2 Schätztheorie

Notation und Terminologie:

- $\mathfrak{L}_1^\vartheta = \mathfrak{L}_1(\Omega, \mathfrak{A}, P^\vartheta)$ und $\mathfrak{L}_2^\vartheta = \mathfrak{L}_2(\Omega, \mathfrak{A}, P^\vartheta)$
- $X = (X_1, \ldots, X_n)$ und $x = (x_1, \ldots, x_n)$
- $g_n(X)$ Schätzvariable zu Schätzfunktion g_n
- $\mathrm{E}^\vartheta, \mathrm{Var}^\vartheta$ 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
- $\overline{x}_n := \dfrac{1}{n} \cdot \sum_{i=1}^{n} x_i$ und $\overline{X}_n := \dfrac{1}{n} \cdot \sum_{i=1}^{n} X_i = S_n/n$

5. Definition Schätzfunktion $g_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ erwartungstreu für γ, falls

$$\forall \vartheta \in \Theta : \mathrm{E}^{\vartheta}(g_n(X)) = \gamma(\vartheta)$$

Folge von Schätzfunktionen $g_n : \mathbb{R}^n \longrightarrow \mathbb{R}$

(i) schwach konsistent für γ, falls

$$\forall \vartheta \in \Theta \forall \varepsilon > 0 \colon \lim_{n \to \infty} P^{\vartheta}(\{|g_n(X) - \gamma(\vartheta)| > \varepsilon\}) = 0$$

(ii) stark konsistent für γ, falls

$$\forall \vartheta \in \Theta : \lim_{n \to \infty} g_n(X) = \gamma(\vartheta) \qquad P^{\vartheta}\text{-f.s.}$$

6. Bemerkung Starke Konsistenz impliziert schwache Konsistenz, siehe Satz VI.38.

Problem: Schätzung des Erwartungswertes, siehe Beispiele 3 und 4.

7. Satz Gelte $X_1 \in \mathfrak{L}_1^\vartheta$ für alle $\vartheta \in \Theta$, und sei

$$\gamma(\vartheta) := \mathrm{E}^\vartheta(X_1).$$

Dann definiert das empirische Mittel $g_n(x) := \overline{x}_n$ eine stark konsistente Folge von erwartungstreuen Schätzern für γ.

Beweis. Für jedes $\vartheta \in \Theta$ gilt

$$\mathrm{E}^{\vartheta}(g_n(X)) = \mathrm{E}^{\vartheta}(\overline{X}_n) = \frac{1}{n} \cdot \sum_{i=1}^{n} \mathrm{E}^{\vartheta}(X_i) = \gamma(\vartheta).$$

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

$$P^p(\{|\overline{X}_n - p| \geq 10^{-2}\}) \leq 9.9 \cdot 10^{-5}$$

$$P^p(\{|\overline{X}_n - p| \geq 10^{-3}\}) \leq 9.9 \cdot 10^{-3}$$

und aufgrund der Hoeffdingsche Ungleichung, siehe Bsp. VI.34

$$P^p(\{|\overline{X}_n - p| \geq 10^{-2}\}) \leq 9.2 \cdot 10^{-2187}$$

$$P^p(\{|\overline{X}_n - p| \geq 10^{-3}\}) \leq 2.7 \cdot 10^{-22}.$$

Problem: Schätzung der Varianz.

9. Satz Gelte $X_1 \in \mathfrak{L}_2^{\vartheta}$ für alle $\vartheta \in \Theta$, und sei

$$\gamma(\vartheta) = \mathrm{Var}^{\vartheta}(X_1).$$

Ferner sei $n > 1$. Dann definiert die empirische Varianz

$$v_n(x) := \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \overline{x}_n)^2$$

eine stark konsistente Folge von erwartungstreuen Schätzern für γ.

Beweis. Setze $\mu(\vartheta) := \mathrm{E}^\vartheta(X_1)$. Es gilt

$$
\begin{aligned}
v_n(x) &= \frac{1}{n-1} \cdot \sum_{i=1}^n ((x_i - \mu(\vartheta)) + (\mu(\vartheta) - \overline{x}_n))^2 \\
&= \frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \mu(\vartheta))^2 - \frac{n}{n-1} \cdot (\mu(\vartheta) - \overline{x}_n)^2
\end{aligned}
$$

