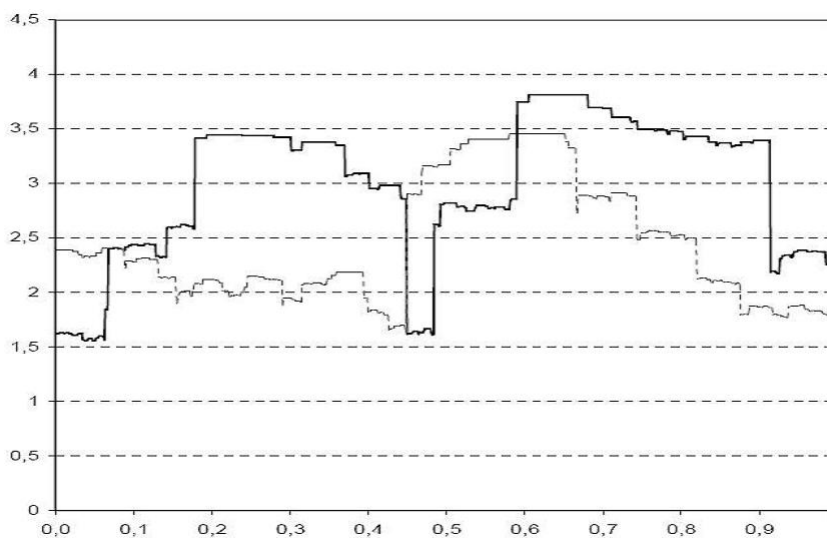


Stochastische Analyse des FIND-Algorithmus

Diplomarbeit

vorgelegt von
Martin Düllmann



Betreuer: Prof. Dr. Gerold Alsmeyer
Institut für Mathematische Statistik
Fachbereich 10 - Mathematik und Informatik
Westfälische-Wilhelms-Universität Münster

Inhaltsverzeichnis

Einleitung	v
1 Metrische Räume	1
1.1 Grundlagen	1
1.2 Der Raum $D[0,1]$ und die Skorokhod-Topologie	3
1.3 Der Raum der beschränkten, messbaren Funktionen	8
1.4 Die Mallows-Metrik	9
2 Gewichtete Verzweigungsprozesse	13
2.1 Bäume	13
2.2 Gewichtete Verzweigungsprozesse	14
2.3 Vorwärts- und Rückwärtsgleichungen	16
3 Die Kontraktionsmethode	19
3.1 Das Prinzip der Kontraktionsmethode	19
3.2 Die Kontraktionsmethode bei Divide-and-Conquer Algorithmen	20
4 Der FIND-Prozess	25
4.1 Die Konstruktion des FIND-Prozesses als gewichteter Verzweigungsprozess	25
4.2 Der FIND-Prozess als zufällige Binärdarstellung einer reellen Zahl	27
5 Konvergenz in Verteilung gegen den FIND-Prozess	29
5.1 Konstruktion des FIND-Prozesses	30
5.2 Schwache Konvergenz gegen den FIND-Prozess	35
5.3 Eigenschaften des FIND-Prozesses	51
6 Ein Markov-Ketten Ansatz für FIND	61
6.1 Konstruktion der begleitenden Markov-Ketten	61
6.2 Schwache Konvergenz und Abschätzungen der Tail-Wahrscheinlichkeiten .	63
6.3 Numerische Ergebnisse	74
Anhang	79
Literaturverzeichnis	79

Einleitung

Hoare hat 1961 in [14] den stochastischen Selektionsalgorithmus **FIND** eingeführt. Dieser dient dazu, aus einer endlichen Menge S , bestehend aus n verschiedenen Elementen, das l -kleinste Element zu finden. Basierend auf den Artikeln [13] und [12] untersucht die vorliegende Arbeit zwei verschiedene Vorgehensweisen zur Analyse der Laufzeit dieses Algorithmus.

Find geht dabei ähnlich vor wie der bekannte Algorithmus Quicksort. Zuerst wird aus der zu untersuchenden Menge S , welche wir ohne Einschränkung als $\{1, \dots, n\}$ annehmen können, ein Pivotelement x nach einem gegebenen Zufallsverfahren ausgewählt. In dieser Arbeit wird davon ausgegangen, dass jedes Element von S mit derselben Wahrscheinlichkeit ausgewählt wird. Durch Vergleiche der restlichen Elemente mit x wird die Menge S in die Mengen $S_{\leq}$, der Elemente kleiner oder gleich x , und $S_{>}$, der Elemente größer als x , aufgeteilt. Sei m die Anzahl der Elemente in $S_{\leq}$.

- Ist $m \geq l$, so liegt l in $S_{\leq}$, und es wird mittels **FIND** rekursiv das l -kleinste Element von $S_{\leq}$ gesucht.
- Ist $m < l$, so liegt l in $S_{>}$, und es wird mittels **FIND** rekursiv das $(l - m)$ -kleinste Element von $S_{>}$ gesucht.

Falls als Pivotelement nicht das größte Element ausgesucht wird, verringert sich die Größe der betrachteten Menge in jedem Rekursionsschritt zumindest um Eins, also terminiert der Algorithmus fast sicher. Diese Version von **FIND** werden wir im Folgenden als *2-Mengen Version* bezeichnen.

Nach einer leichten Modifikation des Algorithmus ergibt sich die *3-Mengen Version* von **FIND**. Hierbei wird S in die drei Mengen $S_{<}$, $S_{=}$ sowie $S_{>}$ aufgeteilt. Dieser Algorithmus terminiert bereits, wenn die Anzahl der Elemente in $S_{<}$ gleich $l - 1$ ist, und gibt dann x als das gesuchte Element aus.

Offensichtlich führt die 3-Mengen Version schneller zu einem Ergebnis, wobei signifikante Unterschiede in der Performance für große n sehr unwahrscheinlich sind. Trotzdem stellt sich die Frage nach der Notwendigkeit der 2-Mengen Version. Hier sei nur kurz darauf hingewiesen, dass der in Kapitel 5 dargestellte Ansatz zur Analyse der Laufzeit von **FIND** eine Betrachtung der 2-Mengen Version erfordert.

Unter Vernachlässigung von Faktoren wie der Rechenleistung des Computers und der Qualität der Programmierung werden wir die Anzahl der benötigten Vergleiche als Maß für die Laufzeit von FIND verwenden. Da die Anzahl der benötigten Vergleiche nicht von der Menge S selbst, sondern nur von n , der Anzahl der Elemente in S , abhängt, werden wir sie im Folgenden mit $C_{n,l}$ für die 2-Mengen Version und $C_{n,l}^*$ für die 3-Mengen Version bezeichnen. $C_{n,l}$ hängt offensichtlich größtenteils von der Aufteilung der Menge S in die Mengen $S_{\leq}$ und $S_{>}$ ab, da im nächsten Schritt die Elemente aus einer der beiden Mengen mit dem nächsten Pivotelement verglichen werden müssen. Wegen der Zufälligkeit der Wahl des Pivotelements ist die Laufzeit $C_{n,l}$ eine Zufallsgröße, die stark von dem Zufallsverfahren zur Wahl des Pivotelements abhängt.

Dementsprechend sind auch andere Verfahren zur Wahl des Pivotelements denkbar, die zu besseren Ergebnissen als die obige Version führen. Hier sei auf die *Median-von- k Version* hingewiesen. Dabei werden zuerst k Elemente aus S zufällig ausgewählt und dann der Median dieser k Elemente bestimmt. Der Median wird dann als Pivotelement verwendet. Alternativ zum Median kann man auch dasjenige Element der k Zahlen wählen, das einen Rang nahe $\frac{l \cdot k}{n}$ hat. Dies führt dazu, dass die Wahrscheinlichkeit, ein Pivotelement nahe des gesuchten Elements l zu finden, wächst und damit im nächsten Rekursionsschritt weniger Vergleiche notwendig sind. Allerdings ist bei beiden Versionen zu beachten, dass die Auswahl des Pivotelements ebenfalls eine gewisse Anzahl an Vergleichen erfordert. Für eine genauere Analyse siehe beispielsweise [18] oder [21].

Nun zurück zur 3-Mengen Version von FIND. Offensichtlich benötigt diese im besten Fall, d.h. wenn l im ersten Schritt als Pivotelement ausgewählt wird, $n - 1$ Vergleiche und im schlechtesten Fall $\sum_{i=1}^{n-1} i = \frac{(n-1)n}{2}$. Als Erwartungswert hat Knuth in [19]

$$E(C_{n,l}^*) = 2(n+3 + (n+1)H_n - (l+2)H_l - (n+2-l)H_{n-l+1})$$

mit $H_n \stackrel{\text{def}}{=} \sum_{i=1}^n \frac{1}{i}$ bestimmt. Der große Unterschied von Average, Best and Worst Case Performance verlangt nach einer detaillierteren Analyse der Laufzeit. Einen Schritt auf diesem Weg hat Devroye in [9] beschritten, indem er exponentielle Schranken für die Tails der Verteilung von $C_{n,l}^*$ bewiesen hat. Genauer zeigt er dort, dass

$$P(C_{n,k}^* > zn) = o(\rho^z)$$

für alle $\rho > \frac{3}{4}$ gilt. Dies zeigt, dass das Auftreten von schlechter Performance für große n unwahrscheinlich ist.

In dieser Arbeit wird es darum gehen, das asymptotische Verhalten von $C_{n,l}$ bzw. $C_{n,l}^*$ für große n zu untersuchen. Es stellt sich beispielsweise die Frage, ob Folgen $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ existieren, so dass $a_n(C_{n,l} - b_n)$ gegen einen nicht-degenerierten Grenzwert konvergiert. Zur positiven Beantwortung dieser Frage sind mehrere Ansätze bekannt. Zwei davon werden wir in dieser Arbeit vorstellen und jeweils Eigenschaften des Grenzprozesses,

welchen wir im Folgenden als *FIND-Prozess* bezeichnen werden, beweisen. Bevor jedoch damit begonnen werden kann, werden noch einige hilfreiche Grundlagen bereitgestellt.

Das **erste Kapitel** dient der Wiederholung einiger Begriffe bezüglich metrischer Räume und stellt danach einige spezielle Räume vor, die im späteren Verlauf der Arbeit eine Rolle spielen werden. Hierbei sei vor allem die Skorokhod-Topologie auf dem Raum der rechtsseitig stetigen Funktionen mit linksseitigen Limiten genannt.

Im **zweiten Kapitel** werden dann gewichtete Verzweigungsprozesse eingeführt, die eine geeignete Struktur zur Beschreibung von rekursiven Algorithmen bieten.

Daran anschließend wird die Kontraktionsmethode vorgestellt. Dies ist eine Methode zum Nachweis der Konvergenz von Folgen von Wahrscheinlichkeitsmaßen und findet Anwendung bei der Analyse von Divide-and-Conquer Algorithmen, wie Quicksort und FIND.

Das **vierte Kapitel** dient dann der Verdeutlichung der Zusammenhänge von gewichteten Verzweigungsprozessen und dem FIND-Prozess und leistet somit einen Beitrag zum Verständnis der späteren detaillierten Analysen von FIND.

Diese beginnen im **fünften Kapitel** mit der Präsentation des von Grübel und Rösler in [13] dargestellten Vorgehens. Wir starten hierbei zuerst mit der Konstruktion des FIND-Prozesses und zeigen dann im zweiten Abschnitt das Hauptresultat dieser Arbeit, die Verteilungskonvergenz von $\frac{1}{n}C_{n,l}$ gegen den FIND-Prozess $Z(t)$ bezüglich der Skorokhod-Topologie, falls l mit n so variiert, dass $\frac{l}{n}$ gegen ein $t \in [0, 1]$ konvergiert. Im dritten Abschnitt werden noch einige Eigenschaften des FIND-Prozesses vorgestellt.

Der zweiten Methode zur Analyse von FIND, zurückgehend auf Grübels Artikel [12], widmet sich das **sechste Kapitel**. Das Vorgehen dabei beruht auf der Konstruktion von geeigneten begleitenden Markov-Ketten und ermöglicht einen deutlich kürzeren Beweis des Hauptresultats aus Kapitel 5. Es werden hier auch Schranken für die Tails von $C_{n,l}^*$ und dem FIND-Prozess bewiesen, die mittels numerischer Berechnungen zu besseren Abschätzungen der Tail-Wahrscheinlichkeiten führen.

Ich danke Herrn Prof. Dr. Gerold Alsmeyer für die Bereitstellung des Themas und die Betreuung während der Anfertigung dieser Diplomarbeit. Weiterhin gilt mein Dank allen Menschen, die auf ihre Weise zur Erstellung dieser Arbeit beigetragen haben, insbesondere meinen Eltern für die Unterstützung während meiner gesamten Studienzzeit.

Hiermit versichere ich, dass ich diese Diplomarbeit selbstständig verfasst und keine anderen als die im Literaturverzeichnis aufgeführten Quellen und Hilfsmittel verwendet habe.

Münster, den 30. Dezember 2009

Martin Düllmann

Kapitel 1

Metrische Räume

In diesem Kapitel werden wir zuerst einige Definitionen für metrische Räume bereitstellen und den Banachschen Fixpunktsatz, der einen wichtigen Bestandteil der Kontraktionsmethode darstellt, beweisen. Im Anschluss werden drei spezielle metrische Räume mit jeweiligen Eigenschaften kurz dargestellt. Dabei werden wir vor allem den Raum der rechtsseitig stetigen Funktionen mit linksseitigen Limiten betrachten, da dieser eine wichtige Rolle in der untersuchten Arbeit [13] spielt. Wir beginnen nun mit den Grundlagen.

1.1 Grundlagen

Eine *Metrik* auf einer Menge X ist eine Funktion $d : X \times X \rightarrow [0, \infty)$, die symmetrisch ist, die Dreiecksungleichung erfüllt und zwei Elementen $x, y \in X$ genau dann den Wert 0 zuordnet, wenn diese gleich sind. Das Tupel (X, d) heißt dann *metrischer Raum*. Einen solchen nennt man *vollständig*, wenn in ihm jede Cauchyfolge konvergiert. Desweiteren heißt er *separabel*, wenn er eine abzählbare, dichte Teilmenge besitzt.

Einen vollständigen, normierten Vektorraum bezeichnet man als *Banachraum*. Jeder abgeschlossene Untervektorraum U eines solchen Banachraums X ist wiederum vollständig, da jede Cauchyfolge in U auch eine solche in X ist, entsprechend konvergiert und der Grenzwert aufgrund der Abgeschlossenheit von U wieder in U liegt.

Eine stetige Abbildung $f : X \rightarrow X$ eines metrischen Raums (X, d) in sich selbst heißt *Kontraktion*, wenn ein $\alpha \in [0, 1)$ existiert, so dass

$$d(f(x), f(y)) \leq \alpha d(x, y)$$

für alle $x, y \in X$ gilt. Definiert man nun für eine Kontraktion f eines metrischen Raums (X, d) und ein beliebiges $x_0 \in X$ die Iterationsfolge

$$x_n \stackrel{\text{def}}{=} f(x_{n-1}) = f_n(x_0),$$

wobei f_n die n -fache Hintereinanderschaltung von f bezeichnet, so schrumpft der Abstand zweier Iterationsfolgen $(f_n(x))_{n \in \mathbb{N}}$ und $(f_n(y))_{n \in \mathbb{N}}$ unter f geometrisch schnell, denn aus der Kontraktionseigenschaft folgt

$$d(f_n(x), f_n(y)) \leq \alpha^n d(x, y). \quad (1.1)$$

Der nun folgende *Banachsche Fixpunktsatz* liefert die Konvergenz jeder solchen Folge gegen einen eindeutigen Fixpunkt, solange der Raum (X, d) auch vollständig ist.

1.1 Satz (Banachscher Fixpunktsatz)

Jede Kontraktion $f : X \rightarrow X$ eines vollständigen metrischen Raums (X, d) besitzt genau einen Fixpunkt $\xi \in X$.

Beweis:

Wir beweisen zunächst die Existenz des Fixpunktes. Wähle hierzu ein beliebiges $x_0 \in X$ und setze $x_n = f(x_{n-1})$ für $n \geq 1$. Dann folgt mit der Kontraktionseigenschaft für alle $k \geq 1$

$$d(x_{k+1}, x_k) = d(f(x_k), f(x_{k-1})) \leq \alpha d(x_k, x_{k-1})$$

für ein $\alpha \in [0, 1)$ und somit nach Iteration

$$d(x_{k+1}, x_k) \leq \alpha^k d(x_1, x_0).$$

Nun ergibt sich für alle $m, n \in \mathbb{N}$ mit Hilfe der geometrischen Reihe

$$d(x_{m+n}, x_m) \leq \sum_{k=m}^{m+n-1} d(x_{k+1}, x_k) \leq \sum_{k=m}^{m+n-1} \alpha^k d(x_1, x_0) \leq \frac{\alpha^m}{1-\alpha} d(x_1, x_0).$$

$(x_n)_{n \in \mathbb{N}}$ bildet also eine Cauchyfolge in X und konvergiert aufgrund der Vollständigkeit des metrischen Raums (X, d) gegen ein $\xi \in X$. Dieser bildet natürlich einen Fixpunkt von f , da f als Kontraktion auch stetig ist.

Für den Beweis der Eindeutigkeit, seien ξ und ψ zwei Fixpunkte von f . Dann folgt wiederum aus der Kontraktionseigenschaft von f , dass

$$d(\xi, \psi) = d(f(\xi), f(\psi)) \leq \alpha d(\xi, \psi)$$

gilt. Somit folgt $\xi = \psi$.

□

1.2 Der Raum $D[0,1]$ und die Skorokhod-Topologie

In diesem Abschnitt wird zuerst der Raum $D[0,1]$ der rechtsseitig stetigen Funktionen mit linksseitigen Limiten eingeführt. Dieser dient in vielen Fällen der Beschreibung von stochastischen Sprungprozessen. Wir werden einige Eigenschaften dieser Funktionen angeben und im Anschluss zwei Metriken auf $D[0,1]$ definieren, die diesen zu einem vollständigen metrischen Raum machen. Dabei werden wir vor allem auf die Skorokhod-Topologie eingehen, da diese in der untersuchten Arbeit [13] eine entscheidende Rolle spielt. Wir haben uns hierbei an dem entsprechenden Kapitel aus [6] orientiert.

1.2 Definition

$D[0,1]$ sei der Raum der càdlàg Funktionen auf $[0,1]$ nach $\mathbb{R}$, d.h. der Raum der Funktionen $x : [0,1] \rightarrow \mathbb{R}$, die rechtsseitig stetig sind und deren linksseitige Limiten existieren, welche also folgende Eigenschaften erfüllen:

- (i) für $0 \leq t < 1$ existiert $x(t+) \stackrel{\text{def}}{=} \lim_{s \downarrow t} x(s)$, und es gilt $x(t+) = x(t)$
- (ii) für $0 < t \leq 1$ existiert $x(t-) \stackrel{\text{def}}{=} \lim_{s \uparrow t} x(s)$.

Für $x \in D[0,1]$ und $T_0 \subseteq [0,1]$ setze

$$w_x(T_0) \stackrel{\text{def}}{=} \sup \left\{ |x(s) - x(t)| : s, t \in T_0 \right\}$$

und für $\delta > 0$ missbräuchlich dieselbe Schreibweise benutzend

$$w_x(\delta) \stackrel{\text{def}}{=} \sup_{0 \leq t \leq 1-\delta} w_x([t, t+\delta]).$$

Bekanntlich sind stetige Funktionen auf kompakten Intervallen bereits gleichmäßig stetig. Das folgende Lemma liefert eine ähnliche Eigenschaft für die Elemente von $D[0,1]$.

1.3 Lemma

Für jedes $x \in D[0,1]$ und jedes $\epsilon > 0$ existieren Punkte $t_0, \dots, t_r$ mit $0 = t_0 < t_1 < \dots < t_r = 1$, so dass gilt:

$$w_x([t_{i-1}, t_i]) < \epsilon \quad \text{für alle } i=1, \dots, r. \quad (1.2)$$

Beweis:

Setze

$$\tau = \sup \left\{ t \in [0,1] : [0,t] \text{ kann in endlich viele Teilintervalle } [t_{i-1}, t_i] \text{ zerlegt werden,} \right. \\ \left. \text{die (1.2) erfüllen} \right\}.$$

Da $x(0) = x(0+)$ gilt, folgt bereits, dass $\tau > 0$ ist. Da aber $x(\tau-)$ existiert, kann auch

$[0, \tau)$ derart zerlegt werden.

Aus $x(\tau) = x(\tau+)$ folgt, dass $\tau < 1$ unmöglich ist. Also gilt schon $\tau = 1$.

□

1.4 Bemerkung

Aus Lemma 1.3 folgt, dass zu jeder gegebenen positiven Zahl $z \in \mathbb{R}$ nur endlich viele Punkte $t \in [0, 1]$ existieren, in denen die Sprünge $|x(t) - x(t-)| > z$ sind. Insbesondere hat x maximal abzählbar viele Unstetigkeitsstellen.

Es folgt also auch, dass x gleichmäßig beschränkt ist, d.h.

$$\sup_{t \in [0,1]} |x(t)| < \infty. \quad (1.3)$$

In der folgenden Definition werden wir eine weitere Kennzahl für càdlàg Funktionen einführen, welche sowohl eine äquivalente Formulierung von Lemma 1.3 als auch eine Charakterisierung der kompakten Mengen in der später eingeführten Skorokhod-Topologie ermöglicht.

1.5 Definition

Für $x \in D[0, 1]$ und $0 < \delta < 1$ setze

$$w'_x(\delta) \stackrel{\text{def}}{=} \inf_{\{t_i\}} \max_{0 < i \leq r} w_x([t_{i-1}, t_i]),$$

wobei das Infimum über alle $r \in \mathbb{N}$ und alle endlichen Mengen $\{t_i\}_{1 \leq i \leq r} \subseteq [0, 1]$ gebildet wird, die Folgendes erfüllen:

$$\begin{aligned} 0 &= t_0 < t_1 < \dots < t_r = 1 \\ t_i - t_{i-1} &> \delta, \quad \text{für } i = 1, \dots, r. \end{aligned}$$

1.6 Bemerkung

Lemma 1.3 lässt sich auch äquivalent über das soeben eingeführte w'_x beschreiben.

Für alle $x \in D[0, 1]$ gilt:

$$\lim_{\delta \rightarrow 0} w'_x(\delta) = 0$$

1.7 Satz

$(D[0, 1], d_\infty)$ ist ein vollständiger metrischer Raum. Dabei ist d_∞ für alle $x, y \in D[0, 1]$ definiert durch

$$d_\infty(x, y) \stackrel{\text{def}}{=} \sup_{t \in [0,1]} |x(t) - y(t)|.$$

Beweis:

Seien $x, y, z \in D[0, 1]$ beliebig. Es ist direkt klar, dass $d_\infty(x, y) \geq 0$ und $d_\infty(x, y) = d_\infty(y, x)$ gelten, sowie $d_\infty(x, y) = 0$ genau dann gilt, wenn $x = y$ ist. Die Dreiecksungleichung ergibt sich aus

$$\begin{aligned} d_\infty(x, y) + d_\infty(y, z) &= \sup_{t \in [0, 1]} |x(t) - y(t)| + \sup_{t \in [0, 1]} |y(t) - z(t)| \\ &\geq \sup_{t \in [0, 1]} |x(t) - z(t)| = d_\infty(x, z). \end{aligned}$$

Also definiert d_∞ eine Metrik auf $D[0, 1]$. Es bleibt noch zu zeigen, dass $(D[0, 1], d_\infty)$ vollständig ist. Bekanntlich ist der Vektorraum $\mathcal{B}([0, 1])$ der beschränkten Funktionen auf $[0, 1]$, ein Banachraum bzgl. d_∞ . Nach Bemerkung 1.4 ist $D[0, 1]$ ein Untervektorraum von $\mathcal{B}([0, 1])$. Gemäß den Anmerkung zu Beginn des Kapitels reicht es nun zu zeigen, dass $D[0, 1]$ als Untervektorraum von $\mathcal{B}([0, 1])$ abgeschlossen bzgl. d_∞ ist. Aus der Analysis ist allerdings bekannt, dass der gleichmäßige Limes einer Folge von càdlàg Funktionen wieder eine càdlàg Funktion ist. Somit liegt der Grenzwert einer Folge aus $D[0, 1]$ wieder in $D[0, 1]$, was wiederum bedeutet, dass $D[0, 1]$ ein bzgl. d_∞ abgeschlossener Untervektorraum von $\mathcal{B}([0, 1])$ ist. □

Die durch d_∞ induzierte Topologie ermöglicht es allerdings nicht, ein Hauptresultat aus [13] zu zeigen. Hierfür wird eine andere Topologie auf $D[0, 1]$, die sogenannte Skorokhod-Topologie, benötigt. Diese wird in folgender Definition eingeführt.

1.8 Definition und Satz

$D[0, 1]$ versehen mit $\rho : D[0, 1] \times D[0, 1] \rightarrow [0, \infty)$,

$$\begin{aligned} \rho(x, y) \stackrel{\text{def}}{=} \inf \left\{ \epsilon > 0 : \text{es existiert ein } \lambda \in \Lambda \text{ mit } \sup_{t \in [0, 1]} |\lambda(t) - t| \leq \epsilon \right. \\ \left. \text{und } \sup_{t \in [0, 1]} |x(t) - y(\lambda(t))| \leq \epsilon \right\} \end{aligned} \quad (1.4)$$

ist ein metrischer Raum. Dabei bezeichne

$$\begin{aligned} \Lambda \stackrel{\text{def}}{=} \left\{ \lambda : [0, 1] \rightarrow [0, 1] : \lambda \text{ ist stetig und streng monoton wachsend mit} \right. \\ \left. \lambda(0) = 0 \text{ und } \lambda(1) = 1 \right\}. \end{aligned}$$

Die durch ρ induzierte Topologie heißt Skorokhod-Topologie. Ein $\lambda \in \Lambda$ wird auch als Zeitwechsel bezeichnet.

Beweis:

Seien $x, y, z \in D[0, 1]$ beliebig. Um zu sehen, dass $\rho(x, y)$ endlich ist, wähle $\lambda(t) \equiv t$ und

beachte die gleichmäßige Beschränktheit von x . Weiter ist klar, dass $\rho(x, y) \geq 0$ gilt. $\rho(x, y) = 0$ impliziert, dass für alle $t \in [0, 1]$ entweder $x(t) = y(t)$ oder $x(t) = y(t-)$ gilt. Sei nun $t_0 \in [0, 1]$ beliebig. Angenommen es gelte $x(t_0) \neq y(t_0)$. Dann muss aber $x(t_0) = y(t_0-)$ gelten, also ist y unstetig in t_0 . Da aber für alle $\epsilon > 0$ entweder $x(t_0 + \epsilon) = y(t_0 + \epsilon)$ oder $x(t_0 + \epsilon) = y(t_0 + \epsilon-)$ gilt und x und y rechtsseitig stetig sind, folgt schon der Widerspruch. Also gilt $x(t_0) = y(t_0)$ und damit $x = y$.

Für den Nachweis der Symmetrie beachte, dass aus $\lambda \in \Lambda$ auch $\lambda^{-1} \in \Lambda$ folgt. Die beiden folgenden Gleichungen gelten, da die Suprema jeweils über alle $t \in [0, 1]$, also insbesondere über $\lambda(t)$ und $\lambda^{-1}(t)$, gebildet werden:

$$\begin{aligned} \sup_{t \in [0,1]} |\lambda^{-1}t - t| &= \sup_{t \in [0,1]} |\lambda t - t| \\ \sup_{t \in [0,1]} |x(\lambda^{-1}t) - y(t)| &= \sup_{t \in [0,1]} |x(t) - y(\lambda t)|. \end{aligned}$$

Also folgt $\rho(x, y) = \rho(y, x)$.

Es bleibt noch die Dreiecksungleichung zu zeigen. Dafür beachte, dass für $\lambda_1, \lambda_2 \in \Lambda$ auch $\lambda_1 \circ \lambda_2 \in \Lambda$ gilt. Damit gelten unter Berücksichtigung der obigen beiden Gleichungen folgende Ungleichungen:

$$\begin{aligned} \sup_{t \in [0,1]} |\lambda_2 \circ \lambda_1(t) - t| &\leq \sup_{t \in [0,1]} |\lambda_2(t) - t| + \sup_{t \in [0,1]} |\lambda_1(t) - t| \\ \sup_{t \in [0,1]} |x(t) - z(\lambda_2 \circ \lambda_1(t))| &\leq \sup_{t \in [0,1]} |x(t) - y(\lambda_1(t))| + \sup_{t \in [0,1]} |y(t) - z(\lambda_2(t))|. \end{aligned}$$

Also gilt auch $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ und damit die Dreiecksungleichung. Insgesamt folgt, dass $(D[0, 1], \rho)$ ein metrischer Raum ist.

□

1.9 Bemerkung

(i) In der durch d_∞ induzierten gleichmäßigen Topologie sind zwei Funktionen ähnlich, wenn sich ihre Graphen bei festgehaltener Abzisse nur durch eine gleichmäßig kleine Veränderung der Ordinate voneinander unterscheiden. In der Skorokhod-Topologie allerdings wird zugleich auch eine gleichmäßig kleine Verzerrung der Abzisse, also der Zeitachse, zugelassen. Physikalisch gesehen drückt diese Definition das Problem der Unmöglichkeit der perfekten Zeitmessung aus.

