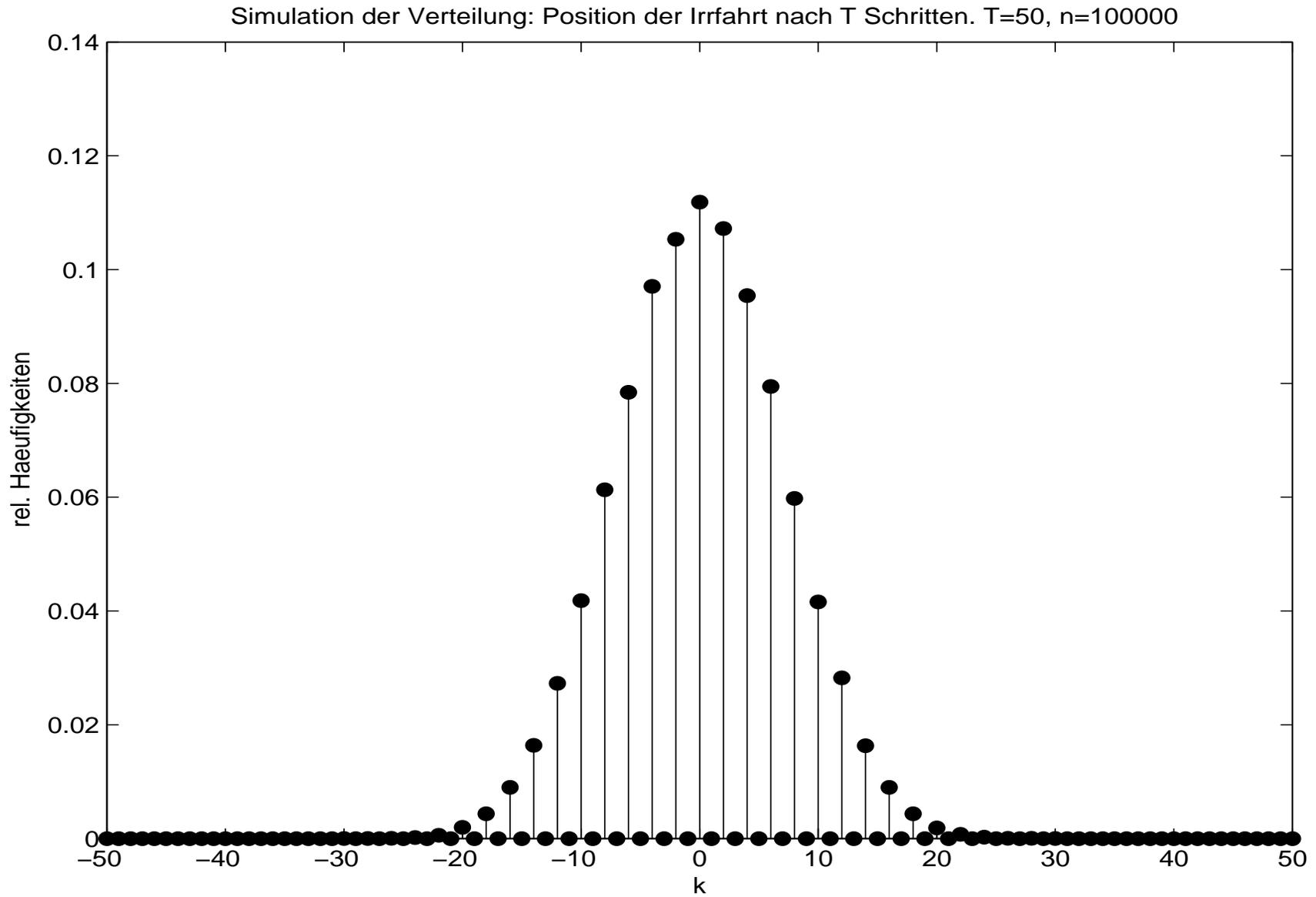
für $x = -T, -T + 2, \dots, T - 2, T$.

$n = 1000$

Simulation der Verteilung: Position der Irrfahrt nach T Schritten. T=50, n=1000



$$n = 100000$$



36. Satz Für $x \in \{-T, -T + 2, \dots, T - 2, T\}$

$$P'(\{S'_T = x\}) = \binom{T}{(x + T)/2} \cdot 2^{-T}.$$

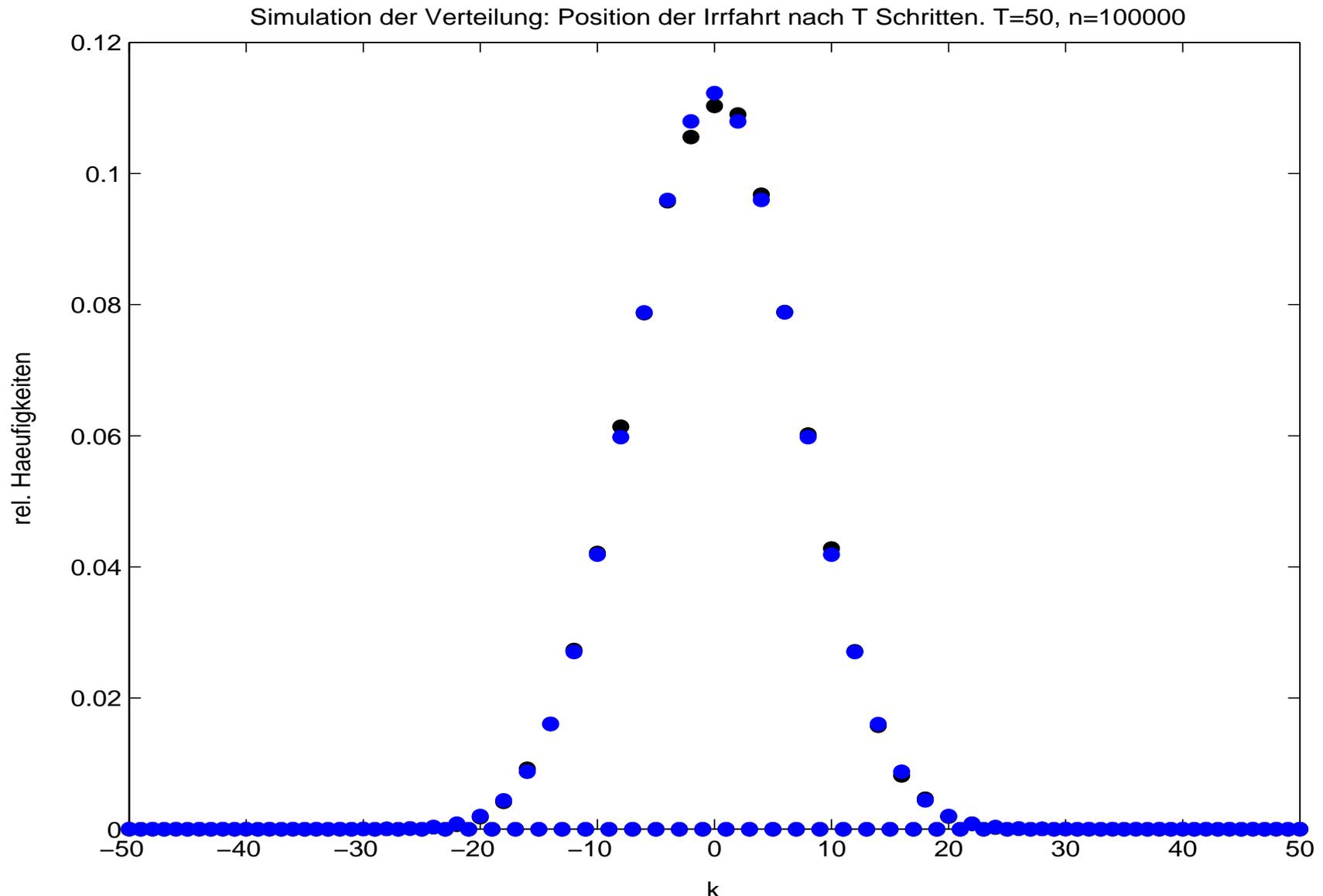
Beweis. Es gilt

$$S'_T = \sum_{j=1}^T Y'_j = 2 \cdot \underbrace{\sum_{j=1}^T \underbrace{(Y'_j + 1)/2}_{=: U'_j}}_{=: R'},$$

und U'_1, \dots, U'_T sind iid mit $U'_1 \sim \mathbf{B}(1, 1/2)$. Es folgt $R' \sim \mathbf{B}(T, 1/2)$ und

$$P'(\{S'_T = x\}) = P'(\{R' = (x + T)/2\})$$

Vergleich der Simulationsergebnisse (schwarz) mit den exakten Werten (blau) für $n = 100000$



Bezeichnung $a_k \approx b_k$ für Folgen von Zahlen $a_k, b_k > 0$, falls

$$\lim_{k \rightarrow \infty} a_k/b_k = 1.$$

37. Korollar Sei $T = 2k$ gerade. Dann

$$P'(\{S'_T = 0\}) \approx \sqrt{2/(\pi T)}.$$

Beweis. Stirlingsche Formel:

$$k! \approx \sqrt{2\pi k} \cdot (k/e)^k,$$

siehe Krengel (2003, p. 76) und TUTORIUM T2:1. Also folgt mit Satz 36

$$\begin{aligned} P'(\{S'_T = 0\}) &= 2^{-T} \cdot \binom{T}{T/2} \\ &\approx 2^{-T} \cdot \frac{\sqrt{2\pi T} \cdot T^T \cdot \exp(-T)}{\left(\sqrt{2\pi T/2} \cdot (T/2)^{T/2} \cdot \exp(-T/2)\right)^2} \\ &= \sqrt{2/(\pi T)}. \end{aligned}$$

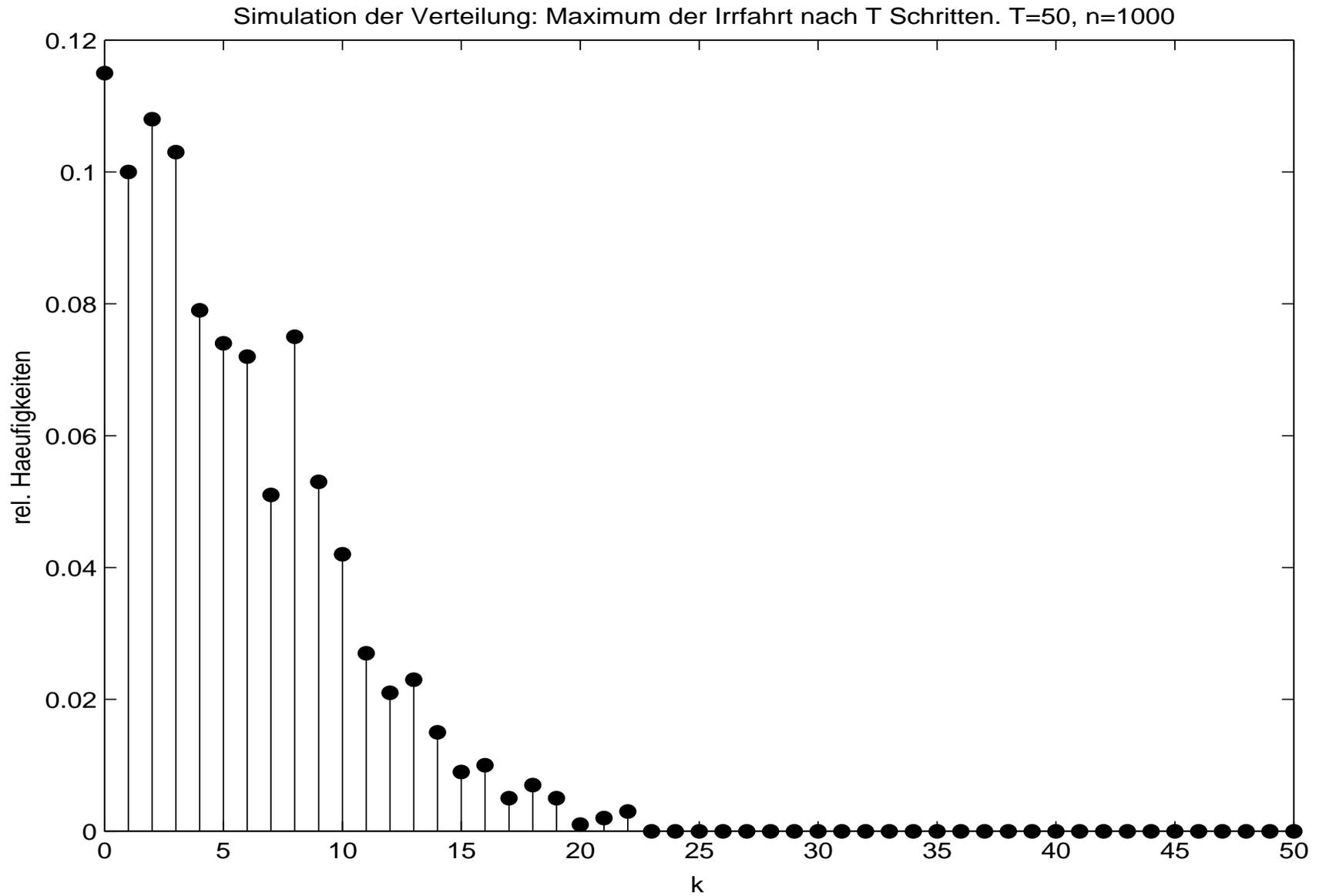
□

38. Beispiel Approximative Berechnung von

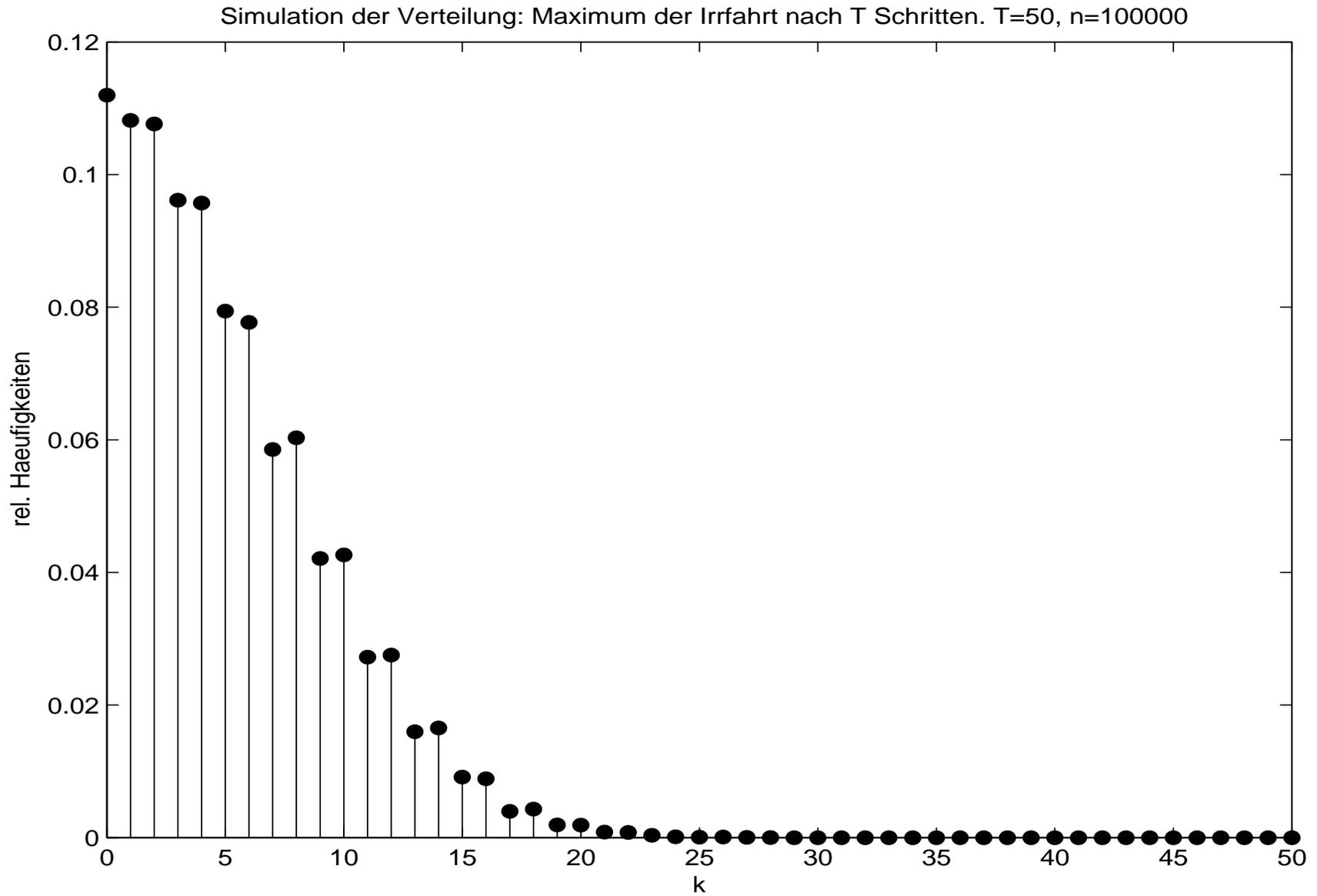
$$P'(\{\max_{t=0,\dots,T} S'_t = x\})$$

für $x = 0, \dots, T$.

$n = 1000$



$n = 100000$



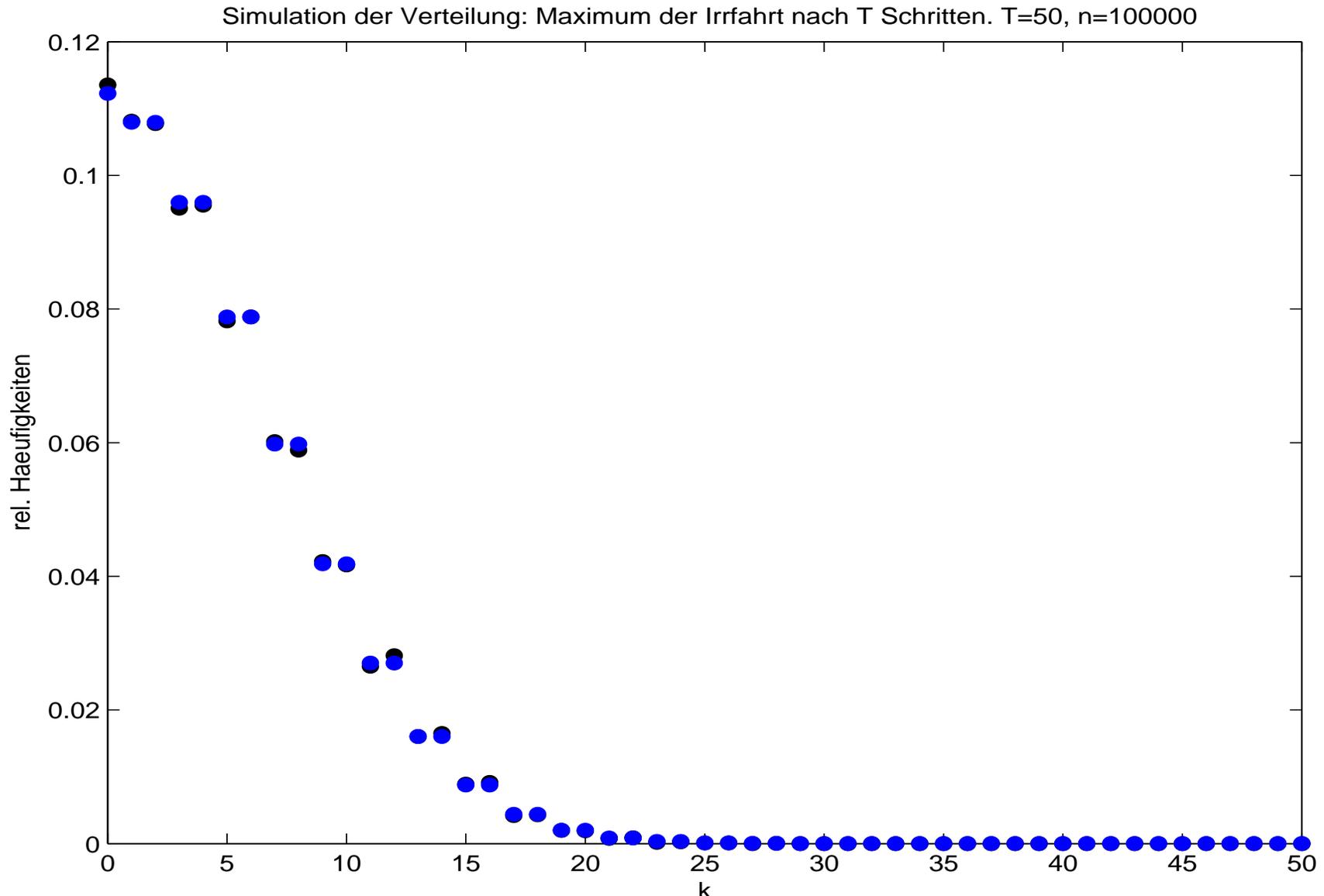
39. Satz Für $x \in \{0, \dots, T\}$

$$\begin{aligned} P'(\{\max_{t=0, \dots, T} S'_t = x\}) \\ = P'(\{S'_T = x\}) + P'(\{S'_T = x + 1\}). \end{aligned}$$

Beweis. ÜBUNG M:G14

□

Vergleich der Simulationsergebnisse (schwarz) mit den exakten Werten (blau) für $n = 100000$



Im folgenden

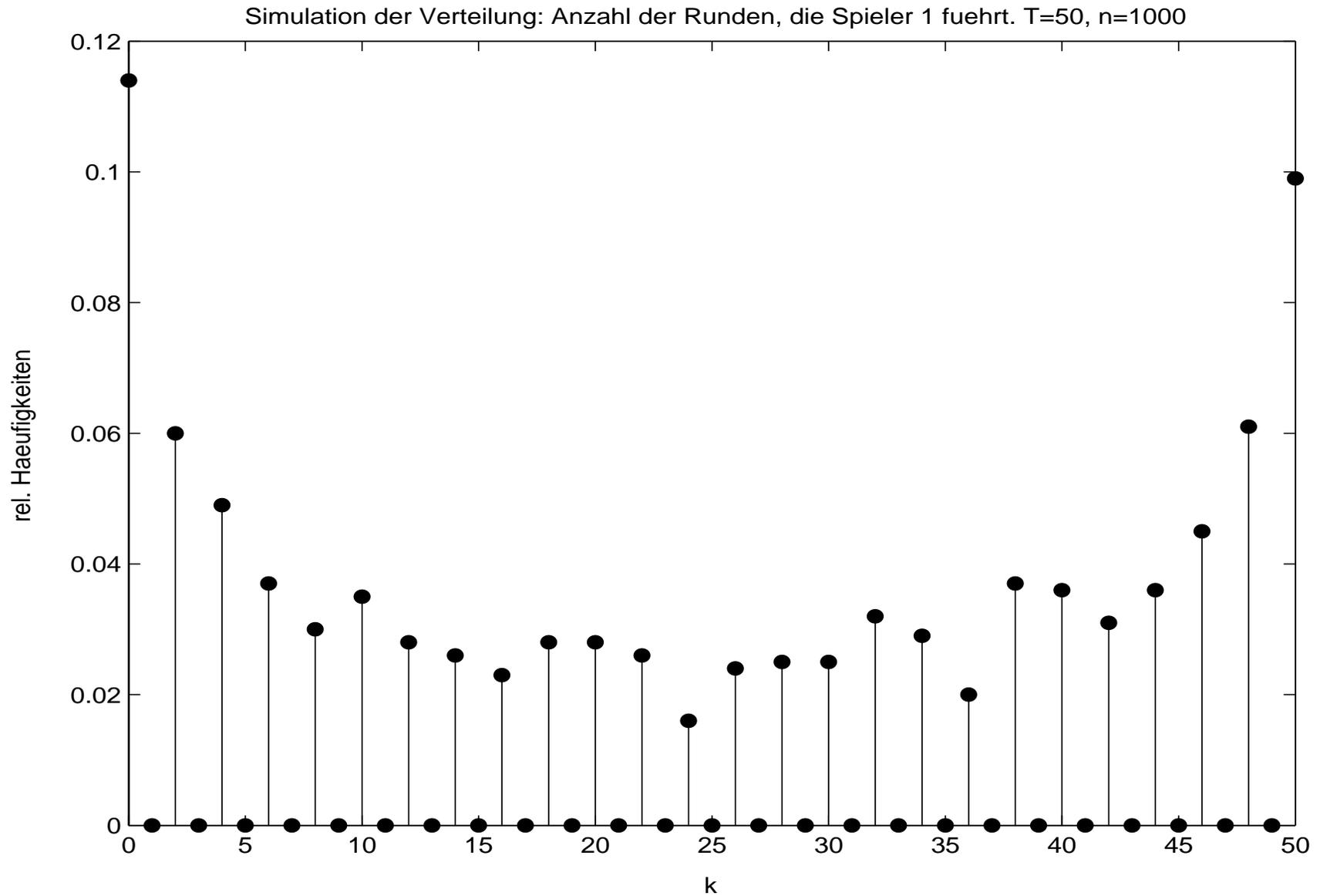
$$X' := |\{t \in \{1, \dots, T\} : S'_t \geq 0 \text{ und } S'_{t-1} \geq 0\}|.$$

40. Beispiel Approximative Berechnung von

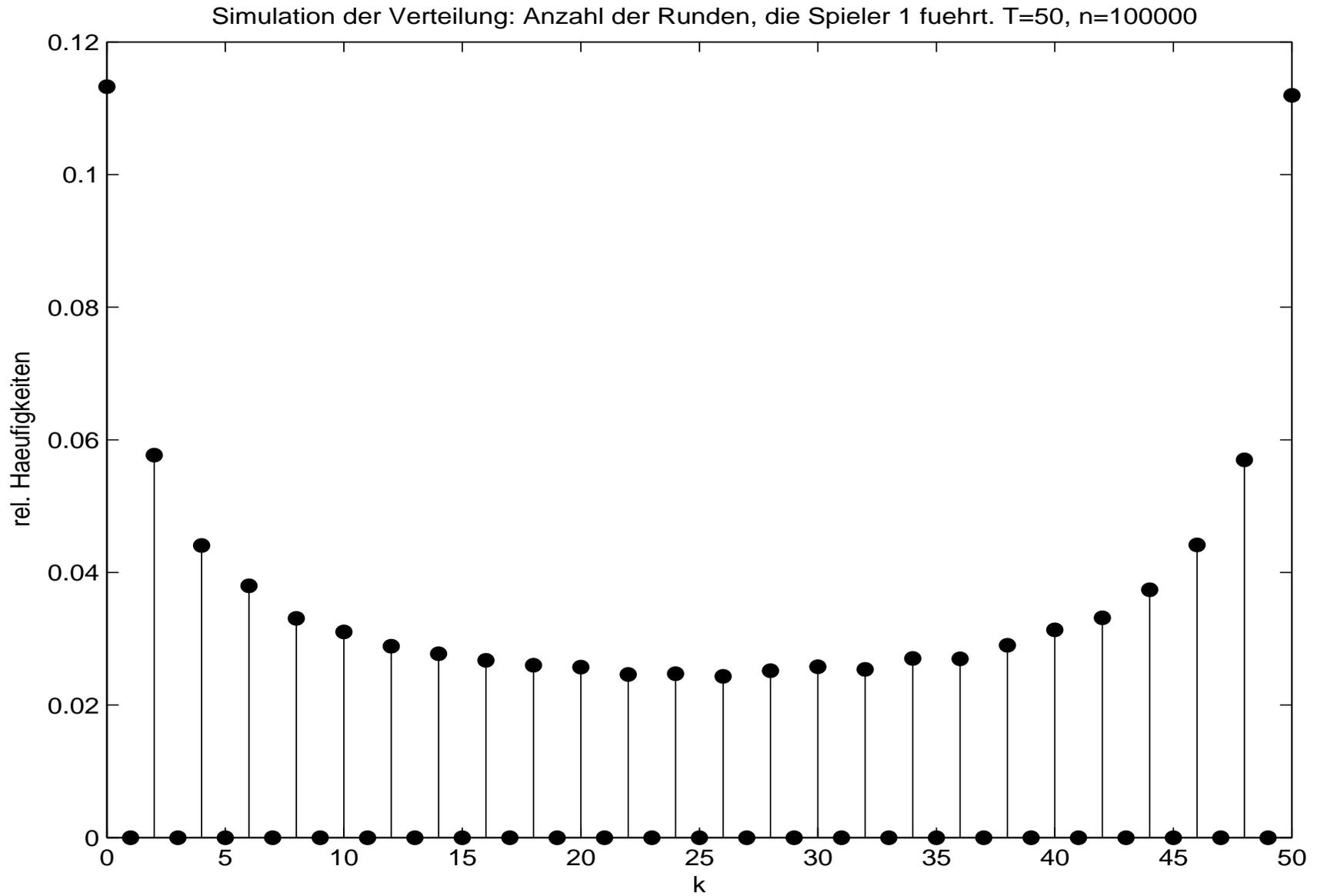
$$P'(\{X' = x\})$$

für $x = 0, \dots, T$.

$n = 1000$



$n = 100000$



41. Bemerkung $P'(\{X' = x\}) = 0$, falls T gerade und x ungerade.

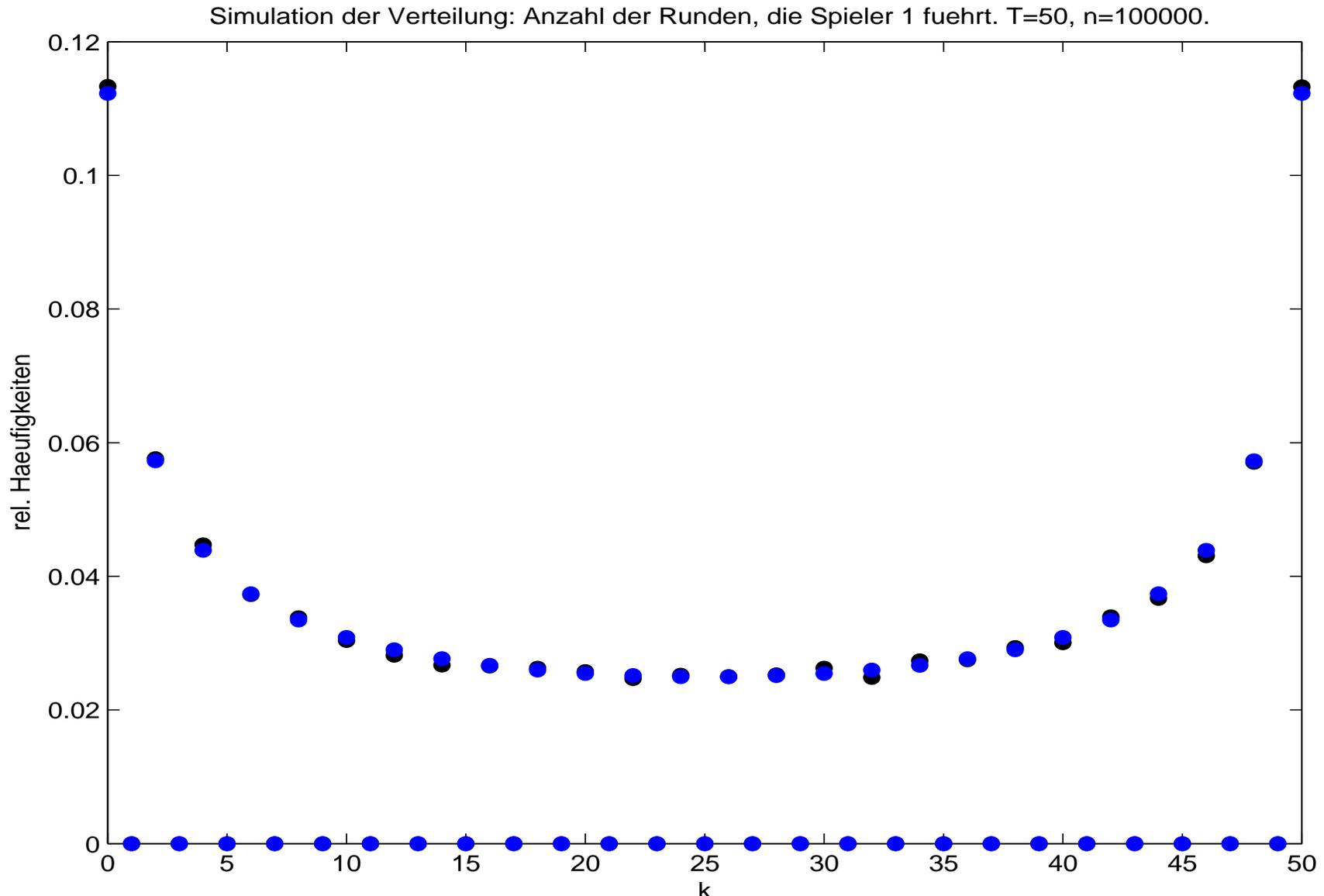
42. Satz Sei T gerade. Für $y \in \{0, \dots, T/2\}$

$$\begin{aligned} P'(\{X' = 2y\}) \\ = P'(\{S'_{2y} = 0\}) \cdot P'(\{S'_{T-2y} = 0\}). \end{aligned}$$

Beweis. TUTORIUM T3.

□

Vergleich der Simulationsergebnisse (schwarz) mit den exakten Werten (blau) für $n = 100000$



43. Korollar Sei T durch 4 teilbar. Dann

$$\frac{P'(\{X' = T\})}{P(\{X' = T/2\})} \approx \sqrt{\pi T/8}.$$

Beweis. Satz 42 und Korollar 37 zeigen

$$\frac{P'(\{X' = T\})}{P(\{X' = T/2\})} = \frac{P'(\{S'_T = 0\})}{\left(P'(\{S'_{T/2} = 0\})\right)^2} \approx \frac{\sqrt{2/(\pi T)}}{4/(\pi T)} = \sqrt{\pi T/8}.$$

44. Beispiel Für $T = 1000$ gilt

$$\sqrt{\pi T/8} = 19,8166 \dots$$

und

$$\frac{P'(\{X' = T\})}{P(\{X' = T/2\})} = 19,8315 \dots$$

Kap. V Verteilungen und absolutstetige Zufallsvariablen

1. Die Borelsche σ -Algebra in \mathbb{R}^d
2. Das d -dimensionale Lebesgue-Maß
3. Verteilungen
4. Absolutstetige Zufallsvariablen
5. Verteilungsfunktionen
6. Dichte-Schätzung

Bisher rigoros studiert: diskrete ZVen, also

$X : \Omega \rightarrow \mathbb{R}$ mit $P(\{X \in D\}) = 1$ für eine abzählbare Menge $D \subset \mathbb{R}$.

