II. Die Abschätzung der maximalen relativen Häufigkeit einer Häufigkeitsverteilung

1. Einführung

Die Ungleichung (17) vom I.Kapitel wirft die Frage auf, wie man unter gewissen Annahmen die maximale relative Häufigkeit $\frac{\lambda_N}{N}$ einer Häufigkeitsverteilung nach oben abschätzen kann

Mit der Abschätzung von $\frac{\lambda_N}{N}$ beschäftigt sich das vorliegende Kapitel, die Notationen des I.Kapitels werden weitestgehend beibehalten.

Es wird vorausgesetzt, dass es eine unendlich große „superpopulation" Ω gibt, auf welcher eine Zufallsvariable Z definiert ist, die eine Verteilung $\mathcal{F}$ und eine Verteilungsfunktion F besitzt.

Betrachtet werden N Werte von Z und deren relative Häufigkeitsverteilung. Die N Werte $z_1, \ldots, z_N$ werden aufgefaßt als Realisationen der ersten N Glieder einer Folge von unabhängigen, identisch gemäß $\mathcal{F}$ verteilten Zufallsgrößen, $\mathcal{X} := (Z_k)_{k \in \mathbb{N}}$ mit $Z_k : \Omega \to \mathbb{R}$ $(k \in \mathbb{N})$.

Damit sind auch die unter den $z_1, \ldots, z_N$ paarweise verschiedenen Werte $a_1, \ldots, a_m$, deren absolute bzw. relative Häufigkeiten (insbesondere λ_N, die maximale absolute Häufigkeit) sowie die Anzahl m Realisationen von Zufallsgrößen $(1 \leq m \leq N)$.

Mit $Y_1, \ldots, Y_m$ werden die zufälligen Häufigkeits-Variablen bezeichnet, die die Anzahlen angeben, mit denen die paarweise verschiedenen unter den Zufallsgrößen $Z_1, \ldots, Z_N$ auftreten. $Y_{m,1}, \ldots, Y_{m,m}$ sind die der Größe nach geordneten Häufigkeits-Variablen (sogenannte „order statistics"). Dann ist λ_N eine Realisation der maximalen „order statistic" $Y_{m,m}$. Das $(1-\alpha)$ - Quantil $y_{1-\alpha}$ der Verteilung der Größe $\frac{1}{N} \cdot Y_{m,m}$ ist mit einem Vertrauen von mindestens $(1-\alpha) \cdot 100\%$ eine obere Schranke von $\frac{\lambda_N}{N}$. Nach den Recherchen des Autors gibt es bisher keine Veröffentlichungen über die Verteilung von $\frac{1}{N} \cdot Y_{m,m}$. Die klassische Methode bestünde darin, ersatzweise eine Beziehung mit einem $(1-\alpha)$ - Quantil der Grenzverteilung von $\frac{1}{N} \cdot Y_{m,m}$ herzustellen - falls eine solche Grenzverteilung existiert und bekannt ist.

In der Literatur ist die Verteilung sogenannter „extreme statistics" eingehend untersucht worden.

Entweder wird vorausgesetzt, dass die Zufallsgrößen $Y_1 , \ldots , Y_m$ unabhängig sind (M.L.Juncosa 1949), oder dass sie identisch verteilt sind („exchangeable variates“ bei S.M. Berman 1962, J.Galambos 1982), oder dass sogar beide Voraussetzungen gelten (C.W.Anderson 1970 und 1980 u.a.).
Bekanntermaßen ist jede der Zufallsgrößen $Y_1 , \ldots , Y_m$ binomialverteilt - aber jede ist binomialverteilt zu einem speziellen Parameter p_i ($i \in \{ 1 , \ldots , m \}$). Außerdem sind die Größen $Y_1 , \ldots , Y_m$ nicht unabhängig, da für jede ihrer Realisationen $N_1 , \ldots , N_m$ stets $\sum_{k=1}^{m} N_k = N$ gilt.
Die $Y_1 , \ldots , Y_m$ formen weder einen stationären Prozeß, noch eine „nonstationary normal sequence“ (M.R.Leadbetter / G.Lindgren / H.Rootzen 1983), schließlich besitzen sie auch nicht die von T.L.Lai / H.Robbins (1978) angegebene „maximally dependence“ - Eigenschaft.

Die dem Autor zugängliche Theorie der „extreme statistics“ konnte auf die Abschätzung von $\frac{\lambda_N}{N}$ nicht angewendet werden.
In der vorliegenden Arbeit wird eine Methode vorgestellt, die verschiedene Schätz-Versionen zuläßt und stets eine Aussage über die Schätz-Genauigkeit liefert.
Zunächst wird $\frac{\lambda_N}{N}$ als Supremum einer gewissen Treppenfunktion f_N mit der Schrittweite $2 \cdot h_N$ dargestellt. Die anschließende Abschätzung von $\frac{\lambda_N}{N}$ richtet sich danach, welch Voraussetzungen über F , h_N sowie die Höhe von $N \in \mathbb{N}$ jeweils zugrundeliegen.
Bei der Abschätzung von $\frac{\lambda_N}{N}$ spielt die Kolmogorov / Smirnov - Distance $\mathbb{D}_N$ eine entscheidende Rolle, für die wiederum zwei Hauptversionen der Abschätzung zur Verfügung stehen:
(i) Die fast sichere Abschätzung für große N nach dem Gesetz vom „Iterierten Logarithmus“ von Smirnov (1944).
(ii) Die Abschätzung mit einer gewissen Mindest-Wahrscheinlichkeit für jedes N nach der Ungleichung von Dvoretzki / Kiefer / Wolfowitz (1956).
Je nach Anwendung von (i) oder (ii) erhält man die entsprechend gültige Version der Abschätzung von $\frac{\lambda_N}{N}$.

Insbesondere für die $\frac{\lambda_N}{N}$ - Abschätzung nach unten ist von Bedeutung, ob die Schrittweite $2 \cdot h_N$ unterhalb der Größe $b_N := \min_{z_i \neq z_j} \left\{ \left| z_i - z_j \right| \right\}$ liegt. Kann man davon nicht ausgehen, bleibt noch die stets gültige Untergrenze $\frac{1}{N}$ (im Falle $m < N$ sogar $\frac{2}{N}$).

Für die angegebenen $\frac{\lambda_N}{N}$ - Abschätzungen nach oben ist $2 \cdot h_N < b_N$ keine notwendige Bedingung; das Nicht-Bestehen dieser Ungleichung hat aber eventuell Auswirkungen auf die Schätzqualität. In diesem Sinne gilt auch die Anwendung einer $\frac{\lambda_N}{N}$ - Schätzversion auf die Ungleichung (17) des I.Kapitels.

Es stellt sich die Frage, wie sich $\frac{\lambda_N}{N}$ verhält, wenn N wächst.

Die Sätze II.2 / 3 stellen einen Zusammenhang her zwischen der maximalen Sprunghöhe der Empirischen Verteilungsfunktion F_N ($= \frac{\lambda_N}{N}$) und der maximalen Sprunghöhe der Grenzverteilungsfunktion F ($=: j(F)$) . Es wird gezeigt, dass $\frac{\lambda_N}{N}$ unter gewissen Bedingungen gegen $j(F)$ strebt.

Im letzten Abschnitt des II.Kapitels wird eine Daten-gestütze Abschätzung von $\frac{\lambda_N}{N}$ vorgestellt, die für den Fall anwendbar ist, dass über F nichts bekannt ist.

2. Die Darstellung der maximalen relativen Häufigkeit $\frac{\lambda_N}{N}$

Es sind $z_1, \ldots, z_N \in \mathbb{R}$ die N Werte einer Zufallsgröße, von denen $a_1, \ldots, a_m$ paarweise verschieden sind $(1 \leq m \leq N$ mit $m, N \in \mathbb{N})$.

Man hat die folgende Häufigkeitstabelle:

Paarweise Verschiedene Werte	a_1	. . .	a_m
Absolute Häufigkeiten	N_1	. . .	N_m

Es gilt:

$N_1 + \ldots + N_m = N$,

$\lambda_N := \max \{ N_1, \ldots, N_m \}$.

Bei der Entwicklung der Abschätzung (17) vom I.Kapitel war selbst nominale Meßbarkeit für das betrachtete Merkmal zugelassen. Unter derartig allgemein gehaltenen Voraussetzungen kann man nicht erwarten, dass für die Unbestimmte $\frac{\lambda_N}{N}$ eine besonders „scharfe" Abschätzung möglich ist.

Setzt man in $B(n, N, \frac{\lambda_N}{N})$ aus (16) vom I.Kapitel für $\frac{\lambda_N}{N}$ die „triviale" obere Schranke $\frac{N-1}{N}$ ein, erhält man wegen Satz I.1 eine nur noch von n und N abhängende obere Schranke für $D(n, N; \vec{N}, \vec{n})$. Unter bestimmten Voraussetzungen kann eine Verbesserung der Abschätzung von $\frac{\lambda_N}{N}$ erreicht werden.

Für zwei Sonderfälle erübrigt sich eine Abschätzung von $\frac{\lambda_N}{N}$:

(i) Im Falle $m = 1$ sind alle N Werte identisch, d.h. $z_1 = \ldots = z_N$.

Es folgt $\frac{\lambda_N}{N} = 1$.

(ii) Im Falle $m = N$ sind alle N Werte paarweise verschieden, d.h. $z_i \neq z_j$ für alle $i \neq j$.

Es folgt $\frac{\lambda_N}{N} = \frac{1}{N}$.

2.1 Die Verteilung $\mathcal{F}$

Die Größen X , N , Ω_N und $z_1 , \dots , z_N$ aus dem I.Kapitel werden jetzt in ein erweitertes Konzept eingebettet.

Mit der Potenzmenge $\wp(\Omega_N)$ und der relativen Häufigkeit H_N ist $(\Omega_N , \wp(\Omega_N) , H_N)$ ein diskreter Wahrscheinlichkeitsraum.

Sei Ω eine Menge mit $\Omega_N \subset \Omega$ und $(\Omega , A , \mathcal{P})$ ein Wahrscheinlichkeitsraum.

Mit Z wird eine Zufallsgröße bezeichnet, $Z : \Omega \to \mathbb{R}$ mit der Verteilung $\mathcal{P}Z^{-1} =: \mathcal{F}$.

Die Größe X wird aufgefaßt als Einschränkung der zufälligen Abbildung Z , $Z|_{\Omega_N} = X$.

Gegeben ist eine Folge $\mathcal{X} := (Z_i)_{i\in\mathbb{N}}$ von unabhängigen, identisch gemäß $\mathcal{F}$ verteilten reellwertigen Zufallsgrößen, $Z_k : \Omega \to \mathbb{R}$, $\mathcal{P}(Z_k)^{-1} = \mathcal{F}$ für jedes $k \in \mathbb{N}$.