(ii) Eine Folge $(x_n)_{n \in \mathbb{N}} \subset D[0, 1]$ konvergiert in der Skorokhod-Topologie gegen ein $x \in D[0, 1]$ genau dann, wenn eine Folge $(\lambda_n)_{n \in \mathbb{N}} \in \Lambda$ existiert, so dass

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{t \in [0,1]} |\lambda_n t - t| &= 0 \quad \text{und} \\ \lim_{n \rightarrow \infty} \sup_{t \in [0,1]} |x_n(\lambda_n t) - x(t)| &= 0 \end{aligned}$$

gilt.

1.10 Satz

Seien $x, x_1, x_2, \dots \in D[0, 1]$.

- (i) Aus der gleichmäßigen Konvergenz von $(x_n)_{n \in \mathbb{N}}$ gegen x , d.h. $x_n \rightarrow x$ bzgl. d_∞ , folgt schon die Konvergenz von $(x_n)_{n \in \mathbb{N}}$ gegen x in der Skorokhod-Topologie.
- (ii) Die Umkehrung von (i) gilt nicht.
- (iii) Aus der Konvergenz von $(x_n)_{n \in \mathbb{N}}$ gegen x in der Skorokhod-Topologie folgt die Konvergenz von $(x_n(t))_{n \in \mathbb{N}}$ gegen $x(t)$ in allen Stetigkeitsstellen t von x .
- (iv) Aus der Konvergenz von $(x_n)_{n \in \mathbb{N}}$ gegen x in der Skorokhod-Topologie und der Stetigkeit von x auf $[0, 1]$ folgt die gleichmäßige Konvergenz von $(x_n)_{n \in \mathbb{N}}$ gegen x .

Beweis:

(i) folgt durch die Wahl von $\lambda_n(t) \equiv t$ für alle $n \in \mathbb{N}$.

(ii) folgt aus folgendem Gegenbeispiel:

$$x_n = \mathbf{1}_{[0, \frac{1}{2} + \frac{1}{n})} \quad \text{und} \quad x = \mathbf{1}_{[0, \frac{1}{2})}.$$

Klar ist, dass $(x_n(t))_{n \in \mathbb{N}}$ nicht gegen $x(t)$ in $t = \frac{1}{2}$ konvergiert. Für die Konvergenz in der Skorokhod-Topologie wähle $\lambda_n(t) = \begin{cases} \frac{n+2}{n} t & \text{für } t \leq \frac{1}{2} \\ \frac{n-2}{n} t + \frac{2}{n} & \text{für } t > \frac{1}{2} \end{cases}$.

(iii) folgt aus der Ungleichung

$$|x_n(t) - x(t)| \leq |x_n(t) - x(\lambda_n(t))| + |x(\lambda_n(t)) - x(t)|,$$

da der erste Summand wegen der Konvergenz in der Skorokhod-Topologie gegen 0 konvergiert und der zweite Summand aufgrund der Stetigkeit von x in t und der Konvergenz von $\lambda_n(t)$ gegen t ebenfalls gegen 0 konvergiert.

(iv) folgt aus (iii) und der Tatsache, dass x auf dem abgeschlossenen Intervall $[0, 1]$ auch gleichmäßig stetig ist.

□

Die beiden folgenden Sätze geben weitere Eigenschaften der Skorokhod-Topologie an. Da diese jedoch im Weiteren keine große Rolle spielen werden, und die Beweise eine ausführlichere Untersuchung der Skorokhod-Topologie notwendig machen würde, verzichten wir hier auf ihre Beweise und verweisen stattdessen auf Kapitel 3 in [6].

1.11 Satz

Eine Menge $A \subset D[0, 1]$ ist schwach relativ folgenkompakt bzgl. der Skorokhod-Topologie genau dann, wenn

$$\sup_{x \in A} \sup_{t \in [0, 1]} |x(t)| < \infty \quad \text{und} \\ \lim_{\delta \rightarrow 0} \sup_{x \in A} w'_x(\delta) = 0$$

gelten.

1.12 Satz

$D[0, 1]$ versehen mit der Skorokhod-Topologie ist ein separabler, vollständiger metrischer Raum.

1.3 Der Raum der beschränkten, messbaren Funktionen

In diesem Abschnitt zeigen wir, dass der Raum der beschränkten, messbaren Funktionen von $[0, 1]$ nach $\mathbb{R}$ zusammen mit der durch die Supremumsnorm induzierten Metrik einen vollständigen, metrischen Raum bildet.

1.13 Definition und Satz

Sei

$$\mathcal{B}_m \stackrel{\text{def}}{=} \left\{ f : [0, 1] \rightarrow \mathbb{R} : f \text{ ist beschränkt und messbar} \right\}$$

der Raum der beschränkten, messbaren Funktionen von $[0, 1]$ nach $\mathbb{R}$. Dann bildet für alle $f, g \in \mathcal{B}_m$

$$d_\infty(f, g) = \sup_{t \in [0, 1]} |f(t) - g(t)|$$

eine vollständige Metrik auf $\mathcal{B}_m$.

Beweis:

Wir verzichten auf den elementaren Nachweis, dass d_∞ eine Metrik bildet. Zum Beweis der Vollständigkeit halten wir fest, dass $\mathcal{B}_m$ ähnlich wie im Beweis von Satz 1.7 auch hier ein Untervektorraum von $\mathcal{B}([0, 1])$ ist. Da der Grenzwert von messbaren Abbildungen wiederum messbar ist¹, folgt die Abgeschlossenheit dieses Untervektorraums bzgl. d_∞ . Damit ergibt sich die Vollständigkeit von $(\mathcal{B}_m, d_\infty)$ erneut aus der Tatsache, dass abgeschlossene Untervektorräume von Banachräumen vollständig sind.

□

¹siehe Satz 8.3 in [2]

1.4 Die Mallows-Metrik

Wir werden in diesem Abschnitt die Mallows-Metrik auf dem Raum der p -fach integrierbaren Wahrscheinlichkeitsmaße einführen und diese dann charakterisieren. Die Mallows-Metrik spielt für die Kontraktionsmethode, welche im dritten Kapitel eingeführt wird, eine zentrale Rolle. Wir werden nun erst einige Notationen angeben und auf bekannte Ergebnisse aus der Wahrscheinlichkeitstheorie hinweisen.

Es sei $1 \leq p < \infty$ und $\mathbb{M}_p$ die Menge aller p -fach integrierbaren Wahrscheinlichkeitsmaße auf $[0, \infty)$. Weiterhin sei ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ gegeben, und $\mathfrak{L}_p(P) = \mathfrak{L}_p(\Omega, \mathfrak{A}, P)$ sei der Vektorraum der p -fach integrierbaren Zufallsgrößen auf $(\Omega, \mathfrak{A}, P)$. Bekanntlich definiert dann $\|X\|_p \stackrel{\text{def}}{=} (E|X|^p)^{\frac{1}{p}}$ eine (Pseudo-)Norm auf $\mathfrak{L}_p$. Durch die so induzierte Metrik $\ell_p(X, Y) = \|X - Y\|_p$ wird $\mathfrak{L}_p$ zu einem vollständigen metrischen Raum, indem man die P fast überall übereinstimmenden Zufallsgrößen miteinander identifiziert. Genauer betrachtet man also den Quotientenraum $L_p(P) \stackrel{\text{def}}{=} \mathfrak{L}_p(P)/\mathfrak{N}(P)$, wobei $\mathfrak{N}(P)$ den Raum aller P fast überall verschwindenden Funktionen bezeichnet.

Zu beliebigen Verteilungen F und G auf $[0, \infty)$ wird ein Paar (X, Y) von Zufallsgrößen auf demselben Wahrscheinlichkeitsraum mit $\mathcal{L}(X) = F$ und $\mathcal{L}(Y) = G$, wobei $\mathcal{L}(X)$ die Verteilung von X bezeichne, eine (F, G) -Kopplung genannt und man schreibt $(X, Y) \sim (F, G)$. Schließlich sei $F^{-1}(y) \stackrel{\text{def}}{=} \{x : F(x) \geq y\}$ die für $0 < y < 1$ definierte *Pseudo-Inverse* der Verteilungsfunktion F . Dann gilt bekanntlich $\mathcal{L}(F^{-1}(U)) = F$, falls U auf dem Intervall $[0, 1]$ gleichverteilt ist, d.h. $\mathcal{L}(U) = U(0, 1)$ gilt.

1.14 Definition und Satz

Die Abbildung $d_p : \mathbb{M}_p \times \mathbb{M}_p \rightarrow [0, \infty)$,

$$d_p(F, G) \stackrel{\text{def}}{=} \inf_{(X, Y) \sim (F, G)} \|X - Y\|_p, \quad (1.5)$$

ist für jedes $1 \leq p < \infty$ eine Metrik auf $\mathbb{M}_p$, die sogenannte Mallows-Metrik.

Das Infimum in (1.5) wird angenommen, und zwar gilt

$$d_p(F, G) = \|F^{-1}(U) - G^{-1}(U)\|_p \quad (1.6)$$

für jede $U(0, 1)$ -verteilte Zufallsgröße U .

Beweis:

Es genügt zu zeigen, dass das Infimum in (1.5) in angegebener Weise angenommen wird. Die Metrikeigenschaften folgen dann direkt aus denen der Metrik ℓ_p . Für den Nachweis, dass das Infimum angenommen wird, siehe Kapitel 4 in [1].

□

Zur Charakterisierung der Konvergenz bzgl. d_p notieren wir folgenden Satz. Dabei bezeichnen $\xrightarrow{d_p}$ die Konvergenz bezüglich der Mallows-Metrik und $\xrightarrow{w}$ die schwache Konvergenz von Verteilungen.

1.15 Satz

Für eine Folge $(F_n)_{n \in \mathbb{N}}$ von Verteilungen in $\mathbb{M}_p$ sind äquivalent:

- (i) $F_n \xrightarrow{d_p} F$.
- (ii) $F_n \xrightarrow{w} F$ und $\int |x|^p F_n(dx) \rightarrow \int |x|^p F(dx) < \infty$.
- (iii) $F_n \xrightarrow{w} F$ und $|x|^p$ ist gleichgradig integrierbar bzgl. der F_n , d.h.

$$\lim_{a \rightarrow \infty} \sup_{n \in \mathbb{N}} \int_{(-a,a)^c} |x|^p F_n(dx) = 0$$

Beweis:

Alle Aussagen des Satzes resultieren direkt aus bekannten Ergebnissen der Wahrscheinlichkeitstheorie.

„(i) $\Rightarrow$ (ii)“ Wegen (1.6) ist (i) äquivalent zur $\mathfrak{L}_p$ -Konvergenz von $F_n^{-1}(U)$ gegen $F^{-1}(U)$ für jede $U(0,1)$ -verteilte Zufallsgröße U . Daraus wiederum folgt bereits (ii).

„(ii) $\Rightarrow$ (iii)“ folgt aus Satz 50.8 in [2], wenn man beachtet, dass $F_n^{-1}(U) \rightarrow F^{-1}(U)$ fast sicher gilt und die gleichgradige Integrierbarkeit von $|x|^p$ bzgl. der F_n der gleichgradigen Integrierbarkeit der Folge $(|F_n^{-1}(U)|^p)_{n \in \mathbb{N}}$ entspricht.

„(iii) $\Rightarrow$ (i)“ Falls $F_n \xrightarrow{w} F$, so folgt aus Satz 36.9 in [2] $F_n^{-1}(U) \rightarrow F^{-1}(U)$ fast sicher, was mit der gleichgradigen Integrierbarkeit von $(|F_n^{-1}(U)|^p)_{n \in \mathbb{N}}$ auch $\|F_n^{-1}(U) - F^{-1}(U)\|_p \rightarrow 0$ und somit (i) liefert.

□

1.16 Satz

Der Raum $(\mathbb{M}_p, d_p)$ ist für jedes $1 \leq p < \infty$ vollständig.

Beweis:

Sei $(F_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in $(\mathbb{M}_p, d_p)$. Also gilt für jedes $U(0,1)$ -verteilte U auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$

$$\lim_{m,n \rightarrow \infty} d_p(F_m, F_n) = \lim_{m,n \rightarrow \infty} \|F_m^{-1}(U) - F_n^{-1}(U)\|_p = 0.$$

Da der Raum $L_p(P)$ vollständig ist, konvergiert $F_n^{-1}(U)$ im p -ten Mittel gegen eine Zufallsgröße $X \in \mathfrak{L}_p(P)$ mit Verteilung F und damit auch F_n schwach gegen F . Es folgt $F_n^{-1}(U) \rightarrow F^{-1}(U)$ fast sicher aus Satz 36.9 in [2] und wegen der Eindeutigkeit des Grenz-

wertes auch $X = F^{-1}(U)$ fast sicher. Insgesamt ergibt sich dann das Gewünschte durch

$$\lim_{n \rightarrow \infty} d_p(F_n, F) = \lim_{n \rightarrow \infty} \|F_n^{-1}(U) - F^{-1}(U)\|_p = \lim_{n \rightarrow \infty} \|F_n^{-1}(U) - X\|_p = 0.$$

□

Kapitel 2

Gewichtete Verzweigungsprozesse

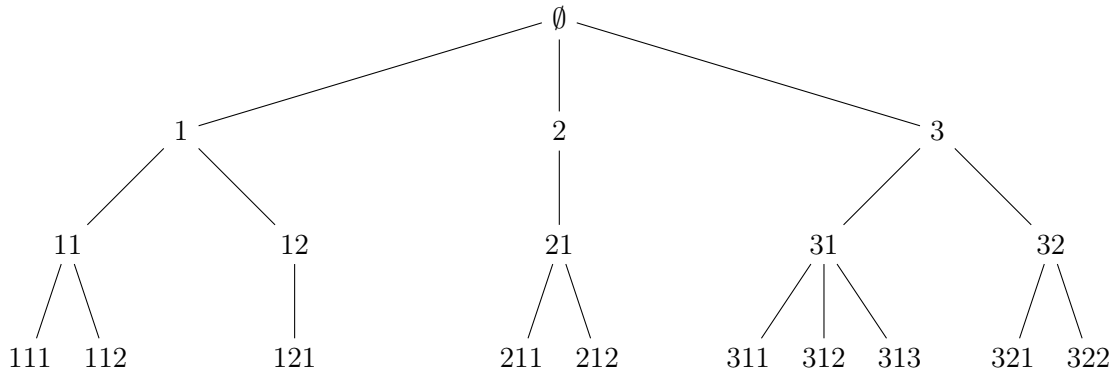
Dieses Kapitel dient der Einführung der gewichteten Verzweigungsprozesse. Dafür werden wir zuerst einige Notationen für Bäume angeben, mit deren Hilfe wir dann die gewichteten Verzweigungsprozesse definieren können. Diese stellen eine geeignete Grundstruktur dar, mit deren Hilfe stochastische Fixpunktgleichungen beschrieben werden können. Die Kontraktionsmethode, welche im folgenden Kapitel vorgestellt wird, ist hierbei ein viel genutztes Analysehilfsmittel.

2.1 Bäume

Wir betrachten den unendlichen *Ulam-Harris-Baum*

$$\mathbb{V} \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} \mathbb{N}^n, \quad \mathbb{N}^0 \stackrel{\text{def}}{=} \{\emptyset\}$$

mit der Wurzel $\emptyset$ und den Knoten $v = (v_1, \dots, v_n)$, für die wir im Folgenden vereinfachend $v_1 \dots v_n$ schreiben werden. Der Name des Knotens $v_1 \dots v_n$ entspricht dem Pfad von der Wurzel $\emptyset$ zum Knoten.



Jeder Knoten $v \in \mathbb{V}$ hat abzählbar viele Nachfolger $vi \stackrel{\text{def}}{=} (v_1, \dots, v_n, i)$. Im Fall $v = \emptyset$ soll dies natürlich $vi \stackrel{\text{def}}{=} (i)$ bedeuten. Für $v = v_1 \dots v_n$ bezeichne $|v| = n$ die Länge von v , was

speziell $|\emptyset| = 0$ bedeutet. Für $m \leq |v|$ benutzen wir die Einschränkung $v|_m = v_1 \dots v_m$. Bei gegebenen $v_1 \dots v_n$ und $w_1 \dots w_m$ verwenden wir die Notation $vw \stackrel{\text{def}}{=} v_1 \dots v_n w_1 \dots w_m$.

Für kleinere Bäume werden allerdings oft andere Bezeichnungen verwendet. Bei binären Bäumen zum Beispiel ist $\mathbb{V} \stackrel{\text{def}}{=} \bigcup_{n \in \mathbb{N}} \{0, 1\}^n$. Formal können diese kleineren Bäume mittels einer Funktion $\varphi : \mathbb{V} \rightarrow \{0, 1\}$ eingebettet werden, die Folgendes erfüllt: Ist v ein Vorfahre von w und $\varphi(v) = 0$, dann ist auch $\varphi(w) = 0$. Anschaulich stellen wir uns v als lebend vor, wenn $\varphi(v) \neq 0$, und tot, wenn $\varphi(v) = 0$ gilt. Der kleinere Baum wird dann nur von den lebenden Individuen gebildet.

2.2 Gewichtete Verzweigungsprozesse

Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum, den wir im Folgenden immer als reichhaltig genug für alle auftretenden Zufallsvariablen annehmen wollen. Seien $T(v) : \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$, $v \in \mathbb{V}$, unabhängig identisch verteilte Zufallsvariablen. Wir verwenden die Notation $T(v) = (T_1(v), T_2(v), \dots)$, wobei $T_i(v) : \Omega \rightarrow \mathbb{R}$ die i -te Projektion von T bezeichnet. Wir stellen uns $T_i(v)$ als Gewicht vor, welches die Kante von v zu vi trägt. Beachte, dass beliebige Abhängigkeiten der Komponenten $T_i(v)$, $i \in \mathbb{N}$, bei fixiertem $v \in \mathbb{V}$ zugelassen sind.

2.1 Definition

Der Prozess $((\Omega, \mathfrak{A}, P), \mathbb{V}, T)$ heißt *gewichteter Verzweigungsprozess*.

Definiere nun rekursiv das Gewicht $l(v) : \Omega \rightarrow \mathbb{R}$ des Astes von der Wurzel zum Knoten v durch

$$l(vi) \stackrel{\text{def}}{=} T_i(v) l(v)$$

für $i \in \mathbb{N}$. Dabei ist der Startwert $l(\emptyset)$ unabhängig von T . Das Gewicht $l(v)$ des Astes hat dann folgende Produktdarstellung

$$\begin{aligned} l(v) &= T_{v_n}(v_1 v_2 \dots v_{n-1}) T_{v_{n-1}}(v_1 v_2 \dots v_{n-2}) \dots T_{v_1}(\emptyset) l(\emptyset) \\ &= T_{v_n}(v|_{n-1}) T_{v_{n-1}}(v|_{n-2}) \dots T_{v_1}(\emptyset) l(\emptyset) \end{aligned}$$

Zur Vereinfachung nehmen wir $l(\emptyset) = 1$ an. Wir nennen ein Knoten bzw. Individuum beobachtbar oder lebend, falls $l(v) \neq 0$ ist, ansonsten nicht beobachtbar oder tot. Beachte, dass jeder Nachkomme eines nicht beobachtbaren Individuums wieder nicht beobachtbar ist.

Im Folgenden werden wir einige Größen vorstellen, die in vielen Anwendungen von gewichteten Verzweigungsprozessen Ziel der Untersuchungen sind. Dafür betrachte zuerst die Gewichte in der n -ten Generation, $l(v)$, mit $|v| = n$. Diese kann man wie folgt über

diskrete Maße beschreiben:

$$\xi_n \stackrel{\text{def}}{=} \sum_{|v|=n} \delta_{l(v)},$$

wobei δ_x das Dirac-Maß in $x \in \mathbb{R}$ bezeichnet. Diese Objekte werden auch in der Theorie der Verzweigungsprozesse untersucht (siehe hierfür [4]).

Weitere Größen, die gerade bei der Analyse von Divide-and-Conquer Algorithmen eine große Rolle spielen, sind das Gewicht

$$Z_n \stackrel{\text{def}}{=} \int x \xi_n(dx) = \sum_{|v|=n} l(v)$$

der n -ten Generation und die Kosten

$$R_n \stackrel{\text{def}}{=} \sum_{|v| \leq n} l(v) C(v)$$

bis zur n -ten Generation. Hierbei sind $(T(v), C(v))$, $v \in \mathbb{V}$ unabhängig identisch verteilte Zufallsvariablen. $l(v)C(v)$ kann man als Kosten abhängig von der Größe und R_n als Gesamtkosten bis zur n -ten Generation auffassen.

In anderen Anwendungen ist man mehr an der maximalen (oder auch minimalen) Höhe des Baumes

$$\sup\{n : l(v) \leq a \quad \forall |v| = n\}$$

für eine geeignete Konstante a interessiert.

2.2 Beispiel

In diesem Beispiel werden wir einen Galton-Watson-Prozess als gewichteten Verzweigungsprozess darstellen. Dafür erlauben wir nur zwei mögliche Zustände für die T_i : lebend ($=1$) und tot ($=0$). Die lebenden Individuen bilden einen zufälligen Baum. In der n -ten Generation leben dann

$$Z_n = \sum_{|v|=n} l(v)$$

Individuen. Z_n bildet so einen Galton-Watson-Prozess mit einem Urahn. Gewöhnlich wird ein solcher Prozess $\tilde{Z}_n$ eingeführt über $\tilde{Z}_1 = 1$ und

$$\tilde{Z}_n = \sum_{i=1}^{\tilde{Z}_{n-1}} X_{n,i}.$$

Die Zufallsgrößen $X_{n,i}$, $n, i \in \mathbb{N}$ sind dabei unabhängig und nehmen Werte in $\mathbb{N}_0$ an. Sie geben die Anzahl der Kinder des i -ten Individuums in der n -ten Generation an. Sie haben

alle die Reproduktionsverteilung

$$p_k = P(X_{n,i} = k) = P\left(\sum_{j=1}^{\infty} T_j(v) = k\right),$$

wobei v das i -te Individuum der n -ten Generation bezeichnet. Dann hat der Prozess $(Z_0 = 1, Z_1, Z_2, \dots)$ dieselbe Verteilung wie der Prozess $(\tilde{Z}_0 = 1, \tilde{Z}_1, \tilde{Z}_2, \dots)$.

2.3 Vorwärts- und Rückwärtsgleichungen

Wie bei Markov-Prozessen kann man auch bei gewichteten Verzweigungsprozessen zwischen einer Vorwärts- und Rückwärtsbetrachtung unterscheiden. Zur Verdeutlichung betrachte die Komposition

$$(f_n \circ f_{n-1} \circ \dots) \circ f_1 = f_n \circ f_{n-1} \circ \dots \circ f_1 = f_n \circ (f_{n-1} \circ \dots \circ f_1)$$

der n Funktionen $f_1, \dots, f_n$. Die linke Seite der Gleichung zeigt die Betrachtung rückwärts in der Zeit, die rechte die Betrachtung vorwärts in der Zeit.

Da beide Betrachtungsweisen ihre Vorteile haben, werden wir nun für die vorher eingeführten Zielgrößen jeweils beide Sichtweisen angeben. Wir beginnen mit der Anzahl ξ_{n+1} der Individuen in der $(n+1)$ -ten Generation. Die Vorwärtsgleichung schaut zuerst auf die n -te Generation ξ_n und betrachtet dann den Nachwuchs jedes Individuums dieser Generation. Es ergibt sich

$$\xi_{n+1} = \sum_{|v|=n} \sum_{j \in \mathbb{N}} \delta_{l(vj)}.$$

Im Gegensatz dazu ergibt sich die Rückwärtsgleichung durch die Betrachtung der Nachfahren der Individuen der ersten Generation. Jedes Individuum in der ersten Generation generiert, unabhängig zu den anderen Individuen der ersten Generation, Nachfahren über n Generationen, die jeweils durch die unabhängigen Kopien $\xi_{n,j}$ von ξ_n beschrieben werden. Allerdings ändert sich dabei der Anfangswert von $l(\emptyset)$ zu $T_j(\emptyset)l(\emptyset)$. Dadurch ergibt sich die Rückwärtsgleichung zu

$$\xi_{n+1}(\cdot) \stackrel{d}{=} \sum_{j \in \mathbb{N}} \xi_{n,j}(T_j(\emptyset)(\cdot)),$$

wobei $\stackrel{d}{=}$ die Gleichheit in Verteilung bezeichnet.

Aus diesem Beispiel erkennt man allerdings noch nicht die Vorteile, die die beiden unterschiedlichen Betrachtungsweisen mit sich bringen. Zur Verdeutlichung betrachten wir nun das Gesamtgewicht Z_{n+1} der $(n+1)$ -ten Generation mit Startwert $l(\emptyset) \equiv 1$. Die

Vorwärtsgleichung ergibt sich dann zu

$$Z_{n+1} = \sum_{|v|=n} \left(l(v) \sum_{j \in \mathbb{N}} T_j(v) \right).$$

Normalisiert man diesen gewichteten Verzweigungsprozess mit dem Faktor m^n , $m = E(\sum_{i \in \mathbb{N}} T_i)$, so ergibt sich ein Martingal. Bekanntermaßen konvergiert jedes nichtnegative Martingal fast sicher, d.h.

$$\frac{Z_n}{m^n} \rightarrow W \quad \text{fast sicher}$$

für eine Zufallsvariable W . Dieses ist ein probabilistisches Resultat für Zufallsgrößen.

Im Vergleich hierzu liefert die Rückwärtsgleichung

$$Z_{n+1} = \sum_{j \in \mathbb{N}} T_j(\emptyset) \sum_{|v|=n} \frac{l(jv)}{T_j(\emptyset)} \quad (2.1)$$

eine Iteration von Verteilungen. Schreibe dafür (2.1) als Verteilungsaussage

$$Z_{n+1} \stackrel{d}{=} \sum_{j \in \mathbb{N}} T_j(\emptyset) Z_{n,j}. \quad (2.2)$$

Die Zufallsvariable

$$Z_{n,j} \stackrel{\text{def}}{=} \sum_{|v|=n} \frac{l(vj)}{T_j(\emptyset)}$$

beschreibt dabei das Gesamtgewicht aller Nachkommen in der $(n+1)$ -ten Generation des j -ten Individuums der ersten Generation. Die Verteilung ist dieselbe wie die von Z_n . Die Iterationsstruktur zeigt sich nun, wenn man den Operator K auf der Menge der Verteilungen auf $\mathbb{R}$ definiert durch

$$K(\mu) \stackrel{d}{=} \sum_{j=1}^{\infty} T_j X_j, \quad (2.3)$$

wobei die $(X_i)_{i \in \mathbb{N}}$ unabhängig sind mit Verteilung μ .

Beachte, dass (2.2) von der Form (2.3) ist, wenn man (2.2) als Aussage für Verteilungen auffasst. Die Verteilung von Z_{n+1} ist die von $K(\mathcal{L}(Z_n))$. Durch Iteration erhalten wir

$$K^n(\mu_0) \stackrel{d}{=} Z_n,$$

wobei μ_0 die Startverteilung von $l(\emptyset)$ bezeichnet. Die Rückwärtsgleichung führt also zu einem funktionalanalytischen Resultat für Verteilungen. Mit Blick auf Gleichung (2.2) stellt sich also die Frage nach Fixpunkten $\nu = K(\nu)$ des Operators K , bzw. nach einer Zufallsgröße W mit

$$W \stackrel{d}{=} \sum_{j \in \mathbb{N}} T_j W_j, \quad (2.4)$$

wobei die $(W_j)_{j \in \mathbb{N}}$ unabhängige Kopien von W sind und T, W ebenfalls unabhängig sind.

Die Gleichung (2.4) gehört offenbar zu der Klasse von Gleichungen des allgemeinen Typs $X \stackrel{d}{=} h(\Gamma, X)$ für eine gegebene messbare Funktion h und eine gegebene Zufallsvariable Γ . Sie werden *stochastische Fixpunktgleichungen* genannt und treten zumeist bei der Beschreibung von Grenzverteilungen stochastischer Prozesse auf, die eine rekursive Struktur besitzen. Eine Methode zur Analyse solcher Prozesse ist die im nächsten Kapitel vorgestellte *Kontraktionsmethode*.

Zuletzt kommen wir noch zu den Gesamtkosten R_n bis zur n -ten Generation. Die Vorwärtsgleichung ist dann

$$R_{n+1} = R_n + \sum_{|v|=n+1} l(v) C(v)$$

und die Rückwärtsgleichung

$$K(\mu) \stackrel{d}{=} \sum_{j \in \mathbb{N}} T_j X_j + C$$

oder intuitiver durch Verwenden von Zufallsvariablen

$$R_{n+1} \stackrel{d}{=} \sum_{j \in \mathbb{N}} T_j(\emptyset) R_{n,j} + C(\emptyset). \quad (2.5)$$

Hierbei sind die Zufallsvariablen $(T, C), R_{n,j}, j \in \mathbb{N}$ unabhängig. Die Verteilung von (T, C) ist gegeben und die $R_{n,j}$ sind unabhängige Kopien von R_n .

Zufallsvariablen von der Gestalt der R_n werden uns im folgenden Kapitel als Laufzeit von Divide-and-Conquer Algorithmen weiter beschäftigen.