Dann gilt für alle $M \in \mathfrak{M}$, siehe Bsp. II.38,

$$\begin{aligned} P(\{X \in M\}) &= P(\{X \in M\} \cap \{X \in D\}) \\ &= P\left(\bigcup_{x \in D} \{X \in M\} \cap \{X = x\}\right) \\ &= \sum_{x \in D} P(\{X \in M\} \cap \{X = x\}) \\ &= \sum_{x \in M \cap D} P(\{X = x\}). \end{aligned}$$

In diesem Kapitel insbesondere ZVen

$X : \Omega \rightarrow \mathbb{R}$ mit

$$\forall x \in \mathbb{R} : P(\{X = x\}) = 0.$$

Bsp.: Wartezeit, Koordinaten von Pfeiltreffer auf Dartscheibe, fehlerhafter Meßwert, ...

Für eine große Klasse solcher ZVen wird die Summation von $x \mapsto P(\{X = x\})$ über $M \cap D$ durch die Integration einer geeigneten Funktion $x \mapsto f_X(x)$ über M ersetzt, also $P(\{X \in M\}) = \int_M f_X(x) dx$.

1 Die Borelsche σ -Algebra in \mathbb{R}^d

1. Beispiel Kontinuierliches „Glücksrad“. Versuch einer stochastischen Modellierung:

(i) $\Omega := [0, 1[$ (Kreislinie der Länge 1)

(ii) $\mathfrak{A} := \mathfrak{P}(\Omega)$

(iii) W 'maß P auf \mathfrak{A} mit folgenden Eigenschaften:

- $P([a, b]) = b - a$ für $0 \leq a < b < 1$
- $P(A) = P(B)$, falls B aus A durch „Rotation“ hervorgeht

Definiere für $\omega, \omega' \in [0, 1[$ und $A \subseteq [0, 1[$

$$\omega \oplus \omega' := \omega + \omega' - \lfloor \omega + \omega' \rfloor,$$

$$\omega \oplus A := \{\omega \oplus a : a \in A\}.$$

Frage: Existiert ein W' -maß P auf $\mathfrak{P}([0, 1[)$ mit

$$\forall A \subseteq [0, 1[\quad \forall \omega \in [0, 1[: P(\omega \oplus A) = P(A)?$$

Antwort: Nein.

Folglich gibt es keine „Gleichverteilung“ auf $\mathfrak{P}([0, 1[)$.

Ausweg: betrachte **kleinere σ -Algebra**.

Beweisskizze. Sei $Q := \mathbb{Q} \cap \Omega$. Betrachte Äquivalenzrelation

$$\omega \sim \omega' :\Leftrightarrow \exists q \in Q : \omega' = \omega \oplus q$$

auf Ω und zugehörige Äquivalenzklassen $[r] = \{\omega \in \Omega : \omega \sim r\}$. Wähle Repräsentantensystem $R \subseteq \Omega$ (Auswahlaxiom), d.h.

$$\forall \omega \in \Omega \quad \exists_1 r \in R : \omega \in [r].$$

Es gilt für $q_1, q_2 \in Q$ mit $q_1 \neq q_2$

$$(q_1 \oplus R) \cap (q_2 \oplus R) = \emptyset.$$

Schließlich erfüllt P mit obigen Eigenschaften

$$1 = P(\Omega) = P\left(\bigcup_{q \in Q} q \oplus R\right) = \sum_{q \in Q} P(q \oplus R) = \sum_{q \in Q} P(R).$$

Widerspruch.

Im folgenden $\Omega \neq \emptyset$ und $\mathcal{E} \subseteq \mathfrak{P}(\Omega)$ sowie

$$\mathbb{A} := \{\mathfrak{A} \subseteq \mathfrak{P}(\Omega) : \mathfrak{A} \text{ } \sigma\text{-Algebra, } \mathcal{E} \subseteq \mathfrak{A}\},$$

$$\sigma(\mathcal{E}) := \bigcap_{\mathfrak{A} \in \mathbb{A}} \mathfrak{A} = \{A \subseteq \Omega : \forall \mathfrak{A} \in \mathbb{A} : A \in \mathfrak{A}\}.$$

Beachte, daß $\mathfrak{P}(\Omega) \in \mathbb{A}$.

2. Lemma $\sigma(\mathcal{E})$ ist die kleinste σ -Algebra, die \mathcal{E} umfaßt, d.h.

- (i) $\sigma(\mathcal{E})$ ist σ -Algebra,
- (ii) $\mathcal{E} \subseteq \sigma(\mathcal{E})$,
- (iii) $\forall \mathfrak{A} \in \mathbb{A} : \sigma(\mathcal{E}) \subseteq \mathfrak{A}$.

Beweis. Ad (i), exemplarisch: Betrachte $A_1, A_2, \dots \in \sigma(\mathcal{E})$. Dann

$$\forall \mathfrak{A} \in \mathbb{A} : A_1, A_2, \dots \in \mathfrak{A},$$

so daß

$$\forall \mathfrak{A} \in \mathbb{A} : \bigcup_{i=1}^{\infty} A_i \in \mathfrak{A},$$

da $\mathfrak{A} \in \mathbb{A}$ σ -Algebra. Dies zeigt $\bigcup_{i=1}^{\infty} A_i \in \sigma(\mathcal{E})$.

Ad (ii): Nach Definition gilt $\forall \mathfrak{A} \in \mathbb{A} : \mathcal{E} \subseteq \mathfrak{A}$, d.h.

$$\forall \mathfrak{A} \in \mathbb{A} \forall E \in \mathcal{E} : E \in \mathfrak{A}.$$

Somit

$$\forall E \in \mathcal{E} : E \in \sigma(\mathcal{E}).$$

Ad (iii): Klar nach Definition. □

3. Definition $\sigma(\mathfrak{E})$ die von \mathfrak{E} erzeugte σ -Algebra (in Ω).

Vgl. erzeugter Untervektorraum.

4. Beispiel Für $\Omega = \{1, 2, 3, 4\}$ und $\mathfrak{E} = \{\{1, 2, 3\}, \{2\}\}$

gilt

$\sigma(\mathfrak{E})$

$$= \{\emptyset, \{2\}, \{4\}, \{1, 3\}, \{2, 4\}, \{1, 2, 3\}, \{1, 3, 4\}, \Omega\}$$

$$= \{A \subseteq \Omega : \{1, 3\} \subseteq A \text{ oder } \{1, 3\} \cap A = \emptyset\}.$$

5. Definition Für $d \in \mathbb{N}$ und

$$\mathfrak{O}_d := \{O \subseteq \mathbb{R}^d : O \text{ offen}\}$$

heißt $\mathfrak{B}_d := \sigma(\mathfrak{O}_d)$ die **Borelsche σ -Algebra** in \mathbb{R}^d . Elemente $B \in \mathfrak{B}_d$ heißen **Borel-Mengen** (in \mathbb{R}^d).

6. Beispiel

(i) $A \subseteq \mathbb{R}^d$ abgeschlossen $\Rightarrow A \in \mathfrak{B}_d$, da A^c offen

(ii) $\mathfrak{M} \subsetneq \mathfrak{B}_1$

7. Lemma

$$A_1, \dots, A_d \in \mathfrak{B}_1 \Rightarrow A_1 \times \dots \times A_d \in \mathfrak{B}_d.$$

Beweis. Siehe Irle (2001, p. 151). Stichwort: Produkt- σ -Algebra. □

8. Bemerkung Es gilt $\mathfrak{B}_d \subsetneq \mathfrak{P}(\mathbb{R}^d)$. Uns werden in dieser Vorlesung jedoch keine Mengen aus $\mathfrak{P}(\mathbb{R}^d) \setminus \mathfrak{B}_d$ begegnen.

Dazu auch Krengel (2003, p. 127).

9. Satz Gilt

$$\begin{aligned}\forall x_1, \dots, x_d \in \mathbb{R} : P(\left] -\infty, x_1 \right] \times \dots \times \left] -\infty, x_d \right]) \\ = Q(\left] -\infty, x_1 \right] \times \dots \times \left] -\infty, x_d \right])\end{aligned}$$

für Wahrscheinlichkeitsmaße P und Q auf \mathfrak{B}_d , so folgt

$$P = Q.$$

Beweis. Siehe Irle (2001, p. 157).

□

2 Das d -dimensionale Lebesgue-Maß

10. Satz Es existiert genau eine σ -additive Abbildung

$\lambda_d : \mathfrak{B}_d \rightarrow [0, \infty]$ mit

$$\forall a_i \leq b_i : \lambda_d([a_1, b_1] \times \cdots \times [a_d, b_d]) = \prod_{i=1}^d (b_i - a_i).$$

Diese erfüllt

$$\lambda_d(a + Q(A)) = \lambda_d(A)$$

für alle $a \in \mathbb{R}^d$, $A \in \mathfrak{B}_d$ und alle orthogonalen Abbildungen

$$Q : \mathbb{R}^d \rightarrow \mathbb{R}^d.$$

11. Definition λ_d gem. Satz 10 heißt **Lebesgue-Maß** auf \mathfrak{B}_d oder d -dimensionales Lebesgue-Maß.

12. Bemerkung Definiere

$$\Omega := [0, 1] \text{ bzw. } [0, 1[,$$

$$\mathfrak{A} := \{A \cap \Omega : A \in \mathfrak{B}_1\},$$

$$P(A) := \lambda_1(A), \quad A \in \mathfrak{A},$$

$$U(\omega) := \omega, \quad \omega \in \Omega.$$

Dann ist $(\Omega, \mathfrak{A}, P)$ ein W'raum, U eine ZV und es gilt $U \sim \mathbf{U}(\Omega)$. Siehe Definitionen IV.25 und V.28.

3 Verteilungen

Im folgenden $(\Omega, \mathfrak{A}, P)$ W'raum und $X_1, \dots, X_d : \Omega \rightarrow \mathbb{R}$.

Setze $X := (X_1, \dots, X_d) : \Omega \rightarrow \mathbb{R}^d$.

13. Lemma Äquivalent sind:

(i) X_1, \dots, X_d Zufallsvariablen

(ii) $\forall A \in \mathfrak{B}_d : \{X \in A\} \in \mathfrak{A}$

Vgl. mit Lemma II.39 im Fall $d = 1$.

Vgl. Begriff der [Meßbarkeit](#) von Abbildungen.

Beweis. „(ii) \Rightarrow (i)“

$$\{X_i \leq c_i\} = \{X_i \leq c_i\} \cap \bigcap_{j \neq i} \{X_j \in \mathbb{R}\} = \{X \in A\}$$

für $A = \mathbb{R}^{i-1} \times]-\infty, c_i] \times \mathbb{R}^{d-i}$. Da A abgeschlossen,

folgt $A \in \mathfrak{B}_d$ und wg. (ii) auch $\{X_i \leq c_i\} \in \mathfrak{A}$.

„(i) \Rightarrow (ii)“ Siehe Irle (2001, p. 151). Stichwort: Produkt-Meßbarkeit. \square

Fortan X_1, \dots, X_d Zufallsvariablen; X heißt dann auch

Zufallsvektor.

Zufallsvektoren dienen zur gemeinsamen Modellierung mehrerer Aspekte

eines Zufallsexperimentes; bisher etwa bei Unabhängigkeit von

Zufallsvariablen und in Bem IV.33.

14. Satz

$$P_X(A) := P(\{X \in A\}), \quad A \in \mathfrak{B}_d,$$

definiert Wahrscheinlichkeitsmaß auf \mathfrak{B}_d .

Beweis. Klar: $P_X \geq 0$ und $P_X(\mathbb{R}^d) = P(\{X \in \mathbb{R}^d\}) = P(\Omega) = 1$.

Für $A_1, A_2, \dots \in \mathfrak{B}_d$ p.d. und $A := \bigcup_{i=1}^{\infty} A_i$ gilt

$$\begin{aligned} P_X(A) &= P\left(\left\{X \in \bigcup_{i=1}^{\infty} A_i\right\}\right) = P\left(\bigcup_{i=1}^{\infty} \underbrace{\{X \in A_i\}}_{\in \mathfrak{A} \text{ p.d.}}\right) \\ &= \sum_{i=1}^{\infty} P(\{X \in A_i\}) = \sum_{i=1}^{\infty} P_X(A_i). \end{aligned}$$

□

15. Definition P_X heißt **Verteilung** von X . Im Falle $d > 1$ heißt P_X auch **gemeinsame Verteilung** von X_1, \dots, X_d und P_{X_1}, \dots, P_{X_d} heißen **(eindim.) Randverteilungen** von X .

Viele Fragestellungen der Stochastik betreffen nicht die konkrete Gestalt des zugrundeliegenden W 'raumes und der betrachteten ZVen sondern nur ihre gemeinsame Verteilung. Bsp.: Unabhängigkeit, siehe Satz 21.

16. Beispiel P Gleichverteilung auf $\Omega := \{0, 1\}^2$

(zweimaliger Münzwurf) und $X_i(\omega) := \omega_i$.

Gemeinsame Verteilung von X_1 und X_2 : für $A \in \mathfrak{B}_2$:

$$\begin{aligned} P_X(A) &= P(\{X \in A\}) = |\{\omega \in \Omega : \omega \in A\}|/4 \\ &= |A \cap \Omega|/4 \end{aligned}$$

Randverteilungen: für $B \in \mathfrak{B}_1$:

$$\begin{aligned} P_{X_1}(B) &= P(\{X \in B \times \mathbb{R}\}) = |\{\omega \in \Omega : \omega_1 \in B\}|/4 \\ &= |\{\omega_1 \in \{0, 1\} : \omega_1 \in B\}|/2 = |B \cap \{0, 1\}|/2 \end{aligned}$$

Analog $P_{X_2}(B) = |B \cap \{0, 1\}|/2$.

Insbesondere: $P_{X_1} = P_{X_2}$, obwohl $X_1 \neq X_2$.

Jetzt $X'_1(\omega) := \omega_1$ und $X'_2(\omega) := \omega_1$. **Randverteilungen:**

$$P_{X'_1} = P_{X'_2} = P_{X_1} = P_{X_2}.$$

Gemeinsame Verteilung: für $A \in \mathfrak{B}_2$:

$$\begin{aligned} P_{X'}(A) &= P(\{X' \in A\}) = |\{\omega \in \Omega : (\omega_1, \omega_1) \in A\}|/4 \\ &= |\{\omega_1 \in \{0, 1\} : (\omega_1, \omega_1) \in A\}|/2 \end{aligned}$$

Also $P_X \neq P_{X'}$. Dies zeigt:

gemeinsame Verteilung ist durch die eindimensionalen
Randverteilungen nicht eindeutig bestimmt

Siehe jedoch Bem. 22.(i).

17. Bemerkung Jedes W -Maß P auf \mathfrak{B}_d ist Verteilung eines Zufallsvektors: betrachte $X(\omega) := \omega$ auf dem W -raum $(\mathbb{R}^d, \mathfrak{B}_d, P)$.

18. Definition X **diskreter Zufallsvektor**, falls

$P(\{X \in D\}) = 1$ für eine abzählbare Menge $D \subset \mathbb{R}^d$.

Siehe Definition III20 im Falle $d = 1$.

19. Bemerkung Berechnung der Verteilung P_X diskreter Zufallsvektoren X prinzipiell durch **Summation** der Funktion $\mathbb{R}^d \rightarrow \mathbb{R}_+ : x \mapsto P(\{X = x\})$. Vgl. Wahrscheinlichkeitsfunktion.
Genauer: für X und D wie oben sowie $A \in \mathfrak{B}_d$ ergibt sich wie auf Seite 190:

$$P_X(A) = P(\{X \in A\}) = \sum_{x \in A \cap D} P(\{X = x\}).$$

Betrachte nochmals Bsp. 16.

20. Beispiel P Gleichverteilung auf $\Omega := \{1, \dots, 6\}^2$
(zweimaliges Würfeln) und

$$X_1(\omega) := \omega_1, \quad X_2(\omega) := \omega_1 + \omega_2.$$

Für

$$D := \{x \in \{1, \dots, 6\} \times \{2, \dots, 12\} \\ : 1 \leq x_2 - x_1 \leq 6\}$$

gilt $P(\{X \in D\}) = 1$ sowie $P(\{X = x\}) = 1/36$ für alle $x \in D$. Also

$$P_X(A) = \sum_{x \in A \cap D} P(\{X = x\}) = \frac{1}{36} \cdot |A \cap D|.$$

21. Satz (X_1, \dots, X_d) genau dann unabhängig, wenn

$\forall A_1, \dots, A_d \in \mathfrak{B}_1 :$

$$\underbrace{P\left(\bigcap_{i=1}^d \{X_i \in A_i\}\right)}_{=P_X(A_1 \times \dots \times A_d)} = \prod_{i=1}^d \underbrace{P(\{X_i \in A_i\})}_{=P_{X_i}(A_i)},$$

kurz: **gemeinsame Verteilung ist Produkt der Randverteilungen.**

Beweis. „ \Leftarrow “ Klar.

„ \Rightarrow “ Siehe Irle (2001, p. 169). Teilaussage in Satz II.46.

Hier Beweis unter der zusätzlichen Annahme, daß

X_1, \dots, X_d diskret.

Wähle abzählbare Menge $D \subset \mathbb{R}$ mit

$$P(\{X \in D^d\}) = P\left(\bigcap_{i=1}^d \{X_i \in D\}\right) = 1.$$

Für $A = A_1 \times \cdots \times A_d$ folgt mit Satz II.46 und Bem. 19:

$$\begin{aligned} P\left(\bigcap_{i=1}^d \{X_i \in A_i\}\right) &= P(\{X \in A\}) = \sum_{x \in A \cap D^d} P(\{X = x\}) \\ &= \sum_{x_1 \in A_1 \cap D} \cdots \sum_{x_d \in A_d \cap D} \prod_{i=1}^d P(\{X_i = x_i\}) \\ &= \prod_{i=1}^d P(\{X_i \in A_i\}) \end{aligned}$$

□

22. Bemerkung

- (i) Falls X_1, \dots, X_d unabhängig, so ist die gemeinsame Verteilung P_X eindeutig durch die Randverteilungen P_{X_1}, \dots, P_{X_d} bestimmt. Vgl. Modellierung in Kap. III.

Beweis. Verwende Definition II.43 und Satz 9. □

- (ii) Diskrete ZVen X_1, \dots, X_d sind genau dann unabhängig, wenn

$$\forall x \in \mathbb{R}^d : P(\{X = x\}) = \prod_{i=1}^d P(\{X_i = x_i\}).$$

Beweis. „ \Rightarrow “ Siehe Satz II.46 bzw. Satz 21. „ \Leftarrow “ Siehe Seite 213.

4 Absolutstetige Verteilungen und Zufallsvariablen

23. Bemerkung Integralbegriffe für Funktionen $f : \mathbb{R}^d \rightarrow \mathbb{R}$

- (i) **Lebesgue-Integral** (Vorlesung Analysis IV). **Spezialfall:**
- (ii) **Uneigentliches Riemann-Integral** (Walther, Analysis II, Springer, 1990, §7.20). **Spezialfall:** Für abgeschlossene Intervalle $B_i \subseteq \mathbb{R}$ und $B := B_1 \times \cdots \times B_d \subseteq \mathbb{R}^d$ sei $f|_B$ stetig. Setze $B^{(K)} := B \cap [-K, K]^d$. Falls $\sup_{K \in \mathbb{N}} \int_{B^{(K)}} |f(x)| dx < \infty$, so gilt

$$\int_B f(x) dx = \lim_{K \rightarrow \infty} \underbrace{\int_{B^{(K)}} f(x) dx}_{\text{Berechnung als iteriertes Integral}}$$

Berechnung als **iteriertes Integral**

24. Definition $f : \mathbb{R}^d \rightarrow \mathbb{R}_+$ **Wahrscheinlichkeitsdichte**, kurz **Dichte**, falls f (Lebesgue)-integrierbar mit

$$\int_{\mathbb{R}^d} f(x) dx = 1.$$

25. Satz Jede Dichte f definiert durch

$$P(A) := \int_A f(x) dx, \quad A \in \mathfrak{B}_d,$$

ein Wahrscheinlichkeitsmaß auf \mathfrak{B}_d .

Vgl. Satz III.3 über Wahrscheinlichkeitsfunktionen. Ausblick: singuläre Verteilungen.

Beweis von Satz 25. Klar: $P \geq 0$ und $P(\mathbb{R}^d) = 1$.

Für $A_1, A_2, \dots \in \mathfrak{B}_d$ p.d. und $A := \bigcup_{i=1}^{\infty} A_i$ gilt

$$\begin{aligned} P(A) &= \int_{\mathbb{R}^d} 1_A(x) \cdot f(x) \, dx = \int_{\mathbb{R}^d} \sum_{i=1}^{\infty} 1_{A_i}(x) \cdot f(x) \, dx \\ &= \sum_{i=1}^{\infty} \int_{\mathbb{R}^d} 1_{A_i}(x) \cdot f(x) \, dx = \sum_{i=1}^{\infty} P(A_i) \end{aligned}$$

aufgrund des Satzes von der monotonen Konvergenz. □

Zur Eindeutigkeit von Dichten:

26. Lemma Seien $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$ integrierbar. Dann sind äquivalent:

$$(i) \quad \forall A \in \mathfrak{B}_d : \int_A f(x) dx = \int_A g(x) dx$$

$$(ii) \quad \lambda_d(\{x \in \mathbb{R}^d : f(x) \neq g(x)\}) = 0$$

Beweis. Folgt aus Meintrup, Schäffler (2005, Satz 2.15). □

Im folgenden: $X = (X_1, \dots, X_d)$ d -dimensionaler Zufallsvektor auf $(\Omega, \mathfrak{A}, P)$.

27. Definition X **absolutstetig verteilt**, falls P_X eine Dichte besitzt. Diese wird ggf. mit f_X bezeichnet.

Nun: Modellierung von Verteilungen durch Vorgabe ihrer Dichten.

28. Definition Sei $B \in \mathfrak{B}_d$ mit Lebesgue-Maß (Länge, Flächeninhalt, Volumen) $\lambda_d(B) \in]0, \infty[$. Zufallsvektor (bzw. -variable) X mit Dichte

$$f_X(x) = \frac{1}{\lambda_d(B)} \cdot 1_B(x)$$

heißt **gleichverteilt** auf B .

Bez.: $X \sim \mathbf{U}(B)$.

29. Bemerkung Für $X \sim \mathbf{U}(B)$ und $A \in \mathfrak{B}_d$:

$$P_X(A) = \frac{1}{\lambda_d(B)} \cdot \int_A 1_B(x) dx = \frac{\lambda_d(A \cap B)}{\lambda_d(B)}$$

30. Beispiel Dichte und Verteilungsfunktion von

$X \sim \mathbf{U}([a, b])$ mit $a < b$:

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{falls } x \in [a, b] \\ 0, & \text{sonst} \end{cases}$$

$$F_X(x) = \begin{cases} 0, & \text{falls } x < a \\ \frac{x-a}{b-a}, & \text{falls } x \in [a, b] \\ 1, & \text{sonst} \end{cases}$$

Vgl. Definition IV.25.

31. Beispiel $X \sim \mathbf{U}(B)$ zur Modellierung von Pfeiltreffer auf Dartscheibe, Glücksrad.

Anwendung: **Zufallszahlen** und **stochastische Simulation**,
siehe Kapitel IV.

32. Definition Zufallsvariable X mit Dichte

$$f_X(x) = \begin{cases} \lambda \cdot \exp(-\lambda x), & \text{falls } x \geq 0 \\ 0, & \text{sonst} \end{cases}$$

für $\lambda > 0$ heißt **exponentialverteilt** mit Parameter λ .

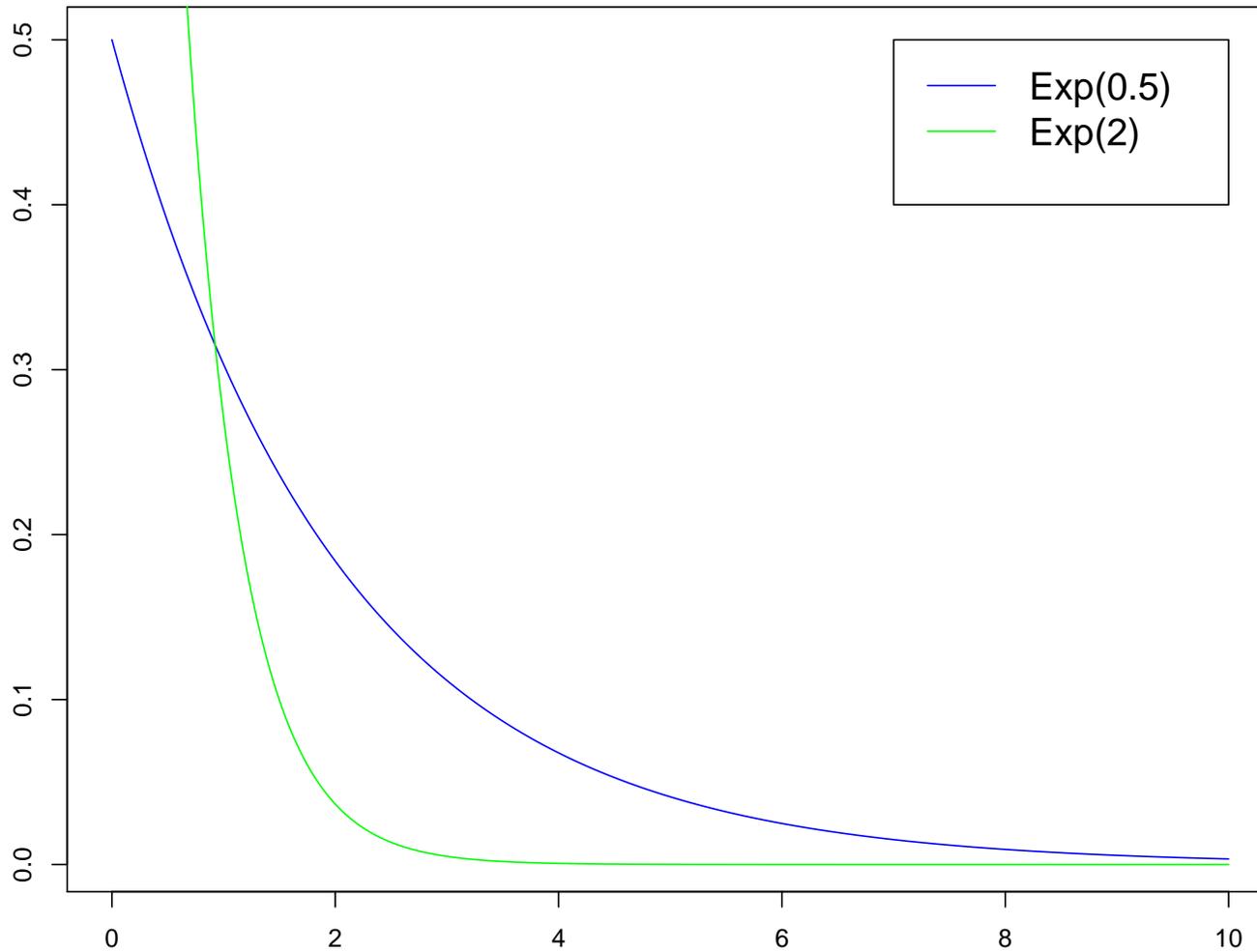
Bez.: $X \sim \mathbf{Exp}(\lambda)$

33. Bemerkung Für $X \sim \mathbf{Exp}(\lambda)$ und $x > 0$:

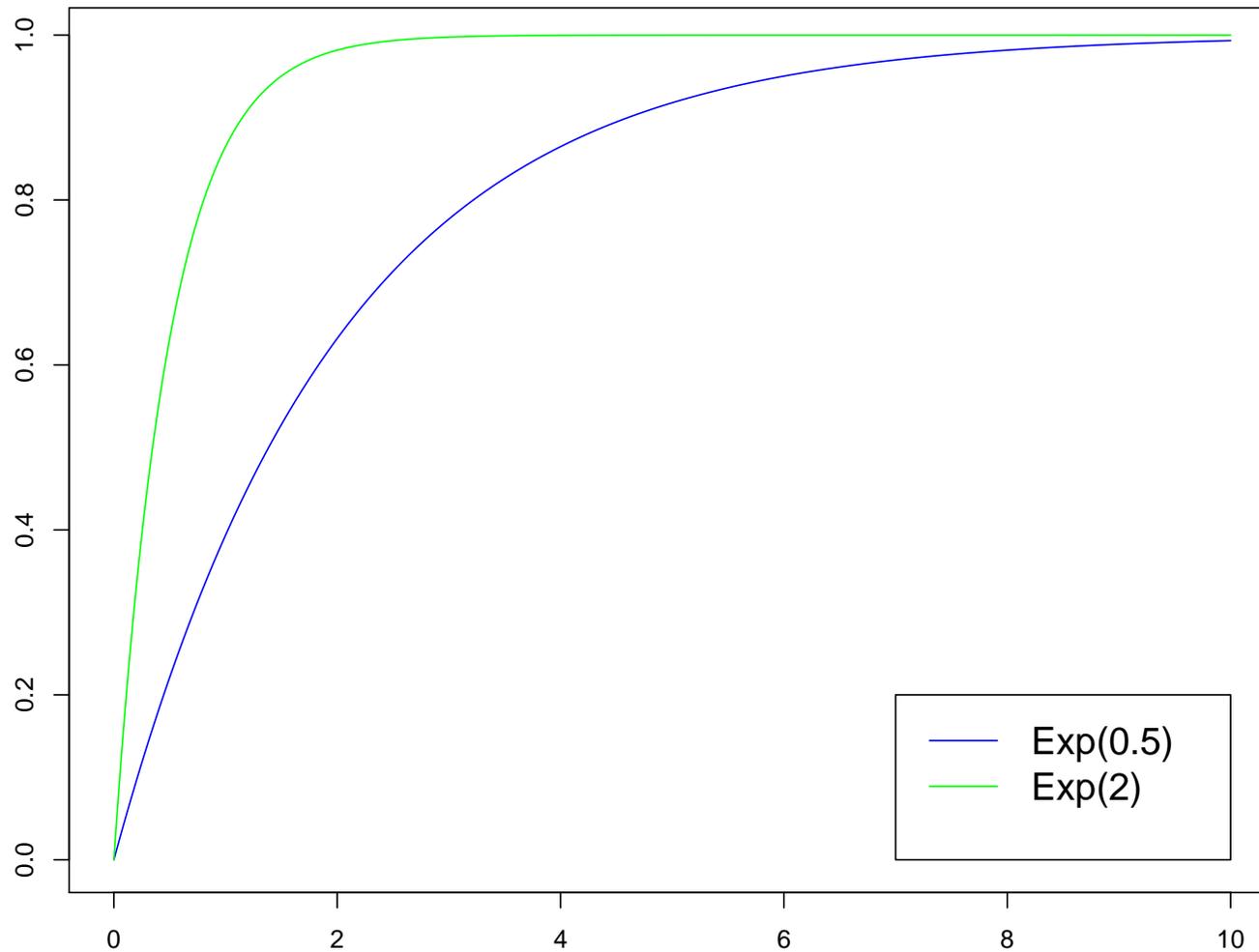
$$F_X(x) = \lambda \cdot \int_0^x \exp(-\lambda y) dy = 1 - \exp(-\lambda x)$$

Klar: $F_X(x) = 0$, falls $x \leq 0$.

34. Beispiel Dichten exponentialverteilter ZVen



35. Beispiel Verteilungsfunktionen exponentialverteilter ZVen



36. Satz Charakterisierung der Exp'-verteilung durch **Gedächtnislosigkeit**

Für ZV X mit

- $P(\{X > 0\}) = 1$ und
- $\forall t > 0 : P(\{X > t\}) > 0$

sind äquivalent:

(i) $\exists \lambda > 0 : X \sim \mathbf{Exp}(\lambda)$

(ii) $\forall s, t > 0 :$

$$P(\{X > t + s\} \mid \{X > t\}) = P(\{X > s\})$$

Beweis. Siehe ÜBUNG M:H21 und vgl. ÜBUNG M:H13 und WInf:H12. \square

37. Beispiel $X \sim \mathbf{Exp}(\lambda)$ zur Modellierung von Lebensdauern, Wartezeiten.

Hier: radioaktiver Zerfall, X Zerfallszeitpunkt. Halbwertszeit $h > 0$ definiert als Median,

$$P(\{X \leq h\}) = \frac{1}{2}.$$

Man erhält

$$h = \ln(2)/\lambda.$$

38. Beispiel $(X_i)_{i \in \mathbb{N}}$ iid, $X_1 \sim \mathbf{Exp}(\lambda)$ mit unbekanntem $\lambda > 0$. Problem: Schätze λ bzw. $h = \ln(2)/\lambda$.

Gem. Kap. IV.9 gilt fast sicher

$$\lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{i=1}^n 1_{]-\infty, h]}(X_i) = \frac{1}{2}.$$

Auf Basis von Realisierungen $x_i = X_i(\omega)$ schätzt man h durch den Median der empirischen Verteilungsfunktion (**empirischer Median**). Siehe Beispiel IV.24.

Warnung.