Mit Satz VI.26 folgt

$$
\begin{aligned}
\mathrm{E}^\vartheta(v_n(X)) &= \frac{n}{n-1} \cdot \gamma(\vartheta) - \frac{n}{n-1} \cdot \mathrm{Var}^\vartheta\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) \\
&= \frac{n}{n-1} \cdot \gamma(\vartheta) - \frac{1}{n-1} \cdot \gamma(\vartheta) = \gamma(\vartheta).
\end{aligned}
$$

Nach dem starken Gesetz der großen Zahlen gilt P^ϑ-f.s.

$$\lim_{n\to\infty} \frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - \mu(\vartheta))^2 = \mathrm{E}^\vartheta (X_1 - \mu(\vartheta))^2 = \mathrm{Var}^\vartheta (X_1)$$

und

$$\lim_{n\to\infty} \frac{1}{n} \cdot \sum_{i=1}^{n} X_i = \mu(\vartheta).$$

Es folgt P^ϑ-f.s.

$$\lim_{n\to\infty} v_n(X) = \gamma(\vartheta).$$

$\square$

10. Bemerkung Für die empirische Varianz gilt

$$\frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \overline{x}_n)^2 = \frac{1}{n-1} \cdot \left(\sum_{i=1}^{n} x_i^2 - n \cdot \overline{x}_n^2 \right) .$$

Beweis. Erster Teil des Beweises von Satz 9 mit $\mu(\vartheta) := 0$. □

11. Beispiel In Beispiel 1 bzw. 3 beträgt die empirische Varianz $0.2498\ldots$.

12. Beispiel Sei $\Theta := [0,1]$ und für $p \in \Theta$ gelte $P^p_{X_1} = \mathbf{B}(1,p)$. Ferner sei

$$\gamma(p) = \sqrt{p \cdot (1-p)}$$

die entsprechende Standardabweichung. Für jede Schätzfunktion g_1 gilt

$$\mathrm{E}^p(g_1(X)) = p \cdot g_1(1) + (1-p) \cdot g_1(0).$$

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

$$R^{\vartheta}(g_n) := \mathrm{E}^{\vartheta}(g_n(X) - \gamma(\vartheta))^2,$$

falls $g_n(X) \in \mathfrak{L}_2^{\vartheta}$.

14. Definition Bias einer Schätzfunktion g_n für γ

$$B^{\vartheta}(g_n) := \mathrm{E}^{\vartheta}(g_n(X)) - \gamma(\vartheta),$$

falls $g_n(X) \in \mathfrak{L}_1^{\vartheta}$.

15. Bemerkung

(i) Erwartungstreue äquivalent zu

$$\forall \vartheta \in \Theta : B^{\vartheta}(g_n) = 0.$$

(ii) Es gilt

$$R^{\vartheta}(g_n) = \mathrm{Var}^{\vartheta}(g_n(X)) + (B^{\vartheta}(g_n))^2.$$

Zu Frage (ii): Optimalität exemplarisch für Problemstellung aus Bsp. 1 bzw. 3.

Bez. $g_n^*(x) := \overline{x}_n$.

16. Beispiel Für alle $p \in \left]0, 1\right[$ gilt

$$B^p(g_n^*) = 0$$

und

$$R^p(g_n^*) = \mathrm{E}^p \left(\overline{X}_n - p\right)^2 = \frac{p \cdot (1-p)}{n}.$$

Für $p \in \,]0,1[$ und $x \in D := \{0,1\}^n$ sei (Likelihood-Funktion)

$$L_x(p) := P^p(\{X = x\}) = p^{k(x)} \cdot (1-p)^{n-k(x)}$$

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 $g_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ für γ erfüllt

$$\forall p \in \,]0,1[\,:\; R^p(g_n) \geq 1/\,\mathrm{E}^p\left((\ell'_X(p))^2\right).$$

Beweis. Für g_n wie oben gilt

$$p = \mathrm{E}^p(g_n(X)) = \sum_{x \in D} g_n(x) \cdot L_x(p)$$

und somit

$$1 = \sum_{x \in D} g_n(x) \cdot L'_x(p) = \sum_{x \in D} g_n(x) \cdot \ell'_x(p) \cdot L_x(p) = \mathrm{E}^p(g_n(X) \cdot \ell'_X(p)).$$

Aus $\sum_{x \in D} L_x(p) = 1$ folgt

$$0 = \sum_{x \in D} L'_x(p) = \sum_{x \in D} \ell'_x(p) \cdot L_x(p) = \mathrm{E}^p(\ell'_X(p)).$$