Kapitel 3

Die Kontraktionsmethode

In diesem Kapitel wird die Kontraktionsmethode vorgestellt. Diese ist eine Methode zum Nachweis der Konvergenz von Folgen von Wahrscheinlichkeitsmaßen bezüglich verschiedener Metriken. Sie findet Anwendung bei der Analyse von rekursiven Algorithmen, gewichteten Verzweigungsprozessen, dem Zentralen Grenzwertsatz und der Untersuchung der Hausdorff-Dimension von zufälligen Cantor-Mengen. Ein wichtiger Vorteil der Kontraktionsmethode ist, dass sie neben Grenzwertresultaten auch Aussagen über die Konvergenzgeschwindigkeit liefert. Erstmals wurde die Kontraktionsmethode von Rösler in [24] zur Analyse des Sortieralgorithmus Quicksort verwendet. In allgemeiner Form haben dann sowohl Rösler in [29] als auch Rachev und Rüschendorf in [23] die Kontraktionsmethode unabhängig voneinander entwickelt. Seitdem ist sie auf viele Algorithmen erfolgreich angewendet worden. Darunter befindet sich auch der in dieser Arbeit untersuchte Selektionsalgorithmus `Find`.

Wir werden nun zuerst die Vorgehensweise der Kontraktionsmethode erklären. Im Anschluss wird dann das Vorgehen für Divide-and-Conquer-Algorithmen präzisiert.

3.1 Das Prinzip der Kontraktionsmethode

Gegeben sei eine rekursiv definierte stochastische Folge. Im Folgenden stellen wir das Grundprinzip der Kontraktionsmethode vor. Diese besteht gewöhnlich aus folgenden Schritten.

1. Finden der korrekten Normalisierung der stochastischen Folge.
Dies geschieht normalerweise durch die Untersuchung der ersten Momente.
2. Bestimmen der Rekursion für die normalisierte Folge.
3. Bestimmen der stochastischen Fixpunktgleichung für die Grenzverteilung.
Typischerweise ist diese Fixpunktgleichung über einen Operator K auf der Menge

aller Wahrscheinlichkeitsmaße, bezeichnet mit $\mathbb{M}$, definiert. In vielen Anwendungen kann sie erraten werden.

4. Wählen einer Metrik auf $\mathbb{M}$, so dass $\mathbb{M}$ zu einem vollständigen metrischen Raum wird und K eine Kontraktion bezüglich dieser Metrik ist.

Dabei sind die im ersten Kapitel beschriebene Mallows-Metrik und die Zolotarev-Metrik häufig verwendete Metriken auf Teilmengen von $\mathbb{M}$. Eine Anwendung des Banachschen Fixpunktsatzes liefert dann die Existenz und Eindeutigkeit eines Fixpunktes des Operators K .

5. Nachweis der Verteilungskonvergenz der normalisierten stochastischen Folge gegen diesen Fixpunkt.

Dies geschieht häufig durch die Einführung einer geeigneten begleitenden Folge von Verteilungen, welche sowohl der ursprünglichen Folge als auch der Struktur der Grenzverteilung ähnelt.

3.2 Die Kontraktionsmethode bei Divide-and-Conquer Algorithmen

Als Divide-and-Conquer Algorithmen werden Algorithmen bezeichnet, die durch sukzessives Herunterbrechen des Grundproblems auf analoge Probleme kleinerer Komplexität zu einer Lösung kommen. Bezeichne $(X_n)_{n \in \mathbb{N}}$ die Laufzeit eines stochastischen Divide-and-Conquer Algorithmus. In unserem Fall des Selektionsalgorithmus Find beispielsweise die Anzahl der Vergleiche, die Find benötigt um das gesuchte Element zu finden. Die Verteilung der Zufallsvariable X_n sei rekursiv gegeben durch

$$X_n \stackrel{d}{=} \sum_{i \in \mathbb{N}} T_{n,i} X_{Z_{n,i}}^{(i)} + C_n \quad (3.1)$$

für $n \geq n_0$. Dabei ist n_0 eine natürliche Zahl. Die Zufallsvariablen (C_n, Z_n, T_n) und $X_j^{(i)}$ für $0 \leq j \leq n-1$ sind für jedes $n \in \mathbb{N}$ stochastisch unabhängig und die $Z_{n,i}$ nehmen Werte in $\{0, \dots, n-1\}$ an. Dabei verwenden wir die Notation $Z_n = (Z_{n,1}, Z_{n,2}, \dots)$ und $T_n = (T_{n,1}, T_{n,2}, \dots)$. Weiter sei die Verteilung von (C_n, Z_n, T_n) gegeben. Die Verteilung von X_j für $0 \leq j \leq n_0-1$ sei im Voraus bekannt und die Verteilung von $X_j^{(i)}$ sei die von X_j und rekursiv gegeben. Aus Gründen der einfacheren Notation setzen wir $X_0 \equiv 0$.

Beachte die Analogie der Gleichungen (3.1) und (2.5). X_n könnte also als Laufzeit bzw. Kosten des Algorithmus bis zum n -ten Iterationsschritt interpretiert werden.

1. Normalisierung des Algorithmus

Gesucht ist eine geeignete Normalisierung Y_n von X_n , so dass Y_n in Verteilung gegen einen Grenzwert W konvergiert. Falls die Varianz von X_n endlich ist und berechnet werden kann,

wähle

$$Y_n = \frac{X_n - a_n}{b_n},$$

wobei $a_n \stackrel{\text{def}}{=} EX_n$ und $b_n \stackrel{\text{def}}{=} \sqrt{\text{Var} X_n}$ gilt. Wegen der Unabhängigkeit der entsprechenden Variablen ergibt sich für die a_n folgende Rekursion

$$a_n = \sum_{i \in \mathbb{N}} ET_{n,i} a_{Z_{n,i}} + c_n, \quad (3.2)$$

wobei $c_n \stackrel{\text{def}}{=} EC_n$ ist. Falls die Varianz von X_n nicht endlich ist, wähle b_n so klein, dass das weiter unten definierte C_n^* gegen einen Grenzwert konvergiert, der falls möglich nicht degeneriert ist.

2. Die Rekursion für den normalisierten Algorithmus

Durch Einsetzen der Gleichungen (3.1) und (3.2) in $Y_n = \frac{X_n - a_n}{b_n}$ ergibt sich

$$\begin{aligned} Y_n &= \frac{\sum_{i \in \mathbb{N}} T_{n,i} X_{Z_{n,i}}^{(i)} + C_n - a_n}{b_n} \\ &= \sum_{i \in \mathbb{N}} \frac{T_{n,i} X_{Z_{n,i}}^{(i)}}{b_n} - \frac{T_{n,i} a_{Z_{n,i}}}{b_n} + \frac{T_{n,i} a_{Z_{n,i}}}{b_n} - \frac{ET_{n,i} a_{Z_{n,i}}}{b_n} + \frac{C_n - c_n}{b_n} \\ &= \sum_{i \in \mathbb{N}} T_{n,i} \frac{b_{Z_{n,i}}}{b_n} Y_{Z_{n,i}}^{(i)} + \sum_{i \in \mathbb{N}} \frac{T_{n,i} a_{Z_{n,i}} - ET_{n,i} a_{Z_{n,i}}}{b_n} + \frac{C_n - c_n}{b_n} \end{aligned} \quad (3.3)$$

Die Gleichungen (3.1) und (3.3) haben die gleiche Struktur, wenn man $T_{n,i}$ und C_n durch

$$\begin{aligned} T_{n,i}^* &\stackrel{\text{def}}{=} T_{n,i} \frac{b_{Z_{n,i}}}{b_n} \\ C_n^* &\stackrel{\text{def}}{=} \sum_{i \in \mathbb{N}} \frac{T_{n,i} a_{Z_{n,i}} - ET_{n,i} a_{Z_{n,i}}}{b_n} + \frac{C_n - EC_n}{b_n} \end{aligned}$$

ersetzt. Der Hauptunterschied der beiden Gleichungen besteht darin, dass der Erwartungswert aller Y_i gleich Null ist.

3. Die stochastische Fixpunktgleichung der Grenzverteilung

In vielen Situationen kann die Grenzverteilung als Fixpunkt eines Operators K geraten werden. Um einen Anhaltspunkt zu erhalten, betrachte die Gleichung (3.3). Falls nun Y_n gegen W , $Z_{n,i}$ gegen ∞ und (C_n^*, T_n^*) gegen (C, T) konvergieren, dann erhalten wir die

Fixpunktgleichung durch

$$\begin{array}{ccccccc}
 Y_n & \stackrel{d}{=} & \sum_{i \in \mathbb{N}} & T_{n,i}^* & Y_{Z_{n,i}}^{(i)} & + & C_n^* \\
 \downarrow & & & \downarrow & \downarrow & & \downarrow \\
 W & \stackrel{d}{=} & \sum_{i \in \mathbb{N}} & T_i & W_i & + & C.
 \end{array} \tag{3.4}$$

Hierbei sind wiederum $(C, T), W_i$ für $i \in \mathbb{N}$ unabhängig und alle W_i haben dieselbe Verteilung wie W . Die Verteilung von W sollte eine Lösung der stochastischen Fixpunktgleichung (3.4) sein.

Alternativ kann man diese Beziehung auch mittels eines Operators K auf einer Teilmenge von $\mathbb{M}$ beschreiben. Definiere hierfür

$$K(\mu) \stackrel{\text{def}}{=} \mathcal{L} \left(\sum_{i \in \mathbb{N}} T_i Y_i + C \right).$$

Hierbei sind ebenfalls wieder $(C, T), Y_i$ für $i \in \mathbb{N}$ unabhängig, die Y_i haben alle die Verteilung μ und (C, T) ist wie in der Grenzgleichung (3.4). Wir nennen K den Grenzoperator. Ein Fixpunkt von K ist ein Wahrscheinlichkeitsmaß ν mit $K(\nu) = \nu$.

4. Wahl der geeigneten Metrik

Wähle eine geeignete Metrik d auf $\mathbb{M}$, so dass $(\mathbb{M}, d)$ vollständig ist und der Operator K eine Kontraktion bildet. Um die Notation zu vereinfachen, wird häufig mathematisch unsauber $d(X, Y) \stackrel{\text{def}}{=} d(\mathcal{L}(X), \mathcal{L}(Y))$ verwendet. Man beachte hierbei, dass zwischen Metriken auf dem Raum der Verteilungen und solchen auf dem Raum der Zufallsvariablen unterschieden werden muss.

5. Verteilungskonvergenz von Y_n gegen W

Definiere die Operatoren K_n , $n \geq n_0$ von $\mathbb{M}^n$ nach $\mathbb{M}$ durch

$$K_n(\mu_0, \dots, \mu_{n-1}) \stackrel{\text{def}}{=} \mathcal{L} \left(\sum_{i \in \mathbb{N}} T_{n,i}^* Y_{Z_{n,i}}^{(i)} + C_n^* \right).$$

Dabei sind $(C_n^*, Z_n, T_n^*), Y_j^{(i)}, 0 \leq j \leq n, i \in \mathbb{N}$ unabhängig, die $Y_j^{(i)}$ haben die Verteilung μ_j und die Verteilung von (C_n^*, Z_n, T_n^*) ist wie zuvor gegeben. Startend mit den Wahr-

scheinlichkeitsmaßen $\mu_0, \dots, \mu_{n_0-1}$ erhalten wir die Folge

$$\begin{aligned}\mu_{n_0} &= K_{n_0}(\mu_0, \dots, \mu_{n_0-1}) \\ \mu_{n_0+1} &= K_{n_0+1}(\mu_0, \dots, \mu_{n_0}) \\ &\vdots \\ \mu_n &= K_n(\mu_0, \dots, \mu_{n-1}).\end{aligned}$$

Falls $\mu_0, \dots, \mu_{n_0-1}$ die Verteilungen von $Y_0, \dots, Y_{n_0-1}$ waren, so ist nun μ_n die Verteilung von Y_n .

Häufig wird für den Nachweis der Verteilungskonvergenz von Y_n gegen W eine begleitende Folge von Wahrscheinlichkeitsmaßen definiert. Diese Folge $(\nu_n)_{n \in \mathbb{N}}$ konvergiere gegen eine bekannte Verteilung ν . Oft wird $\nu_n = \nu$ für alle $n \in \mathbb{N}$ als Fixpunkt des Grenzoperators K_n definiert. Die Kontraktionsmethode ist nun eine Methode, um mittels Kontraktionsargumenten zu zeigen, dass der Abstand von μ_n zu ν_n gegen 0 konvergiert und damit die Konvergenz der Verteilung von Y_n gegen ν .

Kapitel 4

Der FIND-Prozess

Das Ziel dieses Kapitels ist es, den sogenannten **FIND**-Prozess einzuführen. Um die Funktion dieses Prozesses zu verstehen, kommen wir nochmals auf den Untersuchungsgegenstand dieser Arbeit zu sprechen. Wir wollen Aussagen über die Laufzeit des Algorithmus **FIND** treffen. Diese nehmen wir als proportional zu der Anzahl der von **FIND** benötigten Vergleiche an und vernachlässigen Komponenten wie die Rechenleistung oder die Qualität der Programmierung. Es stellt sich also die Frage, wie viele Vergleiche **FIND** benötigt, um das l -kleinste Element einer n -elementigen Menge zu finden.

Die Vorgehensweise hierbei wird sein, zuerst asymptotische Ergebnisse herzuleiten und mit Hilfe dieser auf Aussagen über die Laufzeit von **FIND** bei großen Mengen zu schließen. Dies wird Gegenstand des nächsten Kapitels sein, in dem die Vorgehensweise von Grübel und Rösler in [13] vorgestellt wird. Das Hauptresultat dort wird die Verteilungskonvergenz des Prozesses $(X_{n,l_n})_n$ gegen den **FIND**-Prozess sein, wobei X_{n,l_n} die Anzahl der benötigten Vergleiche zum Finden des l_n -kleinsten Elements einer n -elementigen Menge dividiert durch n bezeichnet und $(l_n)_{n \in \mathbb{N}}$ eine Folge ist, die $\frac{l_n}{n} \xrightarrow{n \rightarrow \infty} t \in [0, 1]$ erfüllt. Obwohl wir den **FIND**-Prozess im bereits angesprochenen nächsten Kapitel ausführlich konstruieren werden, wollen wir zuerst die Idee der Konstruktion und eine Interpretation des **FIND**-Prozesses als gewichteten Verzweigungsprozess angeben.

4.1 Die Konstruktion des FIND-Prozesses als gewichteter Verzweigungsprozess

Zum Finden des l -kleinsten Elements einer n -elementigen Menge mittels **FIND** brauchen wir nicht die gesamte Menge zu ordnen. Es genügt in jedem Rekursionsschritt mit der Teilmenge fortzufahren, welche das gesuchte Element enthält. Die Anzahl der Vergleiche in jedem Rekursionsschritt ist also proportional zur Größe der jeweiligen Teilmengen. Betrachtet man nun die Anzahl der Vergleiche dividiert durch n und lässt n gegen unendlich

laufen, so entsprechen die jeweiligen Teilmengen Teilintervallen des Einheitsintervalls. So gelangt man zu einer stetigen Version von FIND. Die Laufzeit von FIND ist also proportional zu den summierten Längen aller Intervalle, die $\frac{l}{n}$ enthalten.

Wir werden nun eine Möglichkeit angeben, den FIND-Prozess als gewichteten Verzweigungsprozess aufzufassen und erinnern dafür kurz an die Notationen aus Kapitel 2. Wir betrachten den binären Baum $\mathbb{V} = \{0, 1\}^{\mathbb{N}}$. Für $v = v_1 \dots v_n \in \mathbb{V}$ und $m \leq n$ bezeichne $v|_m = v_1 \dots v_m$ die Einschränkung von v auf die ersten m Generationen und v_m die m -te Koordinate von v .

Seien nun $U(v), v \in \mathbb{V}$ unabhängige $U(0,1)$ -verteilte Zufallsgrößen und E die Menge der halboffenen Intervalle $[a, b)$, $0 \leq a < b \leq 1$. $|I|$ bezeichne die Länge des Intervalls I . Definiere die Abbildungen $T(v) = (T_0(v), T_1(v)) : \Omega \times E \rightarrow E \times E$ durch

$$T(v)([a, b)) \stackrel{\text{def}}{=} (T_0(v)([a, b)), T_1(v)([a, b))) = ([a, a + U(v)(b - a)), [a + U(v)(b - a), b)).$$

Definiere weiter für $v \in \mathbb{V}$ und $i \in \{0, 1\}$ die E -wertige Abbildung $l : \mathbb{V} \rightarrow E$ als Gewicht der Äste von der Wurzel bis zum Knoten $v \in \mathbb{V}$ rekursiv durch

$$l(vi) \stackrel{\text{def}}{=} T_i(l(v)) \text{ und } l(\emptyset) \stackrel{\text{def}}{=} [0, 1).$$

Ausführlicher geschrieben meint dies für $v = v_1 v_2 \dots v_n$

$$l(v) = T_{v_n}(v_{|n-1}) \left(T_{v_{n-1}}(v_{|n-2}) \left(\dots (T_{v_1}(\emptyset)([0, 1])) \dots \right) \right).$$

Für das entsprechende $v \in \mathbb{V}$ beschreibt $l(v)$ also das Intervall, in dem sich das gesuchte Element im $|v|$ -ten Rekursionsschritt von FIND befindet.

Definiere für $t \in [0, 1)$ den Prozess $Z_n : \Omega \times [0, 1] \rightarrow \mathbb{R}, n \in \mathbb{N}$ durch

$$Z_n(t) \stackrel{\text{def}}{=} \sum_{|v| \leq n} \mathbf{1}_{\{t \in l(v)\}} |l(v)|$$

und $Z_n(1) = \lim_{t \uparrow 1} Z_n(t)$. Die $Z_n(t)$ addieren also die Länge aller Intervalle bis zum n -ten Rekursionsschritt, die t enthalten.

Der Prozess $Z : \Omega \times [0, 1] \rightarrow \mathbb{R}$ mit

$$Z(t) \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} Z_n(t) = \sum_{v \in \mathbb{V}} \mathbf{1}_{\{t \in l(v)\}} |l(v)|$$

wird im Folgenden als FIND-Prozess bezeichnet.

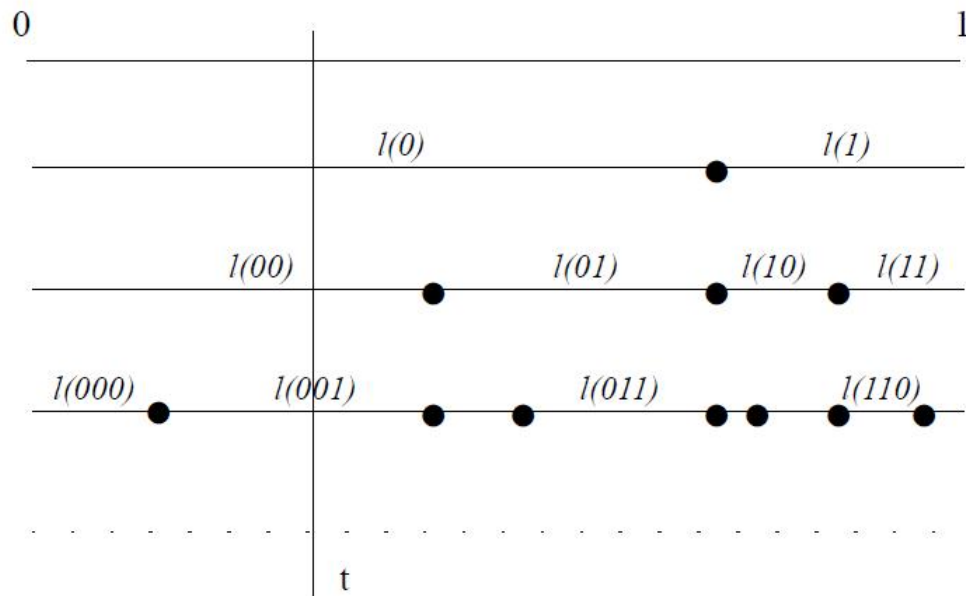
4.2 Der FIND-Prozess als zufällige Binärdarstellung einer reellen Zahl

Nachdem wir die Idee der Konstruktion des FIND-Prozesses im vorherigen Abschnitt erklärt haben, werden wir nun eine Interpretation des FIND-Prozesses als zufällige Binärentwicklung geben und dafür einen weiteren gewichteten Verzweigungsprozess definieren. Wir verwenden die gleichen Notationen wie in Kapitel 2 und dem vorherigen Abschnitt, werden jedoch die Gewichtsfunktionen wie folgt definieren.

Definiere die Gewichte der Kanten durch

$$T_1(v) = U(v), \quad T_2(v) = 1 - U(v), \quad T_3(v) = T_4(v) = \dots = 0.$$

Die Gewichte der Äste von der Wurzel zum Knoten v seien diesmal rekursiv definiert durch $l(vi) = l(v)T_i(v)$, dabei setzen wir $l(\emptyset) = 1$. Wie in Kapitel 2 definiert, bildet $((\Omega, \mathfrak{A}, P), \mathbb{V}, T)$ einen gewichteten Verzweigungsprozess. Im Gegensatz zum gewichteten Verzweigungsprozess im vorherigen Abschnitt bezeichnet $l(v)$ hier nicht das Intervall an sich, sondern nur die Länge des Intervalls. Das Vorgehen zur Konstruktion dieses Prozesses ist in folgender Abbildung verdeutlicht.



Dementsprechend bezeichnet $l(v)$ die Länge des zum Knoten v gehörenden Intervalls. Definiere nun die Abbildung $\psi : [0, 1] \times \Omega \rightarrow \{0, 1\}^{\mathbb{N}}$ durch

$$t = \sum_{n=1}^{\infty} \psi(t)_n l((\psi(t)|_{n-1}, 0)).$$

Beachte, dass $\psi(t)_n$ gleich 0 bzw. 1 ist, falls t im linken bzw. rechten Zweig der n -ten Generation des entsprechenden Knotens der $(n - 1)$ -ten Generation liegt. Damit ist ψ wohldefiniert. In obiger Abbildung ist $\psi(t)_3 = 001$ und $\psi(t)_3 = 1$

Man nennt eine solche Darstellung der reellen Zahl t eine zufällige Binärentwicklung von t . Man beachte die Analogie zur gewöhnlichen Binärdarstellung $t = \sum_{n=1}^{\infty} t_n 2^{-n}$ mit der Gewichtsfunktion $\bar{l}(v) = 2^{-|v|}$.

Definiere nun für $t \in [0, 1)$ und $n \in \mathbb{N}$ den Prozess

$$Z_n(t) \stackrel{\text{def}}{=} 1 + \sum_{j=1}^n l(\psi(t)_{|j}).$$

$Z_n(t)$ summiert die Länge aller halboffenen Intervalle bis zur n -ten Generation, die t enthalten. Analog zum vorherigen Abschnitt ergibt sich nun der FIND-Prozess als $Z(t) \stackrel{\text{def}}{=} \lim_{n \rightarrow \infty} Z_n(t)$ für alle $t \in [0, 1]$, wobei $Z_n(1) \stackrel{\text{def}}{=} \lim_{t \uparrow 1} Z_n(t)$ gilt.

Kapitel 5

Konvergenz in Verteilung gegen den FIND-Prozess

Nachdem im vorherigen Kapitel der FIND-Prozess eingeführt wurde, kommen wir nun zu einem Hauptresultat dieser Arbeit. Wir werden zeigen, dass die vom Algorithmus **FIND** (wir beziehen uns hierbei zuerst auf die 2-Mengen Version von **FIND**) benötigte Anzahl von Vergleichen zum Finden des l -kleinsten Elements einer n -elementigen Menge, geeignet normalisiert, in Verteilung gegen den FIND-Prozess konvergiert. Präziser bedeutet dies, dass die Folge der stochastischen Prozesse

$$X_n(t) \stackrel{\text{def}}{=} \frac{1}{n} C_{n, \lfloor nt \rfloor + 1} \quad \text{für } 0 \leq t < 1, \quad X_n(1) \stackrel{\text{def}}{=} \frac{1}{n} C_{n,n}, \quad (5.1)$$

aufgefasst als zufällige Elemente aus $D[0, 1]$, in Verteilung gegen den FIND-Prozess konvergieren. Dabei bezeichne $\lfloor x \rfloor \stackrel{\text{def}}{=} \max \{z \in \mathbb{Z} : z \leq x\}$ die untere Gaußklammer und $C_{n,l}$ die Anzahl der benötigten Vergleiche, um mittels **FIND** das l -kleinste Element einer n -elementigen Menge zu finden. Die obige Definition der X_n ermöglicht eine simultane Betrachtung aller l . Dies ist notwendig, da sich l mit positiver Wahrscheinlichkeit bei jedem Rekursionsschritt verändern kann. Anschaulich bedeutet dieses Resultat, dass die normalisierte Anzahl der notwendigen Vergleiche $\frac{1}{n} C_{n,l}$ gegen den **FIND**-Prozess konvergiert, falls sich l mit n so verändert, dass $\frac{l}{n}$ gegen einen Grenzwert $t \in [0, 1]$ konvergiert.

Wir werden nun zuerst den **FIND**-Prozess formal einführen und einige Eigenschaften der dabei auftretenden Größen angeben. Daraufhin wird dann im zweiten Abschnitt das oben beschriebene Hauptresultat bewiesen, bevor dann im Anschluss einige Eigenschaften des Grenzprozesses vorgestellt werden. Die Darstellung dieser Ergebnisse beruht auf dem Artikel [13], in dem Grübel und Rösler als Erste eine asymptotische Untersuchung des **FIND**-Algorithmus durchgeführt haben.

5.1 Konstruktion des FIND-Prozesses

Sei $\{U_{k,i} : k \in \mathbb{N}_0, 1 \leq i \leq 2^k\}$ eine Familie von unabhängigen $U(0,1)$ -verteilten Zufallsgrößen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Definiere mit deren Hilfe induktiv eine Familie $\{Y_{k,i} : k \in \mathbb{N}_0, 0 \leq i \leq 2^k\}$ von Zufallsgrößen wie folgt:

$$\begin{aligned} Y_{0,0} &\stackrel{\text{def}}{=} 0, & Y_{0,1} &\stackrel{\text{def}}{=} 1, \\ Y_{k+1,2i} &\stackrel{\text{def}}{=} Y_{k,i} & \text{für } 0 \leq i \leq 2^k, \\ Y_{k+1,2i+1} &\stackrel{\text{def}}{=} U_{k,i+1} Y_{k,i} + (1 - U_{k,i+1}) Y_{k,i+1} & \text{für } 0 \leq i < 2^k. \end{aligned}$$

Dabei verzichten wir zur Vereinfachung der Notation, wenn möglich, auf das explizite Angeben des ω in den Definitionen und Rechnungen.

Die $\{Y_{k,i}\}$ erzeugen also eine Folge von zufälligen Unterteilungen des Einheitsintervalls. Die $(k+1)$ -te Unterteilung geht aus der vorherigen Unterteilung hervor, indem aus jedem Teilintervall $[Y_{k,i}, Y_{k,i+1}]$ mit Hilfe von $U_{k,i+1}$ ein weiterer Punkt gleichverteilt ausgewählt wird.

Da die $U_{k,i}$ gleichverteilt sind, also insbesondere eine stetige Verteilung besitzen, gilt offenbar

$$P(0 = Y_{k,0} < Y_{k,1} < \dots < Y_{k,2^k-1} < Y_{k,2^k} = 1 \quad \text{für alle } k \in \mathbb{N}_0) = 1. \quad (5.2)$$

Es existiert also zu jedem $t \in [0, 1)$ und $k \in \mathbb{N}_0$ eine eindeutige Zufallsgröße $N_k(t) \in \{1, \dots, 2^k\}$, so dass $Y_{k,N_k(t)-1} \leq t < Y_{k,N_k(t)}$, d.h.

$$N_k(t) \stackrel{\text{def}}{=} \min \left\{ l \in \{1, \dots, 2^k\} : Y_{k,l} > t \right\}.$$

Weiterhin setze $N_k(1) = 2^k$. Da die Unterteilungen in dem Sinne aufeinander aufbauen, dass die $(k+1)$ -te Unterteilung jeden Punkt der k -ten Unterteilung enthält und jedes Teilintervall wieder zufällig teilt, folgt

$$N_{k-1}(t) = \begin{cases} \frac{N_k(t)}{2}, & \text{falls } N_k(t) \text{ gerade ist} \\ \frac{N_k(t)+1}{2}, & \text{sonst} \end{cases} \quad (5.3)$$

für alle $t \in [0, 1]$, $k \in \mathbb{N}$. Insbesondere legt $N_k(t)$ bereits alle $N_j(t)$ für $0 \leq j < k$ eindeutig fest.

Definiere $W_{k,i} \stackrel{\text{def}}{=} Y_{k,i} - Y_{k,i-1}$ für $k \in \mathbb{N}_0$ und $1 \leq i \leq 2^k$. Die $W_{k,i}$ geben also die Länge der Teilintervalle der k -ten Partition an. Insbesondere beschreibt $W_{k,N_k(t)}$ die Länge des

Teilintervalls im k -ten Rekursionsschritt, welches t enthält. Setze weiter

$$Z_k(t) \stackrel{\text{def}}{=} \sum_{j=0}^k W_{j,N_j(t)} \text{ und}$$

$$Z(t) \stackrel{\text{def}}{=} \sup_{k \in \mathbb{N}_0} Z_k(t).$$

$Z(t)$ ist der im 4. Kapitel eingeführte FIND-Prozess. Er bildet die Summe der Längen aller Teilintervalle, die t enthalten. Zur Verdeutlichung zeigt Abbildung 5.1 zwei Pfade von Z_{10} . Wir werden, wie bereits angekündigt, im nächsten Abschnitt zeigen, dass die X_n in Verteilung für $n \rightarrow \infty$ gegen Z konvergieren.