39. Lemma Für $\mu \in \mathbb{R}$ und $\sigma > 0$ gilt

$$\int_{-\infty}^{\infty} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx = \sqrt{2\pi\sigma^2}.$$

Beweis. OBdA $\mu = 0$ und $\sigma = 1$ (Substitutionsregel). Es gilt:

$$\begin{aligned} & \left(\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx \right)^2 \\ &= \int_{\mathbb{R}^2} \exp\left(-\frac{x^2 + y^2}{2}\right) d(x, y) \\ &= \int_0^{2\pi} \int_0^{\infty} \exp\left(-\frac{r^2}{2}\right) \cdot r dr d\varphi \\ &= 2\pi \cdot \left(-\exp\left(-\frac{r^2}{2}\right) \right) \Big|_0^{\infty} = 2\pi \end{aligned}$$

□

40. Definition Zufallsvariable X mit Dichte

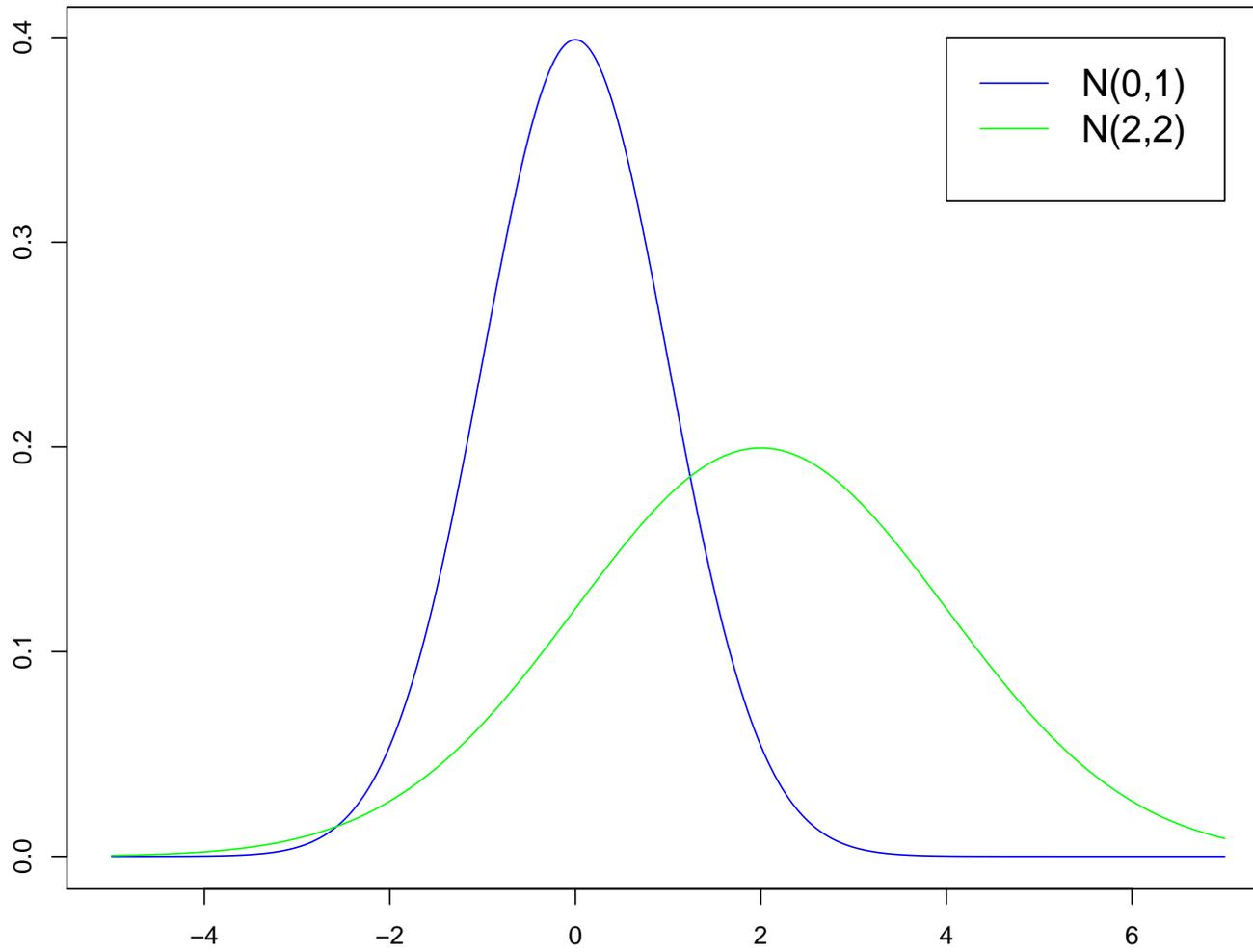
$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

für $\mu \in \mathbb{R}$ und $\sigma > 0$ heißt **normalverteilt** mit Parametern μ und σ^2 .

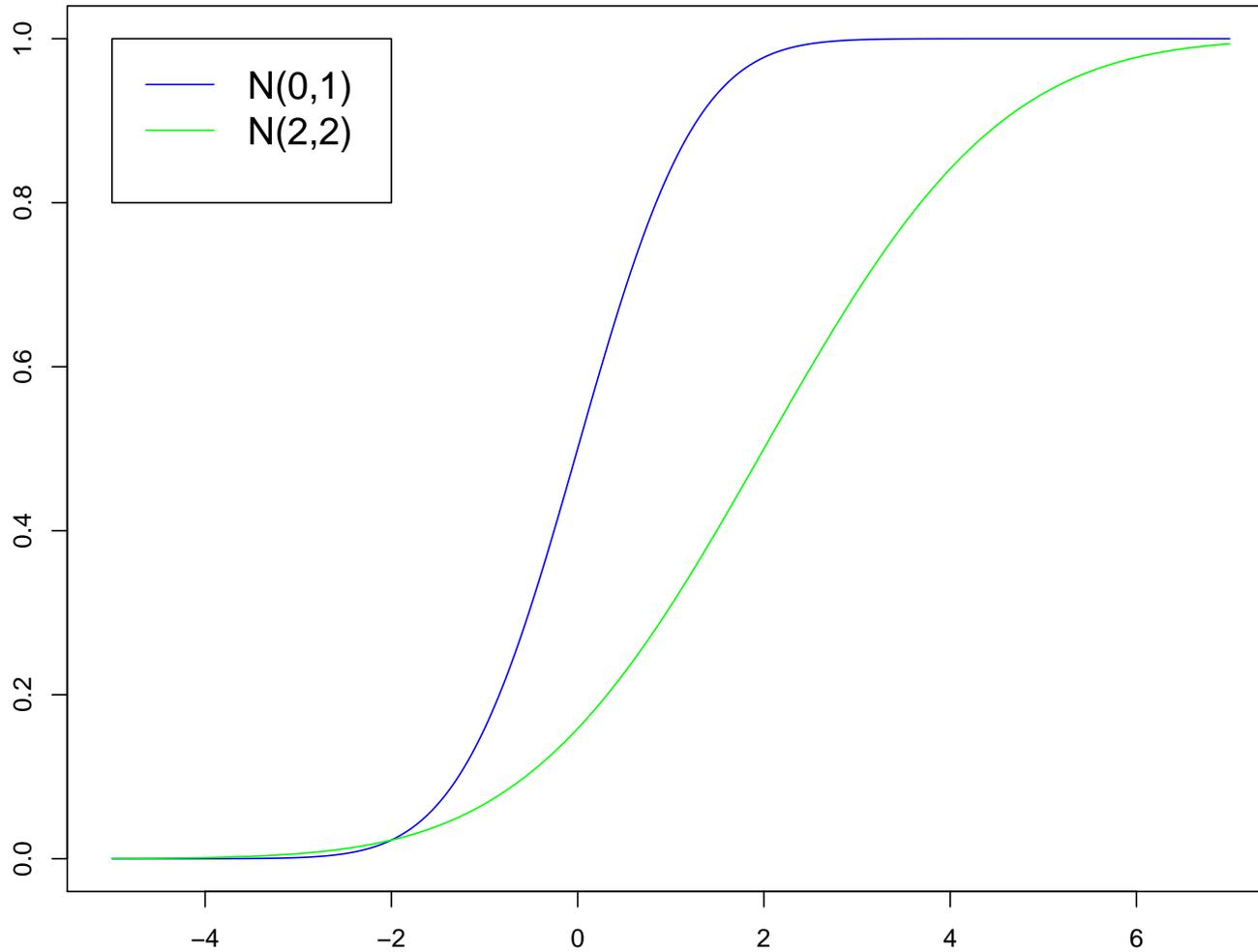
Bez.: $X \sim \mathbf{N}(\mu, \sigma^2)$

Standard-Normalverteilung als Spezialfall: $\mu = 0$ und $\sigma = 1$.

41. Beispiel Dichten normalverteilter ZVen



42. Beispiel Verteilungsfunktionen normalverteilter ZVen



43. Beispiel $X \sim \mathbf{N}(\mu, \sigma^2)$ zur Modellierung (Meß)Fehlern.

Siehe auch Kap. VI.3.

44. Bemerkung Keine explizite Formel für Verteilungsfunktion F_X , falls $X \sim \mathbf{N}(\mu, \sigma^2)$.

Bez.: $\Phi = F_X$, falls $X \sim \mathbf{N}(0, 1)$, also

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{y^2}{2}\right) dy.$$

Zur Berechnung von Φ und entsprechender Quantile:

Numerik, Tabellen, Plots.

Nun speziell: mehrdimensionale Dichten.

Analytisches Hilfsmittel: Satz von Fubini.

45. Lemma Falls f_X Dichte von P_X , so besitzt P_{X_i} die Dichte

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{d-1}} f_X(\bar{x}_1, x_i, \bar{x}_2) d(\bar{x}_1, \bar{x}_2)$$

mit

$$\bar{x}_1 = (x_1, \dots, x_{i-1}), \quad \bar{x}_2 = (x_{i+1}, \dots, x_d).$$

Beweis. Für $A_i \in \mathfrak{B}_1$ sei $A := \mathbb{R}^{i-1} \times A_i \times \mathbb{R}^{d-i}$. Dann

$$\begin{aligned}
 P(\{X_i \in A_i\}) &= P(\{X \in A\}) = \int_A f_X(x) dx \\
 &= \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{A_i} \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f_X(x_1, \dots, x_d) dx_d \cdots dx_1 \\
 &= \int_{A_i} \underbrace{\int_{\mathbb{R}^{d-1}} f_X(\bar{x}_1, x_i, \bar{x}_2) d(\bar{x}_1, \bar{x}_2)}_{=:g(x_i)} dx_i,
 \end{aligned}$$

und g ist eine Dichte. □

46. Beispiel Pfeiltreffer auf Dartscheibe. Hier

$$f_X(x_1, x_2) := \frac{1}{\pi r^2} \cdot 1_K(x_1, x_2)$$

mit

$$K := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq r^2\}$$

Also für $x_1 \in [-r, r]$

$$\begin{aligned} f_{X_1}(x_1) &= \int_{\mathbb{R}} f_X(x_1, x_2) dx_2 = \frac{1}{\pi r^2} \cdot \int_{-\sqrt{r^2-x_1^2}}^{\sqrt{r^2-x_1^2}} 1 dx_2 \\ &= \frac{2}{\pi r^2} \cdot \sqrt{r^2 - x_1^2} \end{aligned}$$

sowie $f_{X_1}(x_1) = 0$, falls $|x_1| > r$. Klar:

$$f_{X_1} = f_{X_2}$$

47. Definition Tensorprodukt $f_1 \otimes \dots \otimes f_d : \mathbb{R}^d \rightarrow \mathbb{R}$ von Abbildungen $f_i : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f_1 \otimes \dots \otimes f_d(x) := f_1(x_1) \cdot \dots \cdot f_d(x_d).$$

Vgl. Abschnitt III.3.

48. Lemma Falls f_1, \dots, f_d Dichten auf \mathbb{R} , so ist $f_1 \otimes \dots \otimes f_d$ Dichte auf \mathbb{R}^d .

Beweis. Klar. Vgl. Lemma III.14. □

49. Satz

(i) Falls X_1, \dots, X_d unabhängig mit Dichten f_{X_i} , so besitzt X die Dichte

$$f_X = f_{X_1} \otimes \dots \otimes f_{X_d}.$$

(ii) Falls X die Dichte

$$f_X = f_1 \otimes \dots \otimes f_d$$

mit eindimensionalen Dichten f_i besitzt, so sind

X_1, \dots, X_d unabhängig mit Dichten $f_{X_i} = f_i$.

Beweis. Ad (i): Gemäß Satz 25 und Lemma 48 definiert

$$Q(A) := \int_A f_{X_1} \otimes \dots \otimes f_{X_d}(x) dx, \quad A \in \mathfrak{B}_d,$$

ein W'maß auf \mathfrak{B}_d . Speziell für $A := A_1 \times \dots \times A_d$ mit $A_i :=]-\infty, b_i]$

$$\begin{aligned} P(\{X \in A\}) &= \prod_{i=1}^d P(\{X_i \in A_i\}) = \prod_{i=1}^d \int_{A_i} f_{X_i}(x_i) dx_i \\ &= \int_{A_1} \dots \int_{A_d} \prod_{i=1}^d f_{X_i}(x_i) dx_d \dots dx_1 = Q(A). \end{aligned}$$

Satz 9 zeigt $P_X = Q$.

Ad (ii): Für $A_1, \dots, A_d \in \mathfrak{B}_1$ und $A := A_1 \times \dots \times A_d$

$$P(\{X \in A\}) = \int_A f_X(x) dx = \int_{A_1} f_1(x_1) dx_1 \cdots \int_{A_d} f_d(x_d) dx_d.$$

Insbesondere

$$P(\{X_i \in A_i\}) = \int_{A_i} f_i(x_i) dx_i,$$

d.h. f_i ist Dichte von X_i , und weiter

$$P\left(\bigcap_{i=1}^d \{X_i \in A_i\}\right) = \prod_{i=1}^d P(\{X_i \in A_i\}).$$

□

50. Bemerkung Mit Satz 49: Modellierung der „unabhängigen Hintereinanderausführung“ von Einzelexperimenten, deren Verteilungen Dichten besitzen.

51. Beispiel Pfeiltreffer auf Dartscheibe, siehe Bsp. 46.

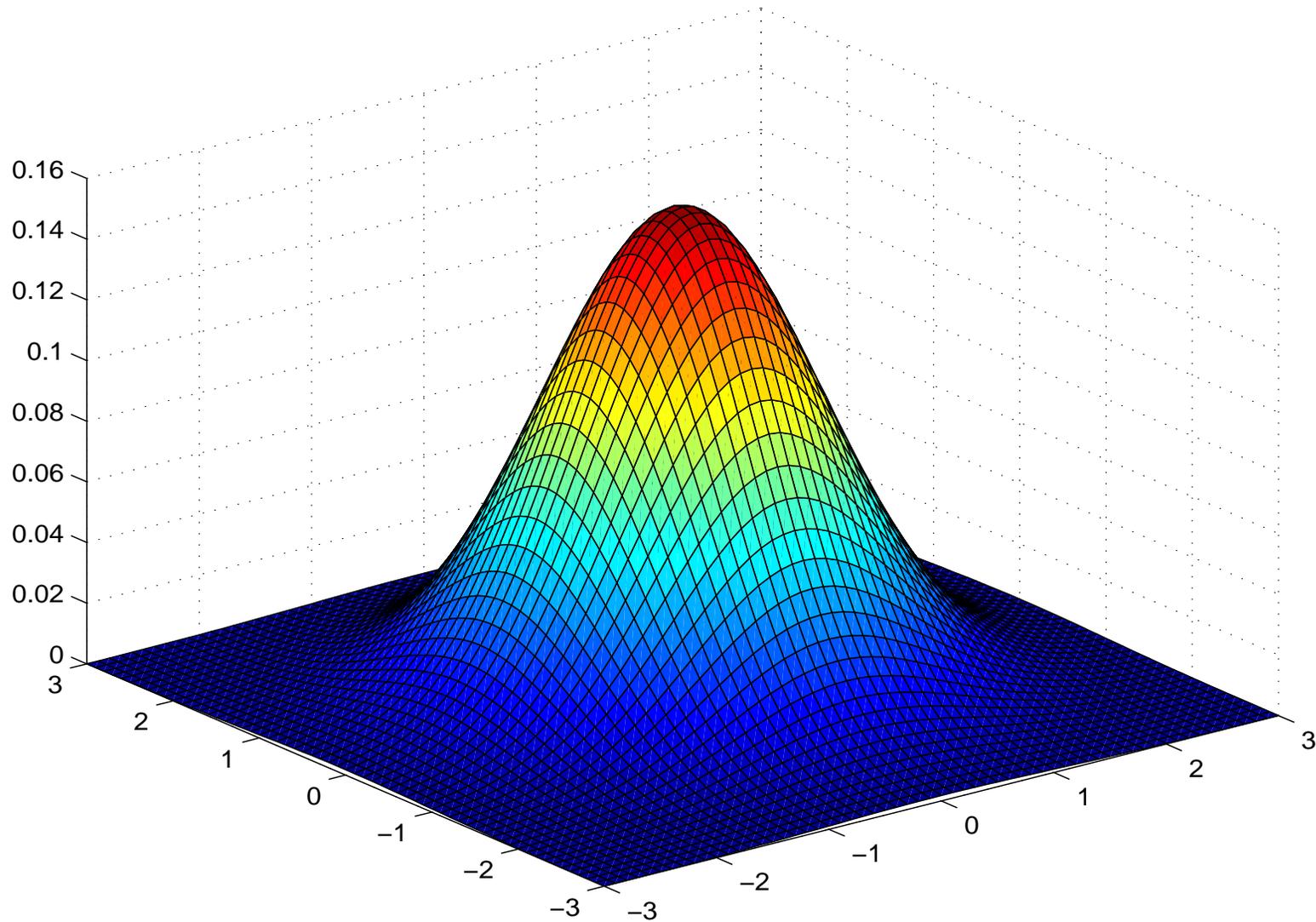
Satz 49 zeigt: X_1, X_2 nicht unabhängig.

52. Definition d -dimensionaler Zufallsvektor X mit Dichte

$$f_X(x) = (2\pi)^{-d/2} \cdot \exp \left(-\frac{1}{2} \sum_{i=1}^d x_i^2 \right)$$

heißt **standard-normalverteilt** (in \mathbb{R}^d).

53. Beispiel Dichte einer 2-dim. normalverteilten ZV



5 Verteilungsfunktionen

54. Satz Für ZV X sind äquivalent:

- (i) F_X stetig differenzierbar
- (ii) X absolutstetig verteilt mit stetiger Dichte f_X

Gilt (i) oder (ii), so folgt $F'_X = f_X$.

Beweis. “(i) \Rightarrow (ii)” Nach dem Hauptsatz der Differential- und Integralrechnung gilt

$$F_X(x) = \int_{-\infty}^x F'_X(u) du.$$

Gem. Satz 9 stimmt P_X mit dem Wahrscheinlichkeitsmaß mit der Dichte F'_X überein.

“(ii) \Rightarrow (i)” Hauptsatz der Differential- und Integralrechnung. □

55. Satz Für ZVen X auf $(\Omega, \mathfrak{A}, P)$ und X' auf $(\Omega', \mathfrak{A}', P')$ sind äquivalent:

i) X, X' identisch verteilt

ii) $P_X = P_{X'}$

Beweis. „(ii) \Rightarrow (i)“: klar. „(i) \Rightarrow (ii)“: Satz 9.

□

6 Dichte-Schätzung

Problem: **Gegeben:** Daten $x_1, \dots, x_n \in \mathbb{R}$ als Resultate „unabhängiger“ Wiederholungen eines Zufallsexperimentes.

Annahme: Einzelexperiment genügt absolutstetiger Verteilung.

Gesucht: Dichte der Verteilung für Einzelexperiment.

Hier die einfachste Methode: Approximation der Dichte durch Treppenfunktion.

Formale Beschreibung:

- $(X_i)_{i \in \mathbb{N}}$ iid-ZVen auf $(\Omega, \mathfrak{A}, P)$, X_1 absolutstetig verteilt mit unbekannter Dichte $f := f_{X_1}$.

Jede dieser ZVen modelliert ein Einzelexperiment.

- Annahme: Daten x_1, \dots, x_n sind eine **Realisierung** der ZVen X_1, \dots, X_n , d.h. für ein $\omega \in \Omega$

$$x_1 = X_1(\omega), \dots, x_n = X_n(\omega)$$

Fixiere Folge $(I_k)_{k \in \mathbb{Z}}$ von beschränkten Intervallen mit

- $\bigcup_{k \in \mathbb{Z}} I_k = \mathbb{R}$,
- $I_k \cap I_\ell = \emptyset$ für $k \neq \ell$.

Definiere stückweise konstante Dichte $\tilde{f} : \mathbb{R} \rightarrow [0, \infty[$ durch

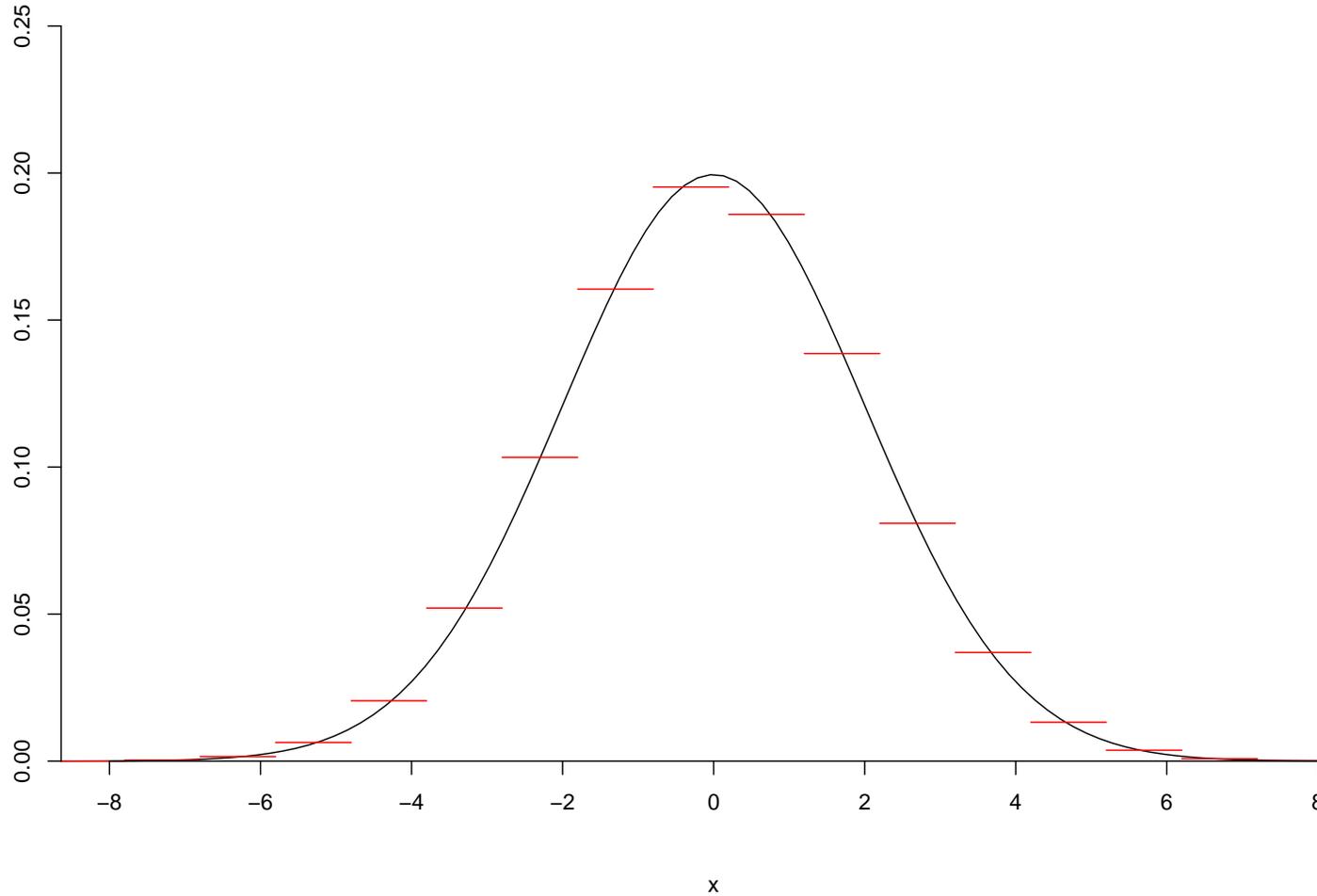
$$\tilde{f}(x) := \sum_{k \in \mathbb{Z}} c_k \cdot 1_{I_k}(x)$$

mit

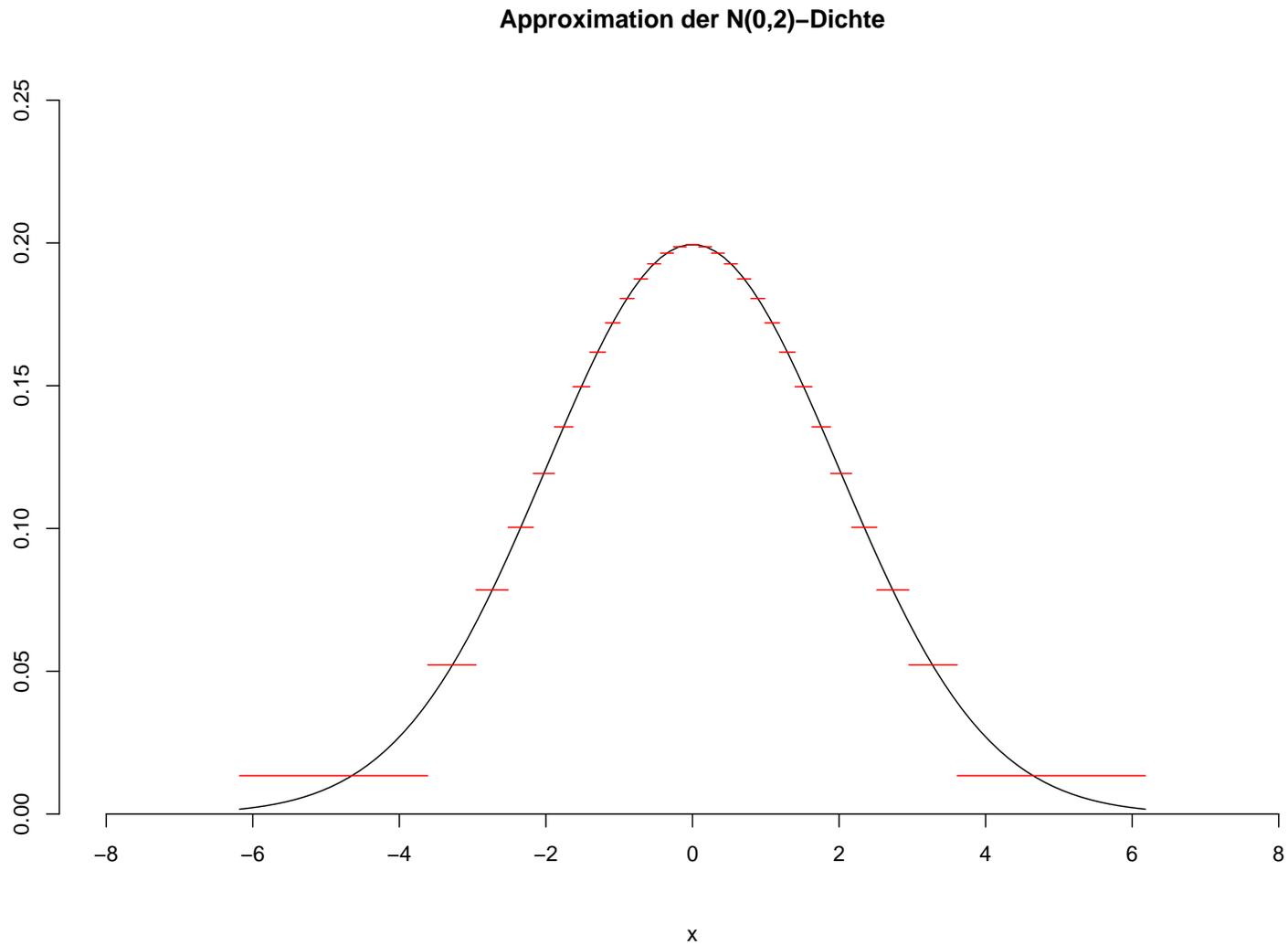
$$c_k := \frac{\int_{I_k} f(u) \, du}{\lambda_1(I_k)} = \frac{P(\{X_1 \in I_k\})}{\lambda_1(I_k)}.$$

56. Beispiel $N(0, 2)$, Intervalle der Länge eins.

Approximation der $N(0,2)$ -Dichte



57. Beispiel $N(0, 2)$, Intervalle nicht-konstanter Länge.



58. Definition Histogramm $f_n(\cdot; x_1, \dots, x_n) : \mathbb{R} \rightarrow [0, \infty[$

zu $x_1, \dots, x_n \in \mathbb{R}$ definiert durch

$$f_n(x; x_1, \dots, x_n) := \sum_{k \in \mathbb{Z}} c_{n,k}(x_1, \dots, x_n) \cdot 1_{I_k}(x)$$

mit

$$c_{n,k}(x_1, \dots, x_n) := \frac{|\{i \in \{1, \dots, n\} : x_i \in I_k\}|}{n \cdot \lambda_1(I_k)}.$$

59. Beispiel PROJEKTOR

60. Bemerkung Histogramm $f_n(\cdot; x_1, \dots, x_n)$ ist stückweise

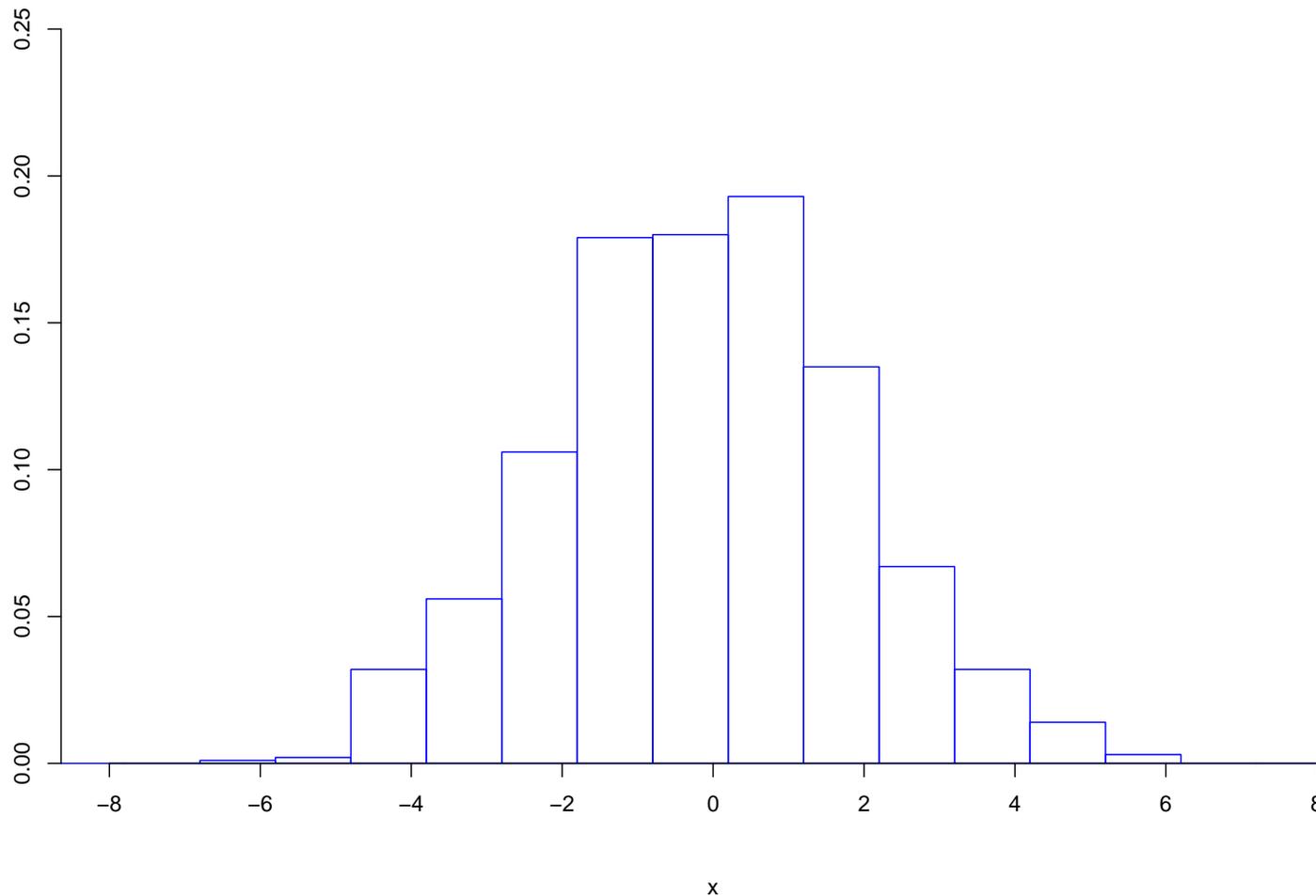
konstante Dichte, die nur auf endlich-vielen Intervallen

ungleich null ist.

61. Beispiel

Histogramm, $n = 1000$, Intervalle der Länge eins.

Histogramm, auf Basis von 1000 Realisierungen von iid $N(0,2)$ -ZV



62. Satz Für alle $x \in \mathbb{R}$ gilt für fast alle ω

$$\lim_{n \rightarrow \infty} f_n(x; X_1(\omega), \dots, X_n(\omega)) = \tilde{f}(x).$$

Verschärfung: fast sicher gleichmäßige Konvergenz.