Also liefert die Cauchy-Schwarzsche Ungleichung

$$1 = \Big(\mathrm{E}^p((g_n(X) - \gamma(p)) \cdot \ell'_X(p))\Big)^2 \le \mathrm{Var}^p(g_n(X)) \cdot \mathrm{E}^p((\ell'_X(p))^2).$$

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

$$\forall p \in \,]0,1[\, : \; R^p(g_n^*)$$
$$= \inf \left\{ R^p(g_n) : g_n \text{ e'treue Schätzfunktion für } \gamma \right\} .$$

Beweis. Man verifiziert

$$\mathrm{E}^p((\ell'_X(p))^2) = \frac{n}{p \cdot (1-p)},$$

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

$$\gamma(\vartheta) := \mathrm{E}^{\vartheta}(X_1) = \vartheta.$$

Setze

$$g_3(x_1, x_2, x_3) := \frac{\max(x_1, x_2, x_3) + \min(x_1, x_2, x_3)}{2}$$

<u>ÜBUNG</u>: 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 $\ell_n, r_n : \mathbb{R}^n \longrightarrow \mathbb{R}$ 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

$$\gamma(\vartheta) := \mathrm{E}^{\vartheta}(X_1)$$

und Intervalle der Form

$$[\overline{x}_n - b_n(x), \overline{x}_n + b_n(x)]$$

mit Borel-meßbaren Abbildungen

$$b_n : \mathbb{R}^n \to [0, \infty[\,.$$

22. Bemerkung b_n definiert genau dann Konfidenzintervall für Erwartungswert zum Niveau $1-\alpha$, wenn

$$\forall \vartheta \in \Theta : \; P^{\vartheta}(\{|\overline{X}_n - \mathrm{E}^{\vartheta}(X_1)| \leq b_n(X)\}) \geq 1-\alpha.$$

Also Tschebyschev-Ungleichung anwendbar, falls (Schranke für) $\sup_{\vartheta \in \Theta} \mathrm{Var}^{\vartheta}(X_1)$ 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 $(\mu, \sigma) \in \Theta$.

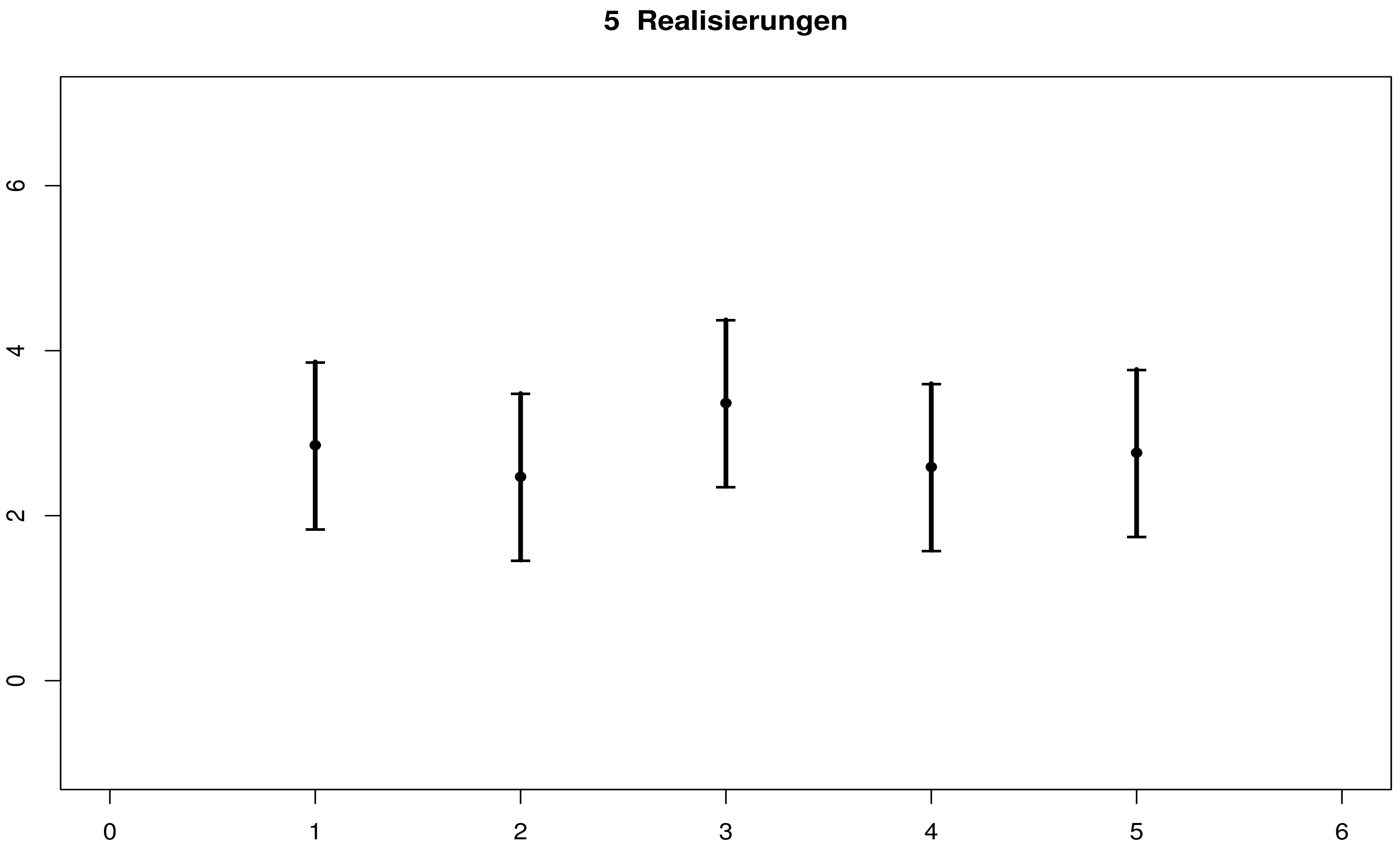
24. Satz Bei bekannter Varianz σ^2 definiert