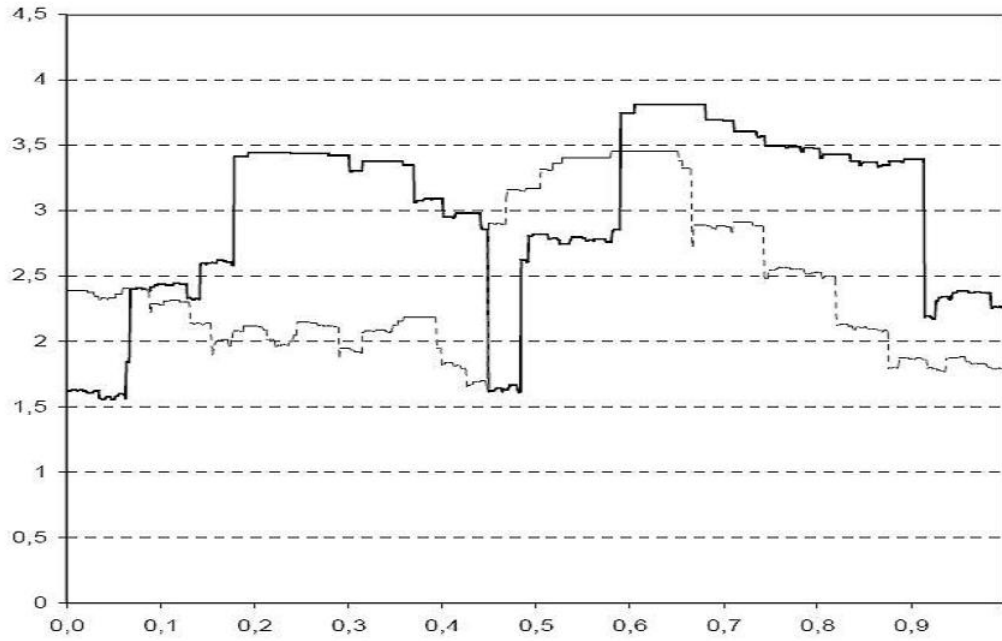


Abbildung 5.1: Zwei Pfade von Z_{10}

Da die $W_{j,N_j(t)}$ positiv sind, wachsen die $Z_k(t)$ monoton in k , und damit existiert $Z(t)$ sogar als Grenzwert der $Z_k(t)$. Es ist allerdings noch nicht klar, dass die Reihe $\sum_{j=0}^{\infty} W_{j,N_j(t)}$ konvergiert. Dies liefert das folgende Lemma, welches eine Abschätzung der $W_{k,i}$ angibt.

Da $Y_{k,N_k(t)-1} \leq t < Y_{k,N_k(t)}$ gilt, sind alle Z_k offensichtlich Elemente von $D[0,1]$ (vgl. Kapitel 1). Wir versehen $D[0,1]$ mit der σ -Algebra, die durch die Projektionen $f \mapsto f(t)$ erzeugt wird.

5.1 Lemma

Sei $M_k \stackrel{\text{def}}{=} \max \{W_{k,i} : 1 \leq i \leq 2^k\}$ die Länge des längsten Intervalls der k -ten Partition. Dann existiert für P -fast alle ω eine Konstante $K(\omega) \in \mathbb{N}$, so dass

$$M_k(\omega) \leq k \left(\frac{2}{3} \right)^{\frac{k}{2}} \quad \text{für alle } k \geq K(\omega)$$

gilt.

Beweis:

Mit $F_k \stackrel{\text{def}}{=} \sum_{i=1}^{2^k} W_{k,i}^2$ gilt $M_k^2 \leq F_k$. Sei weiterhin $\mathcal{F}_k = \sigma(\{U_{j,i} : j < k, 1 \leq i \leq 2^j\})$. Aus der Konstruktion der Unterteilungen folgt

$$W_{k+1,2i-1}^2 + W_{k+1,2i}^2 = W_{k,i}^2 (U_{k,i}^2 + (1 - U_{k,i})^2) \quad \text{für alle } k \in \mathbb{N}_0, 1 \leq i \leq 2^k,$$

denn:

$$\begin{aligned} W_{k+1,2i-1}^2 + W_{k+1,2i}^2 &= (Y_{k+1,2i-1} - Y_{k+1,2i-2})^2 + (Y_{k+1,2i} - Y_{k+1,2i-1})^2 \\ &= 2Y_{k+1,2i-1}^2 - 2Y_{k+1,2i-1}Y_{k+1,2i-2} + Y_{k+1,2i-2}^2 + Y_{k+1,2i}^2 \\ &\quad - 2Y_{k+1,2i}Y_{k+1,2i-1} \\ &= 2(U_{k,i}Y_{k,i-1} + (1 - U_{k,i})Y_{k,i})^2 - 2(U_{k,i}Y_{k,i-1} + (1 - U_{k,i})Y_{k,i})Y_{k,i-1} \\ &\quad + Y_{k,i-1}^2 + Y_{k,i}^2 - 2Y_{k,i}(U_{k,i}Y_{k,i-1} + (1 - U_{k,i})Y_{k,i}) \\ &= 2U_{k,i}^2Y_{k,i-1}^2 + 4U_{k,i}(1 - U_{k,i})Y_{k,i-1}Y_{k,i} + 2(1 - U_{k,i})^2Y_{k,i}^2 \\ &\quad - 2U_{k,i}Y_{k,i-1}^2 + Y_{k,i-1}^2 + Y_{k,i}^2 - 2Y_{k,i-1}Y_{k,i} - 2(1 - U_{k,i})Y_{k,i}^2 \\ &= Y_{k,i}^2(1 + 2U_{k,i}^2 - 2U_{k,i}) + Y_{k,i-1}^2(1 + 2U_{k,i}^2 - 2U_{k,i}) \\ &\quad - 2Y_{k,i-1}Y_{k,i}(1 + 2U_{k,i}^2 - 2U_{k,i}) \\ &= (Y_{k,i} - Y_{k,i-1})^2(1 + 2U_{k,i}^2 - 2U_{k,i}) \\ &= W_{k,i}^2(U_{k,i}^2 + (1 - U_{k,i})^2) \end{aligned}$$

Somit ergibt sich mit Hilfe von $E(U_{k,i}^2 + (1 - U_{k,i})^2) = \frac{2}{3}$ sowie der Tatsache, dass $W_{k,i}$ $\mathcal{F}_k$ -messbar ist,

$$EF_{k+1} = E(E[F_{k+1} | \mathcal{F}_k])$$

$$\begin{aligned}
&= E \left(\sum_{i=1}^{2^{k+1}} E[W_{k+1,i}^2 \mid \mathcal{F}_k] \right) \\
&= E \left(\sum_{i=1}^{2^k} E[W_{k+1,2i-1}^2 + W_{k+1,2i}^2 \mid \mathcal{F}_k] \right) \\
&= E \left(\sum_{i=1}^{2^k} W_{k,i}^2 E(U_{k,i}^2 + (1 - U_{k,i})^2) \right) \\
&= \frac{2}{3} E F_k.
\end{aligned}$$

Aus $F_0 = 1$ ergibt sich iterativ $E F_k \leq \left(\frac{2}{3}\right)^k$. Wegen $\text{Var}(M_k) \leq E(M_k^2) \leq E(F_k)$, folgt mit der Tschebyschev-Ungleichung

$$P \left(M_k \geq k \left(\frac{2}{3} \right)^{\frac{k}{2}} \right) \leq \frac{\text{Var} \left(\frac{M_k}{k} \right)}{\left(\frac{2}{3} \right)^k} \leq \frac{1}{k^2}.$$

Eine Anwendung des Borel-Cantelli-Lemmas liefert dann direkt

$$P \left(\limsup_{k \rightarrow \infty} \left\{ \omega \in \Omega : M_k(\omega) \geq k \left(\frac{2}{3} \right)^{\frac{k}{2}} \right\} \right) = 0$$

bzw. gleichbedeutend

$$P \left(\liminf_{k \rightarrow \infty} \left\{ \omega \in \Omega : M_k(\omega) < k \left(\frac{2}{3} \right)^{\frac{k}{2}} \right\} \right) = 1$$

und damit die Behauptung des Lemmas. □

5.2 Bemerkung

Mit Hilfe des Lemmas und der Feststellung, dass $k \left(\frac{2}{3}\right)^{k/2} \leq \frac{1}{k^2}$ für alle $k \geq 24$ gilt, folgt direkt, dass die Reihe $Z(t) = \sum_{k=0}^{\infty} W_{k,N_k(t)}$ fast sicher konvergiert.

5.3 Lemma

Es gilt $P(Z \in D[0, 1]) = 1$ und $d_{\infty}(Z_k, Z) \rightarrow 0$ fast sicher.

Beweis:

Da alle Z_k fast sicher rechtsseitig stetige Funktionen mit existierenden linksseitigen Limiten sind, gilt dasselbe auch für den Grenzwert Z , und damit $Z \in D[0, 1]$ fast sicher, sofern die andere Behauptung des Lemmas bewiesen ist. Dafür halten wir zuerst fest, dass nach

Satz 1.7 ($D[0, 1], d_\infty$) ein vollständiger metrischer Raum ist. Es reicht also zu zeigen, dass $(Z_k(\omega))_{k \in \mathbb{N}_0}$ für P -fast alle ω eine Cauchyfolge bzgl. d_∞ ist. Betrachte dafür unter Berücksichtigung von Lemma 5.1

$$\begin{aligned} d_\infty(Z_{n+m}, Z_n) &= \sup_{t \in [0, 1]} |Z_{n+m}(t) - Z_n(t)| \\ &= \sup_{t \in [0, 1]} \left| \sum_{k=n+1}^{n+m} W_{k, N_k(t)} \right| \\ &\leq \sum_{k=n+1}^{n+m} k \left(\frac{2}{3} \right)^{\frac{k}{2}} \xrightarrow{n, m \rightarrow \infty} 0 \quad \text{fast sicher.} \end{aligned}$$

□

Nach eventuellem Umdefinieren auf einer P -Nullmenge kann Z als zufälliges Element aus $D[0, 1]$ aufgefasst werden. Es ist jedoch klar, dass die Pfade von Z nicht stetig sind. Allerdings sind die Unstetigkeitsstellen in der abzählbaren Menge der inneren Unterteilungspunkte $\{Y_{k,i} : k \in \mathbb{N}, 1 \leq i \leq 2^k\}$ enthalten.

Das folgende Lemma liefert eine Stetigkeitseigenschaft von Z , die später in Satz 5.9 dazu dienen wird, aus der Verteilungskonvergenz der X_n gegen Z in der Skorokhod-Topologie auf die Verteilungskonvergenz in der gewöhnlichen Topologie zu schließen.¹

5.4 Lemma

Für alle $t \in [0, 1]$ gilt $P(\{\omega \in \Omega : s \mapsto Z(s, \omega) \text{ ist stetig in } t\}) = 1$.

Beweis:

Für $t = 0$ ist nichts zu zeigen, da Z fast sicher rechtsseitig stetig ist. Auch für $t = 1$ ist die Aussage klar, denn für die Unstetigkeit von Z in 1, müsste ein $Y_{k,i} = 1$ sein. Dies geschieht aber nur mit Wahrscheinlichkeit 0.

Sei nun $0 < t < 1$. Zur Konstruktion der Nullmenge, außerhalb derer die Funktion $s \mapsto Z(s, \omega)$ stetig in t ist, werden wir zuerst induktiv zeigen, dass die Verteilung von $Y_{k,i}$ für alle $k \in \mathbb{N}$ und $1 \leq i < 2^k$ Lebesgue-stetig ist. Für den Induktionsanfang mit $k = 1$ halten wir fest, dass $Y_{1,1} = 1 - U_{0,1}$ gilt und $U_{0,1}$ Lebesgue-stetig ist. Für gerade i folgt die Lebesgue-Stetigkeit von $Y_{k+1,i}$ direkt aus der von $Y_{k, \frac{i}{2}}$ durch die Induktionsvoraussetzung. Für ungerade i , also $i = 2j + 1$ für ein $j \in \mathbb{N}_0$, folgt die Lebesgue-Stetigkeit von $Y_{k+1,i}$ aus der Zerlegung

$$Y_{k+1, 2j+1} = U_{k, j+1} Y_{k, j} + (1 - U_{k, j+1}) Y_{k, j+1}.$$

Denn $Y_{k+1,i}$ ist wegen der Unabhängigkeit von $U_{k, j+1}$ zu $Y_{k, j}$ und $Y_{k, j+1}$ eine konvexe Kombination der Lebesgue-stetigen Zufallsgrößen $Y_{k, j}$ und $Y_{k, j+1}$ und somit selbst auch Lebesgue-stetig.

¹vgl. dazu Satz 1.10

Es folgt nun also für alle $k \in \mathbb{N}_0$ und $0 \leq i \leq 2^k$, dass $P(\{\omega \in \Omega : Y_{k,i}(\omega) = t\}) = 0$ gilt, und damit auch $P(A) = 0$ mit

$$A \stackrel{\text{def}}{=} \left\{ \omega \in \Omega : Y_{k,i}(\omega) = t \text{ für ein } k \in \mathbb{N}_0, 0 \leq i \leq 2^k \right\}.$$

Seien nun $\omega \in A^c \cap B^c$ und $\epsilon > 0$ gegeben, wobei B die Nullmenge aus Lemma 5.3 bezeichnet. Wähle nun nach Lemma 5.3 ein $k = k(\omega, \epsilon)$, so dass $|Z_k(s, \omega) - Z(s, \omega)| < \frac{\epsilon}{2}$ für alle $s \in [0, 1]$ gilt. Da $t \notin \{Y_{k,i}(\omega) : k \in \mathbb{N}_0, 0 \leq i \leq 2^k\}$, existiert ein $\delta > 0$, so dass die Abbildung $s \mapsto Z_k(s, \omega)$ konstant auf dem Intervall $(t - \delta, t + \delta)$ ist. Es folgt für alle s aus diesem Intervall

$$\begin{aligned} |Z(s, \omega) - Z(t, \omega)| &\leq |Z(s, \omega) - Z_k(s, \omega)| + |Z_k(s, \omega) - Z_k(t, \omega)| + |Z_k(t, \omega) - Z(t, \omega)| \\ &\leq \frac{\epsilon}{2} + 0 + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

□

5.2 Schwache Konvergenz gegen den FIND-Prozess

Das Ziel dieses Abschnitts ist es, die Verteilungskonvergenz der in (5.1) auf Seite 29 definierten Folge $(X_n)_{n \in \mathbb{N}}$ gegen den FIND-Prozess zu zeigen. Dafür fassen wir beide als zufällige Elemente von $D[0, 1]$ auf. Da die Sprungstellen der X_n in der Menge $\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\}$ liegen, die Verteilung der Sprungstellen von Z jedoch Lebesgue-stetig ist, können wir nicht erwarten, dass die Konvergenz bzgl. der gleichmäßigen Topologie, induziert durch d_∞ , gilt. Vielmehr werden wir die Konvergenz in der Skorokhod-Topologie, die wir im ersten Kapitel eingeführt haben, nachweisen².

Wir werden jetzt zuerst das angesprochene Ergebnis angeben, jedoch vor dessen vollständigen Beweis einige dafür benötigte Lemmata beweisen. Dabei bezeichnet $\xrightarrow{d}$ die Verteilungskonvergenz von Zufallsgrößen.

5.5 Satz

$X_n \xrightarrow{d} Z$ in $(D[0, 1], \rho)$ für $n \rightarrow \infty$.

Beim Beweis des Satzes werden wir wie folgt vorgehen: Wir konstruieren eine Folge $(Z^n)_{n \in \mathbb{N}}$ mit der Eigenschaft $\mathcal{L}(Z^n) = \mathcal{L}(X_n)$ und zeigen, dass diese in der Skorokhod-Topologie gegen Z konvergiert, also $\rho(Z^n, Z) \rightarrow 0$ erfüllt. Hieraus folgt dann die Behauptung des Satzes.

Um eine solche Folge zu erhalten, starten wir mit derselben Menge $\{U_{k,i} : k \in \mathbb{N}_0, 1 \leq i \leq 2^k\}$ von unabhängigen auf $(0, 1)$ gleichverteilten Zufallsgrößen und diskretisieren das Vorgehen

²vgl. Satz und Definition 1.8

zur Konstruktion von Z .

Seien hierfür für alle $n \in \mathbb{N}$ die Zufallsgrößen $\{Y_{k,i}^n : k \in \mathbb{N}_0, 0 \leq i \leq 2^k\}$ rekursiv definiert durch

$$Y_{0,0}^n \stackrel{\text{def}}{=} 0, \quad Y_{0,1}^n \stackrel{\text{def}}{=} 1, \quad Y_{k+1,2i}^n \stackrel{\text{def}}{=} Y_{k,i}^n \quad \text{für alle } 0 \leq i \leq 2^k$$

und für $0 \leq i < 2^k$

$$Y_{k+1,2i+1}^n \stackrel{\text{def}}{=} \begin{cases} Y_{k+1,2i}^n, & \text{falls } Y_{k,i}^n = Y_{k,i+1}^n \\ Y_{k,i}^n + \frac{1}{n} \left(1 + \left\lfloor n(1 - U_{k,i+1}) (Y_{k,i+1}^n - Y_{k,i}^n) \right\rfloor \right), & \text{sonst.} \end{cases}$$

Durch die $Y_{k,i}^n$ erhalten wir wieder für jedes $n \in \mathbb{N}$ eine Folge von aufeinander aufbauenden Unterteilungen des Einheitsintervalls, jedoch diesmal mit der Eigenschaft

$$Y_{k,i}^n \in \left\{ 0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1 \right\} \quad \text{für alle } k \in \mathbb{N}_0, 0 \leq i \leq 2^k.$$

Analog wie beim Vorgehen zur Konstruktion von Z erhalten wir nun zu jedem $t \in [0, 1]$ und $k \in \mathbb{N}_0$ ein eindeutiges

$$N_k^n(t) \stackrel{\text{def}}{=} \min \left\{ l \in \{1, \dots, 2^k\} : Y_{k,l}^n > t \right\}.$$

Setze $N_k^n(1) \stackrel{\text{def}}{=} 2^k$. Für die $N_k^n(t)$ gilt eine analoge Aussage wie (5.3) für die $N_k(t)$, also

$$N_{k-1}^n(t) = \begin{cases} \frac{N_k^n(t)}{2}, & \text{falls } N_k^n(t) \text{ gerade ist} \\ \frac{N_k^n(t)+1}{2}, & \text{sonst} \end{cases} \quad (5.4)$$

für alle $n \in \mathbb{N}$, $t \in [0, 1]$, $k \in \mathbb{N}$.

Sei weiter $W_{k,i}^n \stackrel{\text{def}}{=} Y_{k,i}^n - Y_{k,i-1}^n$ für alle $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ und $1 \leq i \leq 2^k$. Definiere dieses Mal

$$Z_k^n(t) \stackrel{\text{def}}{=} \sum_{j=0}^k \max \left\{ W_{j,N_j^n(t)} - \frac{1}{n}, 0 \right\}, \quad Z^n(t) \stackrel{\text{def}}{=} \sup_{k \in \mathbb{N}} Z_k^n(t).$$

Da die Z_k^n monoton wachsend in k sind, können wir Z^n auch als $\lim_{k \rightarrow \infty} Z_k^n$ schreiben. Zur Konvergenz der Reihe stellen wir nur fest, dass zu jedem n die Zufallsgrößen $W_{j,N_j^n(t)}$ ab einem bestimmten j mit Wahrscheinlichkeit 1 immer kleiner oder gleich $\frac{1}{n}$ sind.

Beachte, dass die stetigen $U_{k,i}$ mittels der unteren Gaußklammer für die diskrete Wahl der Pivotelemente verwendet werden und die Zufallsgröße $nW_{j,N_j^n(t)}$ mit der Größe der mit FIND untersuchten Menge zu Beginn des $(j+1)$ -ten Rekursionsschrittes übereinstimmt. Die Z^n summieren die Größen dieser Mengen auf, solange diese mehr als ein Element enthalten. Mit dieser Konstruktion haben wir also eine Folge Z^n konstruiert, die wie gewünscht $\mathcal{L}(Z^n) = \mathcal{L}(X_n)$ erfüllt.

Wir kommen nun zu den bereits angesprochenen Lemmata, die zum Beweis von obigem Satz benötigt werden. Das Erste liefert eine Abschätzung der Distanz zwischen den ursprünglichen und den obigen diskreten Unterteilungspunkten.

5.6 Lemma

Für alle $n \in \mathbb{N}$ und $k \in \mathbb{N}_0$ gilt:

$$\sup_{0 \leq i \leq 2^k} |Y_{k,i}^n - Y_{k,i}| \leq \frac{k}{n}.$$

Beweis:

Sei $D_{n,k} \stackrel{\text{def}}{=} \sup_{0 \leq i \leq 2^k} |Y_{k,i}^n - Y_{k,i}|$ das gesuchte Supremum. Nach der Definition der $Y_{k,i}$ und $Y_{k,i}^n$ ist klar, dass die Ungleichung für $k = 0$ und beliebiges $n \in \mathbb{N}$ stimmt. Anschließend wollen wir zeigen, dass die $D_{n,k}$ für alle $n \in \mathbb{N}$ die Rekursionsungleichung $D_{n,k+1} \leq D_{n,k} + \frac{1}{n}$ erfüllen. Damit folgt dann die Behauptung per Induktion nach k .

Sei $n \in \mathbb{N}$ beliebig. Definiere zusätzlich induktiv die Zufallsgrößen $\tilde{Y}_{k,i}^n$ durch

$$\tilde{Y}_{k+1,2i+1}^n \stackrel{\text{def}}{=} U_{k,i+1} Y_{k,i}^n + (1 - U_{k,i+1}) Y_{k,i+1}^n$$

für alle $k \in \mathbb{N}_0$ und $0 \leq i \leq 2^k - 1$.

Allgemein gilt für alle $a, b, c, d \in \mathbb{R}$ und $0 \leq \lambda \leq 1$

$$\begin{aligned} |(\lambda a + (1 - \lambda) b) - (\lambda c + (1 - \lambda) d)| &= |\lambda(a - c) + (1 - \lambda)(b - d)| \\ &\leq \max\{|a - c|, |b - d|\}. \end{aligned}$$

Damit folgt

$$\begin{aligned} |Y_{k+1,2i+1} - \tilde{Y}_{k+1,2i+1}^n| &= |(U_{k,i+1} Y_{k,i} + (1 - U_{k,i+1}) Y_{k,i+1}) \\ &\quad - (U_{k,i+1} Y_{k,i}^n + (1 - U_{k,i+1}) Y_{k,i+1}^n)| \\ &\leq \max\{|Y_{k,i} - Y_{k,i}^n|, |Y_{k,i+1} - Y_{k,i+1}^n|\} \\ &\leq D_{n,k}. \end{aligned} \tag{5.5}$$

Wir zeigen nun, dass aus der Definition von $Y_{k+1,2i+1}^n$ schon

$$|Y_{k+1,2i+1}^n - \tilde{Y}_{k+1,2i+1}^n| \leq \frac{1}{n} \tag{5.6}$$

folgt.

Im Fall $Y_{k,i}^n = Y_{k,i+1}^n$ folgt direkt aus der Definition $|Y_{k+1,2i+1}^n - \tilde{Y}_{k+1,2i+1}^n| = 0$. Gilt hingegen $Y_{k,i}^n \neq Y_{k,i+1}^n$, so ist

$$Y_{k+1,2i+1}^n = Y_{k,i}^n + \frac{1}{n} (1 + \lfloor n(1 - U_{k,i+1})(Y_{k,i+1}^n - Y_{k,i}^n) \rfloor).$$

Damit ergibt sich unter Beachtung von $\left| \frac{1}{n} (nx - (1 + \lfloor nx \rfloor)) \right| \leq \frac{1}{n}$ für alle $x \in \mathbb{R}$:

$$\begin{aligned}
\left| Y_{k+1,2i+1}^n - \tilde{Y}_{k+1,2i+1}^n \right| &= \left| Y_{k,i}^n + \frac{1}{n} \left(1 + \lfloor n(1 - U_{k,i+1}) (Y_{k,i+1}^n - Y_{k,i}^n) \rfloor \right) \right. \\
&\quad \left. - U_{k,i+1} Y_{k,i}^n - (1 - U_{k,i+1}) Y_{k,i+1}^n \right| \\
&= \left| (1 - U_{k,i+1}) (Y_{k,i}^n - Y_{k,i+1}^n) \right. \\
&\quad \left. + \frac{1}{n} \left(1 + \lfloor n(1 - U_{k,i+1}) (Y_{k,i+1}^n - Y_{k,i}^n) \rfloor \right) \right| \\
&\leq \frac{1}{n}.
\end{aligned}$$

Aus den beiden Ungleichungen (5.5) und (5.6) ergibt dies nun die gewünschte Beziehung zwischen $D_{n,k+1}$ und $D_{n,k}$, wenn man beachtet, dass die Distanz nur bei ungeraden zweiten Indizes der $Y_{k,i}$ bzw. $Y_{k,i}^n$ wachsen kann, da die $Y_{k+1,i}$ bzw. $Y_{k+1,i}^n$ mit geraden zweiten Index bereits in der k -ten Partition vorhanden sind.

$$\begin{aligned}
D_{n,k+1} &\leq \sup_{0 \leq i \leq 2^{k+1}} \left(\left| Y_{k+1,i}^n - \tilde{Y}_{k+1,i}^n \right| + \left| \tilde{Y}_{k+1,i}^n - Y_{k+1,i} \right| \right) \\
&\leq \sup_{0 \leq i \leq 2^{k+1}} \left| Y_{k+1,2i+1}^n - \tilde{Y}_{k+1,2i+1}^n \right| + \sup_{0 \leq i \leq 2^{k+1}} \left| \tilde{Y}_{k+1,2i+1}^n - Y_{k+1,2i+1} \right| \\
&\leq \frac{1}{n} + D_{n,k}
\end{aligned}$$

Damit folgt die Behauptung nun per Induktion über k .

□

Für fixiertes n erhalten wir nun bei wachsendem k eine Folge von aufeinander aufbauenden Partitionen mit Unterteilungspunkten aus der endlichen, fixierten Menge $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}, 1\}$. Offensichtlich ist $W_{k,N_k^n(t)}^n$ eine fallende Funktion in k . Definiere nun die Zufallsgröße

$$K_n(t) \stackrel{\text{def}}{=} \inf \left\{ k \in \mathbb{N}_0 : W_{k,N_k^n(t)}^n \leq \frac{1}{n} \right\},$$

wobei wie üblich $\inf \emptyset = \infty$ gesetzt wird. Beachte, dass

$$Z^n(t) = Z_{K_n(t)}^n(t) \quad \text{für alle } t \in [0, 1] \quad (5.7)$$

gilt, da $\max \left\{ W_{k,N_k^n(t)}^n - \frac{1}{n}, 0 \right\} = 0$ für alle $k \geq K_n(t)$ gilt.

Diese Größe $K_n(t)$ hat eine direkte Interpretation im Rahmen des Algorithmus FIND. Die Verteilung von $K_n(t)$ ist dieselbe wie die der Anzahl der Rekursionen, die FIND benötigt, um das $(\lfloor nt \rfloor + 1)$ -größte Element aus der Menge $\{1, \dots, n\}$ zu finden. Ist nämlich $nW_{k,N_k^n(t)}^n \leq 1$, so bedeutet dies, dass die von FIND zu untersuchende Menge nur noch ein Element enthält und FIND somit das gesuchte Element gefunden hat.

Wir werden jetzt zuerst ein Lemma beweisen, dass wir benötigen um obere und untere Schranken für $K_n(t)$ herzuleiten.

5.7 Lemma

Sei $a \in \mathbb{N}$ und X eine auf $\{1, 2, \dots, a\}$ gleichverteilte Zufallsgröße. Definiere nun die Zufallsgröße Y durch

$$Y = \begin{cases} (a-1)^2, & \text{falls } X = a \\ (X-1)^2 + (a-X-1)^2, & \text{sonst.} \end{cases}$$

Dann gilt $EY \leq \frac{11}{12}(a-1)^2$.

Beweis:

Für $a = 1$ ist die Behauptung trivial. Sei nun also $a \geq 2$. Unter Berücksichtigung von $\sum_{i=1}^n i = \frac{n(n+1)}{2}$ und $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$ ergibt sich

$$\begin{aligned} EY &= \frac{1}{a} \left((a-1)^2 + \sum_{i=1}^{a-1} (i-1)^2 + (a-i-1)^2 \right) \\ &= \frac{1}{a} \left((a-1)^2 + \sum_{i=1}^{a-1} ((a-1)^2 + 2i^2 - 2ai + 1) \right) \\ &= \frac{1}{a} \left((a-1)^2 + (a-1)^3 - a^2(a-1) + \frac{1}{3}(a(a-1)(2a-1)) + (a-1) \right) \\ &= \frac{1}{a} \left((a-1)^2 + \frac{1}{3}(a-1)(a-2)(2a-3) \right). \end{aligned}$$

Für $a = 2, 3$ lässt sich die Ungleichung direkt nachrechnen, für $a \geq 4$ ergibt sie sich wie folgt:

$$\begin{aligned} \frac{1}{a} \left((a-1)^2 + \frac{1}{3}(a-1)(a-2)(2a-3) \right) &\leq \frac{1}{a} \left((a-1)^2 + \frac{1}{3}(a-1)^2(2a-2) \right) \\ &\leq \frac{1}{a} \left((a-1)^2 + \frac{2}{3}(a-1)^3 \right) \\ &\leq \frac{1}{4}(a-1)^2 + \frac{2}{3}(a-1)^2 \\ &= \frac{11}{12}(a-1)^2. \end{aligned}$$

□

5.8 Lemma

Seien $\tau_n \stackrel{\text{def}}{=} \inf_{0 \leq t \leq 1} K_n(t)$, $T_n \stackrel{\text{def}}{=} \sup_{0 \leq t \leq 1} K_n(t)$. Dann gilt

- (i) $\tau_n \rightarrow \infty$ fast sicher
- (ii) $T_n = O(\log n)$ fast sicher
- (iii) $ET_n = O(\log n)$.