Beweis. Sei $x \in I_k$. Für

$$Z_i := 1_{I_k}(X_i)$$

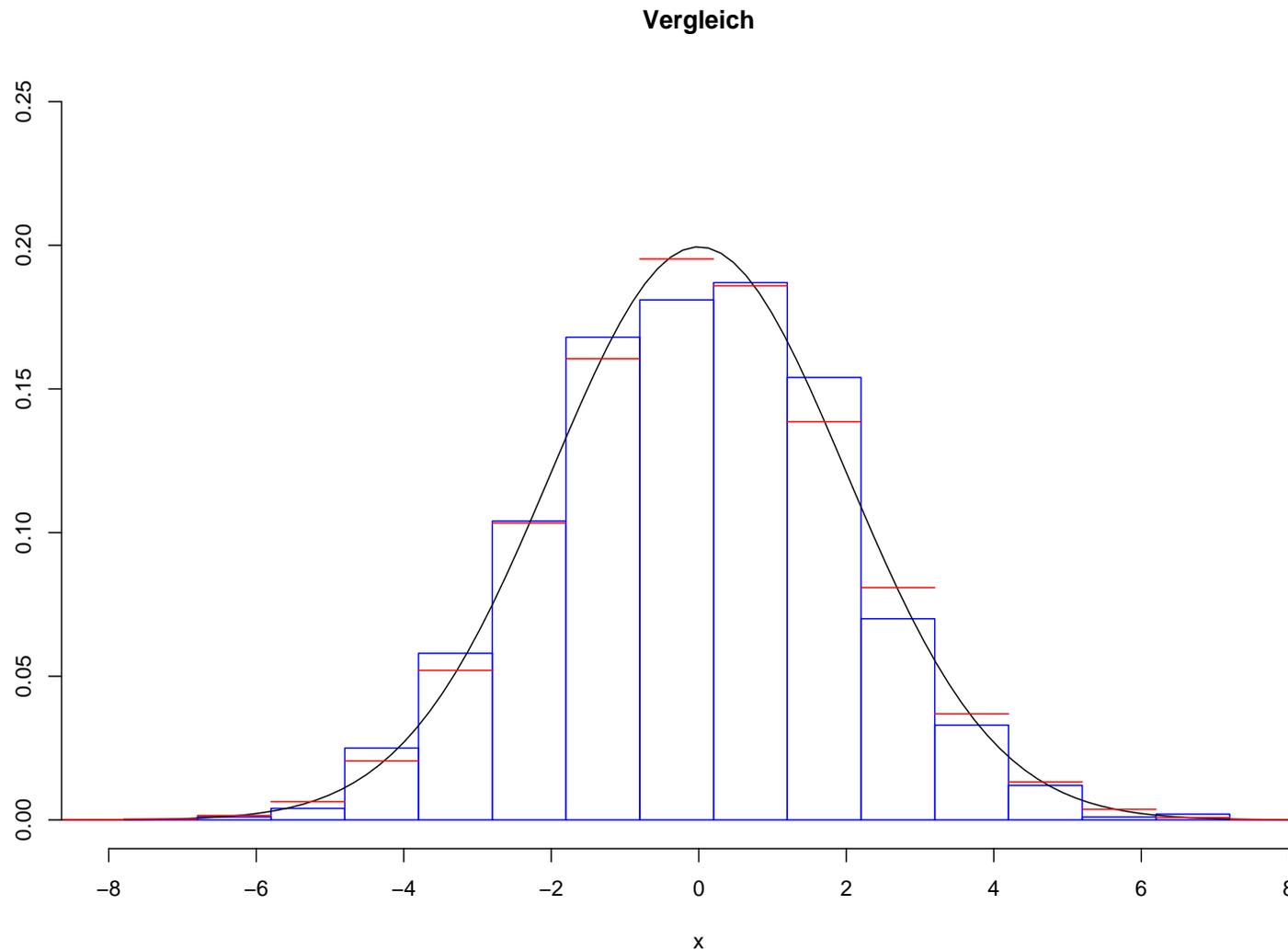
gilt $(Z_i)_{i \in \mathbb{N}}$ iid mit $Z_1 \sim \mathbf{B}(1, p)$ für

$$p = P(\{X_1 \in I_k\}) = \int_{I_k} f(u) du.$$

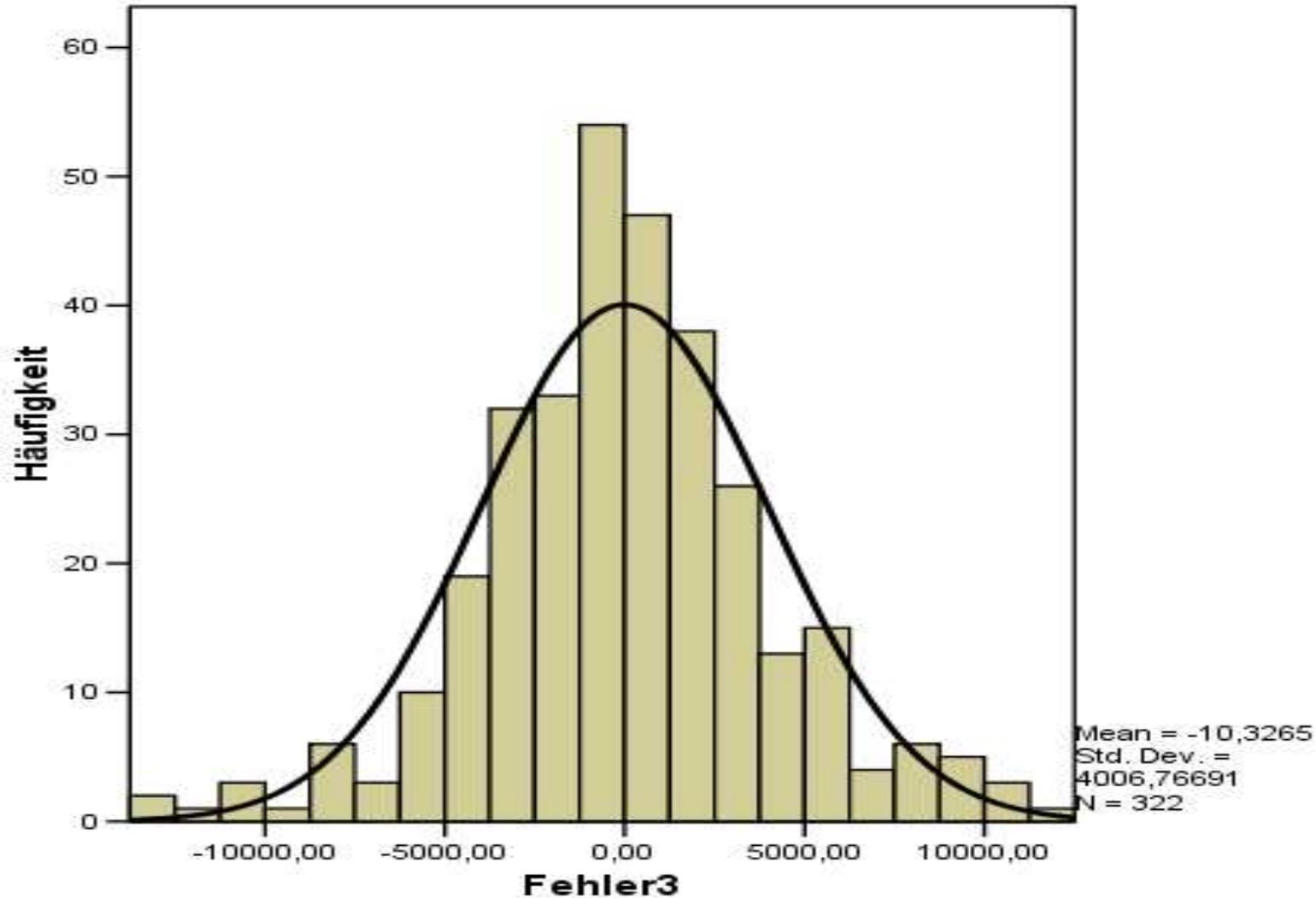
Wende Satz IV.7 an.

□

63. Beispiel Vergleich der Schätzung aus Bsp. 61 mit der Dichte und ihrer Approximation aus Bsp. 56.



64. Beispiel Prognosefehler für Tagesumsätze in Geldautomat, siehe Kapitel I.



65. Bemerkung Fehler der Approximation von f durch \tilde{f} ?

Setze $\delta := \sup_{k \in \mathbb{Z}} \lambda_1(I_k)$. Dann

$$\begin{aligned} \sup_{x \in \mathbb{R}} |f(x) - \tilde{f}(x)| \\ \leq \sup\{|f(s) - f(t)| : s, t \in \mathbb{R}, |s - t| \leq \delta\}. \end{aligned}$$

Anwendbar, falls f gleichmäßig stetig, also speziell falls f stetig und $\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0$.

Kap. VI Grenzwertsätze der Stochastik

1. Erwartungswert und Varianz
2. Gesetze der großen Zahlen
3. Zentraler Grenzwertsatz

1 Erwartungswert und Varianz

Erwartungswert: „**mittlerer Wert**“ einer Zufallsvariablen, „**Schwerpunkt**“ ihrer Verteilung. Allgemeine Definition basiert auf abstraktem Lebesgue-Integral.

1. Bemerkung Die Menge der ZVen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ bildet einen Vektorraum. Auf dem Untervektorraum $\mathfrak{L}_1 = \mathfrak{L}_1(\Omega, \mathfrak{A}, P)$ der **integrierbaren ZVen** erklärt man das Integral

$$E(X) = \int_{\Omega} X(\omega) dP(\omega),$$

genannt **Erwartungswert** von X .

Die Abbildung $E : \mathcal{L}_1 \rightarrow \mathbb{R}$ ist linear und monoton, d.h. für $X, Y \in \mathcal{L}_1$ und $\alpha, \beta \in \mathbb{R}$ gilt:

- $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$
- $X \leq Y \Rightarrow E(X) \leq E(Y)$

Ferner gilt für ZVen X und Y auf $(\Omega, \mathfrak{A}, P)$:

- $|X| \leq Y \wedge Y \in \mathcal{L}_1 \Rightarrow X \in \mathcal{L}_1$
- Falls $X \geq 0$:
$$X \in \mathcal{L}_1 \wedge E(X) = 0 \Leftrightarrow P(\{X > 0\}) = 0$$

Wir betrachten die Spezialfälle:

- (i) X diskret, also $P(\{X \in D\}) = 1$ für abzählbare Menge $D \subset \mathbb{R}$
- (ii) X absolutstetig mit Dichte f_X

Die folgenden Sätze dienen uns als Definitionen, siehe Irle (2001, Kap. 8).

2. Satz Im Fall (i) gilt $X \in \mathfrak{L}_1$ genau dann, wenn $\sum_{x \in D} |x| \cdot P(\{X = x\}) < \infty$. Gegebenenfalls folgt

$$E(X) = \sum_{x \in D} x \cdot P(\{X = x\}).$$

3. Satz Im Fall (ii) gilt $X \in \mathfrak{L}_1$ genau dann, wenn $\int_{-\infty}^{\infty} |x| \cdot f_X(x) dx < \infty$. Gegebenenfalls folgt

$$E(X) = \int_{-\infty}^{\infty} x \cdot f_X(x) dx.$$

4. Bemerkung Integrierbarkeit von X und gegebenenfalls $E(X)$ hängt nur von der Verteilung P_X ab.

5. Bemerkung Für $A \in \mathfrak{A}$ gilt $1_A \in \mathfrak{L}_1$ und $E(1_A) = P(A)$.

6. Beispiel

$$X \sim \mathbf{B}(n, p) \Rightarrow E(X) = n \cdot p$$

$$X \sim \mathbf{G}(p) \Rightarrow E(X) = 1/p$$

$$X \sim \mathbf{P}(\lambda) \Rightarrow E(X) = \lambda$$

$$X \sim \mathbf{U}([a, b]) \Rightarrow E(X) = (a + b)/2$$

$$X \sim \mathbf{Exp}(\lambda) \Rightarrow E(X) = 1/\lambda$$

$$X \sim \mathbf{N}(\mu, \sigma^2) \Rightarrow E(X) = \mu$$

Beweis. Für $X \sim \mathbf{B}(1, p)$

$$\mathbf{E}(X) = 1 \cdot p + 0 \cdot (1 - p) = p.$$

Für $X \sim \mathbf{B}(n, p)$ wegen Bem. 4 und Satz III.26 oBdA

$$X = \sum_{i=1}^n X_i$$

mit X_1, \dots, X_n iid und $X_1 \sim \mathbf{B}(1, p)$. Also

$$\mathbf{E}(X) = \sum_{i=1}^n \mathbf{E}(X_i) = n \cdot \mathbf{E}(X_1) = n \cdot p.$$

Beachte: Unabhängigkeit nicht verwendet.

Sei f_X die Dichte von $X \sim \mathbf{Exp}(\lambda)$. Dann

$$\begin{aligned} \int_{-\infty}^{\infty} |x| \cdot f_X(x) \, dx &= \int_0^{\infty} x \cdot \lambda \cdot \exp(-\lambda x) \, dx \\ &= -x \cdot \exp(-\lambda x) \Big|_0^{\infty} + \int_0^{\infty} \exp(-\lambda x) \, dx \\ &= \frac{-1}{\lambda} \exp(-\lambda x) \Big|_0^{\infty} = \frac{1}{\lambda}. \end{aligned}$$

Für die restlichen Verteilungen: ÜBUNG M:H29.

□

7. Beispiel Sei $(Y_i)_{i \in \mathbb{N}}$ iid mit $Y_1 \sim \mathbf{SB}$. Betrachte die **symmetrische Bernoulli-Irrfahrt** $S_t := \sum_{j=1}^t Y_j$ für $t \in \mathbb{N}_0$ (unendlicher Zeithorizont). Definiere

$$X = \begin{cases} \min\{t \in \mathbb{N} : S_t = 1\}, & \text{falls } \{\dots\} \neq \emptyset \\ 0, & \text{falls } \forall t \in \mathbb{N} : S_t \leq 0. \end{cases}$$

Es gilt $P(\{X > 0\}) = 1$ und für $t \in \mathbb{N}$

$$P(\{X = 2t - 1\}) = \frac{(2t - 2)!}{t! \cdot (t - 1)!} \cdot \frac{1}{2^{2t-1}}.$$

Da $\sum_{t=1}^{\infty} (2t - 1) \cdot P(\{X = 2t - 1\}) = \infty$, folgt $X \notin \mathcal{L}_1$.

Siehe Shreve (2004, Chap. 5.2), Stochastic Calculus for Finance I.

Nun: Hilfsmittel zur Berechnung von Erwartungswerten.

Bezeichnung $(\Omega_j)_{j \in J}$ **abzählbare Partition** von Ω , falls:

- (i) J abzählbar
- (ii) $\Omega_j \in \mathfrak{A}$ für alle $j \in J$
- (iii) $\Omega_j \cap \Omega_\ell = \emptyset$ für $j \neq \ell$
- (iv) $\Omega = \bigcup_{j \in J} \Omega_j$

8. Lemma Sei X ZV auf $(\Omega, \mathfrak{A}, P)$ und sei $(\Omega_j)_{j \in J}$ abzählbare Partition von Ω . Gelte

$$\forall j \in J \exists c_j \in \mathbb{R} \forall \omega \in \Omega_j : X(\omega) = c_j.$$

Dann gilt $X \in \mathfrak{L}_1$ genau dann, wenn

$\sum_{j \in J} |c_j| \cdot P(\Omega_j) < \infty$. Gegebenenfalls folgt

$$\mathbb{E}(X) = \sum_{j \in J} c_j \cdot P(\Omega_j).$$

9. Bemerkung Satz 2 beruht auf der abzählbaren Partition $(\{X = x\})_{x \in D}$.

Beweis von Lemma 8. Für

$$D := \{c_j : j \in J\}$$

gilt

$$P(\{X \in D\}) = 1.$$

Setze $J_x = \{j \in J : c_j = x\}$ für $x \in D$. Dann

$$\begin{aligned} \sum_{x \in D} |x| \cdot P(\{X = x\}) &= \sum_{x \in D} |x| \cdot \sum_{j \in J_x} P(\Omega_j) \\ &= \sum_{x \in D} \sum_{j \in J_x} |c_j| \cdot P(\Omega_j) \\ &= \sum_{j \in J} |c_j| \cdot P(\Omega_j). \end{aligned}$$

Berechnung des Erwartungswertes analog ohne Beträge. □

10. Korollar Falls $(\Omega, \mathfrak{A}, P)$ diskret, gilt $X \in \mathfrak{L}_1$ genau dann, wenn $\sum_{\omega \in \Omega} |X(\omega)| \cdot P(\{\omega\}) < \infty$. Gegebenenfalls folgt

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\}).$$

Beweis. Wähle $J := \Omega$, $\Omega_j := \{j\}$ und $c_j := X(j)$ in Lemma 8. □

11. Definition $g : \mathbb{R}^d \rightarrow \mathbb{R}$ **Borel-meßbar**, falls

$$\forall A \in \mathfrak{B}_1 : g^{-1}(A) \in \mathfrak{B}_d.$$

Vgl. Lemma V.13.

12. Bemerkung

- (i) Menge der Borel-meßbaren Funktionen $\mathbb{R}^d \rightarrow \mathbb{R}$ bildet Vektorraum.
- (ii) Jede stetige Funktion $\mathbb{R}^d \rightarrow \mathbb{R}$ ist Borel-meßbar.
- (iii) $g = 1_A$ mit $A \in \mathfrak{B}_d$ ist Borel-meßbar.

Siehe Irle (2001, Kap. 7).

13. Lemma Transformationssatz

Sei X ein d -dimensionaler Zufallsvektor mit Dichte f_X und $h : \mathbb{R}^d \rightarrow \mathbb{R}$ Borel-meßbar. Genau dann gilt $h(X) \in \mathcal{L}_1$, wenn $\int_{\mathbb{R}^d} |h(x)| \cdot f_X(x) dx < \infty$. Gegebenenfalls folgt

$$\mathbb{E}(h(X)) = \int_{\mathbb{R}^d} h(x) \cdot f_X(x) dx < \infty.$$

Beweis. Siehe Irle (2001, Satz 8.21). □

Lemma 13 gilt ohne die Annahme, daß $h(X)$ absolutstetig verteilt ist.

14. Satz Sind $X, Y \in \mathfrak{L}_1$ unabhängig, so folgt $X \cdot Y \in \mathfrak{L}_1$
und

$$E(X \cdot Y) = E(X) \cdot E(Y).$$

Beweis. Hier Beweis unter der zusätzlichen Annahme, daß X, Y diskret.
Wähle $D \subset \mathbb{R}$ abzählbar mit $P(\{X \in D\}) = P(\{Y \in D\}) = 1$, setze

$$\begin{aligned}\Omega_{(x,y)} &= \{(X, Y) = (x, y)\}, & (x, y) \in D^2, \\ \Omega_* &= \{X \notin D\} \cup \{Y \notin D\}, \\ X' &= 1_D \cdot X \quad \text{und} \quad Y' = 1_D \cdot Y.\end{aligned}$$

Dann

$$P(\{X \cdot Y \neq X' \cdot Y'\}) \leq P(\{X \notin D\} \cup \{Y \notin D\}) = 0.$$

Also $P_{X \cdot Y} = P_{X' \cdot Y'}$ und somit $X \cdot Y \in \mathfrak{L}_1 \Leftrightarrow X' \cdot Y' \in \mathfrak{L}_1$.

Ferner

$$\begin{aligned} & \sum_{(x,y) \in D^2} |x \cdot y| \cdot P(\Omega_{(x,y)}) + 0 \cdot P(\Omega_*) \\ &= \sum_{x \in D} |x| \cdot P(\{X = x\}) \cdot \sum_{y \in D} |y| \cdot P(\{X = y\}) < \infty, \end{aligned}$$

und Lemma 8 zeigt $X' \cdot Y' \in \mathfrak{L}_1$.

Berechnung des Erwartungswertes analog ohne Beträge.

Im Spezialfall absolutstetiger ZVen X, Y verwendet man den Satz von Fubini und Satz V.49.(i). □

Nun: „**Streuungsmaß**“ für Zufallsvariablen.

15. Definition ZV X **quadratisch integrierbar**, falls $X^2 \in \mathfrak{L}_1$.

Bez.: $\mathfrak{L}_2 = \mathfrak{L}_2(\Omega, \mathfrak{A}, P)$ Menge der quadr. int'baren ZVen.

16. Satz \mathfrak{L}_2 ist Untervektorraum von \mathfrak{L}_1 .

Beweis. Verwende $|X| \leq 1 + X^2$ und $(X + Y)^2 \leq 2X^2 + 2Y^2$. \square

17. Bemerkung Quadratische Integrierbarkeit von X und gegebenenfalls $E(X^2)$ hängt nur von der Verteilung P_X ab.

18. Satz Im Fall (i) gilt $X \in \mathfrak{L}_2$ genau dann, wenn $\sum_{x \in D} x^2 \cdot P(\{X = x\}) < \infty$. Gegebenenfalls folgt

$$E(X^2) = \sum_{x \in D} x^2 \cdot P(\{X = x\}).$$

Beweis. Wende Lemma 8 mit $J := D$, $\Omega_j := \{X = j\}$ und $c_j := j^2$ an. □

19. Satz Im Fall (ii) gilt $X \in \mathfrak{L}_2$ genau dann, wenn $\int_{-\infty}^{\infty} x^2 \cdot f_X(x) dx < \infty$. Gegebenenfalls folgt

$$E(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f_X(x) dx.$$

Beweis. Wende Lemma 13 mit $h(x) := x^2$ an. □

20. Definition Für $X \in \mathcal{L}_2$ heißt

$$\text{Var}(X) := \mathbb{E}(X - \mathbb{E}(X))^2$$

die **Varianz** und $\sqrt{\text{Var}(X)}$ die **Standardabweichung** von X .

21. Bemerkung Für $X \in \mathcal{L}_2$ gilt:

$$\text{Var}(X) = 0 \Leftrightarrow X \text{ fast sicher konstant}$$

22. Satz Für $X \in \mathcal{L}_2$ und $\alpha, \beta \in \mathbb{R}$ gilt:

$$(i) \text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

$$(ii) \text{Var}(\alpha \cdot X + \beta) = \alpha^2 \cdot \text{Var}(X)$$

Beweis. Ad (i): Es gilt

$$(X - \mathbb{E}(X))^2 = X^2 - 2 \cdot X \cdot \mathbb{E}(X) + (\mathbb{E}(X))^2.$$

Es folgt

$$\begin{aligned} \mathbb{E}(X - \mathbb{E}(X))^2 &= \mathbb{E}(X^2) - 2 \cdot (\mathbb{E}(X))^2 + (\mathbb{E}(X))^2 \\ &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2. \end{aligned}$$

Ad (ii): Es gilt

$$\alpha \cdot X + \beta - \mathbb{E}(\alpha \cdot X + \beta) = \alpha \cdot (X - \mathbb{E}(X)).$$

Es folgt

$$\text{Var}(\alpha \cdot X + \beta) = \mathbb{E}(\alpha^2 \cdot (X - \mathbb{E}(X))^2) = \alpha^2 \cdot \text{Var}(X).$$

□

23. Bemerkung Für $X, Y \in \mathfrak{L}_2$ gilt $X \cdot Y \in \mathfrak{L}_1$.

Verwende $|X \cdot Y| \leq X^2 + Y^2$.

24. Definition $X, Y \in \mathfrak{L}_2$ **unkorreliert**, falls

$$E(X \cdot Y) = E(X) \cdot E(Y).$$

25. Bemerkung

- (i) $X, Y \in \mathfrak{L}_2$ unabhängig $\Rightarrow X, Y$ unkorreliert,
siehe Satz 14. Umkehrung falsch, siehe ÜBUNG M:H34.
- (ii) $X, Y \in \mathfrak{L}_2$ genau dann unkorreliert, wenn

$$E((X - E(X)) \cdot (Y - E(Y))) = 0.$$

26. Satz Formel von Bienaymé

Falls $X_1, \dots, X_n \in \mathfrak{L}_2$ paarweise unkorreliert:

$$\text{Var} \left(\sum_{i=1}^n X_i \right) = \sum_{i=1}^n \text{Var}(X_i).$$

27. Beispiel Für $X_1 \sim \mathbf{B}(1, 1/2)$ und $X_2 = -X_1$ gilt

$\text{Var}(X_1 + X_2) = 0$ und $\text{Var}(X_1) = \text{Var}(X_2) = 1/4$.

Beweis von Satz 26. Setze $Y_i := X_i - \mathbb{E}(X_i)$ („zentrieren“).

Für $i \neq j$ gilt $\mathbb{E}(Y_i \cdot Y_j) = 0$. Also

$$\begin{aligned}\text{Var} \left(\sum_{i=1}^n X_i \right) &= \text{Var} \left(\sum_{i=1}^n Y_i \right) \\ &= \mathbb{E} \left(\sum_{i=1}^n Y_i \right)^2 = \sum_{i,j=1}^n \mathbb{E}(Y_i \cdot Y_j) \\ &= \sum_{i=1}^n \mathbb{E}(Y_i^2) \\ &= \sum_{i=1}^n \text{Var}(X_i).\end{aligned}$$

□

28. Beispiel

$$X \sim \mathbf{B}(n, p) \Rightarrow \text{Var}(X) = n \cdot p \cdot (1 - p)$$

$$X \sim \mathbf{G}(p) \Rightarrow \text{Var}(X) = (1 - p)/p^2$$

$$X \sim \mathbf{P}(\lambda) \Rightarrow \text{Var}(X) = \lambda$$

$$X \sim \mathbf{U}([a, b]) \Rightarrow \text{Var}(X) = (b - a)^2/12$$

$$X \sim \mathbf{Exp}(\lambda) \Rightarrow \text{Var}(X) = 1/\lambda^2$$

$$X \sim \mathbf{N}(\mu, \sigma^2) \Rightarrow \text{Var}(X) = \sigma^2$$

Beweis. Für $X \sim \mathbf{B}(1, p)$

$$\text{Var}(X) = \mathbf{E}(X^2) - (\mathbf{E}(X))^2 = p - p^2 = p \cdot (1 - p).$$

Für $X \sim \mathbf{B}(n, p)$ wegen Bem. 4, 17 und Satz III.26 oBdA

$$X = \sum_{i=1}^n X_i$$

mit X_1, \dots, X_n iid und $X_1 \sim \mathbf{B}(1, p)$. Mit Satz 26 folgt

$$\text{Var}(X) = n \cdot p \cdot (1 - p).$$

Für $X \sim \mathbf{Exp}(\lambda)$

$$\begin{aligned} & \int_0^{\infty} x^2 \cdot \lambda \cdot \exp(-\lambda x) dx \\ &= -x^2 \cdot \exp(-\lambda x) \Big|_0^{\infty} + \int_0^{\infty} 2x \cdot \exp(-\lambda x) dx \\ &= \frac{2}{\lambda} \cdot \mathbf{E}(X) = \frac{2}{\lambda^2}. \end{aligned}$$

Demnach gilt $\mathbf{E}(X^2) = \frac{2}{\lambda^2}$ und

$$\mathbf{Var}(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

Für die restlichen Verteilungen: ÜBUNG M:H29. □

Nun Abschätzung für die Konzentration einer ZV um ihren Erwartungswert.

29. Satz Tschebyschev-Ungleichung

Für $X \in \mathfrak{L}_2$ und $\varepsilon > 0$

$$P(\{|X - E(X)| \geq \varepsilon\}) \leq \frac{1}{\varepsilon^2} \cdot \text{Var}(X)$$

Beweis. Für $A := \{|X - E(X)| \geq \varepsilon\} \in \mathfrak{A}$ gilt

$$\varepsilon^2 \cdot 1_A \leq (X - E(X))^2 \cdot 1_A \leq (X - E(X))^2.$$

Es folgt

$$\varepsilon^2 \cdot P(A) = \varepsilon^2 \cdot E(1_A) \leq E(X - E(X))^2.$$

□

30. Satz Cauchy-Schwarzsche-Ungleichung

Für $X, Y \in \mathcal{L}_2$

$$|\mathbb{E}(X \cdot Y)| \leq \sqrt{\mathbb{E}(X^2) \cdot \mathbb{E}(Y^2)}.$$

Beweis. ÜBUNG M:H33.

□

2 Gesetze der großen Zahlen

In diesem und dem folgenden Abschnitt

- $(X_i)_{i \in \mathbb{N}}$ Folge von identisch verteilten ZVen auf $(\Omega, \mathfrak{A}, P)$ mit $X_1 \in \mathfrak{L}_1$,
- $S_n := \sum_{i=1}^n X_i$ für $n \in \mathbb{N}$.

Untersucht wird die Konvergenz geeignet normierter
Partialsommen S_n .

Hier zunächst Konvergenz des arithmetischen Mittels S_n/n
gegen $E(X_1)$.

31. Beispiel Gelte

$$P(\{X_1 = 1\}) = p, \quad P(\{X_1 = -1\}) = 1 - p$$

für $p \in]0, 1[$. Dann

$$E(X_1) = 2p - 1$$

und

$$1/2 \cdot (S_n + n) = \sum_{i=1}^n (X_i + 1)/2 \sim \mathbf{B}(n, p).$$

Speziell:

- $p = 1/2$: symmetrische Bernoulli-Irrfahrt mit unendlichem Zeithorizont
- $p = 19/37 = 0.5135\dots$: einfaches Spiel beim Roulette aus Sicht des Kasinos

Somit gilt für jedes $\varepsilon > 0$

$$\begin{aligned}\{|S_n/n - \mathbb{E}(X_1)| \geq \varepsilon\} &= \{|S_n - n \cdot (2p - 1)| \geq n \cdot \varepsilon\} \\ &= \{|(S_n + n)/2 - n \cdot p| \geq n/2 \cdot \varepsilon\},\end{aligned}$$

d.h.

$$P(\{|S_n/n - \mathbb{E}(X_1)| \geq \varepsilon\}) = \sum_{k \in K} \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

mit $K := \{k \in \{0, \dots, n\} : |k - n \cdot p| \geq n/2 \cdot \varepsilon\}$.

Frage: Wie verhalten sich diese Wahrscheinlichkeiten für große Werte n ?

Antwort unter sehr allgemeinen Voraussetzungen im folgenden Satz.

32. Satz Schwaches Gesetz der großen Zahlen

Falls $(X_i)_{i \in \mathbb{N}}$ paarweise unkorreliert, so folgt für jedes $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(\{|S_n/n - E(X_1)| \geq \varepsilon\}) = 0.$$

Beweis. Es gilt $X_i \in \mathfrak{L}_2$ sowie $E(S_n/n) = E(X_1)$ und gem. Satz 26

$$\text{Var}(S_n) = n \cdot \text{Var}(X_1).$$

Also sichert Satz 29

$$P(\{|S_n/n - E(X_1)| \geq \varepsilon\}) \leq \frac{1}{\varepsilon^2} \cdot \text{Var}(S_n/n) = \frac{1}{\varepsilon^2} \cdot \frac{\text{Var}(X_1)}{n}.$$

□

Jakob Bernoulli (1713), Khintchine (1928)

33. Bemerkung Was besagt Satz 32 für die Folge der Verteilungen P_{S_n} ? Es gilt

$$\begin{aligned} & \{|S_n/n - E(X_1)| < \varepsilon\} \\ &= \{(E(X_1) - \varepsilon) \cdot n \leq S_n \leq (E(X_1) + \varepsilon) \cdot n\}. \end{aligned}$$

Also

$$\lim_{n \rightarrow \infty} P(\{(E(X_1) - \varepsilon) \cdot n \leq S_n \leq (E(X_1) + \varepsilon) \cdot n\}) = 1.$$

Beachte: Satz 32 macht keine Aussage über einzelne

Realisierungen $S_1(\omega), S_2(\omega), \dots$ der Folge $(S_n)_{n \in \mathbb{N}}$.

34. Beispiel Sei $X_1 \sim \mathbf{B}(1, p)$. Der Beweis von Satz 32 zeigt

$$P(\{|S_n/n - p| \geq \varepsilon\}) \leq \frac{1}{4 \cdot \varepsilon^2 \cdot n}.$$

Es gilt jedoch auch die für großes n wesentlich bessere

Hoeffdingsche Ungleichung

$$P(\{|S_n/n - p| \geq \varepsilon\}) \leq 2 \cdot \exp(-2 \cdot \varepsilon^2 \cdot n),$$

siehe TUTORIUM T6. Beispielsweise gilt für $\varepsilon = 10^{-2}$ und $n = 10^4$

$$\frac{1}{4 \cdot \varepsilon^2 \cdot n} = 0.25 \quad \text{und} \quad 2 \cdot \exp(-2 \cdot \varepsilon^2 \cdot n) = 0.270 \dots$$

und für $\varepsilon = 10^{-2}$ und $n = 10^5$

$$\frac{1}{4 \cdot \varepsilon^2 \cdot n} = 0.025 \quad \text{und} \quad 2 \cdot \exp(-2 \cdot \varepsilon^2 \cdot n) = 4.12 \dots \cdot 10^{-9}.$$

Nun zur Konvergenz der Realisierungen $S_1(\omega)/1, S_2(\omega)/2, \dots$

35. Satz Starkes Gesetz der großen Zahlen

Falls $(X_i)_{i \in \mathbb{N}}$ unabhängig, so folgt für fast alle $\omega \in \Omega$

$$\lim_{n \rightarrow \infty} S_n(\omega)/n = E(X_1)$$

Beweis. Siehe Irle (2001, Kap. 11) und Vorlesung „Probability Theory“.

Spezialfall in Kapitel IV.2 behandelt. □

Borel (1909), Kolmogorov (1930), Etemadi (1981)

Stochastische Simulation zur Berechnung von $E(X')$:

- Konstruiere iid-ZVen X_1, \dots, X_n auf $(\Omega, \mathfrak{A}, P)$, wobei X' und X_1 identisch verteilt
- Erzeuge eine Realisierung x_1, \dots, x_n der ZVen X_1, \dots, X_n , d.h. für ein $\omega \in \Omega$

$$x_1 = X_1(\omega), \dots, x_n = X_n(\omega)$$

- Approximiere $E(X') = E(X_1)$ durch die entsprechende Realisierung von S_n/n (**Mittelwert**)

$$1/n \cdot \sum_{i=1}^n x_i$$

36. Beispiel Betrachte die ZV

$$X' := \begin{cases} \max\{k \in \mathbb{N}_0 : \sum_{i=1}^k Y_i' \leq 1\}, & \text{falls } \{\dots\} \text{ beschränkt} \\ 0, & \text{sonst,} \end{cases}$$

wobei $(Y_i')_{i \in \mathbb{N}}$ iid mit $Y_1' \sim \mathbf{Exp}(\lambda)$ für $\lambda := 5$.

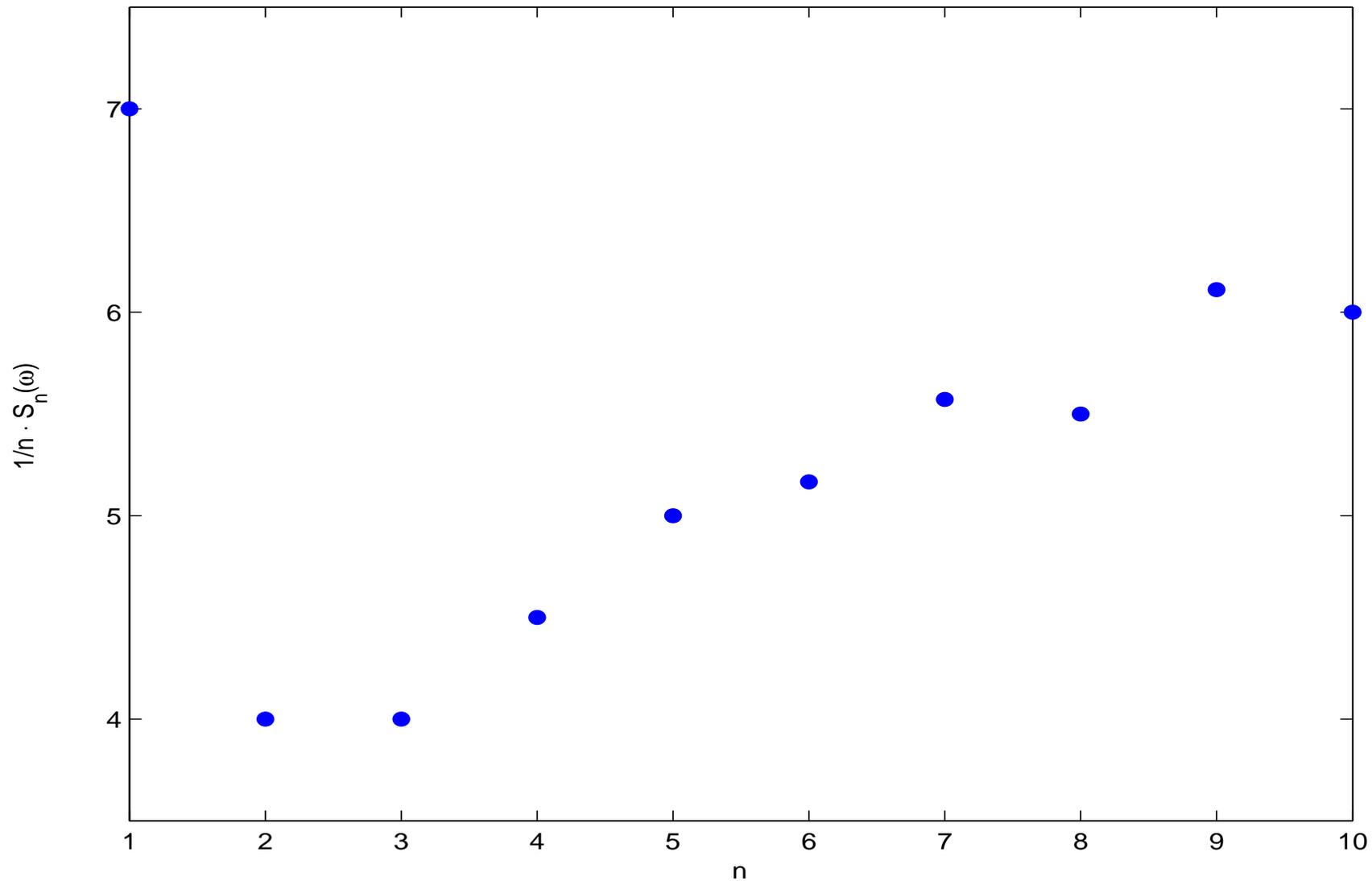
Anwendung: [Warteschlangentheorie](#), Stichwort [Poisson-Prozeß](#).