Sei $\mathrm{IP} = \bigotimes_{\mathbb{N}} \mathcal{F}$ das unendliche Produktmaß, definiert auf der σ-Algebra der Borelmengen des $\mathbb{R}^{\mathbb{N}}$ (H.Bauer 1968, §33; hier mit der speziellen Indexmenge $I = \mathbb{N}$) .

Dann ist $\mathrm{IP}\mathcal{X}^{-1}$ ein Wahrscheinlichkeitsmaß auf der σ-Algebra der Borelmengen des $\mathbb{R}^{\mathbb{N}}$.

Nullmengen bezüglich dieses Maßes spielen noch eine wichtige Rolle.

Basiert die Unabhängigkeit der $Z_1 , \dots , Z_N$ auf dem Ziehen mit Zurücklegen aus Ω , dann muß R-mal gezogen werden mit $N \le R$, um N paarweise verschiedene Objekte zu erhalten, welche die Menge Ω_N bilden. In diesem Fall ist N eine Zufallsgröße.

Der Fall eines zufallsabhängigen N wird in der vorliegenden Arbeit nicht untersucht.

Zieht man genau N-mal mit Zurücklegen aus Ω , erhält man S paarweise verschiedene Objekte, und damit eine Menge Ω_S , die wegen $S \le N$ im allgemeinen zu klein ist.

Daher kommt für die Stichprobenvariablen in den hier betrachteten Fällen ein Ziehen mit Zurücklegen nicht in Betracht, es wird ein Ω mit unendlich-vielen Objekten vorausgesetzt.

Sei $\mathcal{F}_\nu^{\mathcal{X}}$ das auf $Z_1 , \dots , Z_\nu$ basierende Empirische Maß zum Stichprobenumfang $\nu \in \mathbb{N}$ (F.Strobl 1995, S.78).

Betrachtet wird die von rechts stetige Version der Empirischen Verteilungsfunktion IF_ν zum Stichprobenumfang $\nu \in \mathbb{N}$ mit unabhängigen und identisch gemäß $\mathcal{F}$ verteilten Zufallsgrößen $Z_1 , \dots , Z_\nu$:

$$\mathrm{IF}_\nu(t) := \mathcal{F}_\nu^{\mathcal{X}}\big((-\infty \, ; \, t] \big) := \frac{1}{\nu} \cdot \sum_{i=1}^{\nu} 1_{(-\infty\,;\,t]}(Z_i) \quad , \quad t\in\mathbb{R} \; . \tag{1}$$

Eine Realisation von IF_ν entsteht, wenn man in $\mathcal{F}_\nu^{\mathcal{X}}$ die Folge $\mathcal{X} = (Z_i)_{i\in\mathbb{N}}$ ersetzt durch eine Zahlenfolge $z := (z_i)_{i\in N} \in \mathbb{R}^{\mathbb{N}}$:

$$F_\nu(t) := \mathcal{F}_\nu^{z}\big(\,]-\infty\ ;\ t\,]\big) = \frac{1}{\nu}\cdot\sum_{i=1}^{\nu} 1_{]-\infty\,;\,t\,]}(z_i) \quad , \quad t\in\mathbb{R}\ . \tag{2}$$

$\mathcal{F}_\nu^{z}$ wird bezeichnet als das zu $z_1, \ldots, z_\nu$ gehörige Empirische Maß.

Mit den gegebenen Werten $z_1, \ldots, z_N$ ist F_N also genau die Verteilungsfunktion der X-Verteilung $\mathcal{F}(N)$ im I.Kapitel. Diese Werte spielen jetzt praktisch die Rolle der Realisation einer Stichprobe vom Umfang $N \in \mathbb{N}$.

Mit F wird die Verteilungsfunktion von $\mathcal{F}$ bezeichnet, genauer: $\mathcal{F}(\,(-oo\ ;\ t\,]\,) =: F(t)$, $t \in \mathbb{R}$.
Nach dem Satz von Glivenko / Cantelli (R.J.Serfling 1980, S. 61) konvergiert IF_N $\mathrm{IP}\mathcal{X}^{-1}$ - fast sicher gleichmäßig gegen F .

2.2 Die relativen Häufigkeiten

Mit den Realisationen $z_1, \ldots, z_N$ der Zufallsgrößen $Z_1, \ldots, Z_N$ wird eine reellwertige Treppenfunktion konstruiert, die Realisation einer Kern-Dichte-Schätzfunktion ist.
Durch die Formulierung einer milden Zusatz-Bedingung für den minimalen Abstand der paarweise verschiedenen Werte $a_1, \ldots, a_m$ wird erreicht, dass in jedem Teilintervall, in dem die Treppenfunktion constant ist, höchstens ein a_i ($i \in \{1, \ldots, m\}$) liegt.
Dann ist das Maximum dieser Treppenfunktion genau die maximale relative Häufigkeit $\frac{\lambda_N}{N}$.
Konstruiert wird eine Abbildung φ , $\varphi : \mathbb{R} \times \mathbb{R}^{\mathbb{N}} \to \mathbb{R}$, $\varphi(t\,;Z_1, \ldots, Z_N) =: \mathrm{If}_N(t)$.
If_N ist zu jedem festen $t \in \mathbb{R}$ eine Zufallsgröße.
Setzt man die Werte $z_1, \ldots, z_N$ für die $Z_1, \ldots, Z_N$ ein, entsteht eine Realisation f_N von If_N ,
$f_N(t) := \varphi(t\,;z_1, \ldots, z_N)$. Der Graph von f_N liefert ein spezielles Histogramm.
Es werden Teilintervalle der Schrittweite $2\cdot h_N$ definiert:
$J_i(h_N) := [\,a_i - h_N\ ;\ a_i + h_N\,)$, $i \in \{1, \ldots, m\}$.

$$\text{Sei } b_N := \min_{k\neq r}\big\{\,|\,a_k - a_r\,|\,\big\} = \min_{z_i\neq z_j}\big\{\,|\,z_i - z_j\,|\,\big\} \ . \tag{3}$$

Da nach Voraussetzung mindestens zwei der Werte $z_1, \ldots, z_N$ verschieden sind $(2 \le m)$, folgt $0 < b_N$.

Die Schrittweite $2 \cdot h_N$ wird so gewählt, dass zu jedem $i \in \{1, \ldots, m\}$ in das Teilintervall $J_i(h_N)$ genau einer der Werte $a_1, \ldots, a_m$ fällt. Hinreichend dafür ist, dass $0 < h_N < b_N$ gilt.

Im Satz II.1 wird sogar $0 < h_N < \frac{1}{2} \cdot b_N$ vorausgesetzt.

Auf jedem Intervall $J_i(h_N)$ ($i \in \{1, \ldots, m\}$) wird $f_N \equiv c_{i,N}$ definiert derart, dass der Flächeninhalt des Rechtecks mit der Grundseite $J_i(h_N)$ und der Höhe $c_{i,N}$ genau der relativen Häufigkeit des in $J_i(h_N)$ liegenden Wertes $a_i \in \{a_1, \ldots, a_m\}$ entspricht. Ist $t \notin J_i(h_N)$ für alle $i \in \{1, \ldots, m\}$, wird $f_N(t) = 0$ definiert.

Es gilt also konstruktionsgemäß:

$$\frac{\lambda_N}{N} = 2 \cdot h_N \cdot \sup_{t \in \mathbb{R}} f_N(t) \tag{4}$$

Satz II.1

Voraussetzung: $0 < 2 \cdot h_N < b_N$, $1 < m \le N$, $N \in \mathbb{N}$.

Es wird definiert: $J_i(h_N) := [a_i - h_N; a_i + h_N)$, $i \in \{1, \ldots, m\}$.

(i) Für beliebiges $x \in J_i(h_N)$ ist die Differenz $\left[F_N(x+h_N) - F_N(x-h_N)\right]$ genau die relative Häufigkeit $\frac{N_i}{N}$ von a_i in der Häufigkeits-Verteilung $\mathcal{F}(N)$, $i \in \{1, \ldots, m\}$.

(ii) Falls $x \notin J_i(h_N)$ für alle $i \in \{1, \ldots, m\}$, dann ist $\left[F_N(x+h_N) - F_N(x-h_N)\right] = 0$.

Beweis:

(i) N_i ist die absolute Häufigkeit von a_i in der Ausgangsverteilung $\mathcal{F}(N)$, d.h.

$$\frac{N_1}{N} = F_N(a_1)$$

$$\frac{N_i}{N} = F_N(a_i) - F_N(a_{i-1}) \quad \text{für} \quad i \in \{2, \ldots, m\} \quad . \tag{5}$$

Sei $a_1 < \ldots < a_m$ o.B.d.A.

Fallunterscheidung:

(a) Es wird gezeigt, dass $F_N(x+h_N) = F_N(a_i)$ für alle $x \in J_i(h_N)$ und jedes $i \in \{1, \ldots, m\}$.

Sei also $a_i - h_N \leq x < a_i + h_N$.

Es folgt: $a_i \leq x + h_N < a_i + 2 \cdot h_N$.

Gemäß Voraussetzung ist $2 \cdot h_N < a_{i+1} - a_i$ für alle $i \in \{1, \ldots, m-1\}$,

also gilt: $a_i \leq x + h_N < a_i + 2 \cdot h_N < a_{i+1} \Rightarrow (x + h_N) \in [a_i ; a_{i+1})$.

Wegen der Stetigkeit von F_N von rechts gilt: $F_N(t) = F_N(a_i)$ für alle $t \in [a_i ; a_{i+1})$.

Es folgt: $F_N(x+h_N) = F_N(a_i)$ für alle $x \in J_i(h_N)$, $i \in \{1, \ldots, m-1\}$.

$i = m$:

Aus $a_m - h_N \leq x$ folgt: $a_m \leq x + h_N \Rightarrow 1 = F_N(a_m) = F_N(x+h_N)$

(b) Es wird gezeigt, dass $F_N(x-h_N) = F_N(a_{i-1})$ für alle $x \in J_i(h_N)$ und

jedes $i \in \{2, \ldots, m\}$.

Sei $1 < i \leq m$, dann gilt für $a_i - h_N \leq x < a_i + h_N$ zusammen mit $2 \cdot h_N < a_i - a_{i-1}$:

$a_i - 2 \cdot h_N \leq x - h_N < a_i \Rightarrow (x - h_N) \in [a_{i-1} ; a_i)$.

Wegen der Stetigkeit von F_N von rechts gilt: $F_N(t) = F_N(a_{i-1})$ für alle $t \in [a_{i-1} ; a_i)$.

Es folgt $F_N(x-h_N) = F_N(a_{i-1})$.