Beweis:

Zu (i):

Fixiere ein $k \in \mathbb{N}$ und sei $\omega \in \Omega'$ beliebig, wobei

$$\Omega' \stackrel{\text{def}}{=} \left\{ \omega \in \Omega : 0 = Y_{k,0}(\omega) < Y_{k,1}(\omega) < \dots < Y_{k,2^k}(\omega) = 1 \quad \text{für alle } k \in \mathbb{N}_0 \right\}.$$

Wähle nun $n_0(k) \in \mathbb{N}$ so groß, dass

$$\frac{2k+2}{n} \leq \min \left\{ W_{k,i}(\omega) : 1 \leq i \leq 2^k \right\} \quad \text{für alle } n \geq n_0(k)$$

gilt. Mit Hilfe von Lemma 5.6 erhalten wir für alle $1 \leq j \leq 2^k$ und $n \geq n_0(k)$

$$\begin{aligned} W_{k,j}^n(\omega) &= Y_{k,j}^n(\omega) - Y_{k,j-1}^n(\omega) \\ &\geq Y_{k,j}(\omega) - Y_{k,j-1}(\omega) - \frac{2k}{n} \\ &\geq \min \left\{ W_{k,i}(\omega) : 1 \leq i \leq 2^k \right\} - \frac{2k}{n} \\ &\geq \frac{2k+2}{n} - \frac{2k}{n} \\ &= \frac{2}{n}. \end{aligned}$$

Damit folgt aber schon $\tau_n(\omega) \geq k$ für alle $n \geq n_0(k)$, also auch $\tau_n \xrightarrow{n \rightarrow \infty} \infty$ fast sicher.

Zu (ii):

Sei $V_{k,i}^n \stackrel{\text{def}}{=} \max \left\{ W_{k,i}^n - \frac{1}{n}, 0 \right\}$. Wir werden zuerst per Induktion nach k zeigen, dass

$$E \left(\max \left\{ V_{k,i}^n : 1 \leq i \leq 2^k \right\} \right)^2 \leq \left(\frac{11}{12} \right)^k \quad (5.8)$$

gilt. Mit Hilfe dieser Ungleichung werden wir dann im Anschluss

$$P \left(\limsup_{n \rightarrow \infty} T_n > C \log(n) \right) = 0$$

für eine Konstante $0 < C < \infty$ zeigen.

Sei wieder $\mathcal{F}_k = \sigma \left(\{U_{j,i} : j < k, 1 \leq i \leq 2^j\} \right)$. Es ist klar, dass (5.8) für $k = 0$ gilt. Für

den Induktionsschritt von k nach $k+1$ betrachten wir zunächst wie $V_{k+1,2i+1}^n$ und $V_{k+1,2i+2}^n$ aus $Y_{k,i}^n$ und $Y_{k,i+1}^n$ hervorgehen.

Definiere für den Rest des Beweises die Zufallsgröße

$$\xi \stackrel{\text{def}}{=} 1 + \lfloor n(1 - U_{k,i+1})(Y_{k,i+1}^n - Y_{k,i}^n) \rfloor.$$

Dann ist die bedingte Verteilung von ξ gegeben $\mathcal{F}_k$ eine Gleichverteilung auf $\{1, \dots, nW_{k,i+1}^n\}$, also $\mathcal{L}(\xi|\mathcal{F}_k) = U(\{1, \dots, nW_{k,i+1}^n\})$, da $n(1 - U_{k,i+1})$ gleichverteilt auf dem Intervall $(0, n)$ ist.

Für den Fall, dass $\xi = nW_{k,i+1}^n$ ist, gilt $Y_{k+1,2i+1}^n = Y_{k,i+1}^n$ und $(V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 = (V_{k,i+1}^n)^2$, denn:

$$Y_{k+1,2i+1}^n = Y_{k,i}^n + \frac{1}{n} \left(1 + \lfloor n(1 - U_{k,i+1})(Y_{k,i+1}^n - Y_{k,i}^n) \rfloor \right) = Y_{k,i+1}^n$$

und damit

$$\begin{aligned} (V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 &= \left(\max \left\{ Y_{k+1,2i+1}^n - Y_{k+1,2i}^n - \frac{1}{n}, 0 \right\} \right)^2 \\ &\quad + \left(\max \left\{ Y_{k+1,2i+2}^n - Y_{k+1,2i+1}^n - \frac{1}{n}, 0 \right\} \right)^2 \\ &= \left(\max \left\{ \frac{1}{n}\xi - \frac{1}{n}, 0 \right\} \right)^2 \\ &\quad + \left(\max \left\{ Y_{k,i+1}^n - Y_{k,i}^n - \frac{1}{n}\xi - \frac{1}{n}, 0 \right\} \right)^2 \\ &= \left(\max \left\{ W_{k,i+1}^n - \frac{1}{n}, 0 \right\} \right)^2 + \left(\max \left\{ -\frac{1}{n}, 0 \right\} \right)^2 \\ &= (V_{k,i+1}^n)^2. \end{aligned}$$

Gleiches stimmt natürlich auch im Fall $Y_{k,i}^n = Y_{k,i+1}^n$. Die obige Rechnung zeigt bereits, dass für $1 \leq \xi \leq nW_{k,i+1}^n - 1$ gilt:

$$\begin{aligned} V_{k+1,2i+1}^n &= \frac{1}{n}(\xi - 1) \\ V_{k+1,2i+2}^n &= \frac{1}{n}(nW_{k,i+1}^n - \xi - 1). \end{aligned}$$

Interpretiert man ξ gegeben $\mathcal{F}_k$ als eine Zufallsgröße X und $n^2 \left((V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 \right)$ gegeben $\mathcal{F}_k$ als eine Zufallsgröße Y , so lässt sich Lemma 5.7 anwenden, wenn man berücksichtigt, dass $a \stackrel{\text{def}}{=} nW_{k,i+1}^n$ gegeben $\mathcal{F}_k$ ein Element aus $\mathbb{N}$ ist. Insgesamt erhalten wir

dann:

$$\begin{aligned} n^2 E \left[(V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 \middle| \mathcal{F}_k \right] &\leq \frac{11}{12} (nW_{k,i+1}^n - 1)^2 \\ &\leq \frac{11}{12} n^2 (V_{k,i+1}^n)^2, \end{aligned}$$

und damit

$$E \left[(V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 \middle| \mathcal{F}_k \right] \leq \frac{11}{12} (V_{k,i+1}^n)^2.$$

Mit $F_k^n \stackrel{\text{def}}{=} \sum_{i=1}^{2^k} (V_{k,i}^n)^2$ ergibt sich

$$\begin{aligned} E[F_{k+1}^n | \mathcal{F}_k] &= E \left[\sum_{i=1}^{2^{k+1}} (V_{k+1,i}^n)^2 \middle| \mathcal{F}_k \right] \\ &= E \left[\sum_{i=0}^{2^k-1} (V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 \middle| \mathcal{F}_k \right] \\ &= \sum_{i=0}^{2^k-1} E \left[(V_{k+1,2i+1}^n)^2 + (V_{k+1,2i+2}^n)^2 \middle| \mathcal{F}_k \right] \\ &\leq \frac{11}{12} \sum_{i=0}^{2^k-1} (V_{k,i+1}^n)^2 \\ &= \frac{11}{12} F_k^n. \end{aligned}$$

Daraus folgt direkt $EF_{k+1}^n = E(E[F_{k+1}^n | \mathcal{F}_k]) \leq \frac{11}{12} EF_k^n$ und da $EF_0^n = \left(\frac{n-1}{n}\right)^2 \leq 1$ gilt, ergibt sich induktiv das Gewünschte, d.h.

$$E \left(\max \left\{ V_{k,i}^n : 1 \leq i \leq 2^k \right\} \right)^2 \leq EF_k^n \leq \left(\frac{11}{12} \right)^k.$$

Um nun den Beweis von (ii) abzuschließen, sei $\phi(n) \stackrel{\text{def}}{=} \lceil C \log(n) \rceil$ mit einer Konstanten $0 < C < \infty$. Aus der Markov-Ungleichung mit $g(t) = t^2$ und Ungleichung (5.8) folgt

$$\begin{aligned} P \left(\max \left\{ V_{\phi(n),i}^n : 1 \leq i \leq 2^{\phi(n)} \right\} \geq \frac{1}{n} \right) &\leq n^2 E \left(\max \left\{ V_{\phi(n),i}^n : 1 \leq i \leq 2^{\phi(n)} \right\} \right)^2 \\ &\leq n^2 \left(\frac{11}{12} \right)^{\phi(n)} \\ &= \exp \left(\left(2 - C \log \left(\frac{12}{11} \right) \right) \log(n) \right) \end{aligned}$$

Somit erhalten wir für $C > \frac{3}{\log(\frac{12}{11})}$ mit dem Borel-Cantelli-Lemma

$$P\left(\max\left\{V_{\phi(n),i}^n : 1 \leq i \leq 2^{\phi(n)}\right\} \geq \frac{1}{n} \text{ für unendlich viele } n\right) = 0.$$

Zuletzt stellen wir fest, dass für alle $k \in \mathbb{N}_0$ und $n \in \mathbb{N}$

$$T_n > k \iff \max\left\{V_{k,i}^n : 1 \leq i \leq 2^k\right\} \geq \frac{1}{n}$$

gilt, denn:

$$\begin{aligned} T_n > k &\iff \sup_{t \in [0,1]} \inf \left\{ k \in \mathbb{N}_0 : W_{k,N_k^n(t)} \leq \frac{1}{n} \right\} > k \\ &\iff \max\left\{W_{k,i}^n : 1 \leq i \leq 2^k\right\} > \frac{1}{n} \\ &\iff \max\left\{V_{k,i}^n : 1 \leq i \leq 2^k\right\} \geq \frac{1}{n}. \end{aligned}$$

Nun folgt mit $k = \phi(n)$

$$P\left(T_n > C \log(n) \text{ für unendlich viele } n\right) = 0,$$

und damit, dass $T_n = O(\log n)$ fast sicher gilt.

Zu (iii):

Aus obiger Äquivalenz, der Markov-Ungleichung mit $g(t) = t^2$ und der Ungleichung (5.8) auf Seite 40 folgt für alle $k \in \mathbb{N}_0$ und $n \in \mathbb{N}$

$$\begin{aligned} P(T_n > k) &= P\left(\max\left\{V_{k,i}^n : 1 \leq i \leq 2^k\right\} \geq \frac{1}{n}\right) \\ &\leq n^2 E\left(\max\left\{V_{k,i}^n : 1 \leq i \leq 2^k\right\}\right)^2 \\ &\leq n^2 \left(\frac{11}{12}\right)^k. \end{aligned}$$

Mit Hilfe einer aus der WT bekannten Integrationsformel für den Erwartungswert und der geometrischen Reihe liefert dies für alle $k \in \mathbb{N}_0$ und $n \in \mathbb{N}$

$$\begin{aligned} ET_n &= \sum_{l=1}^k P(T_n \geq l) + \sum_{l=k}^{\infty} P(T_n > l) \\ &\leq k + n^2 \sum_{l=k}^{\infty} \left(\frac{11}{12}\right)^l \\ &= k + 12 n^2 \left(\frac{11}{12}\right)^k. \end{aligned}$$

Die Behauptung folgt nun, wenn wir zu jedem $n \in \mathbb{N}$ ein $k(n) = C \log(n)$ mit so großem

C wählen, dass

$$ET_n \leq C \log n + 12n^2 \left(\frac{11}{12} \right)^{C \log n}$$

und somit $ET_n = O(\log(n))$ gilt.

□

Wir kehren nun zum Beweis von Satz 5.5 zurück, d.h. zum Nachweis von $X_n \xrightarrow{d} Z$ in $(D[0, 1], \rho)$.

Beweis:

Wegen der Verteilungsgleichheit von Z^n und X_n für jedes $n \in \mathbb{N}$, genügt zum Beweis von Satz 5.5 der Nachweis der fast sicheren Konvergenz von Z^n gegen Z in der Skorokhod-Topologie, d.h. $\rho(Z^n, Z) \rightarrow 0$ fast sicher. Wie in Bemerkung 1.9 festgestellt, müssen wir dafür eine Folge $(\lambda_n)_{n \in \mathbb{N}} \subset \Lambda$ finden, die fast sicher

$$d_\infty(\lambda_n, \text{id}) \xrightarrow{n \rightarrow \infty} 0 \quad \text{und} \quad d_\infty(Z^n, Z \circ \lambda_n) \xrightarrow{n \rightarrow \infty} 0 \quad (5.9)$$

erfüllt.

Fixiere nun für den Rest des Beweises ein ω aus dem Komplement der Vereinigung aller in diesem Kapitel bisher aufgetauchten Nullmengen. Wenn möglich werden wir zur übersichtlicheren Notation darauf verzichten, dieses ω in den Rechnungen extra mitzuführen. Es sei aber darauf hingewiesen, dass rein formal alle auftretenden Zufallsgrößen natürlich von ω abhängen. Gleiches gilt auch für die als Nächstes definierte Folge von Funktionen $(\lambda_n)_{n \in \mathbb{N}} : [0, 1] \rightarrow [0, 1]$ mit $\lambda_n(0) \stackrel{\text{def}}{=} 0$ und $\lambda_n(1) \stackrel{\text{def}}{=} 1$. Des Weiteren seien die λ_n linear auf den Intervallen $[\frac{m-1}{n}, \frac{m}{n}]$, $1 \leq m \leq n$, und für $1 \leq m < n$ gelte

$$\lambda_n\left(\frac{m}{n}\right) \stackrel{\text{def}}{=} Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})},$$

wobei $I_n(t) \stackrel{\text{def}}{=} N_{K_n(t)}^n(t) - 1$.

Wir werden jetzt zunächst zeigen, dass diese Funktionen wirklich Elemente von Λ sind. Dafür muss nachgewiesen werden, dass die λ_n streng monoton wachsend sind. Wir stellen zuerst fest, dass $Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})}^n = \frac{m}{n}$ gilt. Dies sieht man wie folgt: Die Definition von $K_n(t)$ impliziert, dass $W_{K_n(\frac{m}{n}), N_k^n(\frac{m}{n})}^n \in \{0, \frac{1}{n}\}$ sein muss. Zusammen mit der Definition von $N_k^n(t)$, die sicherstellt, dass $Y_{K_n(\frac{m}{n}), N_k^n(\frac{m}{n})}^n > \frac{m}{n}$ und $Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})}^n \leq \frac{m}{n}$ gilt, liefert dies schon wie gewünscht $Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})}^n = \frac{m}{n}$.

Die strenge Monotonie der λ_n folgt jetzt aus der Tatsache, dass $Y_{k,i}^n < Y_{l,j}^n$ schon $Y_{k,i} < Y_{l,j}$ impliziert. Für $k = l$ ist dies direkt klar. Für $k > l$ mit $k - l = p$ folgt aus Gleichung (5.2), dass $i \leq 2^p \cdot j$ ist und damit $Y_{k,i} < Y_{l,j}$. Der umgekehrte Fall für $k < l$ läuft analog. Die λ_n sind also wirklich Elemente von Λ .

Wir müssen nun noch zeigen, dass die λ_n wirklich die beiden geforderten Eigenschaften (5.9) erfüllen und beginnen mit dem Nachweis von $d_\infty(\lambda_n, \text{id}) \rightarrow \infty$. Der maximale Abstand wird offensichtlich in einem Punkt angenommen, in dem λ_n seine Steigung verändert. Unter Berücksichtigung der Lemmata 5.6 und 5.8 folgt

$$\begin{aligned}
 d_\infty(\lambda_n, \text{id}) &= \max_{1 \leq m < n} \left| \lambda_n \left(\frac{m}{n} \right) - \text{id} \left(\frac{m}{n} \right) \right| \\
 &\leq \max_{1 \leq m < n} \left| \lambda_n \left(\frac{m}{n} \right) - Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})}^n \right| + \underbrace{\max_{1 \leq m < n} \left| Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})}^n - \text{id} \left(\frac{m}{n} \right) \right|}_{=0} \\
 &= \max_{1 \leq m < n} \left| Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})} - Y_{K_n(\frac{m}{n}), I_n(\frac{m}{n})}^n \right| \\
 &\leq \max_{1 \leq m < n} \frac{K_n(\frac{m}{n})}{n} \leq \sup_{t \in [0,1]} \frac{K_n(t)}{n} \\
 &= \frac{T_n}{n} = O\left(\frac{\log n}{n}\right) \xrightarrow{n \rightarrow \infty} 0
 \end{aligned}$$

Schlussendlich nun zum Beweis der zweiten Eigenschaft $d_\infty(Z^n, Z \circ \lambda_n) \rightarrow 0$. Dafür werden wir zuerst die Existenz eines $N(\omega) \in \mathbb{N}$ beweisen, so dass für alle $t \in [0, 1]$ und alle $n \geq N(\omega)$ die Gleichheit

$$N_{K_n(t)}^n(t) = N_{K_n(t)}(\lambda_n(t)) \quad (5.10)$$

erfüllt ist.

Sei also $t \in [0, 1]$ beliebig. Zu jedem $n \in \mathbb{N}$ wähle ein $m(n) \in \mathbb{N}_0$, so dass $\frac{m(n)}{n} \leq t < \frac{m(n)+1}{n}$ gilt. Wegen der Wahl von $m(n)$ ist aufgrund der Definitionen von $K_n(t)$ und $N_k(t)$ klar, dass $K_n(t) = K_n\left(\frac{m(n)}{n}\right)$ und

$$N_{K_n(t)}^n(t) = N_{K_n(t)}^n\left(\frac{m(n)}{n}\right) \quad (5.11)$$

erfüllt sind. Außerdem folgt aus der Definition von λ_n

$$\begin{aligned}
 N_{K_n(t)}\left(\lambda_n\left(\frac{m(n)}{n}\right)\right) &= N_{K_n(t)}\left(Y_{K_n(\frac{m(n)}{n}), I_n(\frac{m(n)}{n})}\right) \\
 &= N_{K_n(t)}\left(Y_{K_n(t), I_n(\frac{m(n)}{n})}\right) \\
 &= I_n\left(\frac{m(n)}{n}\right) + 1 = N_{K_n(t)}^n\left(\frac{m(n)}{n}\right). \quad (5.12)
 \end{aligned}$$

Aufgrund der Stetigkeit der λ_n können wir nun zu dem fixierten ω ein so großes $N(\omega) \in \mathbb{N}$ wählen, dass für alle $n \geq N(\omega)$ fast sicher kein $l \in \{1, \dots, 2^k\}$ mit $\lambda\left(\frac{m(n)}{n}\right) < Y_{K_n(t), l} \leq \lambda_n(t)$ existiert.

Dann gilt aber $N_{K_n(t)}(\lambda_n(t)) = N_{K_n(t)}\left(\lambda_n\left(\frac{m(n)}{n}\right)\right)$ für alle $n \geq N(\omega)$ und zusammen

mit den Gleichungen 5.11 und 5.12 folgt die Behauptung aus Gleichung 5.10 durch

$$N_{K_n(t)}^n(t) = N_{K_n(t)}^n\left(\frac{m(n)}{n}\right) = N_{K_n(t)}\left(\lambda_n\left(\frac{m(n)}{n}\right)\right) = N_{K_n(t)}(\lambda_n(t)).$$

Da gemäß (5.3) alle $N_j(t)$ mit $j < K_n(t)$ schon eindeutig durch $N_{K_n(t)}(t)$ bestimmt sind, und gleiches auch für die $N_j^n(t)$ gilt, folgt insgesamt schon

$$N_j(\lambda_n(t)) = N_j^n(t) \quad \text{für alle } j \leq K_n(t), 0 \leq t \leq 1.$$

Mit Hilfe von Lemma 5.6 und Gleichung (5.2) folgt dann für alle $j \leq K_n(t)$ und $0 \leq t \leq 1$:

$$\begin{aligned} \left| V_{j, N_j^n(t)}^n - W_{j, N_j(\lambda_n(t))} \right| &= \left| \max \left\{ W_{j, N_j^n(t)}^n - \frac{1}{n}, 0 \right\} - W_{j, N_j(\lambda_n(t))} \right| \\ &= \begin{cases} \left| Y_{j, N_j^n(t)-1} - Y_{j, N_j^n(t)} \right|, & \text{falls } j = K_n(t) \\ \left| Y_{j, N_j^n(t)}^n - Y_{j, N_j^n(t)} + \left(Y_{j, N_j^n(t)-1} - Y_{j, N_j^n(t)-1}^n \right) - \frac{1}{n} \right|, & \text{sonst} \end{cases} \\ &\leq \left| Y_{j, N_j^n(t)}^n - Y_{j, N_j^n(t)} \right| + \left| Y_{j, N_j^n(t)-1} - Y_{j, N_j^n(t)-1}^n \right| + \frac{1}{n} \\ &\leq \frac{2j+1}{n}. \end{aligned}$$

Durch Ausnutzen dieser Abschätzung und der Beziehung (5.7) ergibt sich für alle $t \in [0, 1]$

$$\begin{aligned} \left| Z^n(t) - Z_{K_n(t)}(\lambda_n(t)) \right| &= \left| \sum_{j=0}^{K_n(t)} V_{j, N_j^n(t)}^n - W_{j, N_j(\lambda_n(t))} \right| \\ &\leq \sum_{j=0}^{K_n(t)} \frac{2j+1}{n} \\ &\leq \sum_{j=0}^{T_n} \frac{2j+1}{n} = \frac{1}{n} (T_n + 1 + T_n(T_n + 1)). \end{aligned}$$

Diese Schranke ist unabhängig von t , und es gilt $\frac{1}{n} (T_n + 1 + T_n(T_n + 1)) \xrightarrow{n \rightarrow \infty} 0$ nach Lemma 5.8. Damit folgt

$$d_\infty(Z^n, Z_{K_n(t)} \circ \lambda_n) \xrightarrow{n \rightarrow \infty} 0. \quad (5.13)$$

Andererseits gilt $Z_{\tau_n}(t) \leq Z_{K_n(t)}(t) \leq Z(t)$ für alle $t \in [0, 1]$ und zusammen mit $\tau_n \rightarrow \infty$ aus Lemma 5.8 ergibt sich

$$d_\infty(Z_{K_n(t)} \circ \lambda_n, Z \circ \lambda_n) \xrightarrow{n \rightarrow \infty} 0. \quad (5.14)$$

Insgesamt folgt dann aus (5.13) und (5.14) mittels der Dreiecksungleichung die zweite benötigte Eigenschaft der λ_n und damit die Behauptung des Satzes.

□

Nachdem wir nun das Hauptresultat dieses Kapitels bewiesen haben, werden wir nun mit dem nächsten Satz die direkte Beziehung zwischen obigem Resultat und dem eigentlichen Ziel der Untersuchung, der Frage nach der asymptotischen Verteilung der Anzahl der von FIND benötigten Vergleiche zum Finden des l -kleinsten Elements einer n -elementigen Menge, herstellen. Der Satz liefert nicht nur ein Ergebnis für die in diesem Kapitel bisher betrachtete 2-Mengen Version von FIND sondern auch für die 3-Mengen Version.

5.9 Satz

Sei $(l_n)_{n \in \mathbb{N}}$ eine Folge von ganzen Zahlen, $1 \leq l_n \leq n$, mit der Eigenschaft, dass $\frac{l_n}{n}$ gegen einen Grenzwert $t \in [0, 1]$ konvergiert. Dann gilt für $n \rightarrow \infty$

$$\frac{1}{n} C_{n, l_n} \xrightarrow{d} Z(t), \quad \frac{1}{n} C_{n, l_n}^* \xrightarrow{d} Z(t).$$

Beweis:

Wie bereits bei der Konstruktion der Z^n erwähnt, gilt $\mathcal{L}(\frac{1}{n} C_{n, l_n}) = \mathcal{L}(Z^n(\frac{l_n}{n}))$. Für den ersten Teil der Behauptung reicht es also

$$Z^n\left(\frac{l_n}{n}\right) \rightarrow Z(t) \quad \text{fast sicher}$$

zu zeigen. Dies folgt aber direkt aus dem eben bewiesenen Satz 5.5 in Verbindung mit Satz 1.10 und der Stetigkeitseigenschaft von Z aus Lemma 5.4, bzw. ausführlicher:

Nach Lemma 5.4 existiert eine Nullmenge A , so dass die Funktion $s \mapsto Z(s, \omega)$ für alle $\omega \notin A$ stetig in t ist. Des Weiteren existiert nach Satz 5.5 eine Nullmenge B , so dass außerhalb von B $Z^n(\omega)$ in der Skorokhod-Topologie gegen $Z(\omega)$ konvergiert, d.h. dass eine Folge $\lambda_n(\cdot, \omega) \in \Lambda$ existiert mit $d_\infty(\lambda_n(\cdot, \omega), \text{id}) \rightarrow 0$ und $d_\infty(Z^n(\cdot, \omega), Z(\lambda_n(\cdot, \omega), \omega)) \rightarrow 0$. Dann gilt aber für alle $\omega \in A^c \cap B^c$:

$$\begin{aligned} \left| Z^n\left(\frac{l_n}{n}, \omega\right) - Z\left(t, \omega\right) \right| &\leq \left| Z^n\left(\frac{l_n}{n}, \omega\right) - Z\left(\lambda_n\left(\frac{l_n}{n}, \omega\right), \omega\right) \right| \\ &\quad + \left| Z\left(\lambda_n\left(\frac{l_n}{n}, \omega\right), \omega\right) - Z(t, \omega) \right| \\ &\xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Für die zweite Behauptung des Satzes müssen wir nun nur noch

$$\frac{1}{n} (C_{n, l_n} - C_{n, l_n}^*) \xrightarrow{P} 0 \tag{5.15}$$

zeigen. Die Behauptung des Satzes folgt dann mit dem Satz von Slutsky³.

Aus den Definitionen der 2-Mengen und 3-Mengen Version von FIND ist klar, dass sich die

³siehe Satz 36.12 in [2]

linke Seite von (5.15) nur genau dann von 0 unterscheidet, wenn das gesuchte Element l_n bereits als Pivotelement ausgewählt wurde, bevor die Menge S auf ein einziges Element reduziert wurde.

Bezeichne I_{njl} die Indikatorfunktion des Ereignisses, dass im Rekursionsschritt j das gesuchte Element l zufällig aus einer n -elementigen Menge als Pivotelement ausgewählt wird, d.h.

$$I_{njl} = \begin{cases} 1, & \text{falls } Y_{j, N_j^n(\frac{l}{n})-1}^n = \frac{l}{n} \\ 0, & \text{sonst.} \end{cases}$$

Damit ergibt sich dann

$$\frac{1}{n} (C_{n, l_n} - C_{n, l_n}^*) \stackrel{d}{=} \sum_{j=1}^{\infty} V_{j, N_j^n(\frac{l_n}{n})}^n I_{njl_n}.$$

Sei nun wieder $\mathcal{F}_k$ die σ -Algebra die von allen $U_{j,i}$, $0 \leq j < k$, $1 \leq i \leq 2^k$ erzeugt wird.

Da die jeweiligen Pivotelemente gleichverteilt ausgewählt werden, gilt $E[I_{njl_n} | \mathcal{F}_{j-1}] = \left(n W_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n \right)^{-1}$ auf dem Ereignis $K_n(\frac{l_n}{n}) > j$.

Unter Verwendung von $V_{j, N_j^n(\frac{l_n}{n})}^n \leq V_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n \leq W_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n$ und $V_{j, N_j^n(\frac{l_n}{n})}^n = 0$ für $j \geq K_n(\frac{l_n}{n})$ erhalten wir nun mit Hilfe des Satzes der monotonen Konvergenz

$$\begin{aligned} 0 &\leq \frac{1}{n} E(C_{n, l_n} - C_{n, l_n}^*) \leq E \left(\sum_{j=1}^{\infty} V_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n I_{njl_n} \right) \\ &= E \left(\sum_{j=1}^{\infty} E \left[V_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n I_{njl_n} | \mathcal{F}_{j-1} \right] \right) \\ &= E \left(\sum_{j=1}^{\infty} V_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n E[I_{njl_n} | \mathcal{F}_{j-1}] \right) \\ &= E \left(\sum_{j=1}^{K_n(\frac{l_n}{n})} V_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n E[I_{njl_n} | \mathcal{F}_{j-1}] \right) \\ &= E \left(\sum_{j=1}^{K_n(\frac{l_n}{n})} V_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n \left(n W_{j-1, N_{j-1}^n(\frac{l_n}{n})}^n \right)^{-1} \right) \\ &\leq \frac{1}{n} E T_n. \end{aligned}$$

Aus Lemma 5.8 und der Markov-Ungleichung mit $g(t) = t$ folgt schließlich

$$\begin{aligned} P \left(\frac{1}{n} (C_{n, l_n} - C_{n, l_n}^*) > t \right) &\leq \frac{1}{t} E \left(\frac{1}{n} (C_{n, l_n} - C_{n, l_n}^*) \right) \\ &\leq \frac{1}{t n} E T_n \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

□

5.10 Bemerkung

Wir erinnern zuerst an die in der Einleitung erwähnte 3-Mengen Version von FIND und definieren für diese

$$X_n^*(t) \stackrel{\text{def}}{=} \frac{1}{n} C_{n, [nt]+1}^* \quad \text{für } 0 \leq t < 1 \quad \text{und} \quad X_n^*(1) \stackrel{\text{def}}{=} \frac{1}{n} C_{n,n}^*.$$

Im Gegensatz zu X_n konvergiert X_n^* dann nicht in Verteilung bzgl. der Skorokhod-Topologie.