Berechne $E(X')$. Vgl. ÜBUNG M:H20.

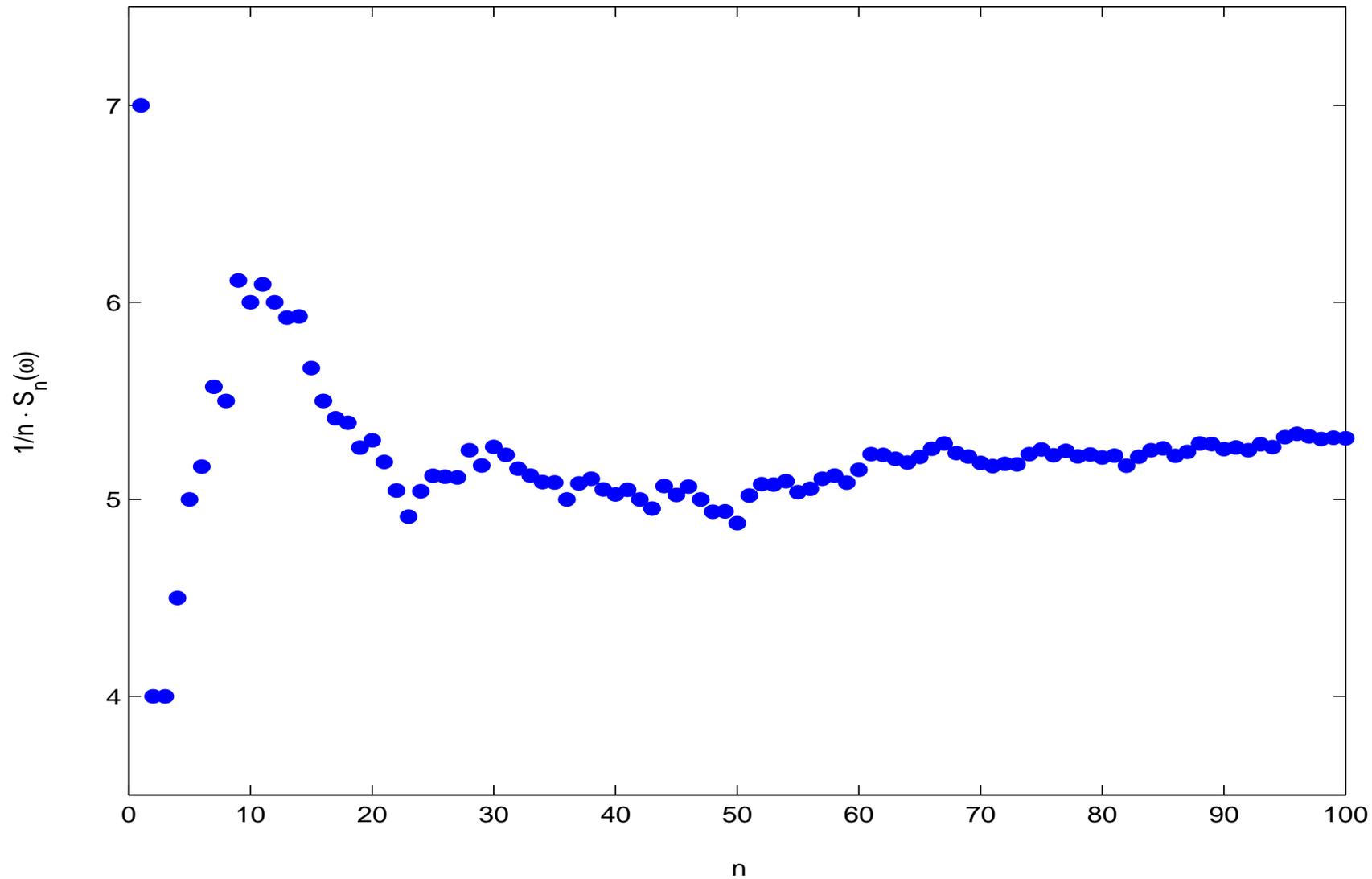
Die ersten 10 Simulationen lieferten die Werte

7, 1, 4, 6, 7, 6, 8, 5, 11, 5.

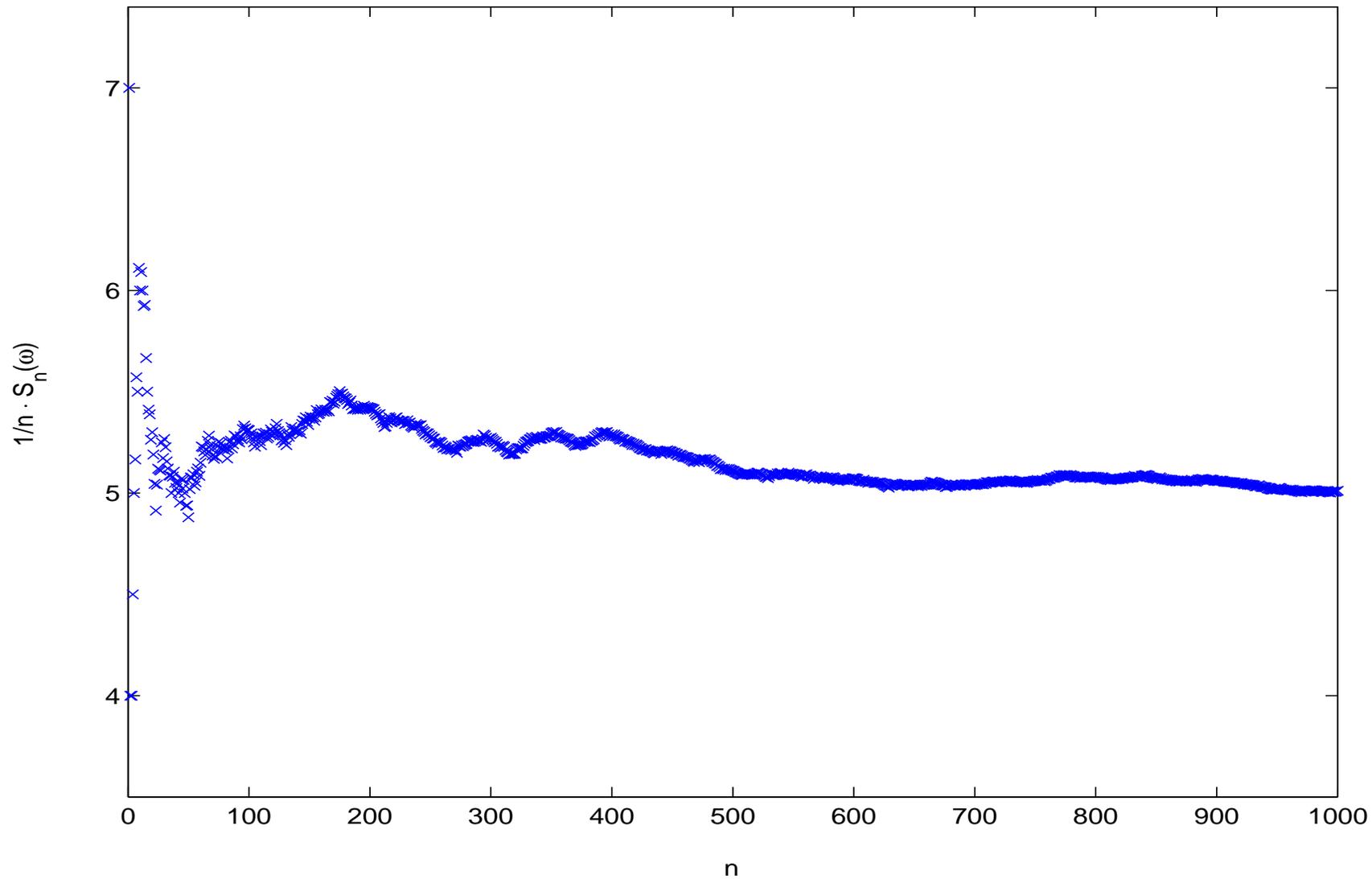
Realisierung von S_n/n für $n = 1, \dots, 10$



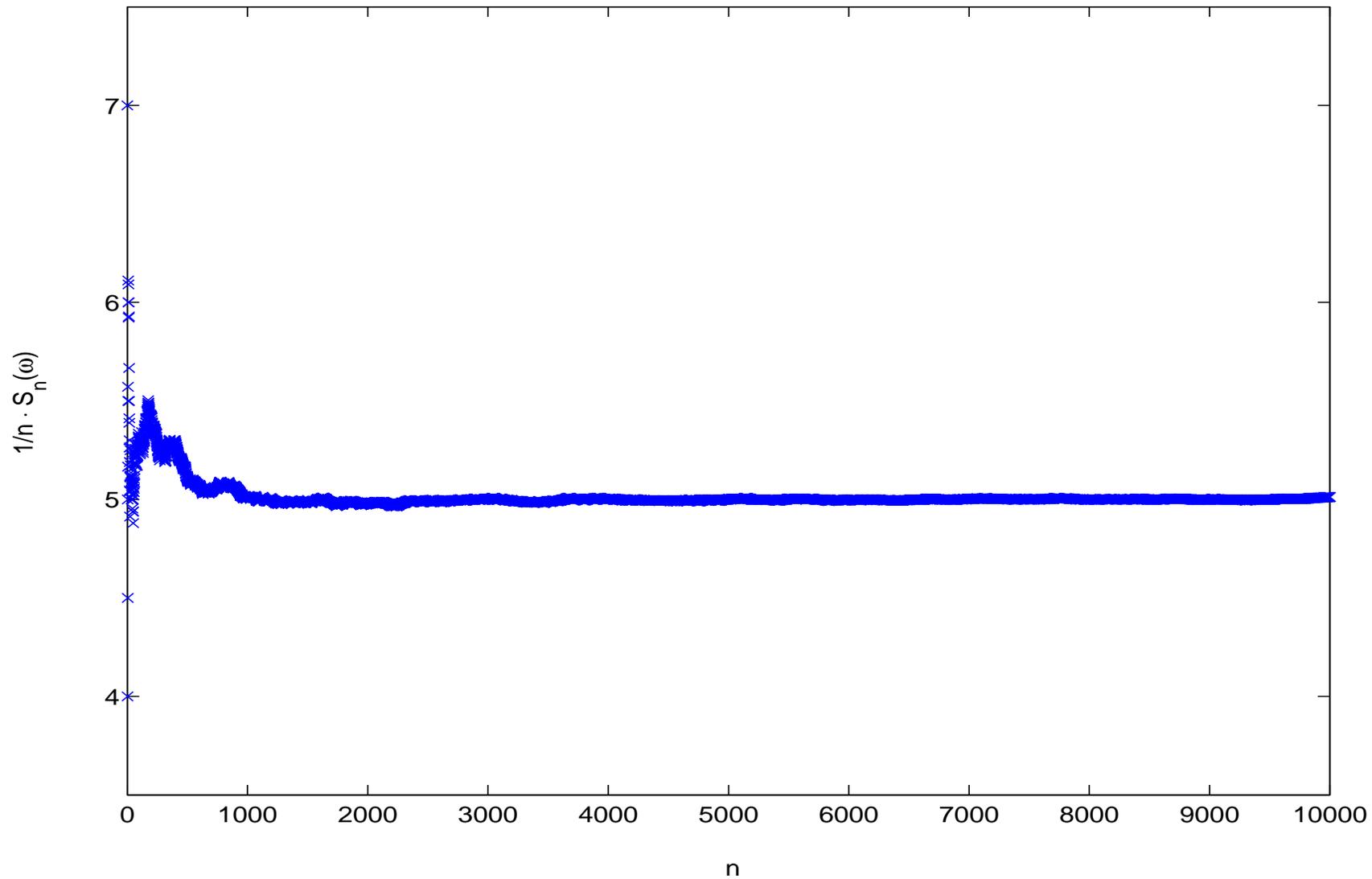
Realisierung von S_n/n für $n = 1, \dots, 100$



Realisierung von S_n/n für $n = 1, \dots, 1000$



Realisierung von S_n/n für $n = 1, \dots, 10000$



37. Bemerkung Was besagt Satz 35 für Folgen von Realisierungen der ZVen S_n ?

Für alle $\varepsilon > 0$ und fast alle $\omega \in \Omega$ gilt

$\exists n_0(\varepsilon, \omega) \forall n \geq n_0(\varepsilon, \omega) :$

$$(\mathbf{E}(X_1) - \varepsilon) \cdot n \leq S_n(\omega) \leq (\mathbf{E}(X_1) + \varepsilon) \cdot n.$$

Vergleich der Konvergenzbegriffe im Starken und Schwachen Gesetz der großen Zahlen.

38. Satz Gilt fast sicher

$$\lim_{n \rightarrow \infty} Y_n = 0$$

für eine Folge von ZVen $(Y_n)_{n \in \mathbb{N}}$ auf $(\Omega, \mathfrak{A}, P)$, so folgt für alle $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P(\{|Y_n| \geq \varepsilon\}) = 0.$$

Beweis. ÜBUNG M:G15. Siehe auch ÜBUNG M:G18. □

Anwendung mit $Y_n := S_n/n - E(X_1)$.

3 Zentraler Grenzwertsatz

Starkes Gesetz der großen Zahlen: fast sichere Konvergenz der **Realisierungen** einer Folge von ZVen.

Nun: Konvergenz der **Verteilungen** einer Folge von ZVen.

Bezeichnung $\mathfrak{C}_X = \{x \in \mathbb{R} : F_X \text{ stetig in } x\}$ **Menge der Stetigkeitspunkte** der Verteilungsfunktion von X .

39. Beispiel

$$X \sim \mathbf{U}([a, b]) \Rightarrow \mathfrak{C}_X = \mathbb{R}$$

$$X \sim \mathbf{N}(\mu, \sigma^2) \Rightarrow \mathfrak{C}_X = \mathbb{R}$$

$$X \sim \mathbf{B}(n, p) \wedge p \in]0, 1[\Rightarrow \mathfrak{C}_X = \mathbb{R} \setminus \{0, \dots, n\}$$

40. Definition Folge $(X_n)_{n \in \mathbb{N}}$ von ZVen **konvergiert in Verteilung** gegen eine ZV X , falls

$$\forall x \in \mathfrak{C}_X : \lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

Bez.: $X_n \xrightarrow{d} X$

41. Beispiel Für $X_n = a_n$ und $X = a$ mit $a, a_n \in \mathbb{R}$ gilt

$$X_n \xrightarrow{d} X \Leftrightarrow \lim_{n \rightarrow \infty} a_n = a.$$

Beachte: Aus $a_n > a$ folgt $F_{X_n}(a) = 0$, während $F_X(a) = 1$.

42. Beispiel

(i) Satz III.37: Gelte $X_n \sim \mathbf{B}(n, p_n)$ mit

$\lim_{n \rightarrow \infty} n \cdot p_n = \lambda > 0$ sowie $X \sim \mathbf{P}(\lambda)$. Dann

$$X_n \xrightarrow{d} X.$$

(ii) ÜBUNG M:H9, WInf:H9: Gelte $X_n \sim \mathbf{H}(n, n_0(n), k)$ mit

$n_0(n) \in \{1, \dots, n\}$ und

$\lim_{n \rightarrow \infty} n_0(n)/n = p \in]0, 1[$ sowie $X \sim \mathbf{B}(k, p)$.

Dann $X_n \xrightarrow{d} X$.

(iii) ÜBUNG M:G16: Sei Y_n Anzahl der Führungszeitpunkte bei symmetrischer Bernoulli-Irrfahrt der Länge n ,
 $X_n = Y_n/n$ und X **Arcussinus-verteilt**, d.h.

$$F_X(x) = 2/\pi \cdot \arcsin(\sqrt{x})$$

für $x \in [0, 1]$. Dann $X_n \xrightarrow{d} X$.

(iv) TUTORIUM T2:3: Satz von de Moivre-Laplace. Gelte

$Y_n \sim \mathbf{B}(n, p)$ mit $p \in]0, 1[$ und

$$X_n := \frac{Y_n - n \cdot p}{\sqrt{n \cdot p(1 - p)}}$$

sowie $X \sim \mathbf{N}(0, 1)$. Dann $X_n \xrightarrow{d} X$.

43. Bemerkung Es existiert eine Metrik ρ auf

$$\mathbb{M} := \{Q : Q \text{ Wahrscheinlichkeitsmaß auf } \mathfrak{B}_1\},$$

so daß

$$X_n \xrightarrow{d} X \Leftrightarrow \lim_{n \rightarrow \infty} \rho(P_{X_n}, P_X) = 0,$$

siehe Vorlesung „Probability Theory“.

44. Lemma Falls F_X stetig und $X_n \xrightarrow{d} X$:

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_{X_n}(x) - F_X(x)| = 0.$$

Beweis. ÜBUNG M:H40.

□

45. Lemma Falls F_X stetig und $X_n \xrightarrow{d} X$:

$$\forall A \in \mathfrak{M} : \lim_{n \rightarrow \infty} P(\{X_n \in A\}) = P(\{X \in A\}).$$

Beweis. Gilt nach Definition für $A =]-\infty, x]$. Für $A = \{x\}$ und $\varepsilon > 0$

$$\begin{aligned} P(\{X_n \in A\}) &\leq P(\{x - \varepsilon < X_n \leq x\}) \\ &= P(\{X_n \leq x\}) - P(\{X_n \leq x - \varepsilon\}). \end{aligned}$$

Somit

$$0 \leq \limsup_{n \rightarrow \infty} P(\{X_n \in A\}) \leq F_X(x) - F_X(x - \varepsilon).$$

Aufgrund der Stetigkeit von F_X in x folgt

$$\lim_{n \rightarrow \infty} P(\{X_n \in A\}) = 0 = P(\{X \in A\}).$$

Im folgenden

- $(X_i)_{i \in \mathbb{N}}$ iid und $X_1 \in \mathcal{L}_2$ mit $\sigma^2 := \text{Var}(X_1) \in]0, \infty[$
- $\mu := \text{E}(X_1)$
- $S_n^* := \frac{S_n - n \cdot \mu}{\sigma \cdot \sqrt{n}}$ standardisierte Summenvariablen
- Z standard-normalverteilte Zufallsvariable

Es gilt $\text{E}(S_n^*) = 0$ und $\text{Var}(S_n^*) = 1$.

46. Satz Zentraler Grenzwertsatz

Für jede Folge $(X_i)_{i \in \mathbb{N}}$ wie oben gilt $S_n^* \xrightarrow{d} Z$.

Beweis. Irle (2001, Kap. 12) und Vorlesung „Probability Theory“. □

47. Bemerkung Es gilt

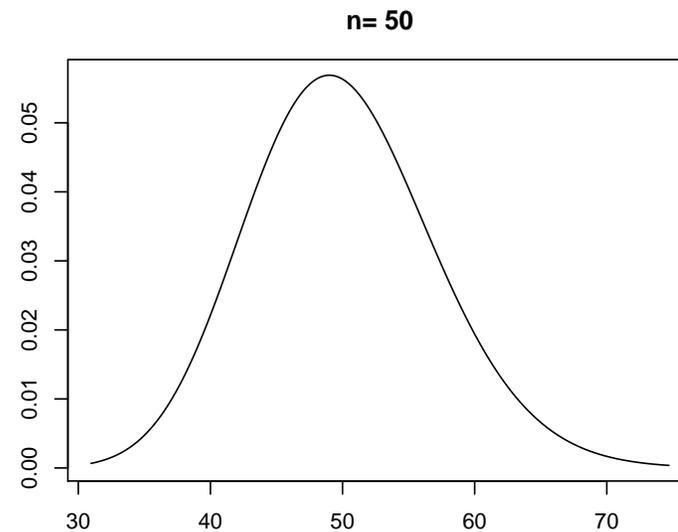
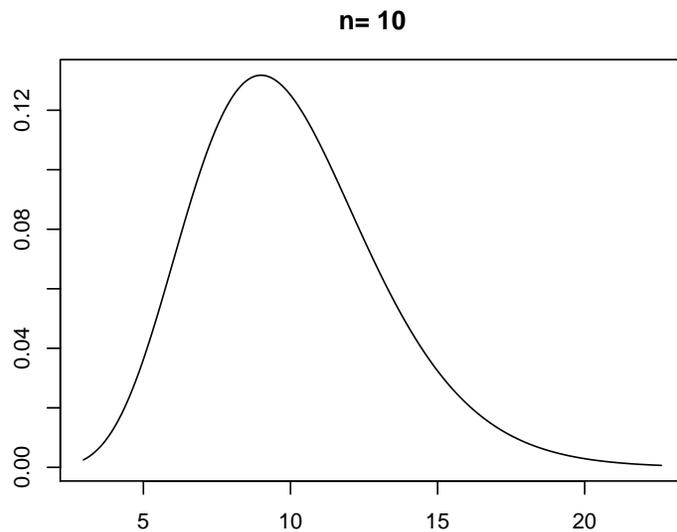
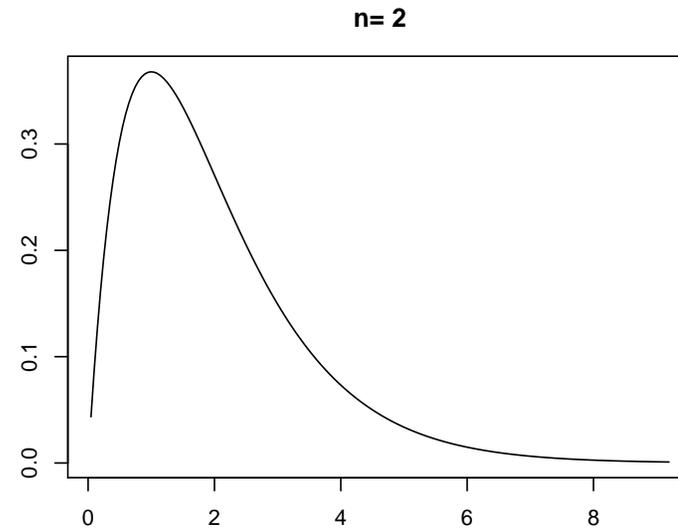
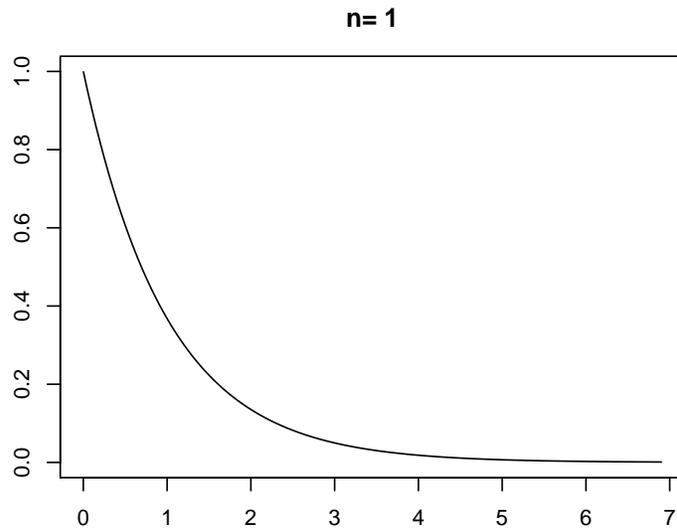
$$S_n^* = \sum_{i=1}^n \underbrace{\frac{X_i - E(X_i)}{\sigma \cdot \sqrt{n}}}_{=: X_i^*}$$

mit $E(X_i^*) = 0$ und $\text{Var}(X_i^*) = 1/n$.

Der zentrale Grenzwertsatz besagt also grob: „Ein Gesamteffekt, der Summe vieler kleiner zentrierter unabhängiger Einzeleffekte ist, ist näherungsweise normalverteilt.“

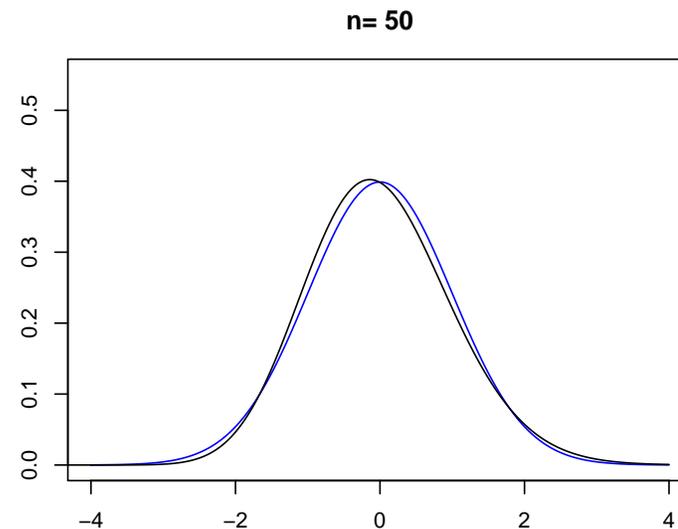
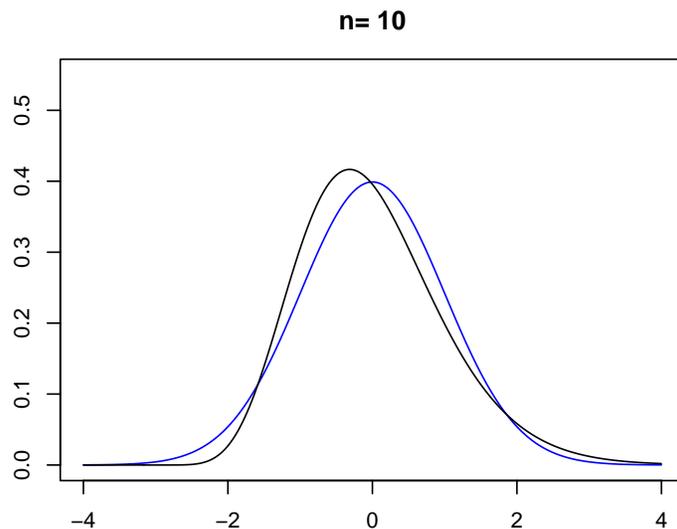
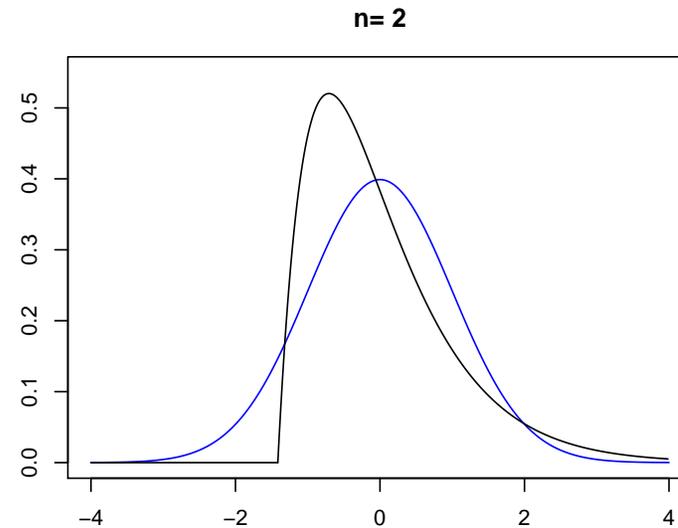
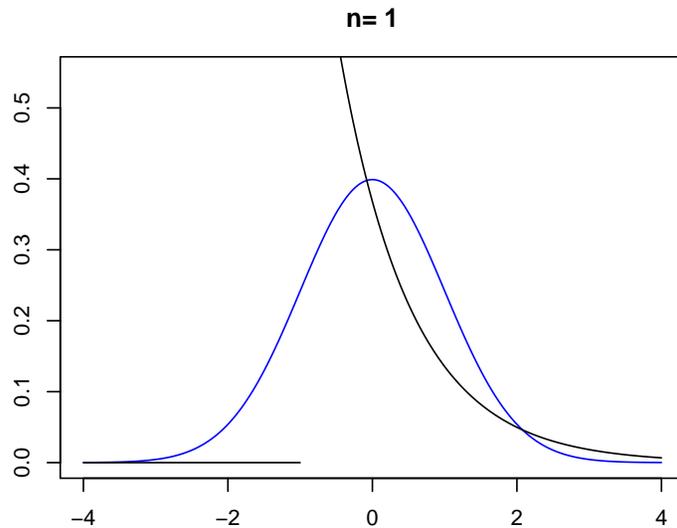
48. Beispiel Dichten von Summen n unabhängiger $\text{Exp}(1)$ -verteilter

ZVen



49. Beispiel Dichten von standardisierten Summen n unabhängiger

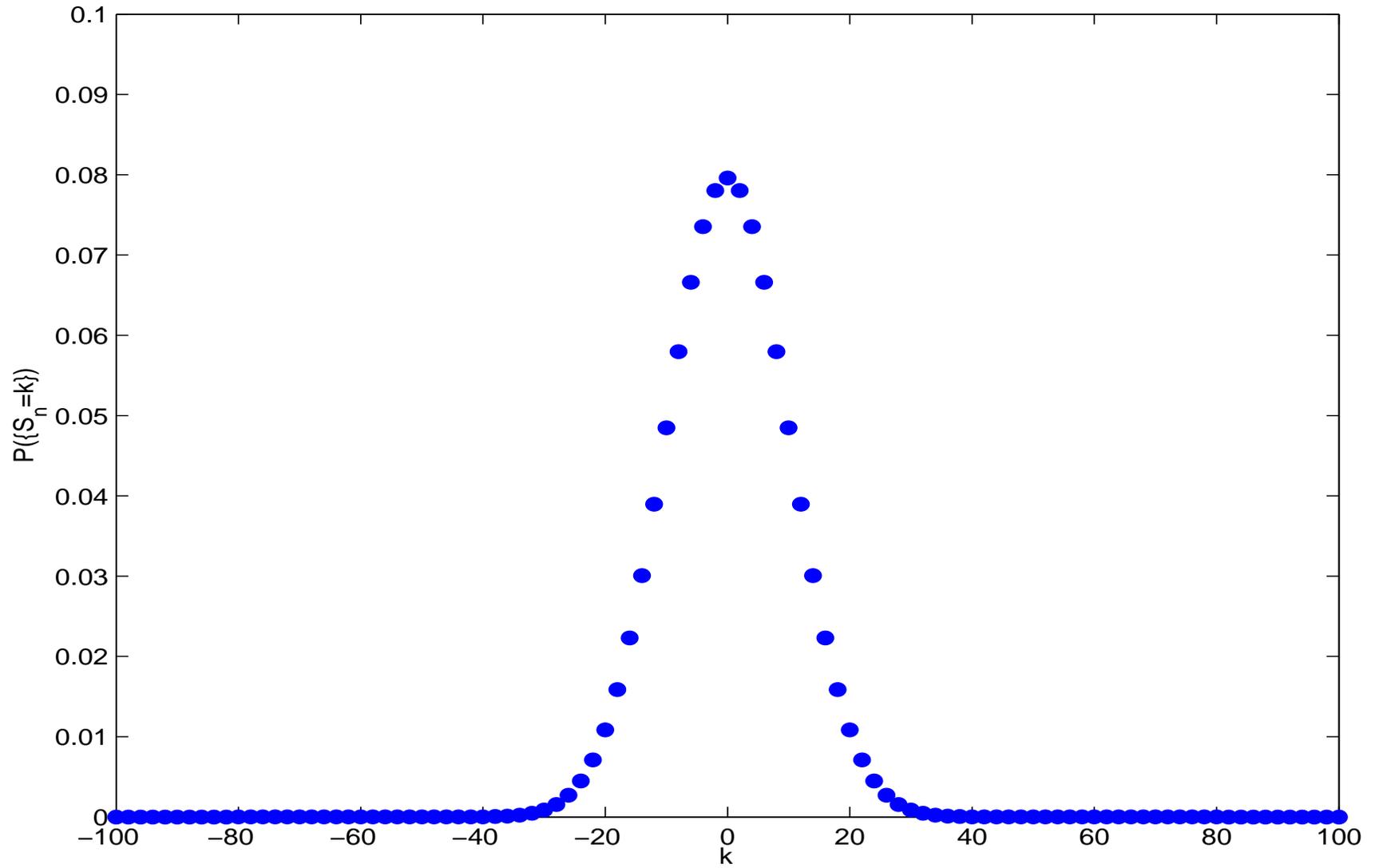
$\text{Exp}(1)$ -verteilter ZVen



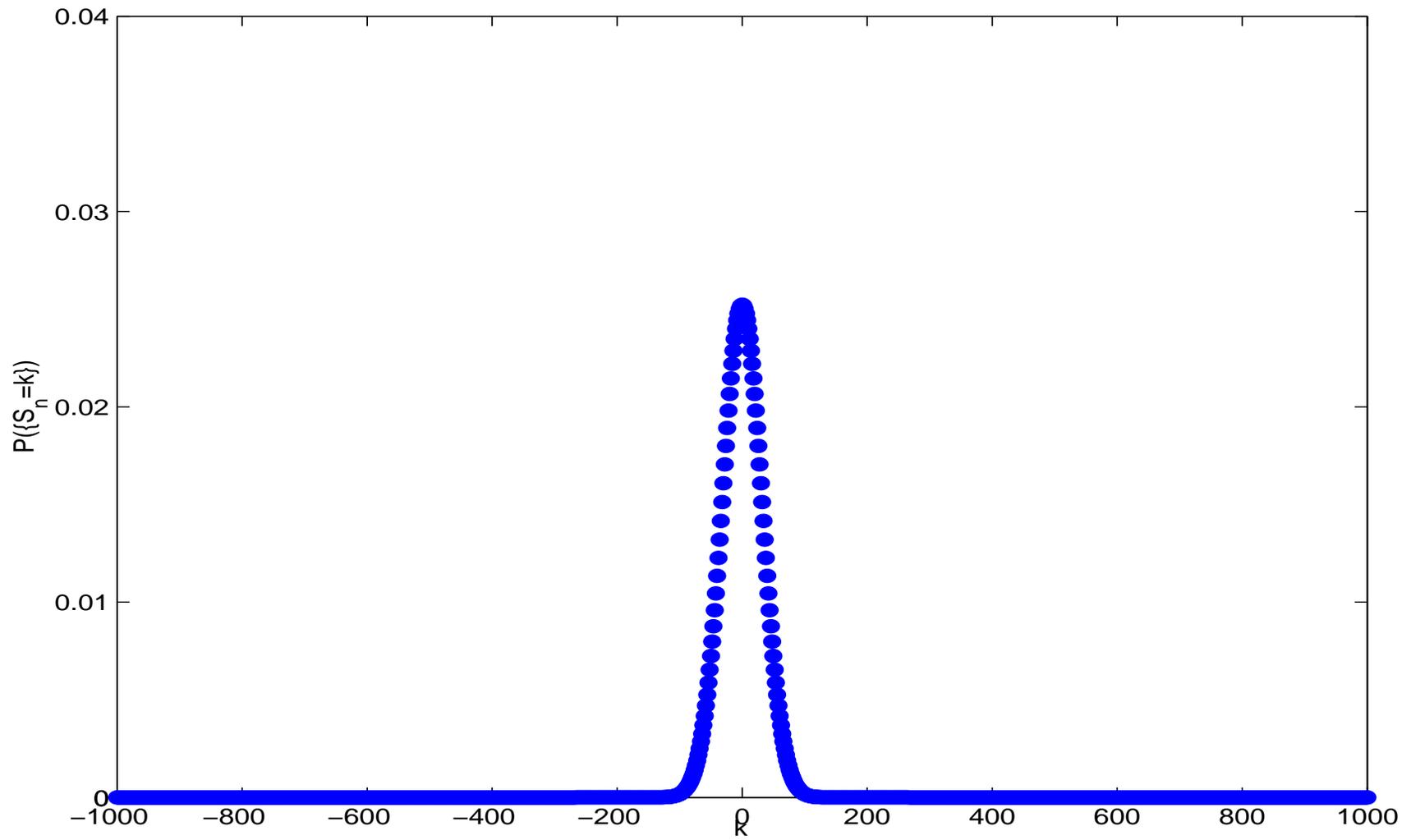
50. Beispiel Sei $\alpha \in]0, 1[$. Betrachte symmetrische Bernoulli-Irrfahrt $(S_n)_{n \in \mathbb{N}}$. Gesucht: Folge $(c_n)_{n \in \mathbb{N}}$ in $]0, \infty[$, so daß

$$\lim_{n \rightarrow \infty} P(\{|S_n| \leq c_n\}) = \alpha.$$

Wahrscheinlichkeitsfunktion von S_n für $n = 100$



Wahrscheinlichkeitsfunktion von S_n für $n = 1000$



Standardisieren mit $\mu = 0$ und $\sigma = 1$:

$$\{|S_n| \leq c_n\} = \{-c_n \leq S_n \leq c_n\} = \{-c_n/\sqrt{n} \leq S_n^* \leq c_n/\sqrt{n}\}.$$

Also für $c_n := c \cdot \sqrt{n}$

$$\begin{aligned} \lim_{n \rightarrow \infty} P(\{|S_n| \leq c_n\}) &= \lim_{n \rightarrow \infty} P(\{-c \leq S_n^* \leq c\}) \\ &= \Phi(c) - \Phi(-c) = 2\Phi(c) - 1. \end{aligned}$$

Somit leistet das $(1 + \alpha)/2$ -Quantil

$$c := \Phi^{-1}((1 + \alpha)/2)$$

der Standard-Normalverteilung das Verlangte.