$$b_n := \sigma/\sqrt{n} \cdot \Phi^{-1}(1 - \alpha/2)$$

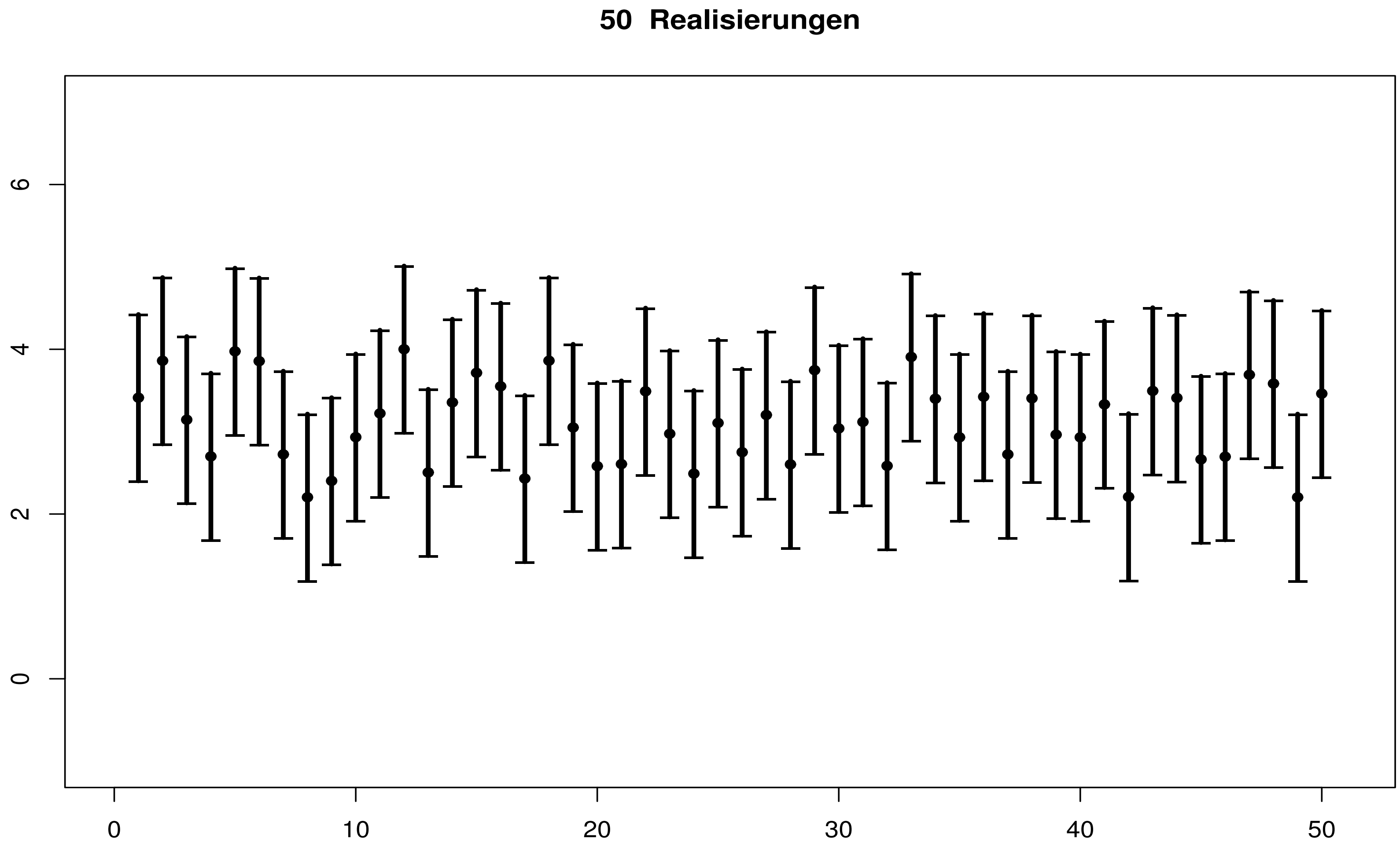
ein Konfidenzintervall für den Erwartungswert zum Niveau $1 - \alpha$.