Beweis:

Bezeichne $I \in \{0, 1, \dots, n\}$ das Element von S , das im ersten Rekursionsschritt als Pivotelement ausgewählt wird. Dann gilt

$$X_n^*(t) = \frac{n-1}{n} \quad \text{auf} \quad \left[\frac{I-1}{n}, \frac{I}{n} \right) \quad (5.16)$$

$$X_n^*(t) \geq \frac{n-1}{n} + \frac{1}{n} \min\{I-1, n-I-1\} \quad \text{auf} \quad \left[\frac{I-1}{n}, \frac{I}{n} \right)^c. \quad (5.17)$$

Weiter sei $A_n \subset D[0, 1]$ die Menge der Funktionen f mit folgender Eigenschaft: Es existiert ein Intervall $[s, s + \frac{1}{n})$, $\frac{1}{4} \leq s \leq \frac{3}{4}$, so dass $f(x) \leq 1$ auf diesem Intervall ist und $f(x) \geq 1, 2$ sonst.

B_n bezeichne das Ereignis, dass $\frac{1}{4} \leq \frac{I-1}{n} \leq \frac{3}{4}$ gilt. Anschaulich ist B_n also gerade das Ereignis, dass im ersten Rekursionsschritt ein Pivotelement aus $[\lceil \frac{n}{4} \rceil + 1, \lfloor \frac{3n}{4} \rfloor + 1]$ ausgewählt wird. Auf B_n gilt dann:

$$\frac{n-1}{n} + \frac{1}{n} \min\{I-1, n-I-1\} \geq \frac{5}{4} - \frac{3}{n},$$

denn $\frac{1}{4} \leq \frac{I-1}{n} \leq \frac{3}{4}$ impliziert $I-1 \geq \frac{n}{4}$ und $n-I-1 \geq \frac{n}{4} - 2$. Daraus wiederum folgt dann

$$\begin{aligned} \frac{n-1}{n} + \frac{1}{n} \min\{I-1, n-I-1\} &\geq \frac{n-1}{n} + \frac{1}{n} \left(\frac{n}{4} - 2 \right) \\ &= \frac{5}{4} - \frac{3}{n}. \end{aligned}$$

Als Nächstes betrachten wir das Ereignis $\{X_n^* \in A_n\}$ für $n \geq 60$. Wir wollen zeigen, dass B_n eine Teilmenge von $\{X_n^* \in A_n\}$ ist. Sei dafür ein $\omega \in B_n$ gegeben. Wir stellen zuerst fest, dass $X_n^*(t, \omega)$ genau dann kleiner oder gleich 1 ist, wenn das gesuchte Element im ersten Rekursionsschritt ausgewählt wird. Im Gegensatz dazu ist $X_n^*(t, \omega)$ genau dann größer oder gleich 1,2, wenn das gesuchte Element nicht im ersten Schritt gefunden wird, denn in diesem Fall gilt für alle $n \geq 60$:

$$X_n^*(t, \omega) \geq \frac{n-1}{n} + \frac{1}{n} \min\{I(\omega)-1, n-I(\omega)-1\} \geq \frac{5}{4} - \frac{3}{n} \geq 1, 2.$$

Damit ist aber klar, dass B_n eine Teilmenge von $\{X_n^* \in A_n\}$ ist. Denn liegt das Pivotele-

ment im ersten Rekursionsschritt zwischen $\left[\left\lceil \frac{n}{4} \right\rceil + 1, \left\lfloor \frac{3n}{4} \right\rfloor + 1\right]$, so existiert ein $s \in \left[\frac{1}{4}, \frac{3}{4}\right]$, so dass $X_n^*(t)$ kleiner oder gleich 1 auf $\left[s, s + \frac{1}{n}\right]$ ist und größer oder gleich 1,2 sonst. Daraus ergibt sich dann unter Beachtung von $\mathcal{L}(I) = U(\{1, \dots, n\})$

$$P(X_n^* \in A_n) \geq P(B_n) = P\left(\left\lceil \frac{n}{4} + 1 \right\rceil \leq I \leq \left\lfloor \frac{3n}{4} + 1 \right\rfloor\right) \geq \frac{1}{2} - \frac{1}{n}$$

Wir nehmen jetzt an, dass X_n^* in Verteilung bzgl. der Skorokhod-Topologie konvergiere. Dann ist aber $(\mathcal{L}(X_n^*))_{n \in \mathbb{N}}$ schwach relativ folgenkompakt, d.h. jede Folge von Elementen aus $(\mathcal{L}(X_n^*))_{n \in \mathbb{N}}$ enthält eine schwach konvergente Teilfolge. Da $D[0, 1]$ versehen mit der Skorokhod-Topologie ein vollständiger metrischer Raum ist⁴, folgt mit dem Satz von Prohorov⁵, dass $(\mathcal{L}(X_n^*))_{n \in \mathbb{N}}$ straff ist, d.h. es existiert zu jedem $\epsilon > 0$ ein kompaktes $K \subset D[0, 1]$ mit $\sup_{n \in \mathbb{N}} P(X_n^* \in K^c) < \epsilon$. Insbesondere existiert also ein kompaktes $K \subset D[0, 1]$ mit $P(X_n^* \in K) \geq \frac{3}{4}$ für alle $n \in \mathbb{N}$. Da für alle $n \geq 60$ $P(X_n^* \in A_n) \geq \frac{29}{60}$ gilt, folgt schon $K \cap A_n \neq \emptyset$ für alle solchen n . Insbesondere finden wir eine Folge $(f_n)_{n \in \mathbb{N}}$, so dass $f_n \in A_n$ gilt und $\{f_n : n \in \mathbb{N}\}$ schwach relativ folgenkompakt ist. Dann folgt mit Satz 1.11

$$\begin{aligned} \sup_{n \in \mathbb{N}} \sup_{t \in [0, 1]} |f_n(t)| &< \infty \\ \lim_{\delta \rightarrow 0} \sup_{n \in \mathbb{N}} w'_{f_n}(\delta) &= 0. \end{aligned} \tag{5.18}$$

Zu jedem $\delta > 0$ existiert ein $n \in \mathbb{N}$ mit $\delta > \frac{1}{n}$. Dann gilt aber für alle $r \in \mathbb{N}$ und für alle möglichen Partitionen $\{t_i\}_{1 \leq i \leq r}$ von $[0, 1]$ mit $0 = t_0 < \dots < t_r = 1$ und $t_j - t_{j-1} > \delta$ für alle $1 \leq j \leq r$, dass in einem der Intervalle $[t_j, t_{j+1})$ sowohl Punkte aus $\left[s, s + \frac{1}{n}\right)$ als auch $\left[s, s + \frac{1}{n}\right)^c$ vorhanden sind. Damit folgt aber für alle $\delta > 0$

$$\sup_{n \in \mathbb{N}} w'_{f_n}(\delta) = \sup_{n \in \mathbb{N}} \inf_{\{t_i\}} \max_{0 < i \leq r} \sup_{s, t \in [t_{i-1}, t_i)} |f_n(s) - f_n(t)| \geq 0, 2.$$

Dies liefert nun den Widerspruch zu (5.18), also konvergiert X_n^* nicht in Verteilung bzgl. der Skorokhod-Topologie.

□

⁴vgl. Satz 1.12

⁵siehe [2], Satz A.10

5.3 Eigenschaften des FIND-Prozesses

Wir werden in diesem Abschnitt einige Eigenschaften des FIND-Prozesses vorstellen. Wir beginnen mit einer Abschätzung der Tail-Wahrscheinlichkeiten der eindimensionalen Marginalverteilungen von Z .

Wir nennen die Zufallsgrößen X stochastisch kleiner als die Zufallsgröße Y im Zeichen $X \stackrel{d}{\leq} Y$ genau dann, wenn $P(X \geq z) \leq P(Y \geq z)$ für alle $z \in \mathbb{R}$ gilt.

5.11 Satz

Sei $(V_i)_{i \in \mathbb{N}}$ eine Folge von unabhängigen $U(\frac{1}{2}, 1)$ -verteilten Zufallsgrößen. Dann gilt für alle $t \in [0, 1]$

$$Z(t) \stackrel{d}{\leq} V$$

mit $V \stackrel{\text{def}}{=} 1 + \sum_{n=1}^{\infty} \prod_{k=1}^n V_k$. Genauer gilt für alle $t \in [0, 1]$ und alle $0 < \eta < 1$

$$P(Z(t) \geq z) = o(\eta^z) \quad \text{für } z \rightarrow \infty.$$

Beweis:

Da $N_{j+1}(t)$ entweder gleich $2N_j(t) - 1$ oder gleich $2N_j(t)$ ist, gilt:

$$\begin{aligned} W_{j+1, N_{j+1}(t)} &\leq \max \left\{ W_{j+1, 2N_j(t)-1}, W_{j+1, 2N_j(t)} \right\} \\ &= W_{j, N_j(t)} \max \left\{ U_{j, N_j(t)}, 1 - U_{j, N_j(t)} \right\}. \end{aligned}$$

Fixiere nun ein $t \in [0, 1]$ und definiere $V_{j+1} \stackrel{\text{def}}{=} \max \left\{ U_{j, N_j(t)}, 1 - U_{j, N_j(t)} \right\}$. Bedingt unter $\mathcal{F}_j$ ist V_{j+1} eine $U(\frac{1}{2}, 1)$ -verteilte Zufallsgröße, insbesondere sind $\mathcal{F}_j$ und V_{j+1} unabhängig. Iterativ erhalten wir dann $W_{j, N_j(t)} \leq V_1 \dots V_j$. Die erste Behauptung des Satzes folgt dann direkt aus der Identität $Z(t) = 1 + \sum_{j=1}^{\infty} W_{j, N_j(t)}$.

Für die zweite Behauptung werden wir eine Abschätzung für die Tail-Wahrscheinlichkeiten von V herleiten. Wähle $\epsilon \in (0, \frac{1}{2})$ und definiere iterativ die Stoppzeiten

$$\begin{aligned} \sigma_0 &= 1 \\ \sigma_1 &\stackrel{\text{def}}{=} \inf \{k \in \mathbb{N} : V_k \leq 1 - \epsilon\} \\ \sigma_{n+1} &\stackrel{\text{def}}{=} \inf \{k > \sigma_n : V_k \leq 1 - \epsilon\} \\ \tau_1 &\stackrel{\text{def}}{=} \sigma_1 \\ \tau_{n+1} &\stackrel{\text{def}}{=} \inf \{k > \sigma_n : V_k \leq 1 - \epsilon\} - \sigma_n. \end{aligned}$$

Da die V_i unabhängig sind, folgt Gleiches auch für die τ_i und es gilt

$$M_\tau(\theta) \stackrel{\text{def}}{=} E \exp(\theta \tau_n)$$

$$\begin{aligned}
&= \sum_{n=1}^{\infty} \exp(\theta n) (2\epsilon)^{n-1} (1 - 2\epsilon) \\
&= \exp(\theta) (1 - 2\epsilon) \sum_{n=0}^{\infty} (\exp(\theta) 2\epsilon)^n \\
&= \frac{\exp(\theta) (1 - 2\epsilon)}{1 - 2\epsilon \exp(\theta)} < \infty \quad \text{für alle } \theta < -\log(2\epsilon).
\end{aligned}$$

Als Nächstes werden wir die Ungleichung

$$V = 1 + \sum_{n=1}^{\infty} \prod_{k=1}^n V_k \leq 1 + \sum_{n=1}^{\infty} (1 - \epsilon)^{n-1} \tau_n$$

beweisen.

Dafür schreiben wir V zuerst als

$$V = 1 + \sum_{n=1}^{\infty} \prod_{k=1}^n V_k = 1 + \sum_{n=1}^{\infty} \sum_{l=\sigma_{n-1}}^{\sigma_n-1} \prod_{k=1}^l V_k.$$

Wir betrachten nun für jedes n den Summanden $\sum_{l=\sigma_{n-1}}^{\sigma_n-1} \prod_{k=1}^l V_k$ und schätzen diesen wie folgt ab: Für jedes l sind in diesem Produkt genau $(n-1)$ der V_k , $k = 1, \dots, l$, kleiner oder gleich $(1 - \epsilon)$. Damit ergibt sich unter Beachtung von $\tau_n = (\sigma_n - 1) - \sigma_{n-1}$

$$\begin{aligned}
V &= 1 + \sum_{n=1}^{\infty} \sum_{l=\sigma_{n-1}}^{\sigma_n-1} \prod_{k=1}^l V_k \\
&\leq 1 + \sum_{n=1}^{\infty} \sum_{l=\sigma_{n-1}}^{\sigma_n-1} (1 - \epsilon)^{n-1} \cdot 1^{l-n+1} \\
&= 1 + \sum_{n=1}^{\infty} (1 - \epsilon)^{n-1} \tau_n.
\end{aligned}$$

Wegen der Unabhängigkeit der τ_i folgt dann mit Hilfe des Satzes der monotonen Konvergenz

$$\begin{aligned}
E(\exp(\theta V)) &\leq E \exp \left(\theta + \theta \sum_{n=1}^{\infty} (1 - \epsilon)^{n-1} \tau_n \right) \\
&= \exp(\theta) \prod_{n=1}^{\infty} M_{\tau} (\theta (1 - \epsilon)^{n-1}) \\
&= \exp \left(\theta + \sum_{n=1}^{\infty} \log \left(\frac{(1 - 2\epsilon) \exp(\theta (1 - \epsilon)^{n-1})}{1 - 2\epsilon \exp(\theta (1 - \epsilon)^{n-1})} \right) \right).
\end{aligned}$$

Da diese Reihe konvergiert⁶, gilt $E \exp(\theta V) < \infty$ für alle $\theta > 0$, wenn man $\epsilon > 0$ klein

⁶siehe dafür die Abschätzung im Anhang

genug wählt. Mit Hilfe der Markov-Ungleichung mit $g(z) = \exp(\theta z)$ folgt

$$P(Z(t) \geq z) \leq P(V \geq z) \leq \frac{E \exp(\theta V)}{\exp(\theta z)} \xrightarrow{z \rightarrow \infty} 0.$$

Damit ergibt sich die Behauptung durch

$$\begin{aligned} P(Z(t) \geq z) &= o(\exp(-\theta z)) \quad \text{für alle } \theta > 0 \\ &= o(\exp(z \log \eta)) \quad \text{für alle } 0 < \eta < 1 \\ &= o(\eta^z) \quad \text{für alle } 0 < \eta < 1. \end{aligned}$$

□

Als direkte Konsequenz des Satzes halten wir fest, dass alle Momente von $Z(t)$ endlich sind. Um die Momente und weitere Eigenschaften des FIND-Prozesses zu berechnen, geben wir folgende stochastische Fixpunktgleichung zur Charakterisierung des Grenzprozesses an.

5.12 Satz

Seien $Z^{(1)}$ und $Z^{(2)}$ unabhängige $D[0, 1]$ -wertige Zufallsvariablen mit $Z^{(i)} \stackrel{d}{=} Z$, $i = 1, 2$. Sei U eine auf $(0, 1)$ gleichverteilte von $Z^{(1)}$ und $Z^{(2)}$ unabhängige Zufallsgröße. Dann gilt folgende stochastische Fixpunktgleichung:

$$Z(\cdot) \stackrel{d}{=} 1 + \mathbf{1}_{[0, U)}(\cdot) U Z^{(1)}\left(\frac{\cdot}{U}\right) + \mathbf{1}_{[U, 1]}(\cdot) (1 - U) Z^{(2)}\left(\frac{\cdot - U}{1 - U}\right). \quad (5.19)$$

Beweis:

Sei $\{U_{k,i} : k \in \mathbb{N}_0, 1 \leq i \leq 2^k\}$ die Familie der $U(0, 1)$ -verteilten Zufallsgrößen, die für die Konstruktion von Z benutzt wird. Definiere $\bar{U}$, $\{\bar{U}_{k,i}^{(1)} : k \in \mathbb{N}_0, 1 \leq i \leq 2^k\}$ und $\{\bar{U}_{k,i}^{(2)} : k \in \mathbb{N}_0, 1 \leq i \leq 2^k\}$ durch $\bar{U} \stackrel{\text{def}}{=} U_{0,1}$, $\bar{U}_{k,i}^{(1)} \stackrel{\text{def}}{=} U_{k+1,i}$, $\bar{U}_{k,i}^{(2)} \stackrel{\text{def}}{=} U_{k+1,2^k+i}$. Die $\{\bar{U}_{k,i}^{(1)}\}$ bilden also die ursprüngliche Unterteilung des Intervalls $[0, U_{0,1}]$ und $\{\bar{U}_{k,i}^{(2)}\}$ die des Intervalls $[U_{0,1}, 1]$.

Sei nun $\bar{Z}^{(j)}$ der Z -Prozess, der mit $\{\bar{U}_{k,i}^{(j)} : k \in \mathbb{N}_0, 1 \leq i \leq 2^k\}$, $j = 1, 2$, konstruiert wird. Offensichtlich ist die gemeinsame Verteilung von $\bar{U}$, $\bar{Z}^{(1)}$ und $\bar{Z}^{(2)}$ die Gleiche wie die von U , $Z^{(1)}$ und $Z^{(2)}$. Per Konstruktion gilt aber obige stochastische Fixpunktgleichung, d.h.

$$Z(t) = 1 + \mathbf{1}_{[0, \bar{U})}(t) \bar{U} \bar{Z}^{(1)}\left(\frac{t}{\bar{U}}\right) + \mathbf{1}_{[\bar{U}, 1]}(t) (1 - \bar{U}) \bar{Z}^{(2)}\left(\frac{t - \bar{U}}{1 - \bar{U}}\right) \quad \text{für alle } t \in [0, 1].$$

□

Wir werden nun mit Hilfe eines Kontraktionsarguments, wie in Kapitel 3, eine explizite Formel für den Erwartungswert von $Z(t)$ angeben.

5.13 Satz

Mit der Konvention $0 \log 0 = 0$ gilt für alle $t \in [0, 1]$:

$$EZ(t) = 2 - 2t \log(t) - 2(1-t) \log(1-t).$$

Beweis:

Sei $\mathcal{B}_m$ wie in Kapitel 1 der Raum der beschränkten, messbaren Funktionen $f : [0, 1] \rightarrow \mathbb{R}$. Dann ist $(\mathcal{B}_m, d_\infty)$ ein vollständiger metrischer Raum⁷. Definiere den Operator $T : \mathcal{B}_m \rightarrow \mathcal{B}_m$ durch

$$T(f(t)) \stackrel{\text{def}}{=} 1 + \int_0^t (1-u) f\left(\frac{t-u}{1-u}\right) du + \int_t^1 u f\left(\frac{t}{u}\right) du.$$

Dann gilt für alle $f, g \in \mathcal{B}_m$

$$\begin{aligned} d_\infty(T(f), T(g)) &= \sup_{t \in [0,1]} \left| \int_0^t (1-u) \left(f\left(\frac{t-u}{1-u}\right) - g\left(\frac{t-u}{1-u}\right) \right) du \right. \\ &\quad \left. + \int_t^1 u \left(f\left(\frac{t}{u}\right) - g\left(\frac{t}{u}\right) \right) du \right| \\ &\leq d_\infty(f, g) \sup_{t \in [0,1]} \left| \int_0^t (1-u) du + \int_t^1 u du \right| \\ &= d_\infty(f, g) \sup_{t \in [0,1]} \left| (-1) \left(t - \frac{1}{2} \right)^2 + \frac{3}{4} \right| \\ &= \frac{3}{4} d_\infty(f, g). \end{aligned}$$

Also ist T eine Kontraktion des vollständigen metrischen Raums $\mathcal{B}_m$ und hat somit einen eindeutigen Fixpunkt. Wir werden nun zeigen, dass sowohl $EZ(t)$ als auch $g(t) \stackrel{\text{def}}{=} 2 - 2t \log(t) - 2(1-t) \log(1-t)$ Fixpunkte von T sind, und somit übereinstimmen müssen. Die Existenz des Erwartungswertes von $Z(t)$ folgt, wie schon erwähnt, aus der oberen Schranke in Satz 5.11, und damit ist die Funktion $f : [0, 1] \rightarrow \mathbb{R}$, definiert durch $f(t) \stackrel{\text{def}}{=} EZ(t)$, beschränkt.

Aus der stochastischen Fixpunktgleichung (5.19) folgt dann durch Bildung des Erwartungswertes, dass f ein Fixpunkt von T ist:

$$\begin{aligned} f(t) = EZ(t) &= 1 + E \left(\mathbf{1}_{[0,U)}(t) U Z^{(1)} \left(\frac{t}{U} \right) \right) + E \left(\mathbf{1}_{(U,1]}(t) (1-U) Z^{(2)} \left(\frac{t-U}{1-U} \right) \right) \\ &= 1 + \int_t^1 u EZ^{(1)} \left(\frac{t}{u} \right) du + \int_0^t (1-u) EZ^{(2)} \left(\frac{t-u}{1-u} \right) du. \end{aligned}$$

⁷siehe Satz 1.13

Also gilt wie behauptet $T(EZ(t)) = EZ(t)$.

Als Letztes bleibt uns nun noch nachzurechnen, dass auch $g(t)$ ein Fixpunkt von T ist.

Dabei werden wir folgende Identitäten von Integralen benötigen:

$$\begin{aligned} \int_0^t (1-u) \log(1-u) du &= \frac{1}{2} \left(-(1-t^2) \log(1-t) + \frac{t^2}{2} - t \right) \\ \int_0^t (t-u) \log(t-u) du &= \frac{1}{2} \left(t^2 \log(t) - \frac{t^2}{2} \right) \\ \int_t^1 u \log(u) du &= \frac{1}{2} \left(-t^2 \log(t) + \frac{t^2}{2} - \frac{1}{2} \right) \\ \int_t^1 (u-t) \log(u-t) du &= \frac{1}{2} \left((1-t)^2 \log(1-t) - \frac{(1-t)^2}{2} \right). \end{aligned}$$

Mit diesen Identitäten ergibt sich

$$\begin{aligned} T(g(t)) &= 1 + \int_0^t (1-u) g\left(\frac{t-u}{1-u}\right) du + \int_t^1 u g\left(\frac{t}{u}\right) du \\ &= 1 + \int_0^t (1-u) \left(2 - 2 \frac{t-u}{1-u} \log\left(\frac{t-u}{1-u}\right) - 2 \left(1 - \frac{t-u}{1-u}\right) \log\left(1 - \frac{t-u}{1-u}\right) \right) du \\ &\quad + \int_t^1 u \left(2 - 2 \frac{t}{u} \log\left(\frac{t}{u}\right) - 2 \left(1 - \frac{t}{u}\right) \log\left(1 - \frac{t}{u}\right) \right) du \\ &= 2 + 2t - 2t^2 - 2t(1-t)(\log(1-t) + \log(t)) \\ &\quad + 2 \int_0^t (1-u) \log(1-u) du - 2 \int_0^t (t-u) \log(t-u) du \\ &\quad + 2 \int_t^1 u \log(u) du - 2 \int_t^1 (u-t) \log(u-t) du \\ &= 2 - 2t \log(t) - 2(1-t) \log(1-t) = g(t). \end{aligned}$$

Wie oben bereits beschrieben folgt dann aus der Eindeutigkeit des Fixpunktes von T die behauptete Formel für den Erwartungswert von $Z(t)$.

□

Dieses Vorgehen zur Berechnung des Erwartungswertes kann verallgemeinert werden, um auch die n -ten Momente von $Z(t)$ für $n \in \mathbb{N}$ zu berechnen. Die Idee hierfür hat Paulsen in [22] vorgestellt. Sie beruht darauf, zuerst eine stochastische Fixpunktgleichung für das n -te Moment von $Z(t)$ zu bestimmen. Darauf aufbauend kann dann eine Differentialgleichung, die von einer Ableitung der n -ten Momentfunktion erfüllt wird, aufgestellt werden.

Wir werden zuerst die n -te Momentfunktion von $Z(t)$ als eindeutigen Fixpunkt einer Kontraktion auf $\mathcal{B}_m$ charakterisieren. Bezeichne also mit

$$m_n(t) \stackrel{\text{def}}{=} EZ^n(t) \quad \text{für alle } t \in [0, 1]$$

die n -te Momentfunktion des FIND-Prozesses $Z(t)$. Wie $EZ(t)$ ist auch $m_n(t)$ beschränkt. Definiere nun für jedes $n \in \mathbb{N}$ den Operator $K_n : \mathcal{B}_m \rightarrow \mathcal{B}_m$ durch

$$K_n(f(t)) \stackrel{\text{def}}{=} \int_t^1 u^n f\left(\frac{t}{u}\right) du + \int_0^t (1-u)^n f\left(\frac{t-u}{1-u}\right) du + b_n(t),$$

wobei

$$b_n(t) \stackrel{\text{def}}{=} \sum_{i=1}^n (-1)^{i-1} \binom{n}{i} m_{n-i}(t).$$

5.14 Satz

K_n ist eine Kontraktion bzgl. d_∞ und m_n ist der eindeutige Fixpunkt von K_n für jedes $n \in \mathbb{N}$. Des Weiteren konvergiert $(K_n^k(f))_{k \in \mathbb{N}}$ exponentiell schnell gegen m_n für jedes $f \in \mathcal{B}_m$.

Beweis:

Für $f, g \in \mathcal{B}_m$ gilt

$$\begin{aligned} d_\infty(K_n(f), K_n(g)) &= \sup_{t \in [0,1]} \left| \int_t^1 u^n \left(f\left(\frac{t}{u}\right) - g\left(\frac{t}{u}\right) \right) du \right. \\ &\quad \left. + \int_0^t (1-u)^n \left(f\left(\frac{t-u}{1-u}\right) - g\left(\frac{t-u}{1-u}\right) \right) du \right| \\ &\leq d_\infty(f, g) \sup_{t \in [0,1]} \int_t^1 u^n du + \int_0^t (1-u)^n du \\ &\leq \frac{1}{n+1} \left(2 - \left(\frac{1}{2}\right)^n \right) d_\infty(f, g). \end{aligned}$$

Also ist K_n eine Kontraktion in $(\mathcal{B}_m, d_\infty)$ und konvergiert somit gegen den eindeutigen Fixpunkt. Insbesondere konvergiert $(K_n^k(f(t)))_{k \in \mathbb{N}}$ für jeden Startpunkt f exponentiell schnell gegen diesen Fixpunkt.

Es bleibt also zu zeigen, dass $m_n(t)$ dieser Fixpunkt ist, d.h.

$$K_n(m_n) = m_n \quad \text{für alle } n \in \mathbb{N}. \quad (5.20)$$

Aus der stochastischen Fixpunktgleichung (5.19) folgt

$$E(Z(t) - 1)^n = \int_t^1 u^n m_n\left(\frac{t}{u}\right) du + \int_0^t (1-u)^n m_n\left(\frac{t-u}{1-u}\right) du.$$

Die Behauptung folgt dann aus dem binomischen Lehrsatz durch

$$E(Z(t) - 1)^n = m_n(t) + \sum_{i=1}^n \binom{n}{i} (-1)^i m_{n-i}(t).$$

□

Nachdem wir die stochastische Fixpunktgleichung für die Momentfunktion von $Z(t)$ gefunden haben, werden wir nun eine gewöhnliche lineare Differentialgleichung herleiten, mit deren Hilfe die n -te Momentfunktion von $Z(t)$ bestimmt werden kann. Nach Substitution von $s = \frac{t}{u}$ bzw. $s = \frac{1-t}{1-u}$ schreibt sich die Fixpunktgleichung (5.20) als

$$m_n(t) = t^{n+1} \int_t^1 \frac{m_n(s)}{s^{n+2}} ds + (1-t)^{n+1} \int_{1-t}^1 \frac{m_n(1-s)}{s^{n+2}} ds + b_n(t). \quad (5.21)$$

An dieser Darstellung erkennt man, dass m_n eine unendlich oft differenzierbare Funktion ist. Um die Ableitungen zu berechnen, definiere den Operator Ψ_n auf dem Raum der stetigen Funktionen auf $[0, 1]$ durch

$$\Psi_n(f(t)) \stackrel{\text{def}}{=} t^n \int_t^1 \frac{f(s)}{s^{n+1}} ds \quad \text{für alle } f \in C([0, 1]).$$

Mit Hilfe der Leibniz-Formel für das Ableiten von Produkten, d.h. $(f \cdot g)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}$, und unter Berücksichtigung von

$$D^k t^n = \frac{1}{t} \frac{n-k+1}{n+1} D^k t^{n+1}, \quad 0 \leq k \leq n$$

ergibt sich

$$D^{n+1} \Psi_n(f(t)) = -\frac{1}{t} D^n(f(t)),$$

wobei D hier den Ableitungsoperator bezeichnet. Eine analoge Behandlung des zweiten Integrals in (5.21) führt zu

$$D^{n+2} m_n(t) = \left(\frac{1}{1-t} - \frac{1}{t} \right) D^{n+1} m_n(t) + D^{n+2} b_n(t).$$

Dies ist eine gewöhnliche lineare Differentialgleichung, mit deren Hilfe die $(n+1)$ -te Ableitung der n -ten Momentfunktion von $Z(t)$ berechnet werden kann und daraus dann rekursiv die n -te Momentfunktion von $Z(t)$.

Zu guter Letzt werden wir nun in diesem Abschnitt die Worst Case Performance des Algorithmus untersuchen. Dazu betrachten wir das Supremum des FIND-Prozesses. Beachte, dass das Supremum ein stetiges Funktional bzgl. der Skorokhod-Topologie ist, so dass das Resultat über die Verteilungskonvergenz weiterhin gilt.