Aus (a) und (b) folgt also zusammen mit (5) für jedes $i \in \{2, \ldots, m\}$ und alle $x \in J_i(h_N)$:

$$\frac{N_i}{N} = F_N(a_i) - F_N(a_{i-1}) = F_N(x+h_N) - F_N(x-h_N) \quad . \tag{6}$$

$i = 1$:

Für $x \in J_1(h_N)$ gilt: $x - h_N < a_1 \Rightarrow F_N(x - h) = 0$.

Daraus erhält man zusammen mit (a) :

$$\frac{N_1}{N} = F_N(a_1) = F_N(x + h) - F_N(x - h) \text{ für alle } x \in J_1(h_N) .$$

(ii) Sei $x \notin J_i(h_N)$ für alle $i \in \{1, \ldots, m\}$.

Fallunterscheidung

1. $x < a_1 - h_N$.

Dann gilt:

$x - h_N < x + h_N < a_1 \Rightarrow F_N(x-h_n) = F_N(x+h_n) = 0$.

2. Sei $a_m + h_N \leq x$.

Dann gilt:

$a_m \leq x - h_N < x + h_N \Rightarrow F_N(x-h_n) = F_N(x+h_n) = 1$

3. Sei $a_{i-1} + h_N \leq x < a_i - h_N$ für ein $i \in \{2, \ldots, m\}$.

Dann gilt:

$a_{i-1} + 2 \cdot h_N \leq x + h_N < a_i \Rightarrow (x + h_N) \in [a_{i-1} ; a_i)$,

$a_{i-1} \leq x - h_N < a_i - 2 \cdot h_N \Rightarrow (x - h_N) \in [a_{i-1} ; a_i)$.

Wegen der Stetigkeit von F_N von rechts gilt:

$F_N(t) = F_N(a_{i-1})$ für alle $t \in [a_{i-1} ; a_i)$.

Daraus erhält man: $F_N(x-h_n) = F_N(x+h_n) = F_N(a_{i-1})$

Damit ist alles gezeigt.

❑

Unter der Voraussetzung $0 < 2 \cdot h_N < b_N$ gilt also mit $J_i(h_N) = [a_i - h_N ; a_i + h_N)$ konstruktionsgemäß für jedes $i \in \{1, \ldots, m\}$:

$$\frac{N_i}{N} = 1_{J_i(h_N)}(t) \cdot \left[F_N(t+h_n) - F_N(t-h_n) \right] \tag{7}$$

$$= 1_{J_i(h_N)}(t) \cdot 2 \cdot h_N \cdot f_N(t) \quad .$$

Für jedes $t \in \mathbb{R}$ gilt:

$$F_N(t+h_n) - F_N(t-h_n) = \frac{1}{N} \cdot \sum_{i=1}^{N} 1_{(t-h_N ; t+h_N]}(z_i)$$

$$= \int_{\mathbb{R}} 1_{(t-h_N ; t+h_N]}(u) \, dF_N(u) \quad . \tag{8}$$

Außerdem gilt:

$$\frac{1}{N} \cdot \sum_{i=1}^{N} 1_{(t-h_N ; t+h_N]}(z_i) = \frac{1}{N} \cdot \sum_{i=1}^{N} 1_{[z_i-h_N ; z_i+h_N)}(t)$$

$$= \frac{1}{N} \cdot \sum_{j=1}^{m} N_j \cdot 1_{[a_j-h_N ; a_j+h_N)}(t) \quad . \tag{9}$$

Mit (7) erhält man unter der Voraussetzung $0 < 2 \cdot h_N < b_N$ definitionsgemäß für die Abbildung f_N :

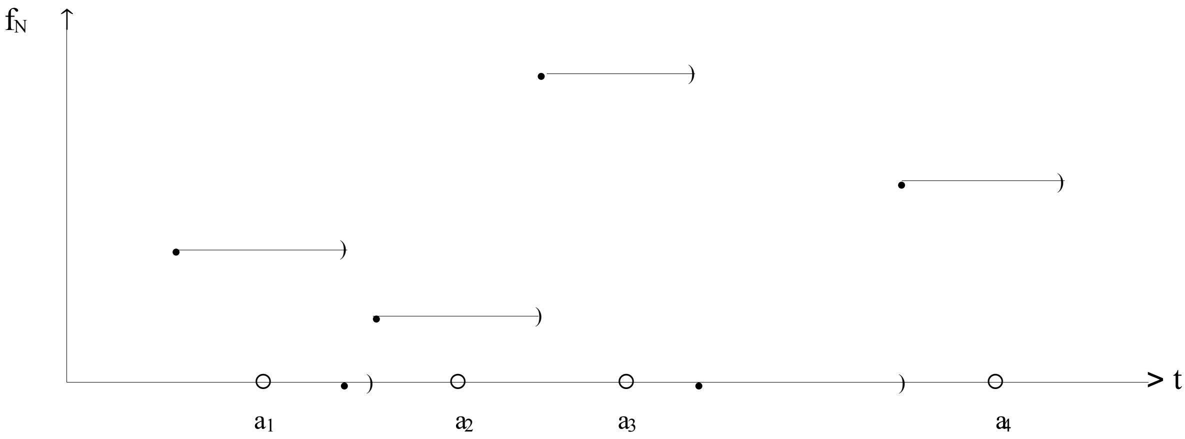
$$1_{J_i(h_N)}(t) \cdot f_N(t) \quad = \quad \frac{1}{2 \cdot h_N} \cdot \frac{N_i}{N} \quad , \quad i \in \{1, \ldots, m\}.$$

Mit (8) ergeben sich daraus für jedes $t \in \mathbb{R}$ die Darstellungen:

$$f_N(t) \quad = \quad \frac{1}{2 \cdot h_N} \cdot \int_{\mathbb{R}} 1_{(t-h_N \,;\, t+h_N]}(u) \; dF_N(u) \tag{10}$$

$$= \quad \frac{1}{2 \cdot h_N \cdot N} \cdot \sum_{j=1}^{m} N_j \cdot 1_{[a_j - h_N \,;\, a_j + h_N)}(t)$$

Zur Veranschaulichung wird ein möglicher f_N - Graph skizziert:



Mit $0 < 2 \cdot h_N < b_N$ folgt wegen (2) und (10):

$$\frac{\lambda_N}{N} \quad = \quad 2 \cdot h_N \cdot \sup_{t \in \mathbb{R}} f_N(t) \quad = \quad \sup_{t \in \mathbb{R}} \left\{ \int_{\mathbb{R}} 1_{(-h_N \,;\, h_N]}(u-t) \; dF_N(u) \right\} \tag{11}$$

$$= \quad \sup_{t \in \mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\} \tag{12}$$

3. Die Abschätzung von $\frac{\lambda_N}{N}$

Die Eigenschaften der Verteilungsfunktion F bestimmen in hohem Maße, welche Abschätzung für $\frac{\lambda_N}{N}$ erreicht werden kann. Denn wegen des Satzes von Glivenko / Cantelli (R.J.Serfling 1980, S. 61) konvergiert die in (2) definierte Verteilungsfunktion F_N mit fast jeder Folge $(z_N)_{N\in\mathbb{N}}$ gleichmäßig gegen F, und $\frac{\lambda_N}{N}$ ist gerade die maximale Sprunghöhe von F_N.

Im Abschnitt 3.1 werden zunächst obere und untere Schranken des Ausdrucks $\sup_{t\in\mathbb{R}}\{F_N(t+h_N)-F_N(t-h_N)\}$ untersucht. Die weitere Abschätzung von $\frac{\lambda_N}{N}$ hängt davon ab, welche Voraussetzungen für F und die Schrittweite $2\cdot h_N$ zugrundeliegen.

Für die Fälle $m = 1$ und $m = N$ sind die Werte von $\frac{\lambda_N}{N}$ bekannt. Im folgenden wird $1 < m < N$ vorausgesetzt ($N \in \mathbb{N}$), so dass stets die Beziehung $\frac{2}{N} \leq \frac{\lambda_N}{N} \leq \frac{N-1}{N}$ gültig ist. Außerdem folgt aus $1 < m$, dass $0 < b_N$ gilt.

In $J_i(h_N) = [a_i - h_N ; a_i + h_N)$ liegt genau einer der Werte $a_1, \ldots, a_m$, sofern $0 < h_N < b_N$ erfüllt ist, $i \in \{1, \ldots, m\}$.

Gilt sogar $0 < 2\cdot h_N < b_N$, folgt aus Satz II.1 mit (7) bis (9):

$$\frac{N_i}{N} = 1_{J_i(h_N)}(x) \cdot \frac{1}{N} \cdot \sum_{j=1}^{m} N_j \cdot 1_{[a_j-h_N\,;\,a_j+h_N)}(t) \ .$$

Wenn $h_N \geq b_N$ ist, kann der Fall eintreten, dass mehr als einer der Werte $a_1, \ldots, a_m$ in einem Intervall $J_i(h_N)$ liegt, und es zu Mehrfachzählungen kommt.

Im Falle $h_N \geq b_N$ gilt daher für jedes $i \in \{1, \ldots, m\}$:

$$\frac{N_i}{N} \leq 1_{J_i(h_N)}(x) \cdot \frac{1}{N} \cdot \sum_{j=1}^{m} N_j \cdot 1_{[a_j-h_N\,;\,a_j+h_N)}(t) \ . \tag{13}$$

Wenn h_N eine beliebige positive Zahl ist, erhält man speziell:

$$\frac{\lambda_N}{N} \leq \sup_{t\in\mathbb{R}}\left\{\frac{1}{N} \cdot \sum_{j=1}^{m} N_j \cdot 1_{[a_j-h_N\,;\,a_j+h_N)}(t)\right\} \tag{14}$$

$$= \sup_{t\in\mathbb{R}}\{F_N(t+h_N)-F_N(t-h_N)\} \ . \tag{15}$$

3.1 Schranken der Größe $\sup_{\mathbf{t} \in \mathbb{R}} \left\{ \mathbf{F_N(t+h_N) - F_N(t-h_N)} \right\}$

Aus (15) folgt:

$$\begin{aligned}
\frac{\lambda_N}{N} &\leq \sup_{t\in\mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\} \\
&= \sup_{t\in\mathbb{R}} \left\{ \left[F_N(t+h_N) - F(t+h_N) \right] + \left[F(t+h_N) - F(t-h_N) \right] + \left[F(t-h_N) - F_N(t-h_N) \right] \right\} \\
&\leq \sup_{t\in\mathbb{R}} \left| F_N(t+h_N) - F(t+h_N) \right| + \sup_{t\in\mathbb{R}} \left| F_N(t-h_N) - F(t-h_N) \right| \\
&\qquad + \sup_{t\in\mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} \\
&\leq 2 \cdot \sup_{x\in\mathbb{R}} \left| F_N(x) - F(x) \right| + \sup_{t\in\mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} . \qquad (16)
\end{aligned}$$

Mit der in (1) definierten Empirischen Verteilungsfunktion IF_N zum Stichprobenumfang $N \in \mathbb{N}$ wird die Kolmogorov / Smirnov - Distance $\mathbb{D}_N$ gebildet, $\mathbb{D}_N := \sup_{x\in\mathbb{R}} \left| \mathrm{IF}_N(x) - F(x) \right|$.