Beweis. PROJEKTOR $\square$

25. Beispiel Konfidenzintervalle für $\alpha := 0,05$, $\sigma := 2$, $n := 15$ (und $\mu := 3$).



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



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 $X'_1, \ldots, X'_n$ iid mit $X'_1 \sim \mathbf{N}(0, 1)$.

Setze

$$\overline{X'}_n := \sum_{i=1}^{n} X'_i/n,$$

$$X' := (X'_1, \ldots, X'_n)$$

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

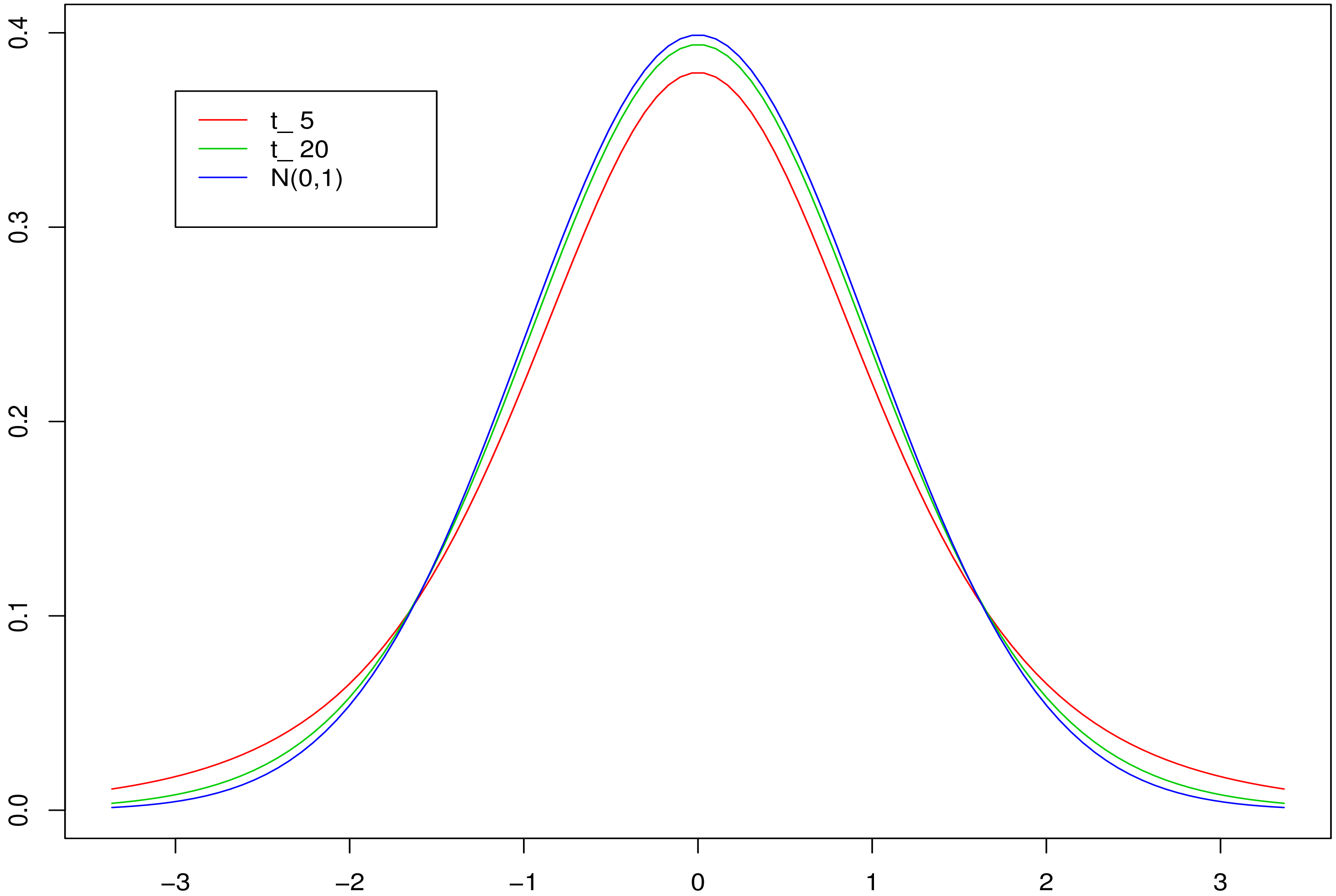
$$\lim_{n\to\infty} f_n(x) = 1/\sqrt{2\pi} \cdot \exp(-x^2/2).$$