5.15 Satz

Sei $M \stackrel{\text{def}}{=} \sup_{t \in [0,1]} Z(t)$. Dann hat M ein endliches zweites Moment und die Verteilung von M ist die eindeutige Verteilung, welche die stochastische Fixpunktgleichung

$$M \stackrel{d}{=} 1 + \max \left\{ UM, (1-U)\tilde{M} \right\}$$

erfüllt, wobei U wieder eine $U(0,1)$ -verteilte Zufallsgröße ist, $\mathcal{L}(\tilde{M}) = \mathcal{L}(M)$ gilt und U , M , und $\tilde{M}$ unabhängig sind.

Beweis:

Zum Beweis werden wir die in Kapitel 3 erklärte Kontraktionsmethode verwenden. Sei wie in Kapitel 1 $\mathbb{M}_2$ die Menge aller Wahrscheinlichkeitsmaße μ auf $[0, \infty)$ mit $\int x^2 \mu(dx) < \infty$ und d_2 die Mallows-Metrik. Definiere den Operator $T : \mathbb{M}_2 \rightarrow \mathbb{M}_2$ durch

$$T(\mu) \stackrel{\text{def}}{=} \mathcal{L}\left(1 + \max\{UX, (1-U)\bar{X}\}\right),$$

wobei $\mathcal{L}(U) = U(0,1)$ sowie $\mathcal{L}(X) = \mathcal{L}(\bar{X}) = \mu$ gelte und U , X , $\bar{X}$ unabhängig seien. Das Bild von T ist offensichtlich eine Teilmenge von $\mathbb{M}_2$, da $0 \leq 1 + \max\{UX, (1-U)\bar{X}\} \leq 1 + X + \bar{X}$ gilt.

Sei $M_k \stackrel{\text{def}}{=} \sup_{t \in [0,1]} Z^k(t)$. Mit der Aufteilung der $U_{k,i}$ wie im Beweis von Satz 5.12 ergibt sich auch für die Z^k

$$Z^{k+1}(\cdot) \stackrel{d}{=} 1 + \mathbf{1}_{[0,U)} U Z^k\left(\frac{\cdot}{U}\right) + \mathbf{1}_{[U,1]} (1-U) \bar{Z}^k\left(\frac{\cdot - U}{1-U}\right),$$

wobei wiederum $\mathcal{L}(U) = U(0,1)$, $\mathcal{L}(\bar{Z}^k) = \mathcal{L}(Z^k) = \mu$ gelte und U , Z^k , $\bar{Z}^k$ unabhängig seien. Somit folgt $M^{k+1} \stackrel{d}{=} 1 + \max\{UM_k, (1-U)\bar{M}_k\}$, insbesondere gilt $\mathcal{L}(M_{k+1}) = T(\mathcal{L}(M_k))$.

Da $M_0 \equiv 1$ ist und das Bild von T wieder eine Teilmenge von $\mathbb{M}_2$ bildet, ist $(\mathcal{L}(M_k))_{k \in \mathbb{N}_0}$ eine Folge in $\mathbb{M}_2$.

Wir wollen nun zeigen, dass T ein Kontraktion des vollständigen metrischen Raums $(\mathbb{M}_2, d_2)$ ist⁸. Denn dann hat T einen eindeutigen Fixpunkt und dieser ist gerade gegeben durch den Grenzwert der Folge $x_0 = 1$, $x_{n+1} = T(x_n)$, d.h. durch $\mathcal{L}(\lim_{k \rightarrow \infty} M_k) = \mathcal{L}(M)$. M ist dann gerade der eindeutige Fixpunkt der Gleichung (5.15).

Zum Nachweis der Kontraktionseigenschaft weisen wir zunächst darauf hin, dass man zu $\mu, \nu \in \mathbb{M}_2$ stets Zufallsvariablen X und Y finden kann, so dass $d_2(\mu, \nu) = \|X - Y\|_2$ gilt⁹. Seien nun μ, ν, X und Y wie gerade beschrieben. Weiterhin seien $(\bar{X}, \bar{Y})$, U so gewählt, dass U , (X, Y) und $(\bar{X}, \bar{Y})$ unabhängig sind und $\mathcal{L}(U) = U(0,1)$ sowie $(\bar{X}, \bar{Y}) \stackrel{d}{=} (X, Y)$ gelte. Mit der allgemein geltenden Ungleichung $(\max\{a, b\} - \max\{c, d\})^2 \leq (a - c)^2 +$

⁸siehe Satz 1.16

⁹siehe Satz 1.14

$(b - d)^2$ ergibt sich

$$\begin{aligned}
d_2(T(\mu), T(\nu))^2 &\leq \|(1 + \max\{UX, (1 - U)\bar{X}\}) - (1 + \max\{UY, (1 - U)\bar{Y}\})\|_2^2 \\
&\leq E(U^2(X - Y)^2 + (1 - U)^2(\bar{X} - \bar{Y})^2) \\
&= EU^2 E(X - Y)^2 + E(1 - U)^2 E(\bar{X} - \bar{Y})^2 \\
&= \frac{2}{3}d_2(\mu, \nu)^2.
\end{aligned}$$

Somit ist T eine Kontraktion und die Behauptung des Satzes folgt wie oben beschrieben. $\square$

Unsers Wissens nach ist keine explizite Lösung dieser stochastischen Fixpunktgleichung bekannt. Es gibt allerdings Arbeiten, die sich gerade mit der Worst Case Performance von **FIND** beschäftigen. Eine dieser Arbeiten, in der einige Eigenschaften der Verteilung von M vorgestellt werden, ist [10]. Wir werden hier im Folgenden einige dieser Ergebnisse ohne Beweis angeben.

Devroye zeigt, dass die Momente $\nu_r \stackrel{\text{def}}{=} E(M - 1)^r$ von $M - 1$ die Carleman-Bedingung

$$\sum_{r=2}^{\infty} \frac{1}{\nu_r^{1/2r}} = \infty$$

für positive Zufallsgrößen erfüllen und M somit eindeutig durch seine Momente festgelegt ist. Des Weiteren hat M eine Lipschitz-stetige Dichte mit Lipschitz-Konstante kleiner gleich 2. Für den Erwartungswert von M ergibt sich

$$3,3862 < 2 + 2\log(2) \leq EM \leq 1 + \frac{5}{\sqrt{2\pi}} + \frac{12e}{5} < 9,5186.$$

Auch eine Abschätzung der Tail-Wahrscheinlichkeiten von M liefert Devroye. Und zwar existieren für alle $A > 0$ Konstanten s und C , so dass

$$\sup_{n \geq 1} P\left(\max_{1 \leq l \leq n} C_{n,l} \geq (z + 1)n\right) \leq P(M \geq z + 1) \leq C \exp(-A(z - s))$$

für alle $z \geq s$ gilt. Die Tails von $\frac{1}{n}T_n$ und S fallen also bei wachsendem Z zumindest exponentiell schnell.

Wir haben zuletzt Kontraktionsargumente benutzt um sowohl die Momente als auch das Supremum von Z zu charakterisieren. Es stellt sich hier natürlich die Frage, ob nicht auch Z selbst als Fixpunkt einer Kontraktion angesehen werden kann. Dies ist tatsächlich möglich. Wir werden einen solchen Ansatz hier nur kurz vorstellen. Für eine ausführlichere Diskussion der Verwendung von Kontraktionsargumenten zur Analyse von **FIND** siehe [16], [27] oder auch [18], wobei die letztgenannte Arbeit auch Ergebnisse für andere Versionen von **FIND**, wie zum Beispiel der Median-von- m -Version, beinhaltet.

Für $u \in (0, 1)$ und $f, g \in D[0, 1]$ definiere $\Psi(u, f, g) \in D[0, 1]$ durch

$$\Psi(u, f, g)(t) \stackrel{\text{def}}{=} 1 + \mathbf{1}_{[0, u)}(t) f\left(\frac{t}{u}\right) + \mathbf{1}_{[u, 1]}(t) g\left(\frac{t-u}{1-u}\right).$$

Sei $\mathcal{M}$ die Menge aller Wahrscheinlichkeitsmaße μ auf $D[0, 1]$ mit $\int \|f\|_\infty^2 \mu(df) < \infty$. Weiter sei $T(\mu)$ die Verteilung von $\Psi(U, X, \bar{X})$, wobei $U, X, \bar{X}$ unabhängig sind und $\mathcal{L}(U) = U(0, 1)$ sowie $\mathcal{L}(X) = \mathcal{L}(\bar{X}) = \mu$ gelte. Man zeigt dann mit einem ähnlichen Argument wie im Beweis von Satz 5.15, dass T eine Abbildung von $\mathcal{M}$ in sich selbst ist. Satz 5.12 besagt, dass der FIND-Prozess Z ein Fixpunkt von T ist. Mit denselben Argumenten wie in dessen Beweis zeigt man, dass $\mathcal{L}(Z^{k+1}) = T(\mathcal{L}(Z^k))$ gilt. Nun muss noch gezeigt werden, dass T eine Kontraktion bzgl. einer geeigneten Metrik auf $\mathcal{M}$ ist. Hierzu definieren wir für $\mu, \nu \in \mathcal{M}$ eine Metrik durch

$$d(\mu, \nu) \stackrel{\text{def}}{=} \inf \left\{ \|X - Y\|_\infty : \mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu \right\}.$$

Für den Beweis fixiere nun ein $\epsilon > 0$ und seien X, Y Prozesse mit $\mathcal{L}(X) = \mu, \mathcal{L}(Y) = \nu$ und $E \|X - Y\|_\infty^2 \leq d(\mu, \nu)^2 + \epsilon$. Seien weiter $U, (\bar{X}, \bar{Y})$ so, dass $U, (X, Y), (\bar{X}, \bar{Y})$ unabhängig sind und $\mathcal{L}(U) = U(0, 1)$ sowie $(\bar{X}, \bar{Y}) \stackrel{d}{=} (X, Y)$ gelte. Dann ergibt sich

$$\begin{aligned} d(T(\mu), T(\nu))^2 &\leq E \left\| \Psi(U, X, Y) - \Psi(U, \bar{X}, \bar{Y}) \right\|_\infty^2 \\ &= E \left\| \mathbf{1}_{[0, u)} U \left(X \left(\frac{\cdot}{U} \right) - Y \left(\frac{\cdot}{U} \right) \right) \right. \\ &\quad \left. + \mathbf{1}_{[u, 1]} (1 - U) \left(\bar{X} \left(\frac{\cdot - U}{1 - U} \right) - \bar{Y} \left(\frac{\cdot - U}{1 - U} \right) \right) \right\|_\infty^2 \\ &\leq E \left(U^2 \|X - Y\|_\infty^2 + (1 - U)^2 \|\bar{X} - \bar{Y}\|_\infty^2 \right) \\ &\leq \frac{2}{3} (d(\mu, \nu)^2 + \epsilon). \end{aligned}$$

Für $\epsilon \downarrow 0$ folgt dann, dass T eine Kontraktion ist, und somit Z als eindeutigen Fixpunkt besitzt.

Im nächsten Kapitel werden wir eine dritte Herangehensweise zur Analyse des Algorithmus FIND vorstellen, bei der begleitende Markov-Ketten konstruiert werden und mit deren Hilfe dann sowohl dieselben Konvergenzresultate als auch weitere Abschätzungen für die Tail-Wahrscheinlichkeiten hergeleitet werden.

Kapitel 6

Ein Markov-Ketten Ansatz für FIND

Wir werden in diesem Kapitel einen weiteren Ansatz zur Analyse des Algorithmus **FIND** vorstellen. Dieser geht zurück auf den 1998 erschienenen Artikel [12] von Grübel. Die grundlegende Idee hierbei ist es, geeignete begleitende Markov-Ketten zu konstruieren und mit deren Hilfe Aussagen über die von **FIND** benötigte Anzahl von Vergleichen herzuleiten. Wir beziehen uns dabei in diesem Kapitel auf die 3-Mengen Version von **FIND**. Mit diesem Ansatz kann man sowohl die bereits im vorherigen Abschnitt bewiesene Konvergenzaussage gegen den **FIND**-Prozess beweisen¹ als auch weitere Schranken zur Abschätzung der Tail-Wahrscheinlichkeiten herleiten. Diese ermöglichen auch nicht asymptotische Abschätzungen und somit einen Vergleich der Ergebnisse aus vorherigem Kapitel mit denen von Devroye aus [9].²

Wir werden im ersten Abschnitt die begleitenden Markov-Ketten konstruieren, bevor wir dann im folgenden Abschnitt die Ergebnisse mit zugehörigen Beweisen präsentieren werden. Zu guter Letzt werden dann numerische Folgerungen dieser Ergebnisse vorgestellt.

6.1 Konstruktion der begleitenden Markov-Ketten

Das Ziel dieser Untersuchung ist wiederum die Bestimmung der Anzahl der benötigten Vergleiche zum Finden des l -kleinsten Elements einer n -elementigen Menge, welche wir wieder mit $C_{n,l}^*$ bezeichnen werden. Wir interessieren uns also für die Paare (n, l) . Betrachtet man nun für $m \in \mathbb{N}_0$ die Zustände (n_m, l_m) nach dem m -ten Rekursionsschritt als Realisierung eines Prozesses $(Z_m)_{m \in \mathbb{N}_0}$, so bildet dieser eine Markov-Kette, da der Zustand nach dem $(m + 1)$ -ten Rekursionsschritt nur vom Zustand nach dem m -ten Rekursionsschritt abhängt und nicht von der gesamten Vergangenheit dieses Prozesses. Falls (n, l) der aktuelle Zustand dieser Markov-Kette ist, dann wird der nächste Zustand gleichverteilt

¹vgl. Satz 5.5

² $P(C_{n,k}^* > zn) = o(\rho^z)$ für alle $\rho > \frac{3}{4}$ bei $z \rightarrow \infty$

aus der Menge

$$\{(1, 1)\} \cup \{(n - k, l - k) : 1 \leq k \leq l - 1\} \cup \{(n - k, l) : 1 \leq k \leq n - l\}$$

ausgewählt. Ist $n, i_1, i_2, \dots, 1, 1, 1 \dots$ die Folge der ersten Komponente der Zustände dieser Markov-Kette, so ergibt sich $C_{n,l}^*$ als Summe der Werte $n - 1, i_1 - 1, i_2 - 1, \dots, 0, 0 \dots$.

Die Idee dieses Vorgehens ist es nach geeigneter Normalisierung die Konvergenz der Folge dieser Markov-Ketten für $n \rightarrow \infty$ gegen eine Markov-Kette mit stetigem Zustandsraum zu zeigen. Für den Nachweis, dass daraus die Konvergenz von $\frac{1}{n} C_{n,l_n}^*$ gegen den FIND-Prozess folgt, wobei l_n eine Folge sei, die dividiert durch n gegen ein $t \in [0, 1]$ konvergiert, werden wir eine geeignete fast sichere Konstruktion einführen, die uns eine pfadweise Analyse ermöglicht.

Nun zur Konstruktion der angesprochenen Markov-Ketten. Wir werden zuerst die Markov-Kette mit stetigen Zustandsraum angeben. Sei hierfür $I \stackrel{\text{def}}{=} \{(x, y) \in \mathbb{R}_+^2 : 0 \leq y \leq x\}$ der stetige Zustandsraum, und für jedes $(x, y) \in I$ sei $P((x, y), \cdot)$ die Verteilung von ξ , wobei ξ definiert ist durch

$$\xi \stackrel{\text{def}}{=} \begin{cases} (x - Ux, y - Ux), & \text{falls } Ux \leq y \\ (Ux, y), & \text{falls } Ux > y \end{cases}$$

mit einer auf $(0, 1)$ gleichverteilten Zufallsgröße U , also $\mathcal{L}(U) = U(0, 1)$. Dies definiert einen Übergangskern P auf I . Sei weiter $Z = (Z_m)_{m \in \mathbb{N}}$ mit $Z_m = (X_m, Y_m)$ die Markov-Kette mit Zustandsraum I und Übergangskern P . Der Übergangsmechanismus von Z ist in Abbildung 6.1 verdeutlicht.

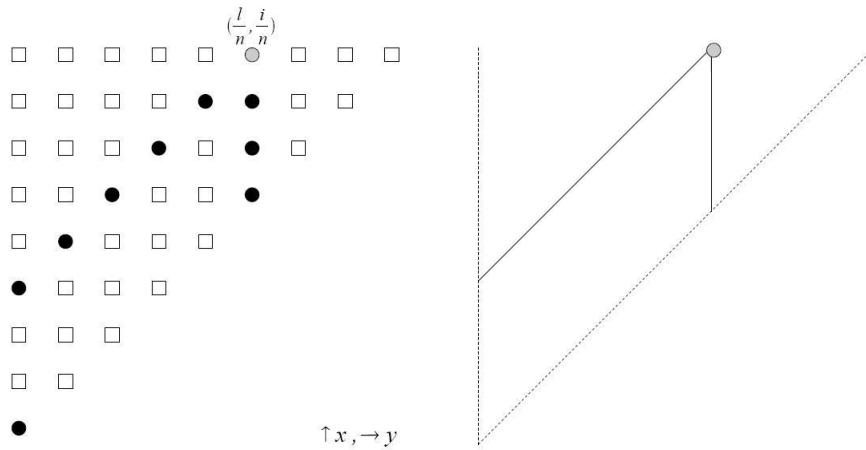


Abbildung 6.1: Übergänge von Z^n und Z . $\circ$: aktueller Zustand $\bullet, -$: mögliche Zustände

Für die Folge der Markov-Ketten definiere zu jedem $n \in \mathbb{N}$ die Menge

$$I_n \stackrel{\text{def}}{=} \left\{ \left(\frac{i}{n}, \frac{l}{n} \right) : i, l \in \mathbb{N}, l \leq i \right\}.$$

Für jedes $(\frac{i}{n}, \frac{l}{n}) \in I_n$ sei $P_n((\frac{i}{n}, \frac{l}{n}), \cdot)$ die Laplace-Verteilung auf

$$\left\{ \left(\frac{1}{n}, \frac{1}{n} \right) \right\} \cup \left\{ \left(\frac{i-k}{n}, \frac{l-k}{n} \right) : 1 \leq k \leq l-1 \right\} \cup \left\{ \left(\frac{i-k}{n}, \frac{l}{n} \right) : 1 \leq k \leq i-l \right\}.$$

Diese definiert einen Übergangskern auf I_n . Sei $Z^{(n)} = (Z_m^{(n)})_{m \in \mathbb{N}_0}$ mit $Z_m^{(n)} = (X_m^{(n)}, Y_m^{(n)})$ die Markov-Kette mit Zustandsraum I_n und Übergangskern P_n .

Um einen Zusammenhang zwischen $C_{n,l}^*$ und den Markov-Ketten herzustellen, schreiben wir $C_{n,l}^*$ wie oben bereits beschrieben als Funktion der $Z^{(n)}$:

$$\mathcal{L} \left(\frac{1}{n} C_{n,l}^* \right) = \mathcal{L} \left[\sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right) \middle| Z_0^{(n)} = \left(1, \frac{l}{n} \right) \right]. \quad (6.1)$$

6.2 Schwache Konvergenz und Abschätzungen der Tail-Wahrscheinlichkeiten

Wir werden zuerst eine stochastische obere Schranke für $\frac{1}{n} C_{n,l}^*$ herleiten, die eine ähnliche Aussage wie Satz 5.11 liefert, jedoch für den nicht asymptotischen Fall. Danach werden wir einen alternativen und deutlich kürzeren Beweis der Konvergenzaussage 5.5 führen.

6.1 Satz

Sei $(V_m)_{m \in \mathbb{N}}$ eine Folge von unabhängigen $U(\frac{1}{2}, 1)$ -verteilten Zufallsgrößen und $W \stackrel{\text{def}}{=} 1 + \sum_{m=1}^{\infty} \prod_{r=1}^m V_r$. Dann gilt

$$\frac{1}{n} C_{n,l}^* \stackrel{d}{\leq} W \quad \text{für alle } n \in \mathbb{N}, 1 \leq l \leq n$$

oder anders ausgedrückt

$$\sup_{1 \leq l \leq n} P(C_{n,l}^* \geq nz) \leq P(W \geq z) \quad \text{für alle } n \in \mathbb{N}, z \geq 0.$$

6.2 Bemerkung

(i) Zufallsgrößen wie W , d.h. Summen von Produkten von unabhängig identisch verteilten Zufallsgrößen, sind in vielen Arbeiten untersucht worden. Beispielsweise haben Goldie und Grübel in [11] gezeigt, dass

$$\lim_{z \rightarrow \infty} \frac{1}{z \log(z)} \log(P(W \geq z)) = -1$$

gilt. Dies impliziert zusammen mit obigem Satz, dass $\sup_{n \in \mathbb{N}, 1 \leq l \leq n} P(C_{n,l}^* > nz)$ mit wachsendem z schneller als exponentiell fällt. Durch eine ausführlichere Untersuchung von W werden wir später in Satz 6.7 eine genauere stochastische obere Schranke für $\frac{1}{n} C_{n,l}^*$ herleiten.

(ii) Die Schranke in Satz 6.1 kann in folgendem Sinne nicht verbessert werden:
Befindet sich das gesuchte Element in jedem Rekursionsschritt in der größeren der beiden Mengen $S_<, S_>$, so erhalten wir für die Worst Case Performance von FIND:

$$\frac{1}{n} \sup_{1 \leq l \leq n} C_{n,l}^* \xrightarrow{n \rightarrow \infty} 1 + \sum_{m=1}^{\infty} \prod_{k=1}^m \max \{U_k, 1 - U_k\}.$$

Dabei hat dieser Grenzwert dieselbe Verteilung wie W .

Nun erst mal zum Beweis von Satz 6.1.

Beweis:

Wir werden zuerst eine geeignete Version der Markov-Ketten konstruieren, welche wir auch in späteren Beweisen noch verwenden werden. Sei also $(U_m)_{m \in \mathbb{N}}$ ein Folge von unabhängigen auf $(0, 1)$ gleichverteilten Zufallsgrößen. Definiere $(Z_m)_{m \in \mathbb{N}_0}$ mit $Z_m = (X_m, Y_m) \in I$ rekursiv durch $Z_0 \stackrel{\text{def}}{=} (1, t)$ und

$$Z_{m+1} \stackrel{\text{def}}{=} \begin{cases} (X_m - U_{m+1} X_m, Y_m - U_{m+1} X_m), & \text{falls } U_{m+1} X_m \leq Y_m \\ (U_{m+1} X_m, Y_m), & \text{falls } U_{m+1} X_m > Y_m. \end{cases}$$

Offensichtlich ist $(Z_m)_{m \in \mathbb{N}_0}$ eine Markov-Kette mit Übergangskern P , die in $(1, t)$ startet. Wir haben also die Folge $(U_m)_{m \in \mathbb{N}}$ genutzt, um eine geeignete Version des späteren Grenzprozesses zu konstruieren.

Wir konstruieren nun mit denselben $(U_m)_{m \in \mathbb{N}}$ für jedes $n \in \mathbb{N}$ einen geeigneten Prozess $(Z_m^{(n)})_{m \in \mathbb{N}_0}$. Dafür benötigen wir die Hilfsfunktionen

$$h_n : \left\{ \frac{k}{n} : 1 \leq k \leq n \right\} \times [0, 1) \rightarrow \left\{ \frac{k}{n} : 1 \leq k \leq n \right\},$$

definiert durch

$$h_n \left(\frac{i}{n}, u \right) \stackrel{\text{def}}{=} \frac{1}{n} \lceil iu \rceil,$$

wobei $\lceil x \rceil \stackrel{\text{def}}{=} \min \{z \in \mathbb{Z} : z \geq x\}$ die obere Gaußklammer bezeichnet.

Damit ist $h_n \left(\frac{i}{n}, U \right)$ Laplace-verteilt auf $\left\{ \frac{k}{n} : 1 \leq k \leq i \right\}$. Definiere für jedes $n \in \mathbb{N}$ den Prozess $(Z_m^{(n)})_{m \in \mathbb{N}_0}$, $Z_m^{(n)} = (X_m^{(n)}, Y_m^{(n)}) \in I_n$ rekursiv durch $Z_0^{(n)} \stackrel{\text{def}}{=} (1, \frac{l}{n})$ und

$$Z_{m+1}^{(n)} \stackrel{\text{def}}{=} \begin{cases} (X_m^{(n)} - \zeta, Y_m^{(n)} - \zeta), & \text{falls } \zeta < Y_m^{(n)}, \\ (\frac{1}{n}, \frac{1}{n}), & \text{falls } \zeta = Y_m^{(n)}, \\ (\zeta - \frac{1}{n}, Y_m^{(n)}), & \text{falls } \zeta > Y_m^{(n)}, \end{cases}$$

wobei $\zeta \stackrel{\text{def}}{=} h_n \left(X_m^{(n)}, U_{m+1} \right)$ das Element bezeichnet, welches im $(m+1)$ -ten Rekursionsschritt als Pivotelement ausgewählt wird.

Wiederum sieht man leicht, dass $(Z_m^{(n)})_{m \in \mathbb{N}_0}$ für jedes $n \in \mathbb{N}$ eine Markov-Kette mit Übergangskern P_n bildet, die in $(1, \frac{1}{n})$ startet.

Wir werden nun die Quotienten von zwei aufeinander folgenden X -Werten untersuchen um eine obere Schranke dieser Quotienten zu erhalten. Diese ermöglichen uns dann eine entsprechende Abschätzung der Tails von W und $\frac{1}{n}C_{n,l}^*$.

Sei hierfür $V_m \stackrel{\text{def}}{=} \max \{U_m, 1 - U_m\}$ für alle $m \in \mathbb{N}$. Beachte, dass $(V_m)_{m \in \mathbb{N}}$ eine unabhängige Folge von U ($\frac{1}{2}, 1$)-verteilten Zufallsgrößen ist. Es folgt direkt für alle $m \in \mathbb{N}$

$$\begin{aligned} \frac{X_{m+1}}{X_m} &= \begin{cases} \frac{X_m - U_{m+1}X_m}{X_m}, & \text{falls } U_{m+1}X_m \leq Y_m \\ \frac{U_{m+1}X_m}{X_m}, & \text{falls } U_{m+1}X_m > Y_m \end{cases} \\ &\leq V_{m+1}. \end{aligned} \quad (6.2)$$

Um eine ähnliche Ungleichung für die diskreten Versionen zu erhalten, benötigen wir die beiden folgenden Ungleichungen:

$$\frac{i-1-\lceil iu \rceil}{i-1} \leq 1-u, \quad \frac{\lceil iu \rceil - 2}{i-1} \leq u \quad \text{für alle } 0 < u < 1, i = 2, 3, \dots, n.$$

Wir werden exemplarisch die erste der beiden Ungleichungen beweisen und verweisen darauf, dass der Beweis der Zweiten ähnlich verläuft. Offensichtlich gilt $\frac{\lceil iu \rceil}{i-1} > \frac{\lceil iu \rceil}{i} \geq u$ und damit ergibt sich bereits $\frac{i-1-\lceil iu \rceil}{i-1} = 1 - \frac{\lceil iu \rceil}{i-1} \leq 1-u$.

Mit Hilfe dieser Ungleichungen folgt leicht

$$\frac{X_{m+1}^{(n)} - \frac{1}{n}}{X_m^{(n)} - \frac{1}{n}} \leq V_{m+1} \quad \text{für alle } m \in \mathbb{N}_0, n \in \mathbb{N} \text{ mit } X_m^{(n)} > \frac{1}{n}, \quad (6.3)$$

denn für $nX_m^{(n)} \geq 2$ ergibt sich

$$\begin{aligned} \frac{X_{m+1}^{(n)} - \frac{1}{n}}{X_m^{(n)} - \frac{1}{n}} &= \begin{cases} \frac{X_m^{(n)} - \zeta - \frac{1}{n}}{X_m^{(n)} - \frac{1}{n}}, & \text{falls } \zeta < Y_m^{(n)} \\ 0, & \text{falls } \zeta = Y_m^{(n)} \\ \frac{\zeta - \frac{2}{n}}{X_m^{(n)} - \frac{1}{n}}, & \text{falls } \zeta > Y_m^{(n)} \end{cases} \\ &= \begin{cases} \frac{nX_m^{(n)} - \lceil nX_m^{(n)} U_{m+1} \rceil - 1}{nX_m^{(n)} - 1}, & \text{falls } \zeta < Y_m^{(n)} \\ 0, & \text{falls } \zeta = Y_m^{(n)} \\ \frac{\lceil nX_m^{(n)} U_{m+1} \rceil - 2}{nX_m^{(n)} - 1}, & \text{falls } \zeta > Y_m^{(n)} \end{cases} \\ &\leq V_{m+1}. \end{aligned}$$

Somit folgt für alle $z \geq 0$ unter Beachtung von $X_0^{(n)} = 1$, der Ungleichung (6.3), sowie der

Aussage (6.1):

$$\begin{aligned}
P(W > z) &= P\left(1 + \sum_{m=1}^{\infty} \prod_{r=0}^{m-1} V_{r+1} > z\right) \\
&\geq P\left(1 + \sum_{m=1}^{\infty} \prod_{r=0}^{m-1} \frac{X_{r+1}^{(n)} - \frac{1}{n}}{X_r^{(n)} - \frac{1}{n}} > z\right) \\
&= P\left(1 + \sum_{m=1}^{\infty} \frac{X_m^{(n)} - \frac{1}{n}}{X_0^{(n)} - \frac{1}{n}} > z\right) \\
&= P\left(1 + \frac{n}{n-1} \sum_{m=1}^{\infty} \left(X_m^{(n)} - \frac{1}{n}\right) > z\right) \\
&\geq P\left(\sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n}\right) > z\right) \\
&= P\left(\frac{1}{n} C_{n,l}^* > z\right).
\end{aligned}$$

□

Wir kommen nun zum bereits angesprochenen Beweis der Verteilungskonvergenz der von FIND benötigten Vergleiche gegen den FIND-Prozess. Der hier vorgestellte Beweis, der im Vergleich zum Beweis von Satz 5.5 deutlich kürzer ist, benutzt die bisher in diesem Kapitel eingeführten Notationen.