Mit der Realisation F_N der Empirischen Verteilungsfunktion IF_N erhält man eine Realisation D_N von $\mathbb{D}_N$,

$$D_N := \sup_{x\in\mathbb{R}} \left| F_N(x) - F(x) \right| . \qquad (17)$$

Mit (16) gilt also:

$$\frac{\lambda_N}{N} \leq 2 \cdot D_N + \sup_{t\in\mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} . \qquad (18)$$

Andererseits erhält man:

$$\begin{aligned}
& F_N(t+h_N) - F_N(t-h_N) \\
&= \left[F_N(t+h_N) - F(t+h_N) \right] + \left[F(t+h_N) - F(t-h_N) \right] + \left[F(t-h_N) - F_N(t-h_N) \right] \\
&\geq -2 \cdot D_N + F(t+h_N) - F(t-h_N) . \qquad (19)
\end{aligned}$$

3.1.1 Die Abschätzung der Kolmogorov / Smirnov - Distance $\mathbb{D}_N$

Man unterscheidet gundsätzlich zwei Versionen der Abschätzung von $\mathbb{D}_N$:

Die fast-sichere Abschätzung für große $N \in \mathbb{N}$, und die Abschätzung mit einer gewissen Mindest-Wahrscheinlichkeit für jedes $N \in \mathbb{N}$.

Während daraus in den Abschnitten 3.3.1 und 3.3.2 jeweils eine andere Abschätzung für $\frac{\lambda_N}{N}$ entwickelt wird, erfolgt in Abschnitt 3.3.3 eine simultane Anwendung beider Versionen.

Die fast-sichere Abschätzung von $\mathbb{D}_N$ für große N

Aus dem zuerst von N.V.Smirnov (1944) bewiesenen „Gesetz vom Iterierten Logarithmus“ (R.J.Serfling 1980, S. 62) folgt, dass für alle hinreichend großen $N \in \mathbb{N}$ $IP\mathcal{X}^{-1}$ - fast sicher gilt:

$$\mathbb{D}_N \leq 2 \cdot \sqrt{\frac{\ln\ln N}{N}} \quad . \tag{20}$$

Die $\mathbb{D}_N$ - Abschätzung mit einer Mindest-Wahrscheinlichkeit für jedes N

Die Ungleichung von A.Dvoretzki / J. Kiefer / J. Wolfowitz (kurz: DKW-Ungleichung, siehe R.J.Serfling 1980, S. 59) besagt:
Es existiert eine reelle Zahl $C > 0$, so dass für alle reellen $d > 0$ und alle $N \in \mathbb{N}$ gilt:

$$1 - \frac{C}{e^{2 \cdot N \cdot d^2}} \leq P\left(\mathbb{D}_N \leq d \right) \quad . \tag{21}$$

Dies kann man per Landau-Symbol auch so formulieren:

$$P\left(\mathbb{D}_N > d \right) = O\left(e^{-2 \cdot N \cdot d^2} \right) \quad \text{für alle reellen } d > 0 \text{ und alle } N \in \mathbb{N} \,. \tag{22}$$

3.2 Die Schrittweite $2 \cdot h_N$

Die Wahl der Schrittweite beeinflußt die Qualität der Abschätzung von $\frac{\lambda_N}{N}$.
Wegen (12) gilt unter der Voraussetzung $0 < 2 \cdot h_N < b_N$ die Identität
$\frac{\lambda_N}{N} = \sup_{t \in \mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\}$. Hieraus lassen sich obere und untere Schranken für $\frac{\lambda_N}{N}$ formulieren, sowie ein Grenzwertsatz bei stückweise stetigem F . Bei Verwendung eines beliebigen $h_N > 0$ ($N \in \mathbb{N}$) gilt wegen (14) und (15) anstelle des obigen Gleichheitszeichens i.a. nur noch ein „ $\leq$ “ . Eine <u>untere</u> Schranke für $\frac{\lambda_N}{N}$ steht mit $\frac{2}{N}$ im Falle $1 < m < N$ stets zur Verfügung.

Die Nicht-Erfülltheit der Bedingung $2 \cdot h_N < b_N$ führt neben einer Verschlechterung der Abschätzung von $\frac{\lambda_N}{N}$ auch dazu, dass der Beweis des o.g. Grenzwertsatzes nicht mehr funktioniert.
Es wird eine relativ milde Bedingung für die Daten angegeben, unter der die Voraussetzung $0 < 2 \cdot h_N < b_N$ mit einem speziellen h_N erfüllt ist.
Die Realisationen $z_1, \ldots, z_N$ der Zufallsgrößen $Z_1, \ldots, Z_N$ werden aufgefaßt als N Glieder einer Folge $(z_i)_{i \in N} \in \mathbb{R}^{\mathbb{N}}$. Gemäß (3) ist für alle $i, j \in \{1, \ldots, N\}$

$$b_N := \min_{z_i \neq z_j} \left\{ \left| z_i - z_j \right| \right\} =: B(z_1, \ldots, z_N). \tag{23}$$

Dann ist $B(z_1, \ldots, z_N)$ eine Realisation von

$$B_N := B(Z_1, \ldots, Z_N) := \min_{Z_i \neq Z_j} \left\{ \left| Z_i - Z_j \right| \right\} \text{ mit } i, j \in \{1, \ldots, N\}.$$

Aus $1 < m$ folgt $0 < b_N$, weil es im Falle $1 < m$ mindestens zwei paarweise verschiedene Werte unter den $z_1, \ldots, z_N$ gibt.
Da hier nur der Fall $1 < m$ betrachtet wird, gibt es zu endlichem N stets ein passendes $h_N > 0$, so dass $0 < 2 \cdot h_N < b_N$ erfüllt ist. Um dies zu konkretisieren, wird das folgende Lemma angegeben.

Lemma II.1

$b_N = B(z_1, \ldots, z_N)$ ist wie in (23) definiert. Es wird $1 < m$ vorausgesetzt.
Mit einem $r > 0$ und beliebigem $N \in \mathbb{N}$ gelte

$$\frac{2}{N^r} \leq b_N . \tag{24}$$

Setze mit einem beliebig gewählten, festen $q > r$ zu jedem $N \in \mathbb{N}$

$$h_N := \frac{1}{N^q} . \tag{25}$$

Dann folgt: $0 < 2 \cdot h_N < b_N$ für alle $N \in \mathbb{N}$.

Beweis:

(i) Wenn die Folge $(b_N)_{N \in \mathbb{N}}$ nicht den Häufungspunkt 0 besitzt, ist $2 \cdot h_N < b_N$ für jedes $N \in \mathbb{N}$ auch ohne die Bedingung (24) erfüllt:

Wegen $1 < m$ ist $0 < b_N$, also gibt es ein $s \in \mathbb{R}^+$, so dass $0 < s < b_N$ für alle $N \in \mathbb{N}$.

Daraus folgt: $2 \le N^q \cdot s < N^q \cdot b_N$ für alle $N \in \mathbb{N}$ => $2 \cdot h_N < b_N$ für alle $N \in \mathbb{N}$.

(ii) Die Folge $(b_N)_{N \in \mathbb{N}}$ besitze den Häufungspunkt 0, dann existiert nicht notwendig eine für alle $N \in \mathbb{N}$ gültige untere Schranke $s \in \mathbb{R}^+$ wie in (i).

Es gilt aber für alle $N > 1$: $h_N = \dfrac{1}{N^q} = \dfrac{1}{N^\delta} \cdot \dfrac{1}{N^{q-\delta}} < \dfrac{1}{N^{q-\delta}}$ mit $r := q - \delta > 0$.

Wegen der Voraussetzung folgt daraus für alle $N \in \mathbb{N}$ mit $1 < m \le N$:

$$2 \cdot h_N < \frac{2}{N^{q-\delta}} = \frac{2}{N^r} \le b_N .$$

❑

Zu den beiden im Beweis von Lemma II.1 unterschiedenen Situationen wird je ein Beispiel angegeben.

(i) Sei $\mathcal{F}$ die Verteilung einer diskreten Zufallsgröße, deren Wertemenge $\mathbb{W}$ keinen endlichen Häufungspunkt besitzt (z.B. Negative Binomialverteilung mit $\mathbb{W} = \{ s , s+1 , \dots \}$ und $s \in \mathbb{N}$, Poisson-Verteilung mit $\mathbb{W} = \mathbb{N}_0$).

Dann existiert ein $\eta > 0$ derart, so dass $| x - y | \ge \eta$ für alle $x , y \in \mathbb{W}$ mit $x \ne y$, d.h. $(b_\nu)_{\nu \in \mathbb{N}}$ besitzt nicht den Häufungspunkt Null.

Daraus folgt:

$$2 \cdot h_N = \frac{2}{N^q} < \eta \le b_N \quad \text{für alle hinreichend großen } N \in \mathbb{N} .$$

(ii) Im folgenden Beispiel konvergiert fast jede Teilfolge von $(b_\nu)_{\nu \in \mathbb{N}}$ gegen Null.

Es wird definiert: $T_N := T(Z_1 , \dots , Z_N) := \dfrac{1}{\binom{N}{2}} \cdot \displaystyle\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} | Z_i - Z_j |$.

T_N heißt „Gini's mean difference“ zum Stichprobenumfang $N \in \mathbb{N}$ (R.J.Serfling 1980, S. 174).

$t_N := T(z_1 , \dots , z_N)$ ist eine Realisation von $T_N = T(Z_1 , \dots , Z_N)$.

Wegen der Definitionen gilt stets

$$B(z_1 , \dots , z_N) \le T(z_1 , \dots , z_N) , \quad N \in \mathbb{N} . \tag{26}$$

Setze $a := E_F\left(\left|Z_1 - Z_2\right|\right) = \int\limits_{IR}\int\limits_{IR} |z_1 - z_2| \, dF(z_1) \, dF(z_2)$,

dann ist a der mittlere Abstand zweier Zufallsgrößen der Folge $\mathcal{X} = (Z_i)_{i \in IN}$.
Es gilt die folgende Implikation (R.J.Serfling 1980, S. 190):

$a < +oo \quad \Rightarrow \quad T_\nu \underset{\nu \to +\infty}{\to} a \quad IP\mathcal{X}^{-1}$ - fast sicher. (27)

Das bedeutet im Falle $a < +oo$:

$t_\nu = T(z_1, \ldots, z_\nu)$ strebt mit $\nu \to +oo$ gegen a . Es sei denn, die Folge $(z_\nu)_{\nu \in IN}$ gehört zu einer Ausnahmemenge $\mathcal{N} \subset IR^{IN}$ mit $IP\mathcal{X}^{-1}(\mathcal{N}) = 0$. (28)

Ist $a = 0$, dann folgt wegen (26) und (27) , dass $(b_\nu)_{\nu \in IN}$ gegen Null konvergiert.