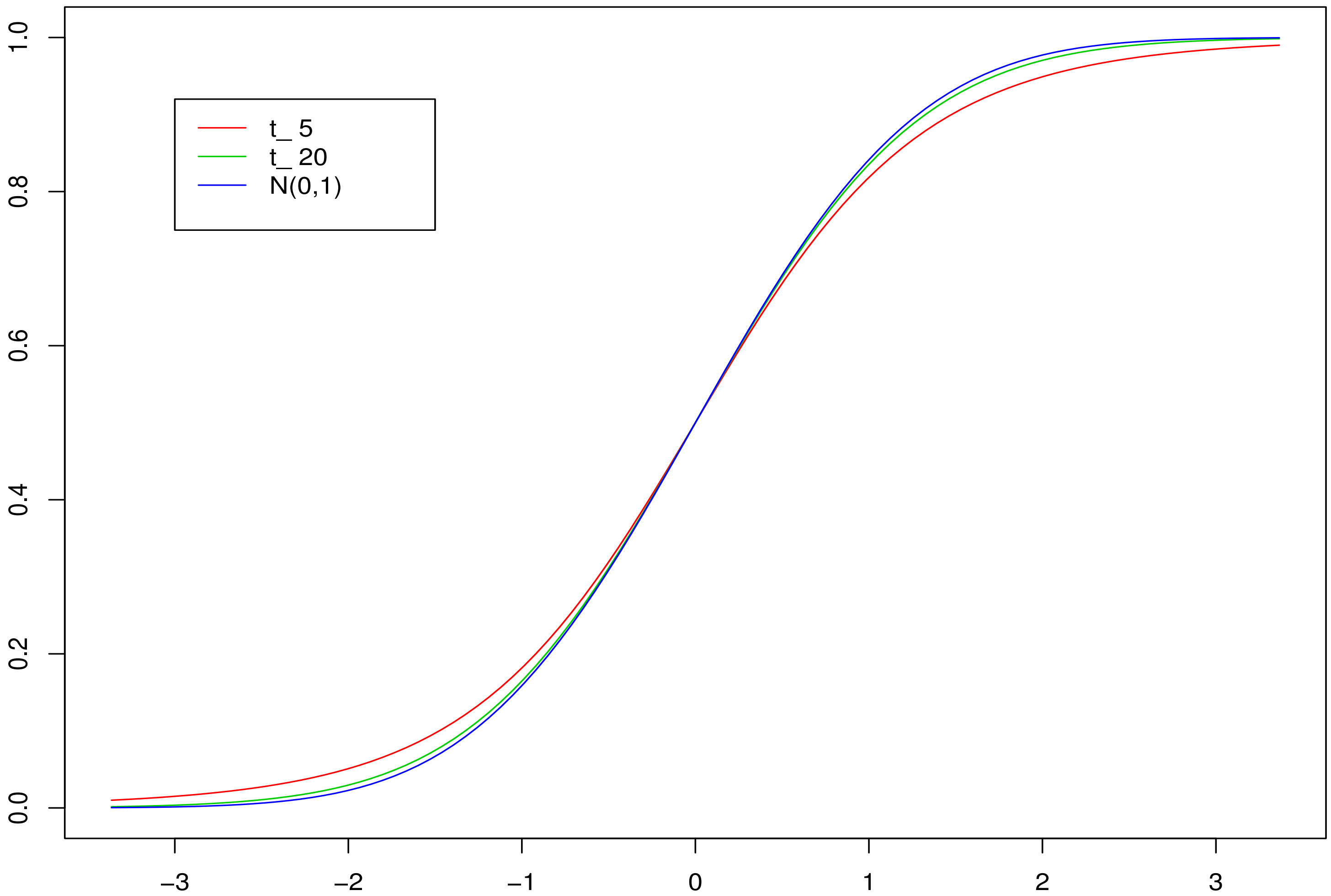
28. Definition Die Verteilung der ZV T_n' heißt t-Verteilung mit $n-1$ Freiheitsgraden.