So gilt näherungsweise $c = 1$ für $\alpha = 0.6826$.

51. Beispiel Überbuchung von Flugverbindungen:

- Kapazität $K \in \mathbb{N}$
- Buchungsanzahl $n \in \mathbb{N}$, wobei die Passagiere unabhängig jeweils mit Wahrscheinlichkeit $p \in]0, 1[$ erscheinen

Gegeben $\alpha \in]0, 1[$. Bestimme $n \in \mathbb{N}$, so daß Überbuchung ungefähr mit Wahrscheinlichkeit α .

Modell: X_1, \dots, X_n iid, $X_1 \sim \mathbf{B}(1, p)$.

Es gilt $\left\{ \sum_{i=1}^n X_i > K \right\} = \{S_n^* > c_n\}$ mit

$$S_n^* := \sum_{i=1}^n (X_i - p) / \sqrt{n \cdot p \cdot (1 - p)},$$

$$c_n := (K - n \cdot p) / \sqrt{n \cdot p \cdot (1 - p)}.$$

Somit ist c_n näherungsweise durch $c_n = \Phi^{-1}(1 - \alpha)$ gegeben.

Für $K := 1000$, $p := 0.9$ und $\alpha := 0.01$ ergibt sich näherungsweise

$c_n = 2.33$ und

$$n = 1086.$$

Hiermit gilt für die erwartete Anzahl nicht beförderter Passagiere

$$E\left(\max\left(\sum_{i=1}^n X_i - K, 0\right)\right) \leq 86 \cdot P\left(\left\{\sum_{i=1}^n X_i > K\right\}\right) \doteq 0.86.$$

Zum Vergleich die exakten Werte

$$E\left(\max\left(\sum_{i=1}^n X_i - K, 0\right)\right) = 0.0287 \dots,$$

$$P\left(\left\{\sum_{i=1}^n X_i > K\right\}\right) = 0.00811 \dots,$$

$$86 \cdot P\left(\left\{\sum_{i=1}^n X_i > K\right\}\right) = 0.697 \dots$$

52. Satz Für $a, b, \mu, \mu_i \in \mathbb{R}$ mit $a \neq 0$ und $\sigma, \sigma_i \in]0, \infty[$ gilt:

(i) Falls $X \sim \mathbf{N}(\mu, \sigma^2)$, dann

$$a \cdot X + b \sim \mathbf{N}(a \cdot \mu + b, a^2 \cdot \sigma^2).$$

(ii) Falls X_1, \dots, X_n unabhängig und $X_i \sim \mathbf{N}(\mu_i, \sigma_i^2)$,

dann $\sum_{i=1}^n X_i \sim \mathbf{N}(\mu, \sigma^2)$ mit $\mu := \sum_{i=1}^n \mu_i$ und

$$\sigma^2 := \sum_{i=1}^n \sigma_i^2.$$

Beweis. ad (i): ÜBUNG WInf:H20.

ad (ii): ÜBUNG M:H39. □

53. Bemerkung Satz 52 zeigt, daß $S_n^* \sim \mathbf{N}(0, 1)$, falls

$$X_1 \sim \mathbf{N}(\mu, \sigma^2).$$

Kap. VII Schätz- und Testprobleme

1. Statistische Modellbildung und statistisches Entscheiden
2. Schätztheorie
3. Testtheorie
4. Lineare Regression

1 Statistische Modellbildung und statistisches Entscheiden

Fortan bezeichnen wir mit $\mathbf{B}(n, p)$, $\mathbf{N}(\mu, \sigma^2)$, \dots auch die entsprechenden Wahrscheinlichkeitsmaße auf \mathfrak{B}_1 .

Grundform wichtiger statistischer Fragestellungen:

- (i) Zufallsexperiment mit unbekannter Verteilung Q
- (ii) **Verteilungsannahme**: $Q \in \mathfrak{P}$ für eine Menge \mathfrak{P} von Wahrscheinlichkeitsmaßen auf \mathcal{B}_d .
- (iii) (a) **Schätzproblem**. Gegeben: Abbildung $\eta : \mathfrak{P} \rightarrow \mathbb{R}$.
Bestimme $\eta(Q)$.
- (b) **Testproblem**. Gegeben $\emptyset \neq \mathfrak{P}_0 \subsetneq \mathfrak{P}$.
Entscheide, ob $Q \in \mathfrak{P}_0$.
- (iv) Verfügbar: **Stichprobe** $x_1, \dots, x_n \in \mathbb{R}^d$ aus n -maliger unabhängiger Wiederholung des Zufallsexperimentes.

1. Beispiel Geschlecht eines Neugeborenen ($1 \triangleq W, 0 \triangleq M$),

siehe Bsp. II.13. Hier $d = 1$ und

(i) $Q = \mathbf{B}(1, p)$, wobei p die Wahrscheinlichkeit, daß
Neugeborenes weiblich.

(ii) $\mathfrak{P} := \{\mathbf{B}(1, p) : p \in]0, 1[\}$

(iii) (a) $\eta(\mathbf{B}(1, p)) := p$

(b) $\mathfrak{P}_0 := \{\mathbf{B}(1, 1/2)\}$ oder

$\mathfrak{P}_0 := \{\mathbf{B}(1, p) : p < 1/2\}$

(iv) Geschlecht bei n Lebendgeburten

Bei Stichprobenumfang $n = 25\,171\,123$ scheint „verlässliche“ Bestimmung von $\eta(Q)$ und Entscheidung, ob $\eta(Q) \in \mathfrak{P}_0$ möglich.

Empirisches Mittel

$$\bar{x}_n := \frac{1}{n} \cdot \sum_{i=1}^n x_i = \frac{12\,241\,392}{25\,171\,123} = 0.4863 \dots$$

legt nahe, daß $\eta(Q)$ ungefähr 0.48 beträgt und $Q \neq \mathbf{B}(1, 1/2)$ gilt.

2. Bemerkung

- Studiert werden auch Varianten dieser Grundform (z.B. abhängige Beobachtungen, \mathbb{R}^k -wertige Abbildungen η).
- Abstrakte Formulierung und Theorie in der Mathematischen Statistik (Gütekriterien, Optimalitätsaussagen, Quantifizierung von Risiken).
- Oft ist \mathfrak{P} in natürlicher Weise parametrisiert, siehe Beispiel 1.

Fortan, der Einfachheit halber, $d = 1$.

Statistisches Experiment formal:

- (i) Familie $(\Omega, \mathfrak{A}, P^\vartheta)_{\vartheta \in \Theta}$ von Wahrscheinlichkeitsräumen
- (ii) Zufallsvariablen $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$, so daß für alle $\vartheta \in \Theta$:
- X_1, \dots, X_n unabhängig bzgl. P^ϑ
 - $P_{X_1}^\vartheta = \dots = P_{X_n}^\vartheta$.

Ferner $P_{X_1}^\vartheta \neq P_{X_1}^{\vartheta'}$ für $\vartheta, \vartheta' \in \Theta$ mit $\vartheta \neq \vartheta'$.

Terminologie: \mathbb{R}^n **Stichprobenraum**, Θ **Parameterraum**.

Verteilungsannahme: $\mathfrak{P} = \{P_{X_1}^\vartheta : \vartheta \in \Theta\}$.

Konkrete Gestalt von Ω , \mathfrak{A} und den W'maßen P^ϑ irrelevant.

Annahme: Daten sind Realisierung

$$x_1 = X_1(\omega), \dots, x_n = X_n(\omega)$$

der ZVen X_1, \dots, X_n für ein $\omega \in \Omega$.

Schätzproblem definiert durch

$$\gamma : \Theta \rightarrow \mathbb{R},$$

also $\eta(P_{X_1}^{\vartheta}) = \gamma(\vartheta)$. Eine Borel-meßbare Abbildung $g_n : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt in diesem Kontext **Schätzfunktion**.

Schätze $\gamma(\vartheta)$ durch

$$g_n(x_1, \dots, x_n) = g_n(X_1(\omega), \dots, X_n(\omega))$$

Ziel: Für **jedes** $\vartheta \in \Theta$ liegen die Werte der **Zufallsvariable**

$$g_n(X_1, \dots, X_n)$$

auf $(\Omega, \mathfrak{A}, P^{\vartheta})$ „nahe“ bei $\gamma(\vartheta)$.

Testproblem definiert durch

$$\emptyset \neq \Theta_0 \subsetneq \Theta,$$

also $\mathfrak{P}_0 = \{P_{X_1}^\vartheta : \vartheta \in \Theta_0\}$. Eine Borel-Menge $R_n \in \mathfrak{B}_n$ heißt in diesem Kontext **Verwerfungsbereich**. Lehne die **Hypothese** Θ_0 (bzw. „ $\vartheta \in \Theta_0$ “) genau dann ab, wenn

$$(x_1, \dots, x_n) = (X_1(\omega), \dots, X_n(\omega)) \in R_n.$$

Ziel: Für jedes $\vartheta \in \Theta_0$ ist die Wahrscheinlichkeit

$$P^\vartheta(\{(X_1, \dots, X_n) \in R_n\})$$

des **Fehlers 1. Art** „klein“, und für jedes $\vartheta \in \Theta \setminus \Theta_0$ ist die Wahrscheinlichkeit

$$P^\vartheta(\{(X_1, \dots, X_n) \notin R_n\})$$

des **Fehlers 2. Art** „klein“. Siehe jedoch Beispiel 43 ff.

Bei Schätz- und Testproblem jeweils worst case-Analyse über alle $\vartheta \in \Theta$.

3. Beispiel Geschlecht eines Neugeborenen. Hier

$\Theta :=]0, 1[$, und für $p \in \Theta$ ist $P_{X_1}^p := \mathbf{B}(1, p)$.

- Schätzproblem: $\gamma(p) := p$,
- Testproblem: $\Theta_0 := \{1/2\}$ oder $\Theta_0 :=]0, 1/2[$.

Als Schätzfunktion für γ bereits betrachtet

$$g_n(x_1, \dots, x_n) := \bar{x}_n = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Naheliegend: Verwerfungsbereiche R_n für $\Theta_0 := \{1/2\}$ von der Form

$$R_n := \{x \in \mathbb{R}^n : |g_n(x) - 1/2| \geq k_n\}.$$

4. Beispiel Analog: **Halbwertszeit**. Hier $\Theta :=]0, \infty[$, und für $\lambda \in \Theta$ ist $P_{X_1}^\lambda := \mathbf{Exp}(\lambda)$.

- Schätzproblem: $\gamma(\lambda) := \ln(2)/\lambda$.
- Testproblem: $\Theta_0 := \{\ln(2)/\lambda_0\}$.

Ausblick: nicht-parametrische Statistik.

2 Schätztheorie

Notation und Terminologie:

- $\mathcal{L}_1^\vartheta = \mathcal{L}_1(\Omega, \mathfrak{A}, P^\vartheta)$ und $\mathcal{L}_2^\vartheta = \mathcal{L}_2(\Omega, \mathfrak{A}, P^\vartheta)$
- $X = (X_1, \dots, X_n)$ und $x = (x_1, \dots, x_n)$
- $g_n(X)$ **Schätzvariable** zu Schätzfunktion g_n
- $E^\vartheta, \text{Var}^\vartheta$ Erwartungswert bzw. Varianz bzgl. P^ϑ
- Eigenschaft gilt P^ϑ -fast sicher, falls sie für alle ω aus einer Menge $A \in \mathfrak{A}$ mit $P^\vartheta(A) = 1$ gilt
- $\bar{x}_n := \frac{1}{n} \cdot \sum_{i=1}^n x_i$ und $\bar{X}_n := \frac{1}{n} \cdot \sum_{i=1}^n X_i = S_n/n$

5. Definition Schätzfunktion $g_n : \mathbb{R}^n \rightarrow \mathbb{R}$ **erwartungstreu** für γ , falls

$$\forall \vartheta \in \Theta : E^{\vartheta}(g_n(X)) = \gamma(\vartheta)$$

Folge von Schätzfunktionen $g_n : \mathbb{R}^n \rightarrow \mathbb{R}$

(i) **schwach konsistent** für γ , falls

$$\forall \vartheta \in \Theta \forall \varepsilon > 0 : \lim_{n \rightarrow \infty} P^{\vartheta}(\{|g_n(X) - \gamma(\vartheta)| > \varepsilon\}) = 0$$

(ii) **stark konsistent** für γ , falls

$$\forall \vartheta \in \Theta : \lim_{n \rightarrow \infty} g_n(X) = \gamma(\vartheta) \quad P^{\vartheta}\text{-f.s.}$$

6. Bemerkung Starke Konsistenz impliziert schwache Konsistenz, siehe Satz VI.38.

Problem: Schätzung des Erwartungswertes, siehe Beispiele 3 und 4.

7. Satz Gelte $X_1 \in \mathcal{L}_1^\vartheta$ für alle $\vartheta \in \Theta$, und sei

$$\gamma(\vartheta) := E^\vartheta(X_1).$$

Dann definiert das empirische Mittel $g_n(x) := \bar{x}_n$ eine stark konsistente Folge von erwartungstreuen Schätzern für γ .

Beweis. Für jedes $\vartheta \in \Theta$ gilt

$$E^{\vartheta}(g_n(X)) = E^{\vartheta}(\bar{X}_n) = \frac{1}{n} \cdot \sum_{i=1}^n E^{\vartheta}(X_i) = \gamma(\vartheta).$$

Die starke Konsistenz ist genau das starke Gesetz der großen Zahlen. \square

8. Beispiel In Beispiel 1 bzw. 3 gilt für jedes $p \in]0, 1[$ aufgrund der Tschebyschev-Ungleichung

$$P^p(\{|\bar{X}_n - p| \geq 10^{-2}\}) \leq 9.9 \cdot 10^{-5}$$

$$P^p(\{|\bar{X}_n - p| \geq 10^{-3}\}) \leq 9.9 \cdot 10^{-3}$$

und aufgrund der Hoeffdingsche Ungleichung, siehe Bsp.

VI.34

$$P^p(\{|\bar{X}_n - p| \geq 10^{-2}\}) \leq 9.2 \cdot 10^{-2187}$$

$$P^p(\{|\bar{X}_n - p| \geq 10^{-3}\}) \leq 2.7 \cdot 10^{-22}.$$

Problem: Schätzung der Varianz.

9. Satz Gelte $X_1 \in \mathcal{L}_2^\vartheta$ für alle $\vartheta \in \Theta$, und sei

$$\gamma(\vartheta) = \text{Var}^\vartheta(X_1).$$

Ferner sei $n > 1$. Dann definiert die **empirische Varianz**

$$v_n(x) := \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

eine stark konsistente Folge von erwartungstreuen Schätzern für γ .

Beweis. Setze $\mu(\vartheta) := \mathbb{E}^\vartheta(X_1)$. Es gilt

$$\begin{aligned}v_n(x) &= \frac{1}{n-1} \cdot \sum_{i=1}^n ((x_i - \mu(\vartheta)) + (\mu(\vartheta) - \bar{x}_n))^2 \\ &= \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \mu(\vartheta))^2 - \frac{n}{n-1} \cdot (\mu(\vartheta) - \bar{x}_n)^2\end{aligned}$$

Mit Satz VI.26 folgt

$$\begin{aligned}\mathbb{E}^\vartheta(v_n(X)) &= \frac{n}{n-1} \cdot \gamma(\vartheta) - \frac{n}{n-1} \cdot \text{Var}^\vartheta \left(\frac{1}{n} \cdot \sum_{i=1}^n X_i \right) \\ &= \frac{n}{n-1} \cdot \gamma(\vartheta) - \frac{1}{n-1} \cdot \gamma(\vartheta) = \gamma(\vartheta).\end{aligned}$$

Nach dem starken Gesetz der großen Zahlen gilt P^ϑ -f.s.

$$\lim_{n \rightarrow \infty} \frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - \mu(\vartheta))^2 = E^\vartheta (X_1 - \mu(\vartheta))^2 = \text{Var}^\vartheta (X_1)$$

und

$$\lim_{n \rightarrow \infty} \frac{1}{n} \cdot \sum_{i=1}^n X_i = \mu(\vartheta).$$

Es folgt P^ϑ -f.s.

$$\lim_{n \rightarrow \infty} v_n(X) = \gamma(\vartheta).$$

□

10. Bemerkung Für die empirische Varianz gilt

$$\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x}_n)^2 = \frac{1}{n-1} \cdot \left(\sum_{i=1}^n x_i^2 - n \cdot \bar{x}_n^2 \right).$$

Beweis. Erster Teil des Beweises von Satz 9 mit $\mu(\vartheta) := 0$. □

11. Beispiel In Beispiel 1 bzw. 3 beträgt die empirische Varianz 0.2498

12. Beispiel Sei $\Theta := [0, 1]$ und für $p \in \Theta$ gelte

$P_{X_1}^p = \mathbf{B}(1, p)$. Ferner sei

$$\gamma(p) = \sqrt{p \cdot (1 - p)}$$

die entsprechende Standardabweichung. Für jede Schätzfunktion g_1 gilt

$$E^p(g_1(X)) = p \cdot g_1(1) + (1 - p) \cdot g_1(0).$$

Also existiert keine erwartungstreue Schätzfunktion für γ .

Analog für $n > 1$.

Stichwort: asymptotische Erwartungstreue.

Fragen

- (i) Wie definiert man die „Güte“ einer Schätzfunktion?
- (ii) Kennt man in Bsp. 1 bzw. 3 „optimale“ Schätzfunktionen?
- (iii) Wie „verlässlich“ ist ein Schätzwert?

Zu Frage (i):

13. Definition **Quadratmittel-Fehler** einer Schätzfunktion g_n für γ

$$R^\vartheta(g_n) := E^\vartheta(g_n(X) - \gamma(\vartheta))^2,$$

falls $g_n(X) \in \mathfrak{L}_2^\vartheta$.

14. Definition **Bias** einer Schätzfunktion g_n für γ

$$B^\vartheta(g_n) := E^\vartheta(g_n(X)) - \gamma(\vartheta),$$

falls $g_n(X) \in \mathcal{L}_1^\vartheta$.

15. Bemerkung

(i) Erwartungstreue äquivalent zu

$$\forall \vartheta \in \Theta : B^\vartheta(g_n) = 0.$$

(ii) Es gilt

$$R^\vartheta(g_n) = \text{Var}^\vartheta(g_n(X)) + (B^\vartheta(g_n))^2,$$

siehe ÜBUNG M:G19.

Zu Frage (ii): Optimalität exemplarisch für Problemstellung aus Bsp. 1 bzw. 3.

Bez. $g_n^*(x) := \bar{x}_n$.

16. Beispiel Für alle $p \in]0, 1[$ gilt

$$B^p(g_n^*) = 0$$

und

$$R^p(g_n^*) = E^p \left(\bar{X}_n - p \right)^2 = \frac{p \cdot (1 - p)}{n}.$$

Für $p \in]0, 1[$ und $x \in D := \{0, 1\}^n$ sei (Likelihood-Funktion, vgl. ÜBUNG M:G7)

$$L_x(p) := P^p(\{X = x\}) = p^{k(x)} \cdot (1 - p)^{n - k(x)}$$

mit $k(x) := |\{i \in \{1, \dots, n\} : x_i = 1\}|$ sowie

(Log-Likelihood-Funktion)

$$\ell_x(p) := \ln(L_x(p)).$$

17. Satz Ungleichung von Fréchet, Cramér, Rao

Jede erwartungstreue Schätzfunktion $g_n : \mathbb{R}^n \rightarrow \mathbb{R}$ für γ erfüllt

$$\forall p \in]0, 1[: R^p(g_n) \geq 1 / E^p \left((\ell'_X(p))^2 \right) .$$

Beweis. Für g_n wie oben gilt

$$p = \mathbb{E}^p(g_n(X)) = \sum_{x \in D} g_n(x) \cdot L_x(p)$$

und somit

$$1 = \sum_{x \in D} g_n(x) \cdot L'_x(p) = \sum_{x \in D} g_n(x) \cdot \ell'_x(p) \cdot L_x(p) = \mathbb{E}^p(g_n(X) \cdot \ell'_X(p)).$$

Aus $\sum_{x \in D} L_x(p) = 1$ folgt

$$0 = \sum_{x \in D} L'_x(p) = \sum_{x \in D} \ell'_x(p) \cdot L_x(p) = \mathbb{E}^p(\ell'_X(p)).$$

Also liefert die Cauchy-Schwarzsche Ungleichung

$$1 = \left(\mathbb{E}^p((g_n(X) - \gamma(p)) \cdot \ell'_X(p)) \right)^2 \leq \text{Var}^p(g_n(X)) \cdot \mathbb{E}^p((\ell'_X(p))^2).$$

18. Bemerkung Im Beweis von Satz 17 wurde die konkrete Verteilungsannahme nicht wesentlich genutzt. Die untere Schranke des Satzes gilt deshalb unter viel allgemeineren Voraussetzungen.

Siehe Krengel (1998, p. 69) und Irle (2001, p. 308).

19. Satz Optimalität des empirischen Mittels

$$\begin{aligned} \forall p \in]0, 1[: R^p(g_n^*) \\ = \inf \{ R^p(g_n) : g_n \text{ e'treue Schätzfunktion für } \gamma \} . \end{aligned}$$

Beweis. Man verifiziert

$$E^p((\ell'_X(p))^2) = \frac{n}{p \cdot (1 - p)},$$

siehe Krengel (1998, p. 70), und wendet Satz 17 und Bsp. 16 an.

□

20. Beispiel Sei $\Theta := \mathbb{R}$ und

$$P_{X_1}^\vartheta := \mathbf{U}([\vartheta - 1/2, \vartheta + 1/2])$$

für $\vartheta \in \Theta$. Zu schätzen ist

$$\gamma(\vartheta) := \mathbf{E}^\vartheta(X_1) = \vartheta.$$

Setze

$$g_3(x_1, x_2, x_3) := \frac{\max(x_1, x_2, x_3) + \min(x_1, x_2, x_3)}{2}$$

ÜBUNG M:H44: g_n ist erwartungstreu für γ mit

$$\forall \vartheta \in \Theta : R^\vartheta(g_n) < R^\vartheta(g_n^*).$$

Zu Frage (iii):

21. Definition Sei $\alpha \in]0, 1[$. Borel-meßbare Abbildungen

$\ell_n, r_n : \mathbb{R}^n \rightarrow \mathbb{R}$ bilden **Konfidenzintervall** für γ zum Niveau $1 - \alpha$, falls

$$\forall \vartheta \in \Theta : P^\vartheta(\{\gamma(\vartheta) \in [\ell_n(X), r_n(X)]\}) \geq 1 - \alpha.$$

Klar: Gesucht sind „möglichst kleine“ Konfidenzintervalle.

Beachte: Nicht $\gamma(\vartheta)$ sondern $[\ell_n(X), r_n(X)]$ ist zufällig.

Im folgenden der Spezialfall

$$\gamma(\vartheta) := \mathbb{E}^{\vartheta}(X_1)$$

und Intervalle der Form

$$[\bar{x}_n - b_n(x), \bar{x}_n + b_n(x)]$$

mit Borel-meßbaren Abbildungen

$$b_n : \mathbb{R}^n \rightarrow [0, \infty[.$$

22. Bemerkung b_n definiert genau dann Konfidenzintervall für Erwartungswert zum Niveau $1 - \alpha$, wenn

$$\forall \vartheta \in \Theta : P^{\vartheta}(\{|\bar{X}_n - E^{\vartheta}(X_1)| \leq b_n(X)\}) \geq 1 - \alpha.$$

Also Tschebyschev-Ungleichung anwendbar, falls (Schranke für) $\sup_{\vartheta \in \Theta} \text{Var}^{\vartheta}(X_1)$ bekannt. Auf diese Weise erhält man jedoch oft zu große Konfidenzintervalle.

23. Beispiel Fortsetzung von Bsp. VI.34 und 8.

Konfidenzintervalle **deterministischer Breite** $2 \cdot b_n$ mittels

- Tschebyschev-Ungleichung. Es gilt

$$\frac{1}{4 \cdot \varepsilon^2 \cdot n} = \alpha \Leftrightarrow \varepsilon = \frac{1}{\sqrt{4 \cdot \alpha \cdot n}}.$$

- Hoeffdingscher Ungleichung. Es gilt

$$2 \cdot \exp(-2 \cdot \varepsilon^2 \cdot n) = \alpha \Leftrightarrow \varepsilon = \sqrt{\frac{\ln(2/\alpha)}{2 \cdot n}}.$$

Auf diese Weise

α	b_n per T-Ungl.	b_n per H-U
0.05	$4.5 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$
0.01	$1.0 \cdot 10^{-3}$	$3.3 \cdot 10^{-4}$
0.001	$3.2 \cdot 10^{-3}$	$3.8 \cdot 10^{-4}$

Nun: Konfidenzintervalle für Erwartungswert unter

Normalverteilungsannahmen. Unterscheide zwei Fälle:

- Varianz $\sigma^2 > 0$ bekannt. Also $\Theta := \mathbb{R}$ und $P_{X_1}^\mu := \mathbf{N}(\mu, \sigma^2)$ für $\mu \in \Theta$.
- Varianz unbekannt. Also $\Theta := \mathbb{R} \times]0, \infty[$ und $P_{X_1}^{(\mu, \sigma)} := \mathbf{N}(\mu, \sigma^2)$ für $(\mu, \sigma) \in \Theta$.

24. Satz Bei bekannter Varianz σ^2 definiert

$$b_n := \sigma / \sqrt{n} \cdot \Phi^{-1}(1 - \alpha/2)$$

ein Konfidenzintervall für den Erwartungswert zum Niveau

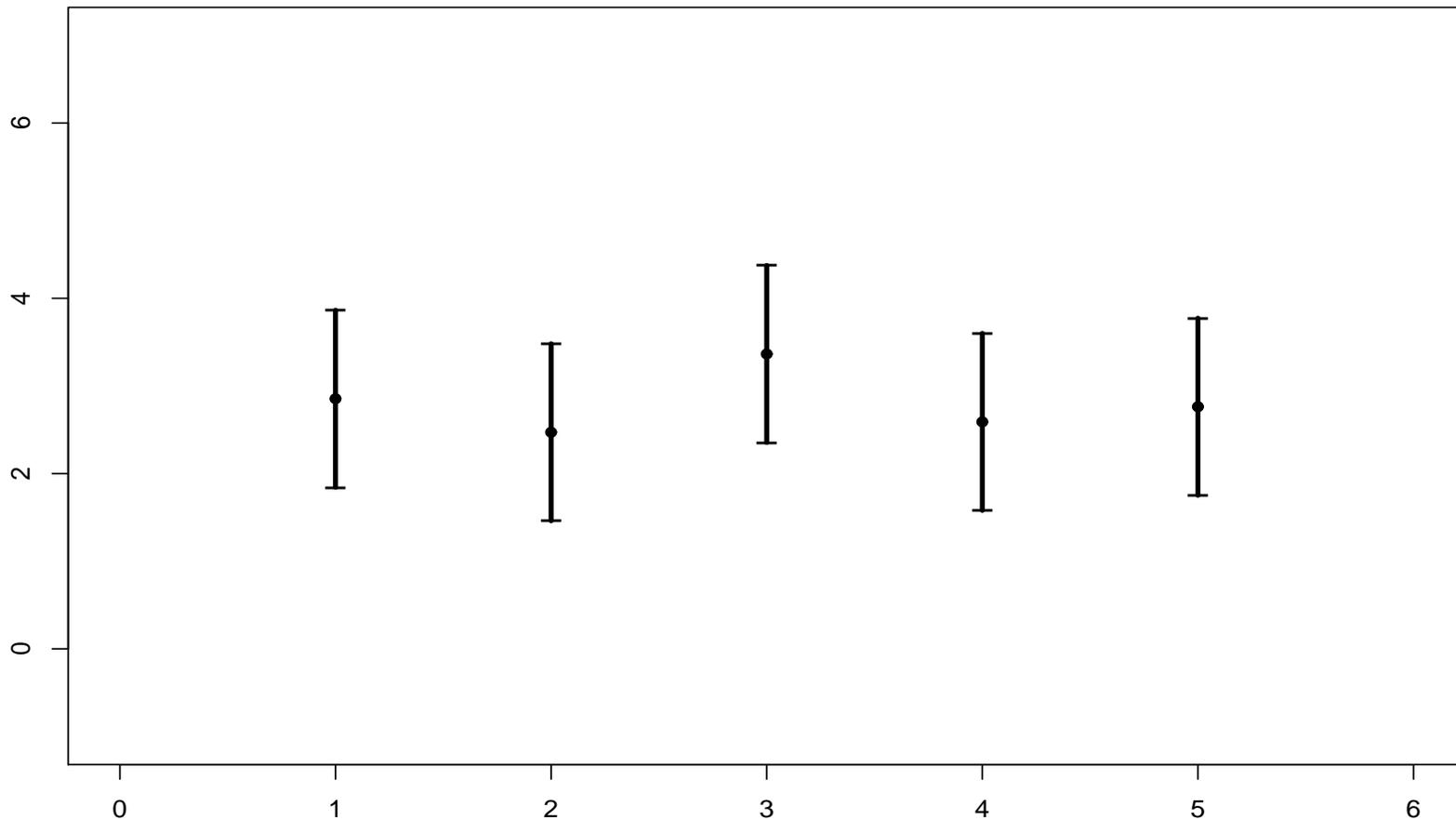
$1 - \alpha$.

Beweis. PROJEKTOR



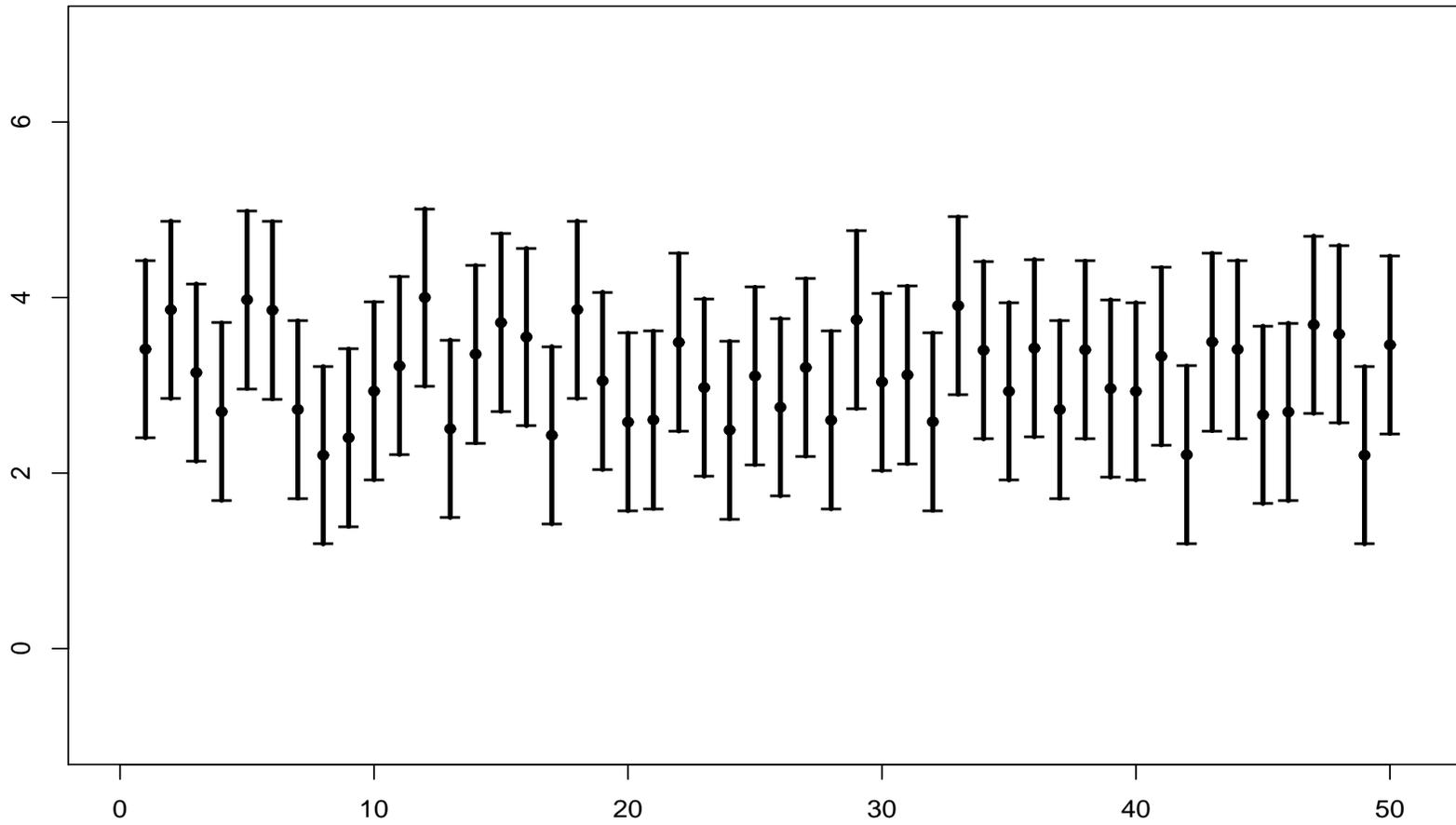
25. Beispiel Konfidenzintervalle für $\alpha := 0,05$, $\sigma := 2$, $n := 15$ (und $\mu := 3$).