Es sei denn, $(z_\nu)_{\nu \in IN}$ ist eine Ausnahmefolge im Sinne von (28).

3.3 Formen der Verteilungsfunktion F und die $\frac{\lambda_N}{N}$ - Abschätzung

Die Fortsetzung der auf (18) und (19) aufbauenden $\frac{\lambda_N}{N}$ - Abschätzung hängt insbesondere von den Eigenschaften der Verteilungsfunktion F sowie von der Schrittweite h_N ab.

3.3.1 Die Verteilungsfunktion F ist stückweise stetig

Mit S (F) wird die Menge der Sprungstellen von F bezeichnet. Da F eine auf ganz IR definierte monoton wachsende Funktion ist, ist S (F) eine abzählbare Menge, und es ist F stetig auf IR \ S (F) (H.Heuser 2000, I S. 240). Es wird vorausgesetzt, dass S (F) keinen endlichen Häufungspunkt besitzt.

Als nächstes wird die Differenz $F(t+h_N) - F(t-h_N)$ nach oben und unten abgeschätzt.

Dazu werden einseitige Funktionen-Grenzwerte sowie die maximale Sprunghöhe einer Funktion definiert: $F(x_0^-) := \lim\limits_{x \uparrow x_0} F(x)$, $F(x_0^+) := \lim\limits_{x \downarrow x_0} F(x)$.

Sei t_0 die Stelle der maximalen Sprunghöhe von F , und j(F) die maximale Sprunghöhe.

Damit gilt: $j(F) = F(t_0^+) - F(t_0^-)$.

Für die Folge der Schrittweiten $\left(2\cdot h_N\right)_{N\in\mathbb{N}}$ wird vorausgesetzt, dass sie gegen Null strebt, und dass $0 < h_{N+1} \leq h_N$ für alle $N \in \mathbb{N}$ gilt.

Desweiteren wird zur Abkürzung zu jedem $N \in \mathbb{N}$ und jedem $t \in \mathbb{R}$ gesetzt:

$$F(t+h_N) - F(t^+) =: \epsilon_N(t) \ , \quad F(t^-) - F(t-h_N) =: \delta_N(t) \ , \tag{29}$$

$$2 \cdot \max\left\{\epsilon_N(t) \ , \ \delta_N(t)\right\} =: A_N(t) \ . \tag{30}$$

Für die Abschätzung von $F(t+h_N) - F(t-h_N)$ sind die Eigenschaften von $A_N(t)$ von grundlegender Bedeutung.

Es gilt:

(i) $0 \leq A_N(t) \leq 2$ für alle $N \in \mathbb{N}$ und alle $t \in \mathbb{R}$.

(ii) $A_N(t) \underset{t\to+\infty}{\downarrow} 0$ und $A_N(t) \underset{t\to-\infty}{\downarrow} 0$ zu jedem festen $N\in \mathbb{N}$ (31)

Daher existiert zu jedem $N \in \mathbb{N}$ ein $t_N \in \mathbb{R}$ mit

$$\sup_{t\in\mathbb{R}}\left\{ A_N(t) \right\} = A_N(t_N) =: \tilde{A}_N \ . \tag{32}$$

Wegen der wachsenden Monotonie von $F(x)$ in $x \in \mathbb{R}$ und der fallenden Monotonie von h_N strebt $A_N(t)$ zu jedem festen $t \in \mathbb{R}$ mit wachsendem $N \in \mathbb{N}$ monoton fallend gegen Null.

Wie das folgende Lemma zeigt, strebt auch die Folge der Suprema $\left(\tilde{A}_N\right)_{N\in\mathbb{N}}$ gegen Null.

Lemma II.2

Für das in (30) definierte $A_N(t)$ gilt:

$A_N(t)$ konvergiert gleichmäßig in $\mathbb{R}$ gegen Null, d.h. $\tilde{A}_N \underset{N\to+\infty}{\to} 0$.

Beweis:

Wie oben gezeigt, existiert $\sup_{t\in\mathbb{R}}\left\{A_N(t)\right\} = A_N(t_N) =: \tilde{A}_N$ zu jedem $N \in \mathbb{N}$,

und es gilt $A_N(t) \geq A_{N+1}(t)$ für jedes $N \in \mathbb{N}$ und für jedes $t \in \mathbb{R}$.

Da die Menge $S(F)$ nach Voraussetzung keinen endlichen Häufungspunkt besitzt, existiert ein $\rho > 0$ derart, so dass für alle $t_1, t_2 \in S(F)$ gilt: $|t_1 - t_2| > 2 \cdot \rho$.

Mit einem $\delta \in (0 ; \rho)$ ist $U_\delta(a) := \{ t : |t - a| < \delta \}$, dann enthält jede Menge $U_\delta(a)$

außer a keine weitere Sprungstelle von F .

Definiert wird die folgende offene Menge:

$$M := \bigcup_{a \in S(F)} U_\delta(a) \quad .$$

Dann ist nach Voraussetzung F stetig auf $\{ \mathbb{R} \setminus M \}$,

und damit ist für jedes $N \in \mathbb{N}$ auch A_N stetig auf $\{ \mathbb{R} \setminus M \}$.

Zu jedem Kompaktum $K := [-k \; ; \; k \,]$ mit $k \in \mathbb{R}^+$ gilt:

$IK := K \cap \{ \mathbb{R} \setminus M \}$ ist kompakt.

Mit dem Satz über die monotone Konvergenz von U.Dini (H.Heuser 2000, I S. 578) folgt:

$$\sup_{t \in IK} \left\{ A_N(t) \right\} \underset{N \to +\infty}{\to} 0 \quad .$$

Wegen (31) läßt sich diese Beziehung ausweiten zu

$$\sup_{t \in \{\mathbb{R} \setminus M\}} \left\{ A_N(t) \right\} \underset{N \to +\infty}{\to} 0 \quad .$$

Weil $\delta \in (0 \, ; \rho)$ beliebig klein gewählt werden darf, bleibt nur noch das Konvergenzverhalten von $A_N(t)$ für $t \in S(F)$ zu betrachten.

Es gilt für alle $t \in \mathbb{R}$ (also insbesondere für alle $t \in S(F)$) :

$A_N(t)$ strebt mit wachsendem $N \in \mathbb{N}$ monoton fallend gegen Null.

$A_N(t)$ ist beschränkt für alle $N \in \mathbb{N}$.

Damit folgt:

$$\sup_{t \in \mathbb{R}} \left\{ A_N(t) \right\} \underset{N \to +\infty}{\to} 0 \quad .$$

❑

Für jedes $N \in \mathbb{N}$ und jedes $t \in \mathbb{R}$ gilt mit den Definitionen in (29) :

$$\begin{aligned} F(t+h_N) - F(t-h_N) &= \Big(F(t+h_N) - F(t^+) \Big) + \Big(F(t^+) - F(t^-) \Big) \\ &\qquad + \Big(F(t^-) - F(t-h_N) \Big) \\ &= \epsilon_N(t) + \Big(F(t^+) - F(t^-) \Big) + \delta_N(t) \; . \end{aligned}$$

Daraus ergibt sich für jedes $N \in \mathbb{N}$ und jedes $t \in \mathbb{R}$:

$$F(t^+) - F(t^-) \le F(t+h_N) - F(t-h_N) \le \tilde{A}_N + F(t^+) - F(t^-)$$

$$\leq \tilde{A}_N + F(t_0^+) - F(t_0^-) = \tilde{A}_N + j(F) .$$

Es folgt:

$$\text{(i)}\ \ j(F) = F(t_0^+) - F(t_0^-) \leq F(t+h_N) - F(t-h_N) \leq \sup_{t\in\mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} \tag{33}$$

$$\text{(ii)}\ \ F(t+h_N) - F(t-h_N) \leq \sup_{t\in\mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} \leq \tilde{A}_N + j(F) . \tag{34}$$

Zusammen mit (15) bis (18) ergibt sich daraus für jedes $N \in \mathbb{N}$ die folgende Ungleichungskette:

$$-2 \cdot D_N + j(F) \leq -2 \cdot D_N + F(t+h_N) - F(t-h_N) \leq F_N(t+h_N) - F_N(t-h_N)$$

$$\leq \sup_{t\in\mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\} \leq 2 \cdot D_N + j(F) + \tilde{A}_N .$$

Also gilt für jedes $N \in \mathbb{N}$ mit beliebiger, monoton fallender Nullfolge $(h_\nu)_{\nu\in\mathbb{N}}$ ($h_\nu > 0$ für alle $\nu \in \mathbb{N}$) mit jedem reellen Datensatz $z_1 , . . . , z_N$ die Abschätzung:

$$-2 \cdot D_N + j(F) \leq \sup_{t\in\mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\} \leq 2 \cdot D_N + j(F) + \tilde{A}_N . \tag{35}$$

Als nächstes werden zwei Nullmengen definiert.

Seien $\mathcal{N}_1 , \mathcal{N}_2 \subset \mathbb{R}^{\mathbb{N}}$ Mengen vom Maße $\mathbb{P}\mathcal{X}^{-1}(\mathcal{N}_1) = 0$ und $\mathbb{P}\mathcal{X}^{-1}(\mathcal{N}_2) = 0$.

Mit der in (2) definierten Realisation F_N der Empirischen Verteilungsfunktion $\mathbb{F}_N$ gelte mit jeder Folge $(z_\nu)_{\nu\in\mathbb{N}} \in \left\{ \mathbb{R}^{\mathbb{N}} \setminus \mathcal{N}_1 \right\}$ zu jedem festen $t \in \mathbb{R}$:

$$F_\nu(t) \underset{\nu\to+\infty}{\longrightarrow} F(t) .$$

Für die in (17) definierte Realisation D_N der Kolmogorov / Smirnov - Distance $\mathbb{D}_N$ gelte mit jeder Folge $(z_\nu)_{\nu\in\mathbb{N}} \in \left\{ \mathbb{R}^{\mathbb{N}} \setminus \mathcal{N}_2 \right\}$ das „Gesetz vom Iterierten Logarithmus“ von N.V.Smirnov (R.J.Serfling 1980, S. 62).