Bez.: t_{n-1}.

29. Beispiel Dichte und Verteilungsfunktion von $\mathbf{t}_5$ und $\mathbf{t}_{20}$.

Zum Vergleich Dichte und Verteilungsfunktion von $\mathbf{N}(0,1)$.




t_ 5
t_ 20
N(0,1)
0.0
0.2
0.4
0.6
0.8
1.0
−3
−2
−1
0
1
2
3

30. Bemerkung Zur Berechnung der Verteilungsfunktion von $\mathbf{t}_n$ und entsprechender Quantile: Numerik, Tabellen, Plots.

31. Lemma Bzgl. $P^{(\mu,\sigma)}$ gilt

$$\frac{\overline{X}_n - \mu}{\sqrt{v_n(X)/n}} \sim \mathbf{t}_{n-1}.$$

Beweis. Setze $X_i' := \dfrac{X_i - \mu}{\sigma}$. Bzgl. $P^{(\mu,\sigma)}$ gilt $X_1', \ldots, X_n'$ iid und $X_1' \sim \mathbf{N}(0,1)$. Ferner

$$\overline{X_n'} = \sum_{i=1}^{n} \frac{X_i'}{n} = \frac{\overline{X_n} - \mu}{\sigma}, \qquad X_i' - \overline{X_n'} = \frac{X_i - \overline{X_n}}{\sigma},$$

$$(n-1) \cdot v_n(X') = \sum_{i=1}^{n} (X_i' - \overline{X_n'})^2 = \frac{1}{\sigma^2} \cdot \sum_{i=1}^{n} (X_i - \overline{X_n})^2 = \frac{(n-1) \cdot v_n(}{\sigma^2}$$

Fazit

$$\frac{\overline{X_n'}}{\sqrt{v_n(X')}/\sqrt{n}} = \frac{\overline{X_n} - \mu}{\sigma} \cdot \frac{\sigma}{\sqrt{v_n(X)}/\sqrt{n}} = \frac{\overline{X_n} - \mu}{\sqrt{v_n(X)}/\sqrt{n}}.$$

$\square$

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
$t_{n-1;1-\alpha/2}$	$2,09\ldots$	$2,01\ldots$	$1,98\ldots$

Zum Vergleich: $\Phi^{-1}(1-\alpha/2) = 1,96\ldots$.

Beweis von Satz 32. Vgl. Beweis von Satz 24.