5 Realisierungen



Stichprobe bei 1. Realisierung: 3, 3.4, 1.7, 6.1, 5.5, 0.6, 2.6, 4.8, 3.7, 2.7, 2.8, 2.2, 2.3, 0.4, 0.4, empir. Mittel: 2.9

50 Realisierungen



Stichprobe bei 1. Realisierung: 6.1 , 6.8 , 3.8 , 3.5 , 1.4 , -0.5 , -0.7 , 6.2 , 5.6 , 2.7 , 6.8 , -0.3 , 5.2 , 2.2 , 2.2 , empir. Mittel: 3.4

Bei unbekannter Varianz naheliegend: Ersetze σ^2 durch die empirische Varianz $v_n(x)$.

Im folgenden: $n > 1$ und X'_1, \dots, X'_n iid mit $X'_1 \sim \mathbf{N}(0, 1)$.

Setze

$$\overline{X}'_n := \sum_{i=1}^n X'_i / n,$$

$$X' := (X'_1, \dots, X'_n)$$

und

$$T'_n := \frac{\overline{X}'_n}{\sqrt{v_n(X')/n}}.$$

Beachte: Nenner fast sicher ungleich Null.

26. Lemma T'_n besitzt die Dichte

$$f_n(x) := \frac{\Gamma(n/2)}{\Gamma((n-1)/2) \cdot \sqrt{\pi \cdot (n-1)}} \cdot \left(1 + x^2/(n-1)\right)^{-n/2}.$$

Beweis. Siehe Irle (2001, Kapitel 20). □

27. Bemerkung Für alle $x \in \mathbb{R}$ gilt

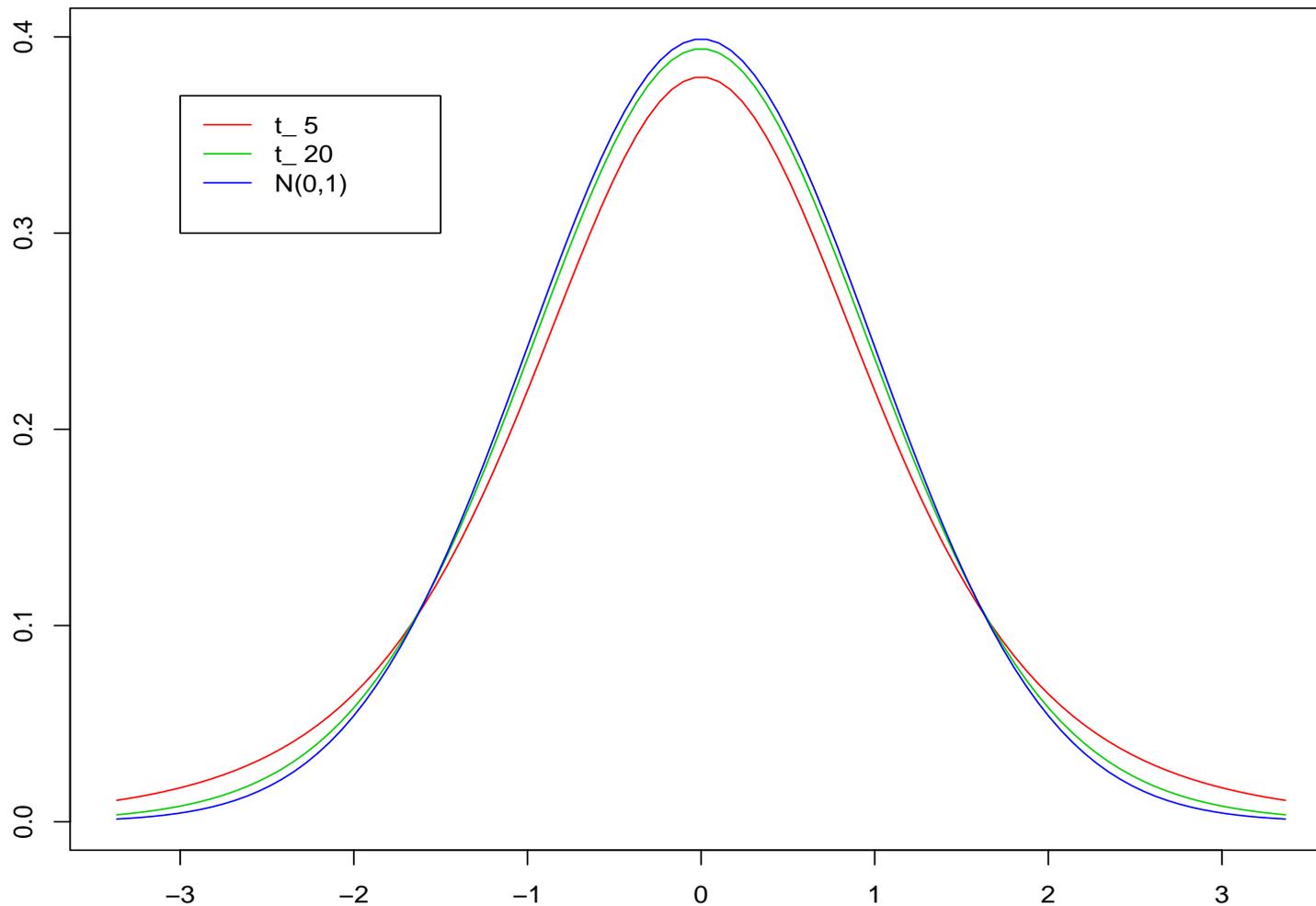
$$\lim_{n \rightarrow \infty} f_n(x) = 1/\sqrt{2\pi} \cdot \exp(-x^2/2).$$

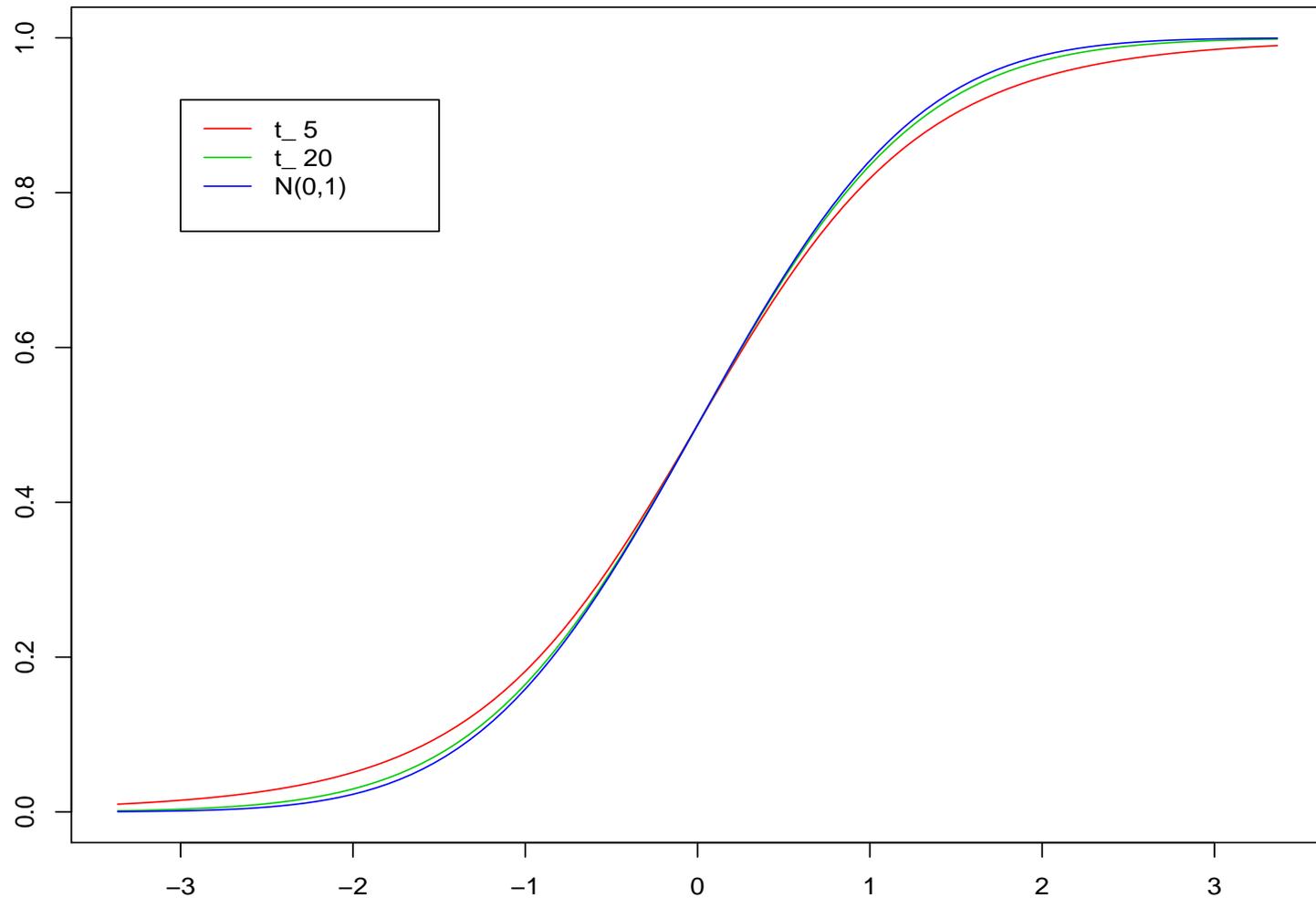
28. Definition Die Verteilung der ZV T'_n heißt **t -Verteilung** mit $n - 1$ Freiheitsgraden.

Bez.: t_{n-1} .

29. Beispiel Dichte und Verteilungsfunktion von t_5 und t_{20} .

Zum Vergleich Dichte und Verteilungsfunktion von $\mathbf{N}(0, 1)$.





30. Bemerkung Zur Berechnung der Verteilungsfunktion von t_n und entsprechender Quantile: Numerik, Tabellen, Plots.

31. Lemma Bzgl. $P^{(\mu, \sigma)}$ gilt

$$\frac{\overline{X}_n - \mu}{\sqrt{v_n(X)/n}} \sim t_{n-1}.$$

Beweis. Setze $X'_i := \frac{X_i - \mu}{\sigma}$. Bzgl. $P^{(\mu, \sigma)}$ gilt X'_1, \dots, X'_n iid und $X'_1 \sim \mathbf{N}(0, 1)$. Ferner

$$\overline{X'_n} = \sum_{i=1}^n \frac{X'_i}{n} = \frac{\overline{X_n} - \mu}{\sigma}, \quad X'_i - \overline{X'_n} = \frac{X_i - \overline{X_n}}{\sigma},$$

$$(n-1) \cdot v_n(X') = \sum_{i=1}^n (X'_i - \overline{X'_n})^2 = \frac{1}{\sigma^2} \cdot \sum_{i=1}^n (X_i - \overline{X_n})^2 = \frac{(n-1) \cdot v_n(X)}{\sigma^2}$$

Fazit

$$\frac{\overline{X'_n}}{\sqrt{v_n(X')}/\sqrt{n}} = \frac{\overline{X_n} - \mu}{\sigma} \cdot \frac{\sigma}{\sqrt{v_n(X)}/\sqrt{n}} = \frac{\overline{X_n} - \mu}{\sqrt{v_n(X)}/\sqrt{n}}.$$

□

32. Satz Bei unbekannter Varianz definiert

$$b_n(x) := \sqrt{v_n(x)/n} \cdot t_{n-1;1-\alpha/2}$$

mit dem $(1 - \alpha/2)$ -Quantil $t_{n-1;1-\alpha/2}$ der t -Verteilung mit $(n - 1)$ -Freiheitsgraden ein Konfidenzintervall für den Erwartungswert zum Niveau $1 - \alpha$.

33. Beispiel Für $\alpha := 0,05$ ergibt sich

n	21	51	101
$t_{n-1;1-\alpha/2}$	2,09 ...	2,01 ...	1,98 ...

Zum Vergleich: $\Phi^{-1}(1 - \alpha/2) = 1,96 \dots$

Beweis von Satz 32. Vgl. Beweis von Satz 24.

Sei $F_n := F_{Z_n}$ die Verteilungsfunktion von

$$Z_n := \frac{\overline{X}_n - \mu}{\sqrt{v_n(X)/n}}.$$

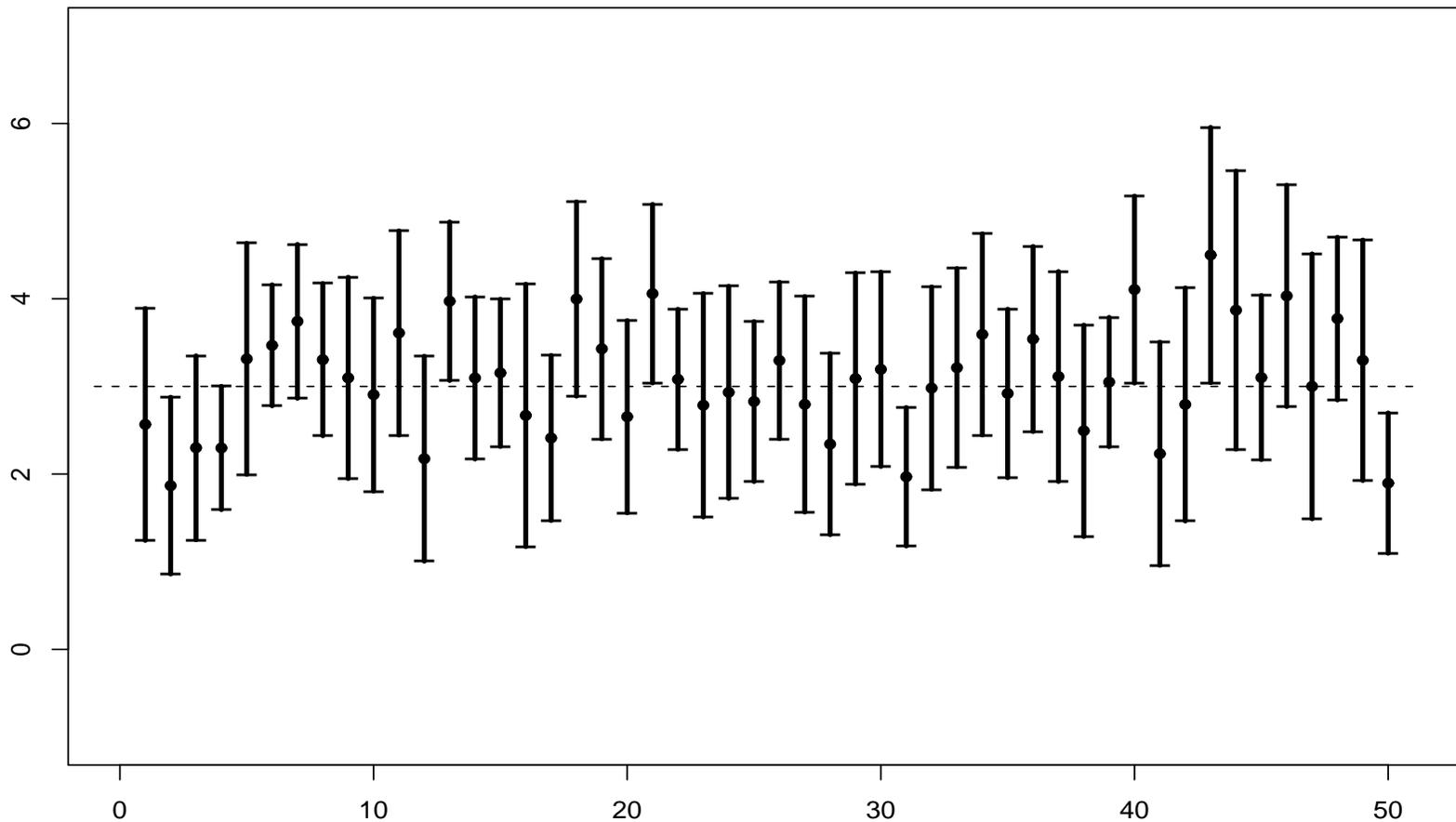
Für alle $\mu \in \mathbb{R}$ und $\sigma > 0$ gilt gem. Lemma 31

$$\begin{aligned} P^{(\mu, \sigma)}(\{|\overline{X}_n - \mu| \leq b_n(X)\}) &= P^{(\mu, \sigma)}(\{|Z_n| \leq t_{n-1; 1-\alpha/2}\}) \\ &= F(t_{n-1; 1-\alpha/2}) - F(-t_{n-1; 1-\alpha/2}) \\ &= 2 \cdot F(t_{n-1; 1-\alpha/2}) - 1 \\ &= 2(1 - \alpha/2) - 1 = 1 - \alpha. \end{aligned}$$

□

34. Beispiel Konfidenzintervalle für $\alpha := 0,05$ $n := 15$ (und $\mu := 3$, $\sigma := 2$).

50 Realisierungen



Stichprobe bei 1. Realisierung: 3.2 , 4.8 , 4.8 , 4.7 , 0.3 , 3.1 , 4.8 , 1.7 , -1.8 , -2.5 , 4.2 , 5 , 2.5 , 1.7 , empir. Mittel: 2.6

Ausblick: **asymptotische Konfidenzintervalle.**

Gelte

$$\forall \vartheta \in \Theta : X_1 \in \mathfrak{L}_2^\vartheta \wedge \text{Var}^\vartheta(X_1) > 0.$$

Setze

$$b_n(x) := \sqrt{v_n(x)/n} \cdot \Phi^{-1}(1 - \alpha/2).$$

35. Satz Für jedes $\alpha \in]0, 1[$

$$\forall \vartheta \in \Theta : \lim_{n \rightarrow \infty} P^\vartheta(\{|\bar{X}_n - E^\vartheta(X_1)| \leq b_n(X)\}) = 1 - \alpha.$$

Beweis. Beruht auf dem Zentralen Grenzwertsatz. □

36. Beispiel Fortsetzung von Bsp. 23. Siehe auch Bsp. 11.

α	b_n per T-Ungl.	b_n per H-U	b_n asymp. per ZGS
0,05	$4,5 \cdot 10^{-4}$	$2,8 \cdot 10^{-4}$	$1,95 \cdot 10^{-4}$
0,01	$1,0 \cdot 10^{-3}$	$3,3 \cdot 10^{-4}$	$2,57 \cdot 10^{-4}$
0,001	$3,2 \cdot 10^{-3}$	$3,8 \cdot 10^{-4}$	$3,27 \cdot 10^{-4}$

3 Testtheorie

37. Beispiel Maschine zur Produktion von Gewinderingen.

Frage: Wird in der Produktion ein Nennmaß μ_0 für den Innendurchmesser eingehalten?

Daten: Innendurchmesser x_1, \dots, x_n von n Gewinderingen.

Annahme: Innendurchmesser sind Realisierungen von unabhängigen normalverteilten ZVen mit Erwartungswert μ und Varianz σ^2 .

1. Fall Präzision der Maschine, gegeben durch σ^2 , bekannt.

2. Fall Präzision der Maschine unbekannt.

Betrachte Testproblem, gegeben durch

- $(\Omega, \mathfrak{A}, P^\vartheta)_{\vartheta \in \Theta}$,
- Hypothese $\emptyset \neq \Theta_0 \subsetneq \Theta$

(„der unbekannte Parameter liegt in Θ_0 “, „ $\vartheta \in \Theta_0$ “),

Stichprobe $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ ist Realisierung von
 $X = (X_1, \dots, X_n)$.

38. Definition Verwerfungsbereich $R_n \in \mathfrak{B}_n$ definiert **Signifikanztest** zum Niveau $\alpha \in]0, 1[$ (für die Hypothese Θ_0), falls

$$\forall \vartheta \in \Theta_0 : P^\vartheta(\{X \in R_n\}) \leq \alpha.$$

Entscheidung: Lehne Hypothese genau dann ab, wenn $x \in R_n$.

Wahrscheinlichkeit des Fehlers 1. Art für jedes $\vartheta \in \Theta$ beschränkt durch α .

39. Bemerkung Um kleine Wahrscheinlichkeiten für Fehler 2. Art zu erreichen, sucht man Signifikanztest zu gegebenem Niveau α , dessen Verwerfungsbereich R_n „möglichst groß“ ist. In dieser Vorlesung keine Optimalitätsaussagen.

40. Bemerkung Formulierung eines Testproblems symmetrisch in Θ_0 (**Hypothese**) und $\Theta \setminus \Theta_0$ (**Alternative**).
Nicht so bei Signifikanztest.

41. Beispiel Signifikanztest zum Niveau α für

- „ $\Theta_0 =$ schuldig“: Freispruch Schuldiger mit Wahrscheinlichkeit höchstens α .
- „ $\Theta_0 =$ unschuldig“: Verurteilung Unschuldiger mit Wahrscheinlichkeit höchstens α .

42. Bemerkung In der Regel (und oBdA) Verwerfungsbereich von der Form

$$R_n := \{g_n \in K_n\}$$

mit Borel-meßbarer Abbildung $g_n : \mathbb{R}^n \rightarrow \mathbb{R}$ und $K_n \in \mathfrak{M}$.

Die ZV $g_n(X)$ heißt in diesem Kontext **Teststatistik** und K_n **kritischer Bereich**.

Entscheidung: Lehne Hypothese genau dann ab, wenn $g_n(x) \in K_n$.

Verwende „plausible“ Teststatistik $g_n(X)$, deren Verteilungen $P_{g_n(X)}^\vartheta$ für $\vartheta \in \Theta_0$ (approximativ) bekannt sind.

43. Beispiel Normalverteilungsannahme mit **bekannter Varianz** $\sigma^2 > 0$. Fortsetzung von Beispiel 37, 1. Fall.

Gegeben: $\mu_0 \in \mathbb{R}$. Entscheide, ob der unbekannte Erwartungswert gleich μ_0 ist.

Formal: $\Theta := \mathbb{R}$, $P_{X_1}^\mu := \mathbf{N}(\mu, \sigma^2)$ und $\Theta_0 := \{\mu_0\}$.

Setze

$$g_n(x) := \frac{\bar{x}_n - \mu_0}{\sigma / \sqrt{n}}.$$

Falls Hypothese korrekt, so ist $g_n(X) \mathbf{N}(0, 1)$ -verteilt.

Also für $k > 0$

$$\begin{aligned} P^{\mu_0}(\{|g_n(X)| \geq k\}) &= 1 - \Phi(k) + \Phi(-k) \\ &= 2(1 - \Phi(k)). \end{aligned}$$

Fazit: $R_n = \{|g_n| \geq k\}$ definiert genau dann einen Signifikanztest zum Niveau α , wenn $2(1 - \Phi(k)) \leq \alpha$, d.h.

$$k \geq \Phi^{-1}(1 - \alpha/2).$$

Klar: Fehlerwahrscheinlichkeiten 2. Art sind monoton wachsend in k . Man wählt also

$$k = \Phi^{-1}(1 - \alpha/2).$$

Speziell für $\alpha = 0,05$ ergibt sich $k = 1,96 \dots$

Entscheidung: Lehne Hypothese Θ_0 genau dann ab, wenn

$$\left| \bar{x}_n - \mu_0 \right| \geq \sigma / \sqrt{n} \cdot \Phi^{-1}(1 - \alpha/2).$$

Vgl. Satz 24.

Bez. zweiseitiger Gauß-Test.

Ausblick: einseitiges Testproblem gegeben durch $\Theta_0 :=]-\infty, \mu_0]$. Siehe

ÜBUNG M:G22.

44. Beispiel Fortsetzung von Bsp. 3, Geschlecht eines Neugeborenen. Widerlege die Annahme, daß Jungen und Mädchen mit gleicher Wahrscheinlichkeit geboren werden.

Verteilungsannahme: $P_{X_1}^\vartheta = \mathbf{B}(1, \vartheta)$ mit $\vartheta \in \Theta :=]0, 1[$.

Hypothese: $\Theta_0 = \{1/2\}$.

Teststatistik

$$g_n(X) := \frac{\bar{X}_n - 1/2}{1/(2\sqrt{n})} = \frac{\sum_{i=1}^n X_i - n \cdot 1/2}{1/2 \cdot \sqrt{n}}.$$

Beachte, daß $n = 25\,171\,123$. Also „ist“ $g_n(X)$ bzgl. $P^{1/2}$ standard-normalverteilt.

Deshalb Gauß-Test, als **asymptotischer α -Niveau-Test**: Lehne Hypothese genau dann ab, wenn

$$\left| \bar{x}_n - 1/2 \right| \geq 1/(2\sqrt{n}) \cdot \Phi^{-1}(1 - \alpha/2).$$

Für $n = 25\,171\,123$ gilt $1/(2\sqrt{n}) = 9.9 \dots 10^{-5}$ und man erhält

α	$1/(2\sqrt{n}) \cdot \Phi^{-1}(1 - \alpha/2)$
10^{-2}	$2,56 \dots 10^{-4}$
10^{-3}	$3,27 \dots 10^{-4}$
10^{-20}	$9,30 \dots 10^{-4}$

Für das empirische Mittel $\bar{x}_n = 0,4863 \dots$ der Daten gilt

$$|\bar{x}_n - 1/2| = 1,37 \dots 10^{-2}.$$

45. Bemerkung Genauer zu Wahrscheinlichkeiten für Fehler 2. Art bei zweiseitigem Gauß-Test. Sei $\mu \neq \mu_0$ und

$$\mu_n := \sqrt{n} \cdot (\mu - \mu_0) / \sigma.$$

Dann ist $g_n(X)$ $\mathbf{N}(\mu_n, 1)$ -verteilt bzgl. P^μ und

$$\begin{aligned} & P^\mu(\{|g_n(X)| < k\}) \\ &= P^\mu(\{-k - \mu_n < g_n(X) - \mu_n < k - \mu_n\}) \\ &= \Phi(k - \mu_n) - \Phi(-k - \mu_n) =: a_n(\mu). \end{aligned}$$

Es gilt:

(i) Für jedes $\mu \neq \mu_0$ konvergiert $(a_n(\mu))_{n \in \mathbb{N}}$ sehr schnell gegen Null, siehe ÜBUNG M:H45.

(ii) Für jedes $n \in \mathbb{N}$

$$\sup_{\mu \in \Theta \setminus \Theta_0} a_n(\mu) = 1 - \alpha.$$

Vgl. Seite 333.

46. Definition **Operationscharakteristik** f eines durch Verwerfungsbereich $R_n \in \mathfrak{B}_n$ gegebenen Tests

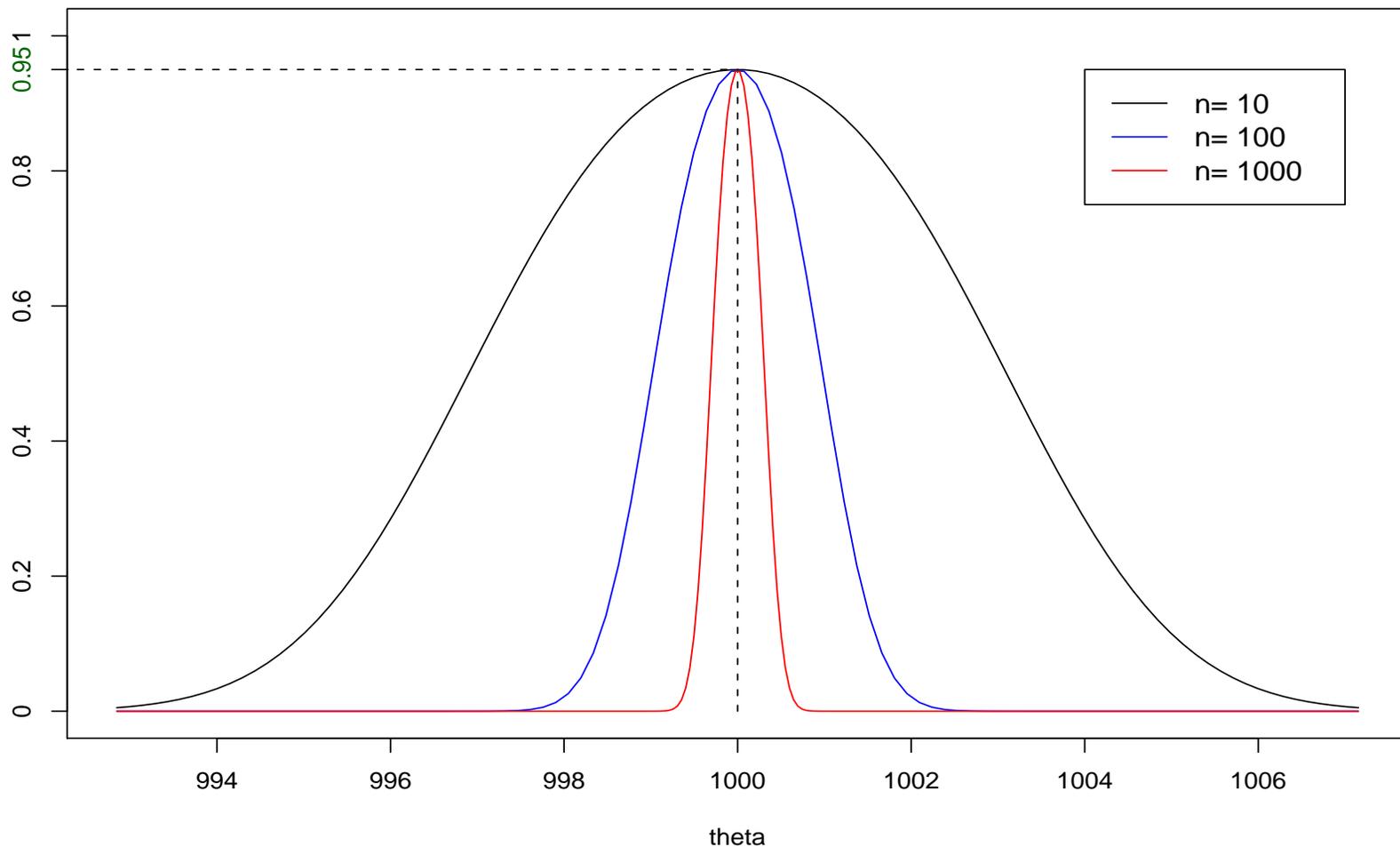
$$f : \Theta \rightarrow [0, 1]$$

$$\vartheta \mapsto P^\vartheta(\{X \notin R_n\}).$$

47. Bemerkung

- (i) $f|_{\Theta \setminus \Theta_0}$: Wahrscheinlichkeiten für Fehler 2. Art.
- (ii) Für Signifikanztest zum Niveau α : $f|_{\Theta_0} \geq 1 - \alpha$.

48. Beispiel Operationscharakteristiken des zweiseitigen Gauß-Tests für $\Theta_0 := \{1000\}$, $\sigma^2 := 25$ und $\alpha := 0,05$.



Nun: **Normalverteilungsannahme** mit **unbekannter Varianz**.

Fortsetzung von Bsp. 37, 2. Fall.

Gegeben: $\mu_0 \in \mathbb{R}$. Entscheide, ob der unbekannte Erwartungswert gleich μ_0 ist.

Formal: $\Theta := \mathbb{R} \times]0, \infty[$, $P_{X_1}^{(\mu, \sigma)} := \mathbf{N}(\mu, \sigma^2)$ und

$\Theta_0 := \{\mu_0\} \times]0, \infty[$.

Betrachte die Teststatistik

$$g_n(X) := \frac{\bar{X}_n - \mu_0}{\sqrt{v_n(X)/n}},$$

wähle $t_{n-1;1-\alpha/2}$ als $(1 - \alpha/2)$ -Quantil der t -Verteilung mit $n - 1$ Freiheitsgraden und definiere

$$R_n := \{|g_n| \geq t_{n-1;1-\alpha/2}\}.$$

Vgl. Satz 32.

49. Satz **zweiseitiger t -Test**

Unter obiger Verteilungsannahme definiert der

Verwerfungsbereich R_n einen Signifikanztest zum Niveau α .

Beweis. Wie in Bsp. 43. Unter Verwendung von Lemma 31.



Nun: Zufallsexperimente mit Werten in einer **endlichen Menge**

$M \subset \mathbb{R}^d$. Teste,

- (i) ob Zufallsexperiment einer gegebenen Verteilung genügt,
- (ii) ob die Komponenten eines vektorwertigen Zufallsexperimentes unabhängig sind.

50. Beispiel

- (i) Teste, ob ein Würfel fair ist.
- (ii) Teste, ob Einkommen und politische Präferenz unabhängig sind.

Im folgenden:

- $\emptyset \neq M \subseteq \mathbb{R}^d$ endlich.
- $\Theta = \{(p_k)_{k \in M} \in [0, 1]^M : \sum_{k \in M} p_k = 1\}$ Menge der **Wahrscheinlichkeitsfunktionen** auf M .
- Für $\mathbf{p} = (p_k)_{k \in M} \in \Theta$ ist die Verteilung von X_1 unter $P^{\mathbf{p}}$ gegeben durch

$$P_{X_1}^{\mathbf{p}}(\{k\}) = p_k.$$

Der Einfachheit halber: X_1, \dots, X_n nehmen nur Werte aus M an.

Zunächst Fragestellung (i), also

$$\Theta_0 = \{\mathbf{p}^0\}$$

für feste Wahrscheinlichkeitsfunktion $\mathbf{p}^0 \in \Theta$ mit $p_m^0 > 0$ für alle $m \in M$.

51. Beispiel $M := \{1, \dots, 6\}$ und $p_k^0 := 1/6$ für $k \in M$ bei der Frage, ob Würfel fair. Siehe auch Bsp. 44.

Im folgenden oBdA $M = \{1, \dots, m\}$ mit $m \geq 2$.

Betrachte die **absoluten Häufigkeiten** der Werte $1, \dots, m$ in Stichprobe. Definiere $H : M^n \rightarrow \mathbb{N}_0^m$ durch

$$H_k(x_1, \dots, x_n) := \sum_{i=1}^n 1_{\{k\}}(x_i)$$

für $k = 1, \dots, m$.

Naheliegend: Verwerfung der Hypothese Θ_0 , falls „Abstand“ von $1/n \cdot H(x_1, \dots, x_n)$ und \mathbf{p}^0 „groß“.

Dazu: Bestimmung der **Verteilung** des Zufallsvektors $H(X)$.