Es wird gesetzt: $\mathcal{N} := \mathcal{N}_1 \cup \mathcal{N}_2$. (36)

Versionen der $\frac{\lambda_N}{N}$ - Abschätzung

Aus (35) ergeben sich verschiedene Möglichkeiten für die Abschätzung von $\frac{\lambda_N}{N}$, je nach dem, ob $2 \cdot h_N < b_N$ vorausgesetzt ist, und welche Methode der $\mathbb{D}_N$ - Abschätzung gewählt wird.

(A) Die Schrittweite ist eine beliebige, monoton fallende, echt positive Nullfolge.

Wegen der Voraussetzung $m < N$ und wegen (15) gilt:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq \sup_{t \in \mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\} .$$

Aus (35) folgt also für jedes $N \in \mathbb{N}$ die Abschätzung:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq 2 \cdot D_N + j(F) + \tilde{A}_N . \tag{37}$$

(A.1) Aus (37) und (20) erhält man mit beliebigen Werten $z_1, \ldots, z_N \in \mathbb{R}$, die nicht zu einer Ausnahmefolge $(z_\nu)_{\nu \in \mathbb{N}} \in \mathcal{N}$ gehören, für alle hinreichend großen $N \in \mathbb{N}$:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq 4 \cdot \sqrt{\frac{\ln \ln N}{N}} + \tilde{A}_N + j(F) . \tag{38}$$

(A.2) Die Anwendung von (21) mit $d = \frac{1}{N^\theta}$, $0 < \theta < ½$ und einer nicht von F abhängigen reellen Konstanten $C > 0$ liefert für jedes $N \in \mathbb{N}$:

$$1 - \frac{C}{e^{2 \cdot N^{1-2\cdot\theta}}} \leq P\left(\mathbb{D}_N \leq \frac{1}{N^\theta} \right) \tag{39}$$

$$= P\left(2 \cdot \mathbb{D}_N + j(F) + \tilde{A}_N \leq \frac{2}{N^\theta} + j(F) + \tilde{A}_N \right) \tag{40}$$

Wegen (37) folgt daraus mit einer nicht von F abhängigen reellen Konstanten $C > 0$ und beliebig gewähltem $\theta \in (0; ½)$ für jedes $N \in \mathbb{N}$:
Zur Abschätzung

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq \frac{2}{N^\theta} + j(F) + \tilde{A}_N \tag{41}$$

darf man ein Vertrauen von mindestens $1 - \frac{C}{e^{2 \cdot N^{1-2\cdot\theta}}}$ haben.

(B) Die Schrittweite ist eine monoton fallende, echt positive Nullfolge und erfüllt zusätzlich die Bedingung $2 \cdot h_N < b_N$.

Wegen $0 < 2 \cdot h_N < b_N$ folgt dann aus (35) und (12) für alle $N \in \mathbb{N}$ die Abschätzung:

$$- 2 \cdot D_N + j(F) \leq \frac{\lambda_N}{N} \leq 2 \cdot D_N + j(F) + \tilde{A}_N . \tag{42}$$

Für große $N \in \mathbb{N}$ ist diese $\frac{\lambda_N}{N}$ - Untergrenze näher bei $\frac{\lambda_N}{N}$ als die „natürliche" untere Grenze $\frac{2}{N}$.

(B.1) Aus (42) und (20) erhält man mit beliebigen Werten $z_1, \ldots, z_N \in \mathbb{R}$, die nicht zu einer Ausnahmefolge $(z_\nu)_{\nu \in \mathbb{N}} \in \mathcal{N}$ gehören, für alle hinreichend großen $N \in \mathbb{N}$:

$$- 4 \cdot \sqrt{\frac{\ln \ln N}{N}} + j(F) \leq \frac{\lambda_N}{N} \leq 4 \cdot \sqrt{\frac{\ln \ln N}{N}} + \tilde{A}_N + j(F) . \tag{43}$$

(B.2) Die Anwendung von (21) mit $d = \frac{1}{N^\theta}$, $0 < \theta < ½$ und einer nicht von F abhängigen reellen Konstanten $C > 0$ liefert für jedes $N \in \mathbb{N}$:

$$1 - \frac{C}{e^{2 \cdot N^{1-2\cdot\theta}}} \leq P\left(- \frac{1}{N^\theta} \leq - \mathbb{D}_N \right)$$

$$= P\left(- \frac{2}{N^\theta} + j(F) \leq - 2 \cdot \mathbb{D}_N + j(F) \right) .$$

Daraus folgt wegen (40):

$$1 - \frac{C}{e^{2 \cdot N^{1-2\cdot\theta}}} \leq$$

$$P\left(\left\{ 2 \cdot \mathbb{D}_N + j(F) + \tilde{A}_N \leq \frac{2}{N^\theta} + j(F) + \tilde{A}_N \right\} \wedge \left\{ - \frac{2}{N^\theta} + j(F) \leq - 2 \cdot \mathbb{D}_N + j(F) \right\} \right)$$

Zusammen mit (41) erhält man daraus mit einer nicht von F abhängigen reellen Konstanten $C > 0$ für jedes $N \in \mathbb{N}$ die folgende Aussage:

Zur Abschätzung

$$- \frac{2}{N^\theta} + j(F) \leq \frac{\lambda_N}{N} \leq \frac{2}{N^\theta} + j(F) + \tilde{A}_N \tag{44}$$

darf man ein Vertrauen von mindestens $1 - \dfrac{C}{e^{2 \cdot N^{1-2\cdot\theta}}}$ haben.

Bemerkungen zu den Versionen der $\dfrac{\lambda_N}{N}$ - Abschätzung

(a) Ist die Verteilung $\mathcal{F}$ diskret, dann ist die Verteilungsfunktion F eine Treppenfunktion.

Bei monoton fallender Nullfolge $(h_\nu)_{\nu \in \mathbb{N}}$ ($h_\nu > 0$ für alle $\nu \in \mathbb{N}$) gilt mit den Definitionen

(29) und (30) für $A_N(t)$:

$$\sup_{t \in \mathbb{R}} \left\{ A_N(t) \right\} = \tilde{A}_N = 0 \quad \text{für jedes hinreichend große } N \in \mathbb{N} .$$

(b) Ist die Verteilungsfunktion F in $\{ \mathbb{R} \setminus S(F) \}$ sogar Lipschitz-stetig zur Konstanten $L > 0$,

dann erhält man eine Aussage über die Mindest-Konvergenzgeschwindigkeit von

$\sup_{t \in \mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\}$ gegen $j(F)$.

Mit den Definitionen (29) und (30) gilt:

$A_N(t) \leq 2 \cdot L \cdot h_N$ für alle $t \in \mathbb{R}$ und alle $N \in \mathbb{N}$, also folgt

$$\sup_{t \in \mathbb{R}} \left\{ A_N(t) \right\} = \tilde{A}_N \leq 2 \cdot L \cdot h_N \quad \text{für alle } N \in \mathbb{N} .$$

Mit (33) und (34) erhält man daraus für alle $N \in \mathbb{N}$:

$$j(F) \leq \sup_{t \in \mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} \leq \tilde{A}_N + j(F) \leq 2 \cdot L \cdot h_N + j(F) \text{, also}$$

$$0 \leq \sup_{t \in \mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} - j(F) \leq 2 \cdot L \cdot h_N .$$

Die Vorgabe h_N ermöglicht die Angabe der o.g. Mindest-Konvergenzgeschwindigkeit.

(c) Die angegebenen Obergrenzen für $\dfrac{\lambda_N}{N}$ sind unter relativ milden Bedingungen für F besser

als die „triviale“ Obergrenze $\dfrac{N-1}{N}$.

Gemäß (35) und (15) gilt mit beliebiger, monoton fallender, echt positiver Nullfolge $(h_\nu)_{\nu \in \mathbb{N}}$

und für jedes $N \in \mathbb{N}$:

$$\frac{\lambda_N}{N} \leq 2 \cdot D_N + j(F) + \tilde{A}_N .$$

(c.1) Unter der Voraussetzung

$$j(F) \;<\; 1 - \left[\frac{1}{N} + 4 \cdot \sqrt{\frac{\ln\ln N}{N}} + \tilde{A}_N \right] \tag{45}$$

gilt für die $\frac{\lambda_N}{N}$ - Obergrenze aus (38) bzw. (43) :

$$4 \cdot \sqrt{\frac{\ln\ln N}{N}} + j(F) + \tilde{A}_N \;<\; \frac{N-1}{N} \;.$$

Die rechte Seite von (45) kommt mit wachsendem N der 1 zwar relativ langsam, aber doch beliebig nahe. Außerdem gilt stets $0 \leq j\,(F) \leq 1$. Daher stellt (45) für nicht zu kleine $N \in \mathbb{N}$ eine relativ milde Bedingung dar.

(c.2) Unter der Voraussetzung

$$j(F) \;<\; 1 - \left[\frac{1}{N} + \frac{2}{N^{\theta}} + \tilde{A}_N \right] \tag{46}$$

gilt für die $\frac{\lambda_N}{N}$ - Obergrenze aus (41) bzw. (44) :

$$\frac{2}{N^{\theta}} + j(F) + \tilde{A}_N \;<\; \frac{N-1}{N} \;.$$

Auch (46) stellt eine relativ milde Bedingung für F dar.

Als nächstes wird ein Satz formuliert, der eine Beziehung herstellt zwischen der maximalen Sprunghöhe von F_N (= $\frac{\lambda_N}{N}$) und der maximalen Sprunghöhe der Grenzfunktion F (= j (F)).

Satz II.2

Es wird definiert:

$\tilde{A}_N$ wie in (32) , D_N wie in (17) , $\mathcal{N}$ wie in (36) , $b_N = B\,(z_1, \ldots, z_N)$ wie in (23) .

Mit $\mathcal{M}_p \subset \mathbb{R}$ wird die Menge aller Folgen $(z_\nu)_{\nu \in \mathbb{N}}$ bezeichnet, für die gilt:

(a) $(z_\nu)_{\nu \in \mathbb{N}} \in \left\{ \mathbb{R}^{\mathbb{N}} \setminus \mathcal{N} \right\}$.

(b) Es gibt ein $p > 0$, so dass $\frac{1}{N^p} \leq b_N$ für alle hinreichend großen $N \in \mathbb{N}$.

Die Werte $z_1, \ldots, z_N$ mögen zu einer Folge aus $\mathcal{M}_p$ gehören.

Die Sprungstellenmenge S(F) besitze keinen endlichen Häufungspunkt.

Dann gilt mit der maximalen Sprunghöhe j (F) von F :

$$\left| \frac{\lambda_N}{N} - j(F) \right| = o(1) \ , \ N \to +\infty . \tag{47}$$

Beweis:

Setze $h_N = \frac{1}{N^q}$ mit einem $q > p$, $N \in \mathbb{N}$.

Wegen Lemma II.1 ist dann $0 < 2 \cdot h_N < b_N$ erfüllt.