Sei $F := F_{Z_n}$ die Verteilungsfunktion von

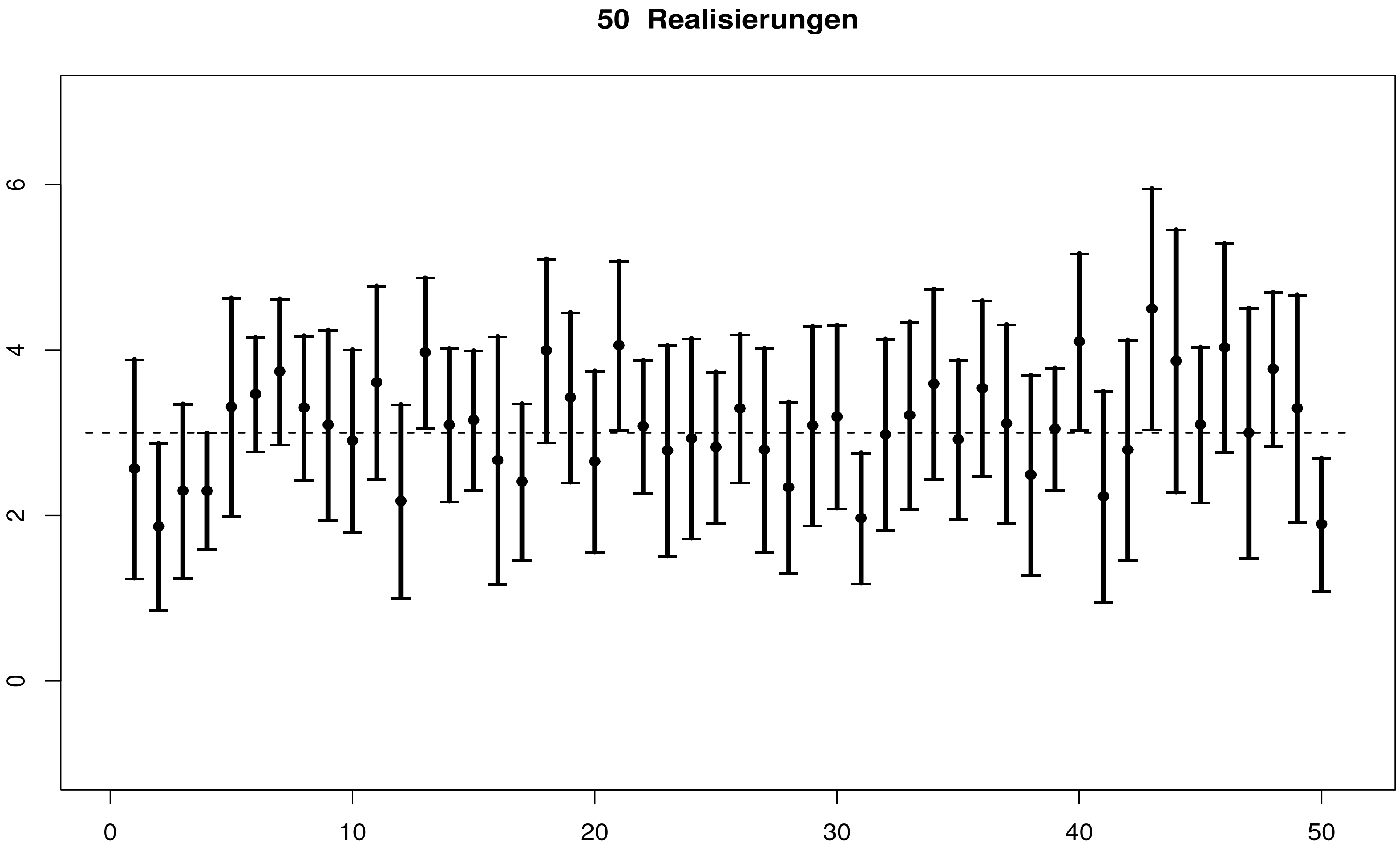
$$Z_n := \frac{\overline{X}_n - \mu}{\sqrt{v_n(X)/n}}.$$

Für jedes $\mu \in \mathbb{R}$ gilt gem. Lemma 31

$$\begin{aligned} P^\mu(\{|\overline{X_n} - \mu| \leq b_n(X)\}) &= P^\mu(\{|Z_n| \leq t_{n-1;1-\alpha/2}\}) \\ &= F(t_{n-1;1-\alpha/2}) - F(-t_{n-1;1-\alpha/2}) \\ &= 2 \cdot F(t_{n-1;1-\alpha/2}) - 1 \\ &= 2(1-\alpha/2) - 1 = 1 - \alpha. \end{aligned}$$

$\square$

34. Beispiel Konfidenzintervalle für $\alpha := 0,05$ $n := 15$ (und $\mu := 3$, $\sigma := 2$).



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 , 1.7 , empir. Mittel: 2.6

Ausblick: asymptotische Konfidenzintervalle.

Gelte

$$\forall \vartheta \in \Theta : X_1 \in \mathfrak{L}_2^\vartheta \wedge \mathrm{Var}^\vartheta(X_1) > 0.$$

Setze

$$b_n(x) := \sqrt{v_n(x)/n} \cdot \Phi^{-1}(1 - \alpha/2).$$

35. Satz Für jedes $\alpha \in \,]0, 1[$

$$\forall \vartheta \in \Theta : \lim_{n \to \infty} P^\vartheta(\{|\overline{X}_n - \mathrm{E}^\vartheta(X_1)| \le b_n(X)\}) = 1 - \alpha.$$

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
$0,05$	$4,5 \cdot 10^{-4}$	$2,8 \cdot 10^{-4}$	$1,95 \cdot 10^{-4}$
$0,01$	$1,0 \cdot 10^{-3}$	$3,3 \cdot 10^{-4}$	$2,57 \cdot 10^{-4}$
$0,001$	$3,2 \cdot 10^{-3}$	$3,8 \cdot 10^{-4}$	$3,27 \cdot 10^{-4}$