52. Beispiel Für $M := \{1, 2\}$ gilt

$$H_1(X) = \sum_{i=1}^n \underbrace{1_{\{1\}}(X_i)}_{=: X'_i}.$$

Aus $P_{X'_1}^{\text{P}} = \mathbf{B}(1, p_1)$ folgt $H_1(X) \sim \mathbf{B}(n, p_1)$ bzgl. P^{P} .

Klar: $H_2(X) = n - H_1(X)$. Also für $k \in \{0, \dots, n\}$

$$P^{\text{P}}(\{H(X) = (k, n - k)\}) = \binom{n}{k} \cdot p_1^k \cdot p_2^{n-k}.$$

53. Bemerkung Analog folgt allgemein $H_k(X) \sim \mathbf{B}(n, p_k)$

bzgl. P^{P} . Beachte: die Komponenten $H_k(X)$ des

Zufallsvektors $H(X)$ sind nicht unabhängig.

54. Satz Für $\mathbf{p} \in \Theta$ und $h \in \mathbb{N}_0^m$ gilt

$$P^{\mathbf{p}}(\{H(X) = h\}) = n! \cdot \prod_{k=1}^m \frac{p_k^{h_k}}{h_k!}$$

falls $\sum_{k=1}^m h_k = n$.

Andernfalls gilt $P^{\mathbf{p}}(\{H(X) = h\}) = 0$.

Beweis. Klar: $P^{\mathbf{p}}(\{H(X) = h\}) = 0$, falls $\sum_{k=1}^m h_k \neq n$.

Im folgenden gelte $\sum_{k=1}^m h_k = n$.

Für $x \in M^n$ mit $H(x) = h$

$$P^{\mathbb{P}}(\{X = x\}) = p_{x_1} \cdots p_{x_n} = \prod_{k=1}^m p_k^{h_k}.$$

Anzahl Stichproben mit absoluten Häufigkeiten h_k

$$\begin{aligned} & |\{x \in M^n : H(x) = h\}| \\ &= \binom{n}{h_1} \cdot \binom{n - h_1}{h_2} \cdots \binom{n - (h_1 + \cdots + h_{m-1})}{h_m} \\ &= \frac{n!}{h_1! \cdots h_m!}. \end{aligned}$$

Fazit:

$$P^{\mathbb{P}}(\{H(X) = h\}) = \sum_{x \in M^n, H(x)=h} P^{\mathbb{P}}(\{X = x\}) = n! \cdot \prod_{k=1}^m \frac{p_k^{h_k}}{h_k!}$$

55. Definition m -dimensionaler Zufallsvektor Y auf $(\Omega, \mathfrak{A}, P)$ **multinomialverteilt** mit Parametern $n, m \in \mathbb{N}$ und $\mathbf{p} \in [0, 1]^m$, wobei $\sum_{k=1}^m p_k = 1$, falls

$$P(\{Y = y\}) = \begin{cases} n! \cdot \prod_{k=1}^m \frac{p_k^{y_k}}{y_k!}, & \text{falls } y \in \mathbb{N}_0^m \text{ mit } \sum_{k=1}^m y_k = n \\ 0, & \text{andernfalls.} \end{cases}$$

Bez.: $Y \sim \mathbf{M}(n, m, \mathbf{p})$.

56. Bemerkung Satz 54 zeigt: Bzgl. $P^{\mathbf{p}}$ gilt

$H(X) \sim \mathbf{M}(n, m, \mathbf{p})$.

57. Beispiel 12 Würfe eines fairen Würfels. Für $n := 12$,
 $m := 6$ und $\mathbf{p} := (1/6, \dots, 1/6)$

$$\begin{aligned} & \mathbf{M}(12, 6, \mathbf{p})(\{(3, 0, 0, 2, 6, 1)\}) \\ &= \frac{12!}{3! \cdot 2! \cdot 6!} \cdot 6^{-12} = 2,54 \dots 10^{-5} \end{aligned}$$

und

$$\begin{aligned} & \mathbf{M}(12, 6, \mathbf{p})(\{(2, 2, 2, 2, 2, 2)\}) \\ &= \frac{12!}{2^6} \cdot 6^{-12} = 3,43 \dots 10^{-3}. \end{aligned}$$

58. Bemerkung Satz 54 ermöglicht prinzipiell die Konstruktion eines **Signifikanztests** zum Niveau α für $\Theta_0 = \{\mathbf{p}^0\}$: Wähle (möglichst kleine endliche) Menge $A \subset \mathbb{N}_0^m$ mit

$$P^{\mathbf{p}^0}(\{H(X) \in A\}) \geq 1 - \alpha$$

und verwende den **Verwerfungsbereich**

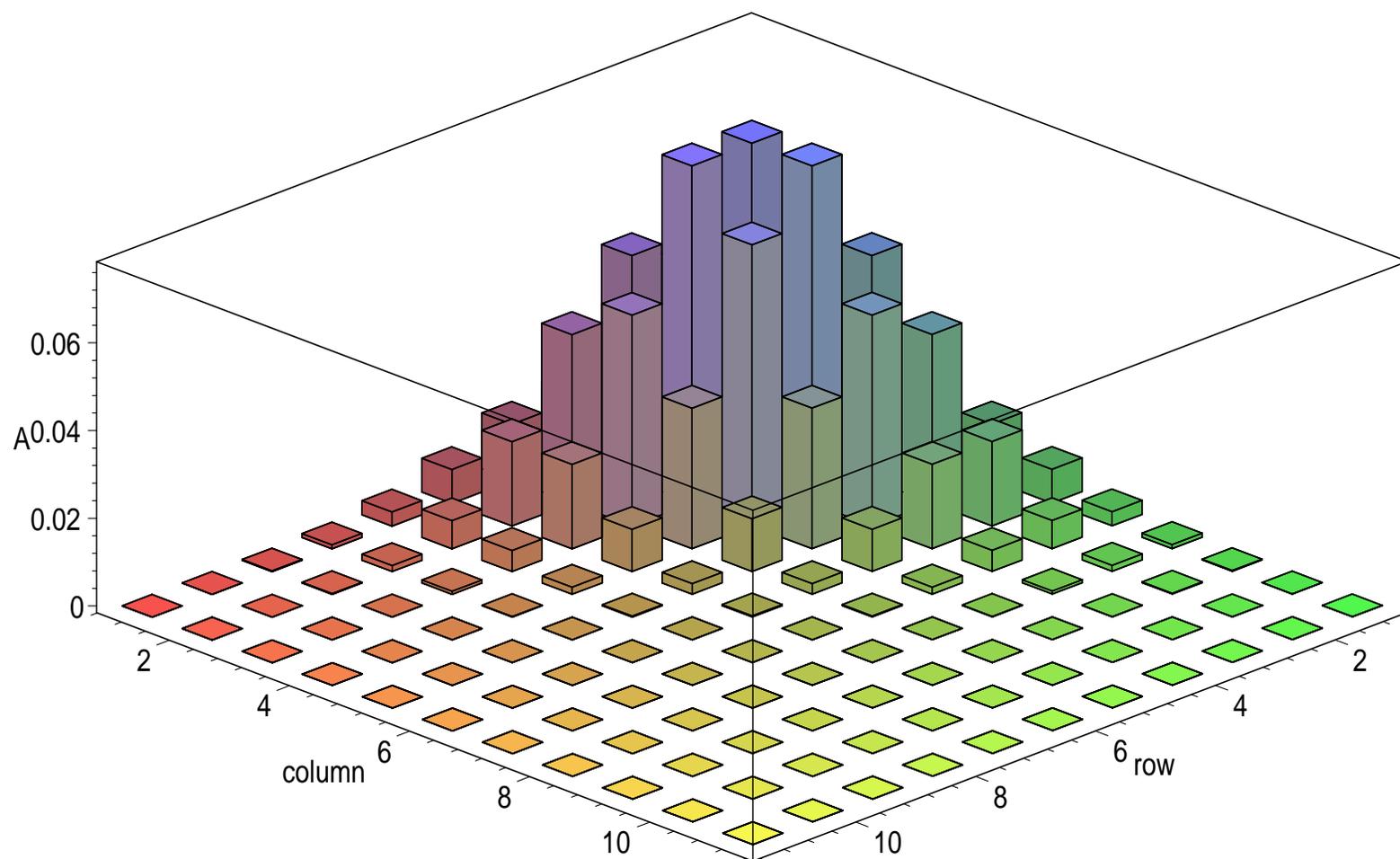
$$R_n := \{H \notin A\}.$$

Allg. Prinzip bei diskreten Teststatistiken.

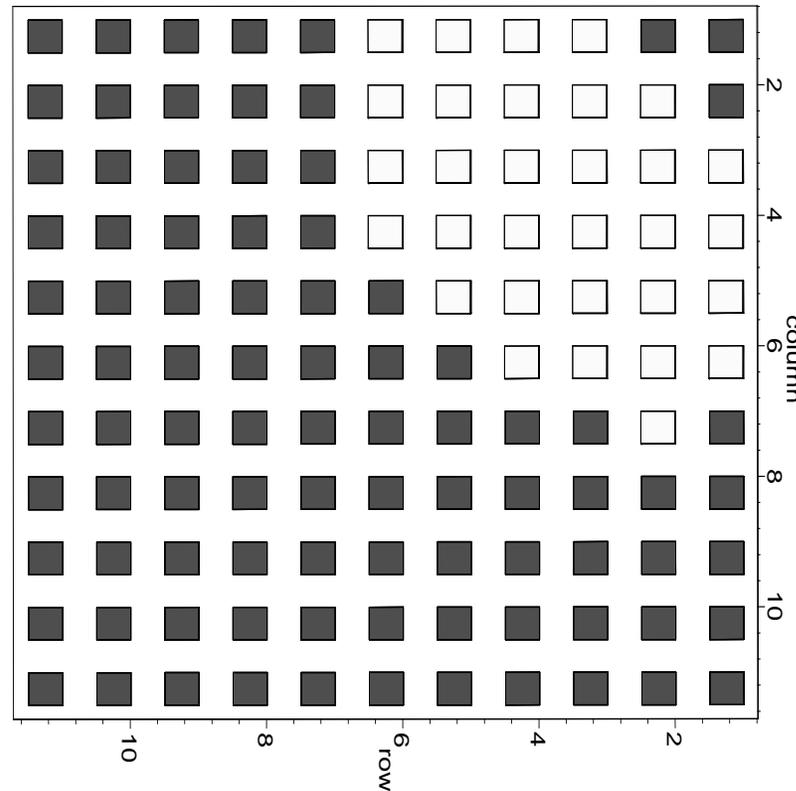
59. Beispiel Multinomialverteilung mit $m = 3$. Dargestellt ist

$$(h_1, h_2) \mapsto M(n, m, \mathbf{p})(\{(h_1, h_2, n - (h_1 + h_2))\})$$

für $n = 10$, $\mathbf{p} = (1/4, 1/4, 1/2)$.



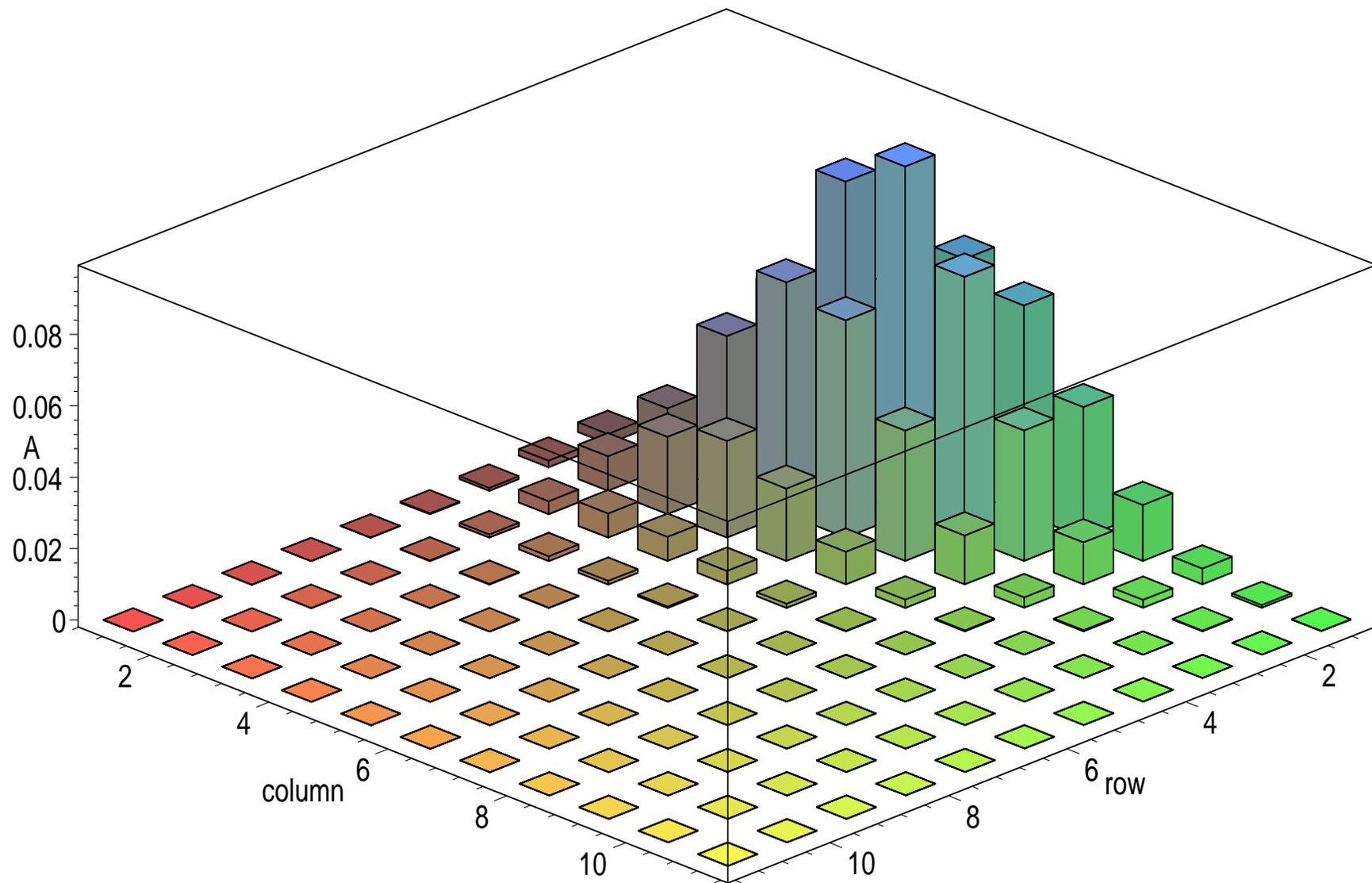
Eine optimale Wahl des Verwerfungsbereiches für $\alpha = 0,05$.



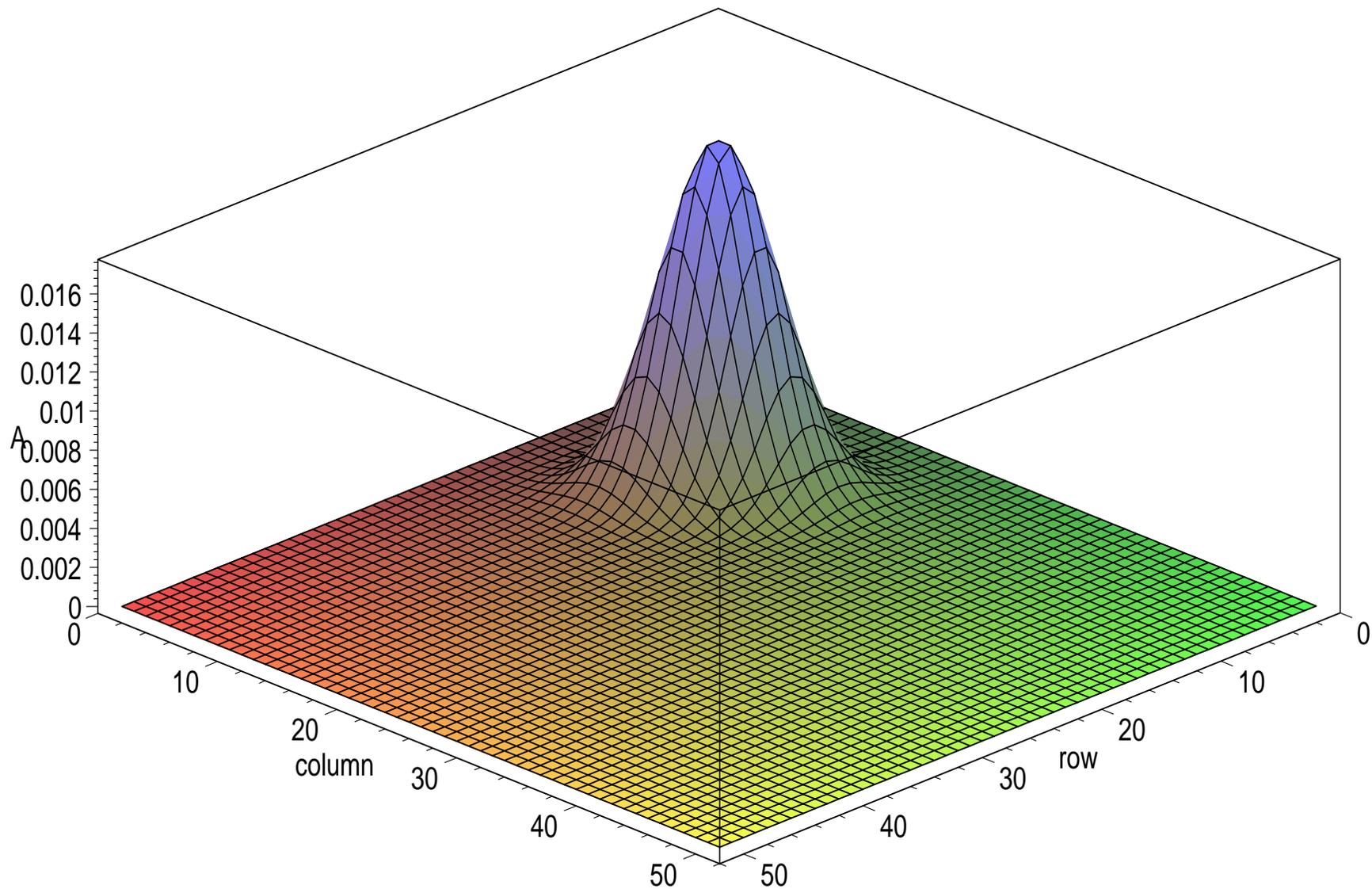
Verwirf die Hypothese genau dann, wenn (h_1, h_2) „schwarz markiert“.

Auf diese Weise: $P^{P^0}(\{X \in R_n\}) = 0,0488 \dots$

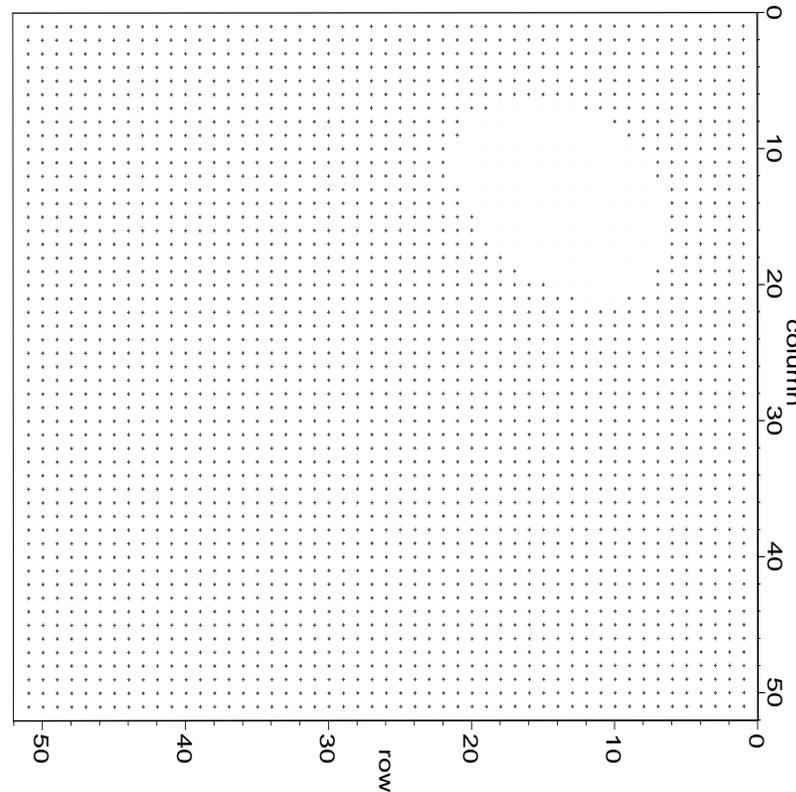
Multinomialverteilung: $n = 10$, $m = 3$, $\mathbf{p} = (1/8, 3/8, 1/2)$



Multinomialverteilung: $n = 50$, $m = 3$, $\mathbf{p} = (1/4, 1/4, 1/2)$. (Mit Interpolation.)



Eine optimale Wahl des Verwerfungsbereiches für $\alpha = 0,05$.



Verwirf die Hypothese genau dann, wenn (h_1, h_2) „schwarz markiert“.

Auf diese Weise: $P^{\mathbb{P}^0}(\{X \in R_n\}) = 0,0497 \dots$

Nachteile der Vorgehensweise gem. Bem. 58: Abhängigkeit von n , m und \mathbf{p}^0 und hoher Rechenaufwand, falls n groß.

Deshalb: **asymptotischer** α -Niveau-Test, siehe auch Seite 383.

Beachte: H und X hängen von n ab.

Definiere **Teststatistik** $Q(H(X))$ mit $Q : \mathbb{N}_0^m \rightarrow \mathbb{R}$ definiert durch

$$Q(h) = \sum_{k=1}^m \frac{(h_k - n \cdot p_k^0)^2}{n \cdot p_k^0}.$$

Untersuche Konvergenz der Verteilung von $Q(H(X))$ bzgl. $P^{\mathbf{p}^0}$ für $n \rightarrow \infty$.

Motivation (partiell): starkes Gesetz der großen Zahlen zeigt, daß

$H_k(X)/n$ $P^{\mathbf{p}^0}$ -f.s. gegen $P^{\mathbf{p}^0}(\{X_1 = k\}) = p_k^0$ konvergiert.

60. Definition Sei $d \in \mathbb{N}$. ZV Y mit Dichte

$$f_Y(x) = \frac{1}{2^{d/2} \cdot \Gamma(d/2)} \cdot x^{d/2-1} \cdot \exp(-x^2/2),$$

falls $x > 0$, und $f_Y(x) = 0$ andernfalls heißt χ^2 -verteilt mit d Freiheitsgraden.

Bez.: $Y \sim \chi_d^2$.

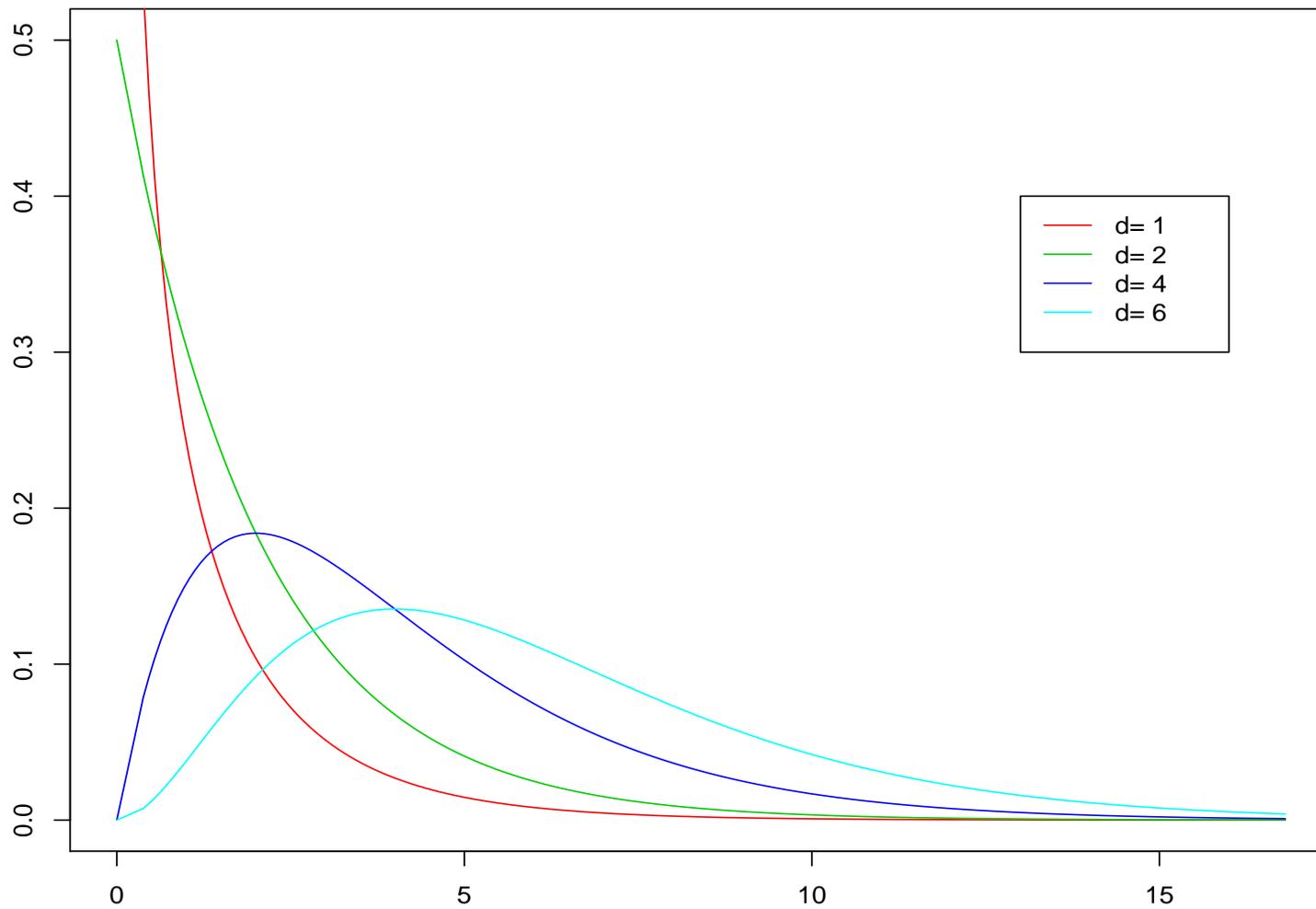
61. Bemerkung Für iid-ZVen Y_1, \dots, Y_d mit $Y_1 \sim \mathbf{N}(0, 1)$

und

$$Y := \sum_{i=1}^d Y_i^2$$

gilt $Y \sim \chi_d^2$. Siehe Krengel (1998, §13).

62. Beispiel Dichte von χ_d^2 für $d = 1, 2, 4, 6$.



63. Satz Gelte

$$Z_n \sim \mathbf{M}(n, m, \mathbf{p})$$

für $m \geq 2$ und $\mathbf{p} \in [0, 1]^m$, wobei $\sum_{k=1}^m p_k = 1$, sowie

$$Y \sim \chi_{m-1}^2.$$

Dann

$$Q(Z_n) \xrightarrow{d} Y.$$

Beweis. Siehe Krengel (1998, §13).

□

Wähle $\chi_{m-1;1-\alpha}^2$ als $(1 - \alpha)$ -Quantil der χ^2 -Verteilung mit $m - 1$ Freiheitsgraden und definiere

$$R_n := \{Q \circ H \geq \chi_{m-1;1-\alpha}^2\}.$$

Vgl. Satz 32.

64. Satz χ^2 -Anpassungstest

Unter der Verteilungsannahme von Seite 393 definiert der Verwerfungsbereich R_n einen Signifikanztest zum asymptotischen Niveau α für $\Theta_0 = \{\mathbf{p}^0\}$, d.h.

$$\lim_{n \rightarrow \infty} P^{\mathbf{p}^0}(\{X \in R_n\}) = \alpha.$$

Beweis. Sei F die Verteilungsfunktion von χ_{m-1}^2 . Es gilt

$$\{X \in R_n\} = \{Q(H(X)) \geq \chi_{m-1;1-\alpha}^2\}.$$

Da F stetig, folgt mit Satz 63

$$\lim_{n \rightarrow \infty} P^{\mathbf{P}^0}(\{X \in R_n\}) = 1 - F(\chi_{m-1;1-\alpha}^2) = \alpha.$$

□

65. Beispiel Fortsetzung von Bsp. 59 mit $\Theta_0 = \{\mathbf{p}^0\}$ für $m := 3$ und für $\mathbf{p}^0 := (1/4, 1/4, 1/2)$.

Wähle $\alpha := 0,05$. Somit $\chi_{m-1;1-\alpha}^2 = 5,99 \dots$

Für $n := 10$ und $H(x_1, \dots, x_n) := (0, 1, 9)$ ergibt sich

$$Q(H(x_1, \dots, x_n)) = 33/5 = 6,6 \geq \chi_{m-1;1-\alpha}^2,$$

so daß die Hypothese Θ_0 verworfen wird.

Für $n := 50$ und $H(x_1, \dots, x_n) := (9, 15, 26)$ ergibt sich

$$Q(H(x_1, \dots, x_n)) = 38/25 = 1,52 < \chi_{m-1;1-\alpha}^2,$$

so daß die Hypothese Θ_0 nicht verworfen wird.

66. Bemerkung Anwendung des χ^2 -Anpassungstests im Falle **absolutstetiger Verteilungen** auf \mathbb{R}^d .

Hypothese: unbekannte Verteilung von X_1 ist gleich P_0 .

Wähle $m \in \mathbb{N}$ und p.d. Mengen $B_k \in \mathfrak{B}_d$ mit

$\bigcup_{k=1}^m B_k = \mathbb{R}^d$. Definiere

$$X'_i := \sum_{k=1}^m k \cdot 1_{B_k}(X_i)$$

und betrachte $\Theta_0 = \{\mathbf{p}^0\}$ mit

$$p_k^0 := P_0(B_k).$$

67. Beispiel Frage: Liefert Ihr Zufallszahlengenerator auf $[0, 1]$ gleichverteilte Zufallszahlen?

Also $P_0 := \mathbf{U}([0, 1])$. Wähle

$$B_1 :=]-\infty, 1/m], \quad B_m :=]1 - 1/m, \infty[$$

und für $k = 2, \dots, m - 1$

$$B_k :=](k - 1)/m, k/m].$$

Somit für $k = 1, \dots, m$

$$p_k^0 := P_0(B_k) = 1/m.$$

Nun Fragestellung (ii):

Sind die Komponenten eines 2-dimensionalen
Zufallsvektors unabhängig?

Ohne Beschränkung der Allgemeinheit

$$M = \underbrace{\{1, \dots, m^{(1)}\}}_{=:M^{(1)}} \times \underbrace{\{1, \dots, m^{(2)}\}}_{=:M^{(2)}}$$

Also wie oben

$$\Theta = \left\{ \mathbf{p} = (p_{k,\ell})_{(k,\ell) \in M} \in [0, 1]^M : \sum_{k,\ell \in M} p_{k,\ell} = 1 \right\},$$

$$P_{X_1}^{\mathbf{p}}(\{(k, \ell)\}) = p_{k,\ell}$$

Hier

$$\Theta_0 = \{ \mathbf{p} \in \Theta : \exists \text{ W'funktionen } \mathbf{p}^{(i)} \text{ auf } M^{(i)} : \\ \forall (k, \ell) \in M : p_{k,\ell} = p_k^{(1)} \cdot p_\ell^{(2)} \}$$

Wie oben $H : M^n \rightarrow \mathbb{N}_0^{m^{(1)} \times m^{(2)}}$ Matrix der absoluten Häufigkeiten (**Kontingenztafel**).

Bez.: für $h \in \mathbb{N}_0^{m^{(1)} \times m^{(2)}}$, $k \in M^{(1)}$ und $\ell \in M^{(2)}$

$$h_{\bullet,\ell} = \sum_{k=1}^{m^{(1)}} h_{k,\ell}, \quad h_{k,\bullet} = \sum_{\ell=1}^{m^{(2)}} h_{k,\ell}.$$

Analog für ZVen.

68. Beispiel Siehe Lehn, Wegmann (2004, p. 172). Hier gilt

$$m^{(1)} = 5, m^{(2)} = 4.$$

Definiere Teststatistik $Q(H(x)) = Q_n(H_n(x))$ mit

$$Q : \mathbb{N}_0^{m^{(1)} \times m^{(2)}} \rightarrow \mathbb{R} \text{ durch}$$

$$Q(h) = \sum_{k=1}^{m^{(1)}} \sum_{\ell=1}^{m^{(2)}} \frac{(h_{k,\ell} - h_{k,\bullet} \cdot h_{\bullet,\ell}/n)^2}{h_{k,\bullet} \cdot h_{\bullet,\ell}/n}.$$

Motivation (partiell): starkes Gesetz der großen Zahlen zeigt

- $H_{k,\ell}(X)/n$ konvergiert $P^{\mathbb{P}}$ -f.s. gegen $p_{k,\ell}$,
- $H_{k,\bullet}(X) \cdot H_{\bullet,\ell}(X)/n^2$ konvergiert $P^{\mathbb{P}}$ -f.s. gegen $p_{k,\bullet} \cdot p_{\bullet,\ell}$.

Für $\mathbf{p} \in \Theta_0$ gilt $p_{k,\ell} = p_{k,\bullet} \cdot p_{\bullet,\ell}$, und somit konvergiert

$$H_{k,\ell}(X)/n - H_{k,\bullet}(X) \cdot H_{\bullet,\ell}(X)/n^2 \text{ } P^{\mathbb{P}}\text{-f.s. gegen } 0.$$

69. Satz Für $\mathbf{p} \in \Theta_0$ gilt

$$Q_n(H_n(X)) \xrightarrow{d} Y,$$

wobei $Y \sim \chi^2_{(m^{(1)}-1) \cdot (m^{(2)}-1)}$.

Beweis. Siehe Krengel (1998, §13). □

Wähle $c = \chi^2_{(m^{(1)}-1) \cdot (m^{(2)}-1); 1-\alpha}$ als $(1 - \alpha)$ -Quantil der χ^2 -Verteilung mit $(m^{(1)} - 1) \cdot (m^{(2)} - 1)$ Freiheitsgraden und definiere

$$R_n = \{Q_n \circ H_n \geq c\}$$

70. Satz χ^2 -Unabhängigkeitstest

Unter obigen Verteilungsannahmen definiert der Verwerfungsbereich R_n einen Signifikanztest zum asymptotischen Niveau α , d.h.

$$\forall \mathbf{p} \in \Theta_0 : \lim_{n \rightarrow \infty} P^{\mathbf{p}}(\{X \in R_n\}) = \alpha.$$

Beweis. Wie für Satz 64.

□

71. Beispiel Fortsetzung von Bsp. 68 mit $\alpha = 0,1$.

Es folgt $c = 21,06$, und ferner gilt $Q(H(x)) = 14,40$, so daß Θ_0 nicht verworfen wird.