Mit (43) folgt daher für alle hinreichend großen $N \in \mathbb{N}$:

$$-4 \cdot \sqrt{\frac{\ln\ln N}{N}} + j(F) \leq \frac{\lambda_N}{N} \leq 4 \cdot \sqrt{\frac{\ln\ln N}{N}} + \tilde{A}_N + j(F) \ . \tag{48}$$

Gemäß Lemma II.2 ist $\left(\tilde{A}_\nu \right)_{\nu \in \mathbb{N}}$ eine Nullfolge.

Damit ist alles gezeigt.

❑

Bemerkung zu Satz II.2

Die Satzaussage (47) gilt auch, wenn die Voraussetzung (b) ersetzt wird durch die Voraussetzung $S(F) = \emptyset$. Darauf wird im nächsten Abschnitt eingegangen.

3.3.2 Die Verteilungsfunktion F ist stetig in $\mathbb{R}$

Satz II.3 (Modifikation von Satz II.2)

Es gelten die Definitionen und Voraussetzungen von Satz II.2 ;

die Schrittweiten-Bedingung (b) wird aber ersetzt durch die Voraussetzung,

dass F in ganz $\mathbb{R}$ stetig ist.

Dann gilt:

$$\frac{\lambda_N}{N} = o(1) \ , \ N \to +\infty . \tag{49}$$

Beweis:

Bei stetiger Verteilungsfunktion F ist $j(F) = 0$. Dann folgt aus (38) für alle hinreichend großen $N \in \mathbb{N}$:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq 4 \cdot \sqrt{\frac{\ln\ln N}{N}} + \tilde{A}_N \quad . \tag{50}$$

Wegen Lemma II.2 ist $\left(\tilde{A}_\nu\right)_{\nu \in \mathbb{N}}$ eine Nullfolge, und damit folgt (49).

❑

Bei stetiger Verteilungsfunktion F und nicht-constantem Merkmal sind die Werte $z_1, \ldots, z_N$ in der Regel paarweise verschieden, d.h. es ist $\lambda_N = 1$ der Regelfall und somit eine Abschätzung von $\frac{\lambda_N}{N}$ nur für Ausnahmefälle erforderlich.

Es sollen aber trotzdem zwei spezielle Fälle erwähnt werden,

(1.) Ist F Lipschitz-stetig mit einer globalen Lipschitz-Konstanten $L > 0$, läßt sich (50) konkretisieren. Es gilt für jedes $N \in \mathbb{N}$ $(1 < m < N)$:

$$\sup_{t \in \mathbb{R}} \left\{ F(t+h_N) - F(t-h_N) \right\} \leq L \cdot 2 \cdot h_N \quad .$$

Mit $h_N = \frac{1}{N^q}$ $(q > 0)$ erhält man daraus zusammen mit (16) für jedes $N \in \mathbb{N}$:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq 2 \cdot D_N + 2 \cdot L \cdot \frac{1}{N^q} \quad . \tag{51}$$

Je nachdem, ob die Zufallsgröße $\mathbb{D}_N$ fast sicher oder mit einer Mindest-Wahrscheinlichkeit abgeschätzt wird, erhält man aus (51) mit der Realisation D_N eine entsprechend gültige Abschätzung von $\frac{\lambda_N}{N}$ $(1 < m < N)$.

Aus (20) und (51) folgt, dass für alle hinreichend großen $N \in \mathbb{N}$ gilt:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq 4 \cdot \sqrt{\frac{\ln\ln N}{N}} + 2 \cdot L \cdot \frac{1}{N^q} \quad . \tag{52}$$

Es sei denn, die Werte $z_1, \ldots, z_N$ gehören zu einer Ausnahmefolge $(z_i)_{i \in \mathbb{N}} \in \mathcal{N}$ mit $\mathcal{N}$ wie in (36) definiert.

Die Anwendung von (21) mit $d = \frac{1}{N^{\theta}}$, $0 < \theta < ½$ und einer nicht von F abhängigen reellen Konstanten $C > 0$ liefert zusammen mit (51) für jedes $N \in \mathbb{N}$:

Zur Abschätzung $\frac{2}{N} \leq \frac{\lambda_N}{N} \leq \frac{2}{N^q} + 2 \cdot L \cdot \frac{1}{N^q}$ (53)

darf man ein Vertrauen von mindestens $1 - \frac{C}{e^{2 \cdot N^{1-2\cdot\theta}}}$ haben.

(2.) Ist $\mathcal{F}$ absolut stetig (bezüglich des Lebesgue-Maßes im $\mathbb{R}^1$) und besitzt eine gleichmäßig stetige Dichte f , ist die $\frac{\lambda_N}{N}$ - Abschätzung auf ein Theorem von E.A.Nadaraya (1965, S. 186) zurückführbar.

Es gilt gemäß (15) mit beliebigem $h_N > 0$ für jedes $N \in \mathbb{N}$:

$$\begin{aligned}
\frac{\lambda_N}{N} &\leq \sup_{t\in\mathbb{R}} \left\{ F_N(t+h_N) - F_N(t-h_N) \right\} \\
&= \sup_{t\in\mathbb{R}} \left\{ \int_{\mathbb{R}} 1_{(t-h_N\,;\,t+h_N]}(u)\, dF_N(u) \right\} \\
&= 2\cdot h_N \cdot \sup_{t\in\mathbb{R}} \left\{ \frac{1}{2\cdot h_N} \cdot \int_{\mathbb{R}} 1_{(-1\,;\,+1]}\left(\frac{u-t}{h_N}\right) dF_N(u) \right\} \\
&= 2\cdot h_N \cdot \sup_{t\in\mathbb{R}} \left\{ \frac{1}{2\cdot h_N} \cdot \sum_{i=1}^{N} 1_{(-1;+1]}\left(\frac{z_i - t}{h_N}\right) \right\} .
\end{aligned}$$

Mit einem beliebig gewählten, festen $q \in (0\,;\,½)$ in $h_\nu = \frac{1}{\nu^q}$ folgt aus dem Nadaraya-Theorem:

$$\sup_{t\in\mathbb{R}} \left\{ \frac{1}{2\cdot h_\nu} \cdot \sum_{i=1}^{\nu} 1_{(-1;+1]}\left(\frac{z_i - t}{h_\nu}\right) \right\} \underset{\nu\to+\infty}{\longrightarrow} \sup_{t\in\mathbb{R}} \left\{ f(t) \right\} .$$

Dann gibt es eine positiv-reelle Nullfolge $(c_\nu)_{\nu\in\mathbb{N}}$ und eine Konstante $K \in \mathbb{R}^+$, so dass für jedes $N \in \mathbb{N}$ gilt $(1 < m < N)$:

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq 2\cdot h_N \cdot \left[\sup_{t\in\mathbb{R}} \left\{ f(t) \right\} + c_N \right] \leq \frac{K}{N^q} . \tag{54}$$

3.3.3 Die $\frac{\lambda_N}{N}$ - Abschätzung durch einen weiteren Datensatz bei beliebigem F

Wie bisher sind $Z_1, \dots, Z_N$ unabhängige und identisch gemäß $\mathcal{F}$ verteilte Zufallsgrößen, $z_1, \dots, z_N$ sind deren Realisationen mit den darunter paarweise verschiedenen Werten $a_1, \dots, a_m$ und der maximalen relativen Häufigkeit $\frac{\lambda_N}{N}$ $(1 < m < N,\ N \in \mathbb{N})$.

Wie in (2) definiert, ist $\quad F_N(t) = \frac{1}{N} \cdot \sum_{i=1}^{N} 1_{(-\infty\,;\,t]}(z_i) \quad , \quad t \in \mathbb{R}$.

An die Verteilungsfunktion F sind im folgenden keine Bedingungen geknüpft (insbesondere wird j (F) nicht benötigt) , so dass also die Methoden der $\frac{\lambda_N}{N}$ - Abschätzung aus den Abschnitten 3.3.1 und 3.3.2 nicht anwendbar sind.

Es wird eine Abschätzung von $\frac{\lambda_N}{N}$ vorgestellt, die auf einem weiteren Datensatz basiert.

Seien $U_1, \dots, U_\nu$ ebenfalls unabhängige und identisch gemäß $\mathcal{F}$ verteilte Zufallsgrößen, und $u_1, \dots, u_\nu$ deren Realisationen mit den darunter paarweise verschiedenen Werten $c_1, \dots, c_\mu$ und der maximalen relativen Häufigkeit $\frac{\ell_\nu}{\nu}$ $(1 \le \mu \le \nu,\ \nu \in \mathbb{N})$. Für die Realisation der zugehörigen Empirischen Verteilungsfunktion gilt entsprechend

$$G_\nu(t) = \frac{1}{\nu} \cdot \sum_{i=1}^{\nu} 1_{(-\infty\,;\,t]}(u_i) \quad , \quad t \in \mathbb{R} \ .$$

In Analogie zu F_N gilt mit $\tilde{\chi} := (U_i)_{i \in \mathbb{N}}$:

Es gibt eine Menge $\mathcal{W} \subset \mathbb{R}^{\mathbb{N}}$ mit $\mathrm{IP}(\tilde{\chi})^{-1}(\mathcal{W}) = 0$, so dass nach dem Satz von Glivenko / Cantelli G_ν mit $\nu \to +\infty$ und allen Folgen $(u_i)_{i \in \mathbb{N}} \in \left\{ \mathbb{R}^{\mathbb{N}} \setminus \mathcal{W} \right\}$ gleichmäßig gegen F konvergiert.

Mit den Realisationen D_N und D_ν der Kolmogorov / Smirnov - Distance folgt:

$$\begin{aligned}
\sup_{x \in \mathbb{R}} \left\{ F_N(x) - G_\nu(x) \right\} &= \sup_{x \in \mathbb{R}} \left\{ F_N(x) - F(x) + F(x) - G_\nu(x) \right\} \\
&\le \sup_{x \in \mathbb{R}} \left\{ F_N(x) - F(x) \right\} + \sup_{x \in \mathbb{R}} \left\{ G_\nu(x) - F(x) \right\} \\
&= D_N + D_\nu \ .
\end{aligned}$$

Daraus ergeben sich mit beliebigen $x, y \in \mathbb{R}$ die beiden Beziehungen:

$$G_\nu(x) - D_N - D_\nu \;\le\; F_N(x) \;\le\; G_\nu(x) + D_N + D_\nu \;, \tag{55}$$

$$-G_\nu(y) - D_N - D_\nu \;\le\; -F_N(y) \;\le\; -G_\nu(y) + D_N + D_\nu \;. \tag{56}$$

Setze $x = t + h_N$ und $y = t - h_N$, es wird zur Abkürzung definiert:

$$W_N(h_N\,;t) \;:=\; F_N(t+h_N) - F_N(t-h_N) \;,$$
$$V_\nu(h_N\,;t) \;:=\; G_\nu(t+h_N) - G_\nu(t-h_N) \;.$$

Damit folgt aus (55) und (56) :

$$V_\nu(h_N\,;t) - 2\cdot D_N - 2\cdot D_\nu \;\le\; W_N(h_N\,;t) \;\le\; V_\nu(h_N\,;t) + 2\cdot D_N + 2\cdot D_\nu \,. \tag{57}$$

Wegen $0 \le W_N(h_N; t) \le 1$ und $0 \le V_\nu(h_N; t) \le 1$ existieren $t_1, t_2 \in \mathbb{R}$, so dass gilt:

$$\sup_{t\in\mathbb{R}} \left\{ W_N(h_N\,;t) \right\} \;=:\; W_N(h_N\,;t_1) \;,$$

$$\sup_{t\in\mathbb{R}} \left\{ V_\nu(h_N\,;t) \right\} \;=:\; V_\nu(h_N\,;t_2) \;.$$

Damit folgt aus (57) :

$$V_\nu(h_N\,;t_2) - 2\cdot D_N - 2\cdot D_\nu \;\le\; W_N(h_N\,;t_2) \;\le\; W_N(h_N\,;t_1) \;,$$

$$W_N(h_N\,;t_1) \;\le\; V_\nu(h_N\,;t_1) + 2\cdot D_N + 2\cdot D_\nu \;\le\; V_\nu(h_N\,;t_2) + 2\cdot D_N + 2\cdot D_\nu \;,$$

und es gilt daher:

$$\sup_{t\in\mathbb{R}}\left\{ V_\nu(h_N;t)\right\} - 2\cdot D_N - 2\cdot D_\nu \;\le\; \sup_{t\in\mathbb{R}}\left\{ W_N(h_N;t)\right\} \;\le\; \sup_{t\in\mathbb{R}}\left\{ V_\nu(h_N;t)\right\} + 2\cdot D_N + 2\cdot D_\nu \;.$$

Wie in (23) wird definiert für $i, j \in \{1, \ldots, \nu\}$:

$$\beta_\nu \;:=\; \min_{u_i \ne u_j} \left\{ \left| \, u_i - u_j \, \right| \right\} \;.$$

Dann gilt wie in (12) :

Sofern $0 < 2 \cdot h_N < \beta_\nu$ ist, folgt die Identität

$$\frac{l_\nu}{\nu} \;=\; \sup_{t\in\mathbb{R}} \left\{ V_\nu(h_N\,;t) \right\} \;.$$

Da man die Daten $u_1, \ldots, u_\nu$ kennt, ist auch β_ν bekannt. Daher kann die Schrittweite h_N so gewählt werden, dass $0 < 2 \cdot h_N < \beta_\nu$ erfüllt ist.

Es folgt:

$$\frac{l_\nu}{\nu} - 2\cdot D_N - 2\cdot D_\nu \quad \le \quad \sup_{t\in\mathbb{R}}\left\{ W_N(h_N;t)\right\} \quad \le \quad \frac{l_\nu}{\nu} + 2\cdot D_N + 2\cdot D_\nu \ . \tag{58}$$

Für die $\frac{\lambda_N}{N}$ - Abschätzung nach unten wird wieder danach unterschieden, ob $2\cdot h_N < b_N$ erfüllt ist, oder nicht.

Unter der Voraussetzung, dass $2\cdot h_N < b_N$ gültig ist, folgt aus (58) zusammen mit (12) für beliebige $N \in \mathbb{N}$:

$$\frac{l_\nu}{\nu} - 2\cdot D_N - 2\cdot D_\nu \quad \le \quad \frac{\lambda_N}{N} \quad \le \quad \frac{l_\nu}{\nu} + 2\cdot D_N + 2\cdot D_\nu \ . \tag{59}$$

Ist unbekannt, ob $2\cdot h_N < b_N$ erfüllt ist, wird die linke Seite von (59) durch $\frac{2}{N}$ ersetzt, und es folgt für beliebige $N \in \mathbb{N}$ $(1 < m < N)$ die Abschätzung:

$$\frac{2}{N} \le \frac{\lambda_N}{N} \le \frac{l_\nu}{\nu} + 2\cdot D_N + 2\cdot D_\nu \ . \tag{60}$$

Die Abschätzung von D_N und D_ν kann getrennt erfolgen, jeweils durch Anwendung von (20) oder (21) wie im Abschnitt 3.3.1 , und man erhält aus (59) und (60) entsprechende Versionen der $\frac{\lambda_N}{N}$ - Abschätzung.

Die Abschätzung von D_N und D_ν kann aber auch simultan durch Kombination von (20) und (21) erfolgen. Daraus ergibt sich zusammen mit (60) die fogende $\frac{\lambda_N}{N}$ - Abschätzung.

Seien $r, s \in \{N, \nu\}$ mit $r \ne s$.

Wegen (20) und (39) gilt mit einer nicht von F abhängigen reellen Konstanten $C > 0$ und einem beliebig gewählten $\theta \in (0 ; ½)$ für jedes s und alle hinreichend großen r :

$$\begin{aligned} 1 - \frac{C}{e^{2\cdot r^{1-2\cdot\theta}}} &\le P\left(\mathbb{D}_r \le \frac{1}{r^\theta} \right) = P\left(2\cdot\mathbb{D}_r + 2\cdot\mathbb{D}_s \le \frac{2}{r^\theta} + 2\cdot\mathbb{D}_s \right) \\ &\le P\left(2\cdot\mathbb{D}_r + 2\cdot\mathbb{D}_s \le \frac{2}{r^\theta} + 4\cdot\sqrt{\frac{\ln\ln s}{s}} \right) \\ &= P\left(\frac{l_\nu}{\nu} + 2\cdot\mathbb{D}_r + 2\cdot\mathbb{D}_s \le \frac{l_\nu}{\nu} + \frac{2}{r^\theta} + 4\cdot\sqrt{\frac{\ln\ln s}{s}} \right) . \end{aligned}$$

Das heißt, es existiert ein $C > 0$, so dass man für jedes s und für alle hinreichend großen r zur Abschätzung

$$\frac{2}{N} \leq \frac{\lambda_N}{N} \leq \frac{\ell_\nu}{\nu} + \frac{2}{r^\theta} + 4 \cdot \sqrt{\frac{\ln \ln s}{s}} \tag{61}$$

ein Mindest-Vertrauen von $1 - \dfrac{C}{e^{2 \cdot r^{1-2\cdot\theta}}}$ haben darf, $\theta \in (0\,;\,½)$ beliebig wählbar.

Bemerkungen

(a) Ein Vorteil der durch Kombination von (20) und (39) konstruierten Abschätzung von D_N und D_ν liegt darin, dass die Beziehung (61) durch Wahl von s und r danach ausgerichtet werden kann, ob man sie für jedes $N \in \mathbb{N}$ und alle hinreichend großen $\nu \in \mathbb{N}$ benötigt, oder umgekehrt.

Wählt man $r = N$ und $s = \nu$, dann gilt (61) für jedes $N \in \mathbb{N}$ und alle hinreichend großen $\nu \in \mathbb{N}$; wählt man $r = \nu$ und $s = N$, dann gilt (61) für jedes $\nu \in \mathbb{N}$ und alle hinreichend großen $N \in \mathbb{N}$.

(b) Die Werte $u_1, \ldots, u_\nu$ des zusätzlichen Datensatzes dürfen teilweise oder ganz mit den Unbestimmten $z_1, \ldots, z_N$ zusammenfallen.

(c) Ein besonderer Fall liegt vor, wenn $\nu = N$ ist und für jedes $i \in \{1, \ldots, N\}$ ein $j \in \{1, \ldots, N\}$ existiert, so dass $U_i = Z_j$ gilt.

In diesem besonderen Fall sind die $U_1, \ldots, U_N$ Bootstrap-Variablen.

Mit $\dfrac{\mathbb{L}_N}{N}$ wird die Zufallsgröße bezeichnet, deren Realisation $\dfrac{\lambda_N}{N}$ ist. Unter der Voraussetzung $2 \cdot h_N < B(Z_1, \ldots, Z_N)$ gilt wegen (12) die Darstellung:

$$\frac{\mathbb{L}_N}{N} = \sup_{t \in \mathbb{R}} \left\{ IF_N(t + h_N) - IF_N(t - h_N) \right\} \quad , \; IF_N \text{ wie in (1) definiert.} \tag{62}$$

Dann ist $\dfrac{\ell_N}{N}$ die Realisation der mit $\dfrac{\mathbb{L}^*_N}{N}$ symbolisierten Bootstrap-Version von $\dfrac{\mathbb{L}_N}{N}$.

Da die Ungleichungskette (59) unter den Voraussetzungen $2 \cdot h_N < b_N$ und $0 < 2 \cdot h_N < \beta_\nu$ für alle möglichen Datensätze $z_1, \ldots, z_N$ und $u_1, \ldots, u_\nu$ gilt, kann man sie im vorliegenden Sonderfall für jedes $N \in \mathbb{N}$ mittels der oben definierten Zufallsgrößen unter der Voraussetzung $2 \cdot h_N < B(Z_1, \ldots, Z_N)$ so darstellen:

$$\frac{\mathbb{L}_N^*}{N} - 4\cdot\mathbb{D}_N \quad \leq \quad \frac{\mathbb{L}_N}{N} \quad \leq \quad \frac{\mathbb{L}_N^*}{N} + 4\cdot\mathbb{D}_N \quad . \tag{63}$$

Aus (63) ergibt sich für jedes $N \in \mathbb{N}$ unter der Voraussetzung $2 \cdot h_N < B(Z_1, \ldots, Z_N)$:

$$\left| \frac{\mathbb{L}_N}{N} - \frac{\mathbb{L}_N^*}{N} \right| \quad \leq \quad 4\cdot\mathbb{D}_N \quad . \tag{64}$$

Zusammen mit (20) erhält man daraus $IP\chi^{-1}$- fast sicher für alle hinreichend großen $N \in \mathbb{N}$:

$$\left| \frac{\mathbb{L}_N}{N} - \frac{\mathbb{L}_N^*}{N} \right| \quad \leq \quad 8 \cdot \sqrt{\frac{\ln\ln N}{N}} \quad . \tag{65}$$

Wegen (43) und (50) ist $\frac{\mathbb{L}_N}{N}$ unter der Voraussetzung $2 \cdot h_N < B(Z_1, \ldots, Z_N)$ mit einer monoton fallenden, echt positiven Nullfolge h_N eine stark konsistente Schätzfunktion für die maximale Sprunghöhe $j(F)$, und wegen (64) gilt das auch für die Bootstrap-Version $\frac{\mathbb{L}_N^*}{N}$.