6.3 Satz

Sei $(l_n)_{n \in \mathbb{N}}$ eine Folge von ganzen Zahlen mit $1 \leq l_n \leq n$, so dass $\frac{l_n}{n}$ für $n \rightarrow \infty$ gegen einen Grenzwert $t \in [0, 1]$ konvergiert. Dann gilt

$$\frac{1}{n} C_{n,l}^* \xrightarrow{d} \mathcal{L}\left(\sum_{m=0}^{\infty} X_m \middle| Z_0 = (1, t)\right).$$

Anders ausgedrückt besagt der Satz, dass $\frac{1}{n} C_{n,l_n}^*$ in Verteilung gegen den FIND-Prozess konvergiert.

Beweis:

Mit Blick auf die Gleichung (6.1) ergibt sich die Behauptung, falls wir

$$\sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n}\right) \xrightarrow{n \rightarrow \infty} \sum_{m=0}^{\infty} X_m \quad \text{fast sicher} \quad (6.4)$$

zeigen.

Wir halten zuerst fest, dass wir $X_{m+1}^{(n)}$ und X_{m+1} als Funktionen von $X_m^{(n)}$, $Y_m^{(n)}$ und U_{m+1} , bzw. X_m , Y_m und U_{m+1} schreiben können, also $X_{m+1}^{(n)} = \Phi_n(X_m^{(n)}, Y_m^{(n)}, U_{m+1})$

und $X_{m+1} = \Phi(X_m, Y_m, U_{m+1})$ mit

$$\begin{aligned} \Phi_n(x, y, u) &\stackrel{\text{def}}{=} \left(x - \frac{1}{n} \lceil nux \rceil \right) \mathbf{1}_{(0,y)} \left(\frac{1}{n} \lceil nux \rceil \right) \\ &\quad + \left(\frac{1}{n} \lceil nux \rceil - \frac{1}{n} \right) \mathbf{1}_{(y,1)} \left(\frac{1}{n} \lceil nux \rceil \right) \\ &\quad + \frac{1}{n} \mathbf{1}_{\{y\}} \left(\frac{1}{n} \lceil nux \rceil \right) \end{aligned}$$

und

$$\Phi(x, y, u) \stackrel{\text{def}}{=} (x - ux) \mathbf{1}_{(0,y]}(ux) + ux \mathbf{1}_{(y,1)}(ux).$$

Falls nun (x_n, y_n) gegen $(x, y) \in I$ mit $ux \neq y$ konvergiert, so konvergiert auch $\Phi_n(x_n, y_n, u)$ gegen $\Phi(x, y, u)$. Zusammen mit einer ähnlichen Analyse für die zweite Komponente $Y_m^{(n)}$ bzw. Y_m folgt dann für alle $m \in \mathbb{N}_0$ die Implikation

$$\left(X_m^{(n)}, Y_m^{(n)} \right) \rightarrow (X_m, Y_m) \text{ fast sicher} \Rightarrow \left(X_{m+1}^{(n)}, Y_{m+1}^{(n)} \right) \rightarrow (X_{m+1}, Y_{m+1}) \text{ fast sicher.}$$

Da aber nach Konstruktion $\left(X_0^{(n)}, Y_0^{(n)} \right) \rightarrow (X_0, Y_0)$ gilt, folgt schon

$$X_m^{(n)} - \frac{1}{n} \xrightarrow{n \rightarrow \infty} X_m \text{ fast sicher für alle } m \in \mathbb{N}.$$

Um mit Hilfe des Satzes von der majorisierten Konvergenz die behauptete Konvergenzaussage (6.4) beweisen zu können, brauchen wir noch eine Majorante von $\sum_{m=0}^{\infty} X_m^{(n)} - \frac{1}{n}$. Aus dem Satz der monotonen Konvergenz und der Unabhängigkeit der $(V_m)_{m \in \mathbb{N}_0}$ ergibt sich für den Erwartungswert von W

$$E(W) = 1 + \sum_{m=1}^{\infty} \prod_{r=1}^m E(V_r) = \sum_{m=0}^{\infty} \left(\frac{3}{4} \right)^m = 4.$$

Damit ist W fast sicher endlich. Die beiden Abschätzungen für die Quotienten (6.2) und (6.3) der X bzw. $X^{(n)}$ liefern die Integrierbarkeit von $\sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right)$ und $\sum_{m=0}^{\infty} X_m$. Wir zeigen nur den aufwendigeren Nachweis für $\sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right)$.

$$\begin{aligned} W &= 1 + \sum_{m=1}^{\infty} \prod_{r=1}^m V_r \geq \sum_{m=1}^{\infty} \prod_{r=1}^{\infty} \frac{X_r^{(n)} - \frac{1}{n}}{X_{r-1}^{(n)} - \frac{1}{n}} \\ &= 1 + \frac{n}{n-1} \sum_{m=1}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right) \geq 1 + \sum_{m=1}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right) \\ &= \sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right). \end{aligned}$$

Es folgt dann nach Anwendung des Satzes der majorisierten Konvergenz

$$\lim_{n \rightarrow \infty} \sum_{m=0}^{\infty} \left(X_m^{(n)} - \frac{1}{n} \right) = \sum_{m=0}^{\infty} \lim_{n \rightarrow \infty} \left(X_m^{(n)} - \frac{1}{n} \right) = \sum_{m=0}^{\infty} X_m \quad \text{fast sicher}$$

und damit die Behauptung des Satzes.

□

Sei $Q_{x,y} \stackrel{\text{def}}{=} \mathcal{L} \left(\sum_{m=0}^{\infty} X_m \middle| Z_0 = (x, y) \right)$. Dann sind die Verteilungen $Q_{1,t}$ für $0 \leq t \leq 1$ von besonderer Bedeutung, da sie als Grenzwert in dem vorangegangenen Satz auftreten. Wegen Satz 6.1 impliziert obige Verteilungskonvergenz die Aussage $Q_{1,t} \stackrel{d}{\leq} W$, wobei W wie im Satz 6.1 definiert ist. Dasselbe Resultat hatten wir auch schon im vorherigen Kapitel erhalten, siehe hierfür Satz 5.11. Im nächsten Satz werden wir nun eine ähnliche stochastische untere Schranke für $Q_{1,t}$ beweisen.

6.4 Satz

Sei $W' \stackrel{\text{def}}{=} 1 + \sum_{m=1}^{\infty} \prod_{r=1}^m U_r$, wobei $(U_r)_{r \in \mathbb{N}}$ eine Folge von unabhängigen $U(0,1)$ -verteilten Zufallsgrößen ist. Dann gilt

$$Q_{1,t} \stackrel{d}{\geq} W' \quad \text{für alle } t \in [0, 1].$$

Beweis:

Beachte, dass $X \stackrel{d}{\leq} Y$ genau dann gilt, wenn $Eh(X) \leq Eh(Y)$ für alle nicht fallenden, messbaren Funktionen $h : \mathbb{R}_+ \rightarrow [0, 1]$ gilt.

Wir behaupten zuerst, dass mit $\xi_m \stackrel{\text{def}}{=} \mathbf{1}_{(0,5,1)} \left(\frac{Y_m}{X_m} \right)$ folgende Ungleichung für alle $m \in \mathbb{N}_0$ gilt:

$$\frac{X_{m+1}}{X_m} \geq (1 - \xi_m) U_{m+1} + \xi_m (1 - U_{m+1}). \quad (6.5)$$

Für den Beweis dieser Behauptung werden wir eine Fallunterscheidung vornehmen. Sei also zuerst $\frac{Y_m}{X_m} \geq U_{m+1}$. Dann folgt aus der Konstruktion der X_m , dass $\frac{X_{m+1}}{X_m} = 1 - U_{m+1}$ gilt. Damit ergibt sich dann unter Beachtung der Definition von ξ_m

$$\begin{aligned} \frac{X_{m+1}}{X_m} = 1 - U_{m+1} &\geq \begin{cases} U_{m+1}, & \text{falls } \frac{Y_m}{X_m} \leq 0,5 \\ 1 - U_{m+1}, & \text{falls } \frac{Y_m}{X_m} > 0,5 \end{cases} \\ &= (1 - \xi_m) U_{m+1} + \xi_m (1 - U_{m+1}). \end{aligned}$$

Nun zum zweiten Fall $\frac{Y_m}{X_m} < U_{m+1}$. Dieses Mal folgt $\frac{X_{m+1}}{X_m} = U_{m+1}$ und dann weiter mit

der Definition von ξ_m , dass

$$\begin{aligned} \frac{X_{m+1}}{X_m} = U_{m+1} &\geq \begin{cases} U_{m+1}, & \text{falls } \frac{Y_m}{X_m} \leq 0,5 \\ 1 - U_{m+1}, & \text{falls } \frac{Y_m}{X_m} > 0,5 \end{cases} \\ &= (1 - \xi_m) U_{m+1} + \xi_m (1 - U_{m+1}) \end{aligned}$$

gilt. Die hiermit bewiesene Ungleichung (6.5) liefert nun für beliebiges $m \in \mathbb{N}_0$

$$\begin{aligned} &P(X_{m+1} > z \mid Z_0, \dots, Z_m) \\ &\geq P\left(\left((1 - \xi_m) U_{m+1} + \xi_m (1 - U_{m+1})\right) X_m > z \mid Z_0, \dots, Z_m\right) \\ &= P\left(U_{m+1} X_m \mathbf{1}_{(0, 0,5]} \left(\frac{Y_m}{X_m}\right) + (1 - U_{m+1}) X_m \mathbf{1}_{(0,5,1)} \left(\frac{Y_m}{X_m}\right) > z \mid Z_0, \dots, Z_m\right) \\ &= P(V > z) \quad \text{für alle } z \geq 0, \end{aligned}$$

mit $\mathcal{L}(V) = U(0, X_m)$. Geschrieben als Aussage für Verteilungen erhalten wir

$$\mathcal{L}(X_{m+1} \mid Z_0, \dots, Z_m) \stackrel{d}{\geq} U(0, X_m) \quad (6.6)$$

für alle $m \in \mathbb{N}_0$.

Sei nun $(U'_m)_{m \in \mathbb{N}}$ eine Folge von unabhängigen $U(0, 1)$ -verteilten Zufallsgrößen. Definiere $(X'_m)_{m \in \mathbb{N}_0}$ durch $X'_0 \stackrel{\text{def}}{=} X_0$ und $X'_{m+1} \stackrel{\text{def}}{=} U'_{m+1} X_m$ für alle $m \in \mathbb{N}_0$.

Wir werden als nächstes $\sum_{m=0}^M \alpha_m X_m \stackrel{d}{\geq} \sum_{m=0}^M \alpha_m X'_m$ für alle $\alpha_1, \dots, \alpha_M \geq 0$ und $M \in \mathbb{N}_0$ per Induktion nach M beweisen.

Der Induktionsanfang für $M = 0$ ist trivialerweise erfüllt. Sie $h : \mathbb{R}_+ \rightarrow [0, 1]$ nicht fallend und messbar. Dann gilt für beliebige $\alpha_1, \dots, \alpha_{M+1} \geq 0$ unter Verwendung des Satzes von Fubini für bedingte Verteilungen und der Ungleichung (6.6):

$$\begin{aligned} &Eh\left(\sum_{m=0}^{M+1} \alpha_m X_m\right) \\ &= E\left(E\left(h\left(\sum_{m=0}^{M+1} \alpha_m X_m\right) \mid Z_0, \dots, Z_M\right)\right) \\ &= \int \int h\left(\sum_{m=0}^M \alpha_m X_m + \alpha_{M+1} X_{M+1}\right) d\mathcal{L}(X_{M+1} \mid Z_0, \dots, Z_M) d\mathcal{L}(Z_0, \dots, Z_M) \\ &\geq \int \int_0^1 h\left(\sum_{m=0}^M \alpha_m X_m + u \alpha_{M+1} X_M\right) du d\mathcal{L}(Z_0, \dots, Z_M) \end{aligned}$$

$$\begin{aligned}
 &= \int_0^1 Eh \left(\sum_{m=0}^{M-1} \alpha_m X_m + (\alpha_M + u \alpha_{M+1}) X_M \right) du \\
 &\geq \int_0^1 Eh \left(\sum_{m=0}^{M-1} \alpha_m X'_m + (\alpha_M + u \alpha_{M+1}) X'_M \right) du \\
 &= Eh \left(\sum_{m=0}^{M+1} \alpha_m X'_m \right).
 \end{aligned}$$

Die Induktionsvoraussetzung für $\alpha_1, \dots, \alpha_{M-1}$ und $\tilde{\alpha}_M \stackrel{\text{def}}{=} \alpha_M + u \alpha_{M+1}$ ging hierbei in der letzten Ungleichung ein. Es folgt nun mit Hilfe des Satzes der monotonen Konvergenz für alle $z \geq 0$, indem alle $\alpha_m = 1$ gesetzt werden:

$$P \left(\sum_{m=0}^{\infty} X_m > z \right) = \lim_{M \rightarrow \infty} P \left(\sum_{m=0}^M X_m > z \right) \geq \lim_{M \rightarrow \infty} P \left(\sum_{m=0}^M X'_m > z \right) = P \left(\sum_{m=0}^{\infty} X'_m > z \right).$$

Da $\mathcal{L}(\sum_{m=0}^{\infty} X'_m) = \mathcal{L}(W')$ und $X_0 = X'_0 = 1$ gilt, folgt die Behauptung des Satzes aus obiger Ungleichung.

□

Wir kommen nun zu einer Integralgleichung, welche eine Beziehung zwischen den Verteilungen $Q_{1,t}$, $0 \leq t \leq 1$, herstellt. Diese kann auch für die Berechnungen von Eigenschaften der Grenzverteilung genutzt werden, wie zum Beispiel zur Berechnung des Erwartungswertes. Zur Formulierung des Satzes benötigen wir allerdings noch folgende Notation. Für $c > 0$ definiere die Funktionen $T_c : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ durch $T_c(x) \stackrel{\text{def}}{=} cx$. Bezeichne $T_c(Q)$ das Bild der Verteilung Q unter T_c , d.h. falls X eine Zufallsgröße mit Verteilung Q ist, dann ist $T_c(X)$ die Verteilung von cX . Mit $\mu * \nu$ bezeichnen wir die Faltung der Maße μ und ν .

6.5 Satz

Für alle $t \in [0, 1]$ gilt

$$Q_{1,t} = \int_0^t \delta_1 * T_{1-u} \left(Q_{1, \frac{t-u}{1-u}} \right) du + \int_t^1 \delta_1 * T_u \left(Q_{1, \frac{t}{u}} \right) du.$$

Beweis:

Mit Hilfe des Satzes von Fubini für bedingte Verteilungen, der Markov-Eigenschaft des Prozesses $(Z_m)_{m \in \mathbb{N}_0}$ und unter Beachtung der Konstruktion des Prozesses $(Z_m)_{m \in \mathbb{N}_0}$ ergibt sich

$$Q_{x,y} = \mathcal{L} \left(\sum_{m=0}^{\infty} X_m \middle| Z_0(x, y) \right)$$

$$\begin{aligned}
&= \int \mathcal{L} \left(x + \sum_{m=1}^{\infty} X_m \middle| Z_1 \right) d\mathcal{L} (Z_1 | Z_0 = (x, y)) \\
&= \int_0^{\frac{y}{x}} \delta_x * Q_{x-ux, y-ux} du + \int_{\frac{y}{x}}^1 \delta_x * Q_{ux, y} du.
\end{aligned}$$

Außerdem gilt $Q_{cx, cy} = T_c(Q_{x, y})$ für alle $(x, y) \in I$ und $c > 0$, denn:

Sei W eine Zufallsgröße mit $\mathcal{L}(W) = \mathcal{L} \left(\sum_{m=0}^{\infty} X_m \middle| Z_0 = (x, y) \right)$. Wegen

$$X_{m+1} = \begin{cases} X_m (1 - U_{m+1}), & \text{falls } U_{m+1} X_m \leq Y_m \\ X_m U_{m+1}, & \text{falls } U_{m+1} X_m > Y_m \end{cases}$$

gilt dann

$$T_c(Q_{x, y}) = \mathcal{L}(cW) = \mathcal{L} \left(c \sum_{m=0}^{\infty} X_m \middle| Z_0 = (x, y) \right) = \mathcal{L} \left(\sum_{m=0}^{\infty} X_m \middle| Z_0 = (cx, cy) \right) = Q_{cx, cy}.$$

Insgesamt folgt somit die Behauptung durch:

$$\begin{aligned}
Q_{1, t} &= \int_0^t \delta_1 * Q_{1-u, t-u} du + \int_t^1 \delta_1 * Q_{u, t} du \\
&= \int_0^t \delta_1 * T_{1-u} \left(Q_{1, \frac{t-u}{1-u}} \right) du + \int_t^1 \delta_1 * T_u \left(Q_{1, \frac{t}{u}} \right) du.
\end{aligned}$$

□

6.6 Bemerkung

Obiger Satz kann auch zur Berechnung der Momente der Grenzverteilung genutzt werden.

Für $m_1(t) \stackrel{\text{def}}{=} \int x Q_{1, t}(dx)$ gilt

$$m_1(t) = 1 + \int_0^t (1-u) m_1 \left(\frac{t-u}{1-u} \right) du + \int_t^1 u m_1 \left(\frac{t}{u} \right) du.$$

Diese Gleichung wird wie im vorherigen Kapitel ausführlich beschrieben von $m_1(t) = 2 - 2t \log(t) - 2(1-t) \log(1-t)$ gelöst. Für die Berechnungen der höheren Momente siehe [22] oder die Ausführungen im 5. Kapitel.

Als letzte Aussage dieses Kapitels werden wir nun eine weitere obere Schranke für die Tail-Wahrscheinlichkeiten von $C_{n, l}$ beweisen. Diese wird die Aussage der Bemerkung 6.2 spezifizieren.

6.7 Satz

Für alle $z \geq 0$, $n \in \mathbb{N}$ und $1 \leq l \leq n$ gilt

$$P(C_{n,l}^* \geq nz) \leq \exp(z + z \log 4 - z \log z).$$

Beweis:

Wir halten zuerst fest, dass die Behauptung offensichtlich für alle $z \leq 4e \approx 10,873$ gilt. Wir können uns also im Folgenden auf die $z > 4e$ beschränken.

Der Beweis basiert auf einem Argument von Goldie und Grübel in [11]. Diese zeigen dort, dass die momenterzeugende Funktion $M(\theta) = E \exp(\theta W)$ von W aus Satz 6.1 die Ungleichung $M(\theta) \leq \exp(C \exp(\theta))$ erfüllt, falls C so gewählt ist, dass

$$2 \int_{\frac{1}{2}}^1 \exp(C \exp(\theta u)) du \leq \exp(C \exp(\theta) - \theta) \quad (6.7)$$

für alle $\theta > 0$ gilt.

Fixiere ein $z > 4e$. Unter der Bedingung, dass (6.7) für $C = 4$ gilt, folgt unter Beachtung von $\log(\frac{z}{4}) > 0$ mit Satz 6.1 und der Markov-Ungleichung mit $g(t) = \exp(t \log(\frac{z}{4}))$ die Behauptung mittels

$$\begin{aligned} P(C_{n,l}^* \geq nz) &\leq P(W \geq z) \\ &\leq \frac{M(\log(\frac{z}{4}))}{\exp(z \log(\frac{z}{4}))} \\ &\leq \frac{\exp(4 \exp(\log(\frac{z}{4})))}{\exp(z \log(\frac{z}{4}))} \\ &= \exp(z + z \log(4) - z \log(z)). \end{aligned}$$

Wir müssen also noch zeigen, dass (6.7) für $C = 4$ und alle $\theta > 0$ gilt. Wir bezeichnen mit $f_l(\theta)$ und $f_r(\theta)$ die linke bzw. rechte Seite der Ungleichung (6.7) mit $C = 4$. Wegen $f_l(0) = f_r(0)$ reicht es aus

$$\frac{d}{d\theta}(\theta f_l(\theta)) \leq \frac{d}{d\theta}(\theta f_r(\theta)) \quad (6.8)$$

für alle $\theta > 0$ zu zeigen.

Zuerst betrachten wir die linke Seite dieser Ungleichung. Durch die Substitution $x = \exp(4 \exp(\theta u))$ erhalten wir

$$\begin{aligned} \theta f_l(\theta) &= 2\theta \int_{\frac{1}{2}}^1 \exp(4 \exp(\theta u)) du \\ &= 2 \int_{\exp(4 \exp(\frac{\theta}{2}))}^{\exp(4 \exp(\theta))} \frac{1}{\log(x)} dx, \end{aligned}$$

so dass sich

$$\begin{aligned}\frac{d}{d\theta}(\theta f_l(\theta)) &= \frac{2 \exp(4 \exp(\theta)) \cdot 4 \exp(\theta)}{\log(\exp(4 \exp(\theta)))} - \frac{2 \exp(4 \exp(\frac{\theta}{2})) \cdot 2 \exp(\frac{\theta}{2})}{\log(\exp(4 \exp(\theta)))} \\ &= 2 \exp(4 \exp(\theta)) - \exp\left(4 \exp\left(\frac{\theta}{2}\right)\right)\end{aligned}$$

ergibt. Sei nun $y \stackrel{\text{def}}{=} \exp\left(\frac{\theta}{2}\right)$, nach Multiplikation mit $y^2 \exp(-4y^2)$ liefert dies nun

$$\begin{aligned}y^2 \exp(-4y^2) \frac{d}{d\theta}(\theta f_l(\theta)) &= y^2 \exp(-4y^2) (2 \exp(4y^2) - \exp(4y)) \\ &= 2y^2 - y^2 \exp(4y(1-y)).\end{aligned}$$

Nun betrachten wir die rechte Seite von (6.7). Dieses Mal erhalten wir nach Multiplikation mit $y^2 \exp(-4y^2)$

$$\begin{aligned}y^2 \exp(-4y^2) \frac{d}{d\theta}(\theta f_r(\theta)) &= y^2 \exp(-4y^2) \left(\exp(4y^2 - 2 \log(y)) \right. \\ &\quad \left. + 2 \log(y) \exp(4y^2 - 2 \log(y)) (4y^2 - 1) \right) \\ &= 1 + 8y^2 \log(y) - 2 \log(y).\end{aligned}$$

Wir sehen also, dass (6.8) genau dann gilt, wenn

$$g(y) \stackrel{\text{def}}{=} -2y^2 + y^2 \exp(4y(1-y)) + 1 + 8y^2 \log(y) - 2 \log(y) \geq 0$$

für alle $y > 1$ erfüllt ist. Nach elementarer Berechnung der Ableitung von g sehen wir, dass $g(1) = g'(1) = 0$ gilt. Das Gewünschte, d.h. $g(y) \geq 0$ für alle $y > 1$, folgt also, wenn $g''(y) \geq 0$ für alle $y > 1$ gilt. Wir haben

$$g''(y) = 20 + p(y) \exp(4y(1-y)) + 16 \log(y) + \frac{2}{y^2},$$

mit $p(y) \stackrel{\text{def}}{=} 64y^4 - 64y^3 - 24y^2 + 16y + 2$.

Mittels einer elementaren Untersuchung von $p(y)$ erkennen wir, dass $p(y) \geq p(1) = -6$ für alle $y > 1$ gilt. Dann folgt aber aus $\exp(4y(1-y)) \leq 1$ für alle $y > 1$, dass $g''(y) \geq 0$ für alle $y > 1$ gilt. Wie oben beschrieben folgt nun hieraus die Behauptung des Satzes.

□

6.3 Numerische Ergebnisse

- (1) Sei $F(t, x) \stackrel{\text{def}}{=} Q_{1,t}([0, x])$ die Verteilungsfunktion von $Q_{1,t}$. Die Gleichung (6.5) in Satz 6.5 führt zur Integralgleichung

$$F(t, x) = \int_0^t F\left(\frac{t-u}{1-u}, \frac{\max\{x-1, 0\}}{1-u}\right) du + \int_t^1 F\left(\frac{t}{u}, \frac{\max\{x-1, 0\}}{u}\right) du. \quad (6.9)$$

Es erscheint schwierig eine direkte Lösung dieser Gleichung anzugeben. Allerdings kann (6.9) dazu genutzt werden, eine numerische Lösung zu bestimmen. Nach Diskretisierung und Anwendung der Trapezregel, d.h. $\int_a^b f(x) \approx (b-a) \frac{f(a)+f(b)}{2}$, erhalten wir eine approximative Version von (6.9), welche per Iteration zu einer ausreichend genauen Lösung der Gleichung führt. Tabelle 6.1, die wir aus [12] übernommen haben, zeigt die 0,5-, 0,9- und 0,99-Quantile von $Q_{1,t}$ für einige t -Werte. Beachte dass, $Q_{1,t} = Q_{1,1-t}$ gilt.

t	Quantile von $Q_{1,t}$		
	0,50	0,90	0,99
0,00	1,89	2,97	4,11
0,10	2,52	3,82	5,15
0,25	3,01	4,41	5,81
0,50	3,27	4,72	6,16

Tabelle 6.1: Quantile von $Q_{1,t}$

Wir können solche Quantile auch auf andere Weise berechnen. Mit der expliziten Formel für den Erwartungswert $m_1(t)$ von $Q_{1,t}$ aus Satz 5.13 können die Quantile unter Benutzung der Markov-Ungleichung berechnet werden.

$$Q_{1,t}([\alpha m_1(t), \infty)) \leq \frac{m_1(t)}{\alpha m_1(t)} = \frac{1}{\alpha}.$$

Das 0,99-Quantil erhalten wir also für $\alpha = 100$. Für $t = 0,5$ ergibt sich eine Schranke von 338,63. Diese ist ungemein höher als der Wert 6,16 aus Tabelle 6.1.

Als weitere Berechnungsalternative kann die Abschätzung in Satz 6.7 dienen. Hier ergibt sich eine Schranke von 14,84.

Die numerische Lösung führt also zu einer genaueren Berechnung der Quantile.

- (2) Die stochastischen oberen und unteren Schranken aus den Sätzen 6.1 und 6.4 sind in den Tails sehr nah beieinander. Folgendes Diagramm, welches wir aus [12] übernommen haben, zeigt die Verteilungsfunktionen dieser beiden Schranken und die von $Q_{1,0,5}$.

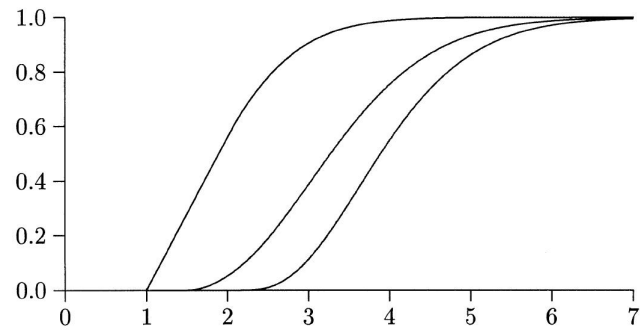


Abbildung 6.2: Verteilungsfunktionen von $Q_{1,0,5}$ und der stochastischen Schranken

Anhang

Nebenrechnung (zum Beweis von Satz 5.11)

Sei $\epsilon \in (0, \frac{1}{2})$ beliebig. Dann konvergiert die Reihe

$$\sum_{i=1}^{\infty} \log \left(\frac{(1-2\epsilon) \exp(\theta (1-\epsilon)^{n-1})}{1-2\epsilon \exp(\theta (1-\epsilon)^{n-1})} \right)$$

aus dem Beweis von Satz 5.11 für alle $\theta < -\log(2\epsilon)$.

Beweis:

Für den Beweis werden wir folgende elementare Abschätzungen für die Exponential- und Logarithmusfunktion benötigen:

- (i) $\log(x) \leq x - 1$ für alle $x > 0$,
- (ii) $\exp(x) \leq \frac{1}{1-x}$ für alle $x < 1$,
- (iii) $-\log(x) \leq \frac{1}{x} - 1$ für alle $x > 0$,
- (iv) $\log(x) \geq 1 - \frac{1}{x}$ für alle $x > 0$.

Wir wählen nun ein $N \in \mathbb{N}$ so groß, dass für alle $n \geq N$ gilt:

- (a) $-\log(2\epsilon) (1-\epsilon)^{n-1} < 1$,
- (b) $1 - 2\epsilon + (1-\epsilon)^{n-1} \log(2\epsilon) > 0$,
- (c) $(n^2 + 1)(1-\epsilon)^{n-1} \leq 2\epsilon$.

Damit sind in folgender Rechnung bei allen auftretenden Brüchen jeweils Zähler und Nenner positiv, und es ergibt sich:

$$\begin{aligned} \sum_{n=N}^{\infty} \log \left(\frac{(1-2\epsilon) \exp(\theta (1-\epsilon)^{n-1})}{1-2\epsilon \exp(\theta (1-\epsilon)^{n-1})} \right) &\stackrel{(i)}{\leq} \sum_{n=N}^{\infty} \frac{(1-2\epsilon) \exp(\theta (1-\epsilon)^{n-1})}{1-2\epsilon \exp(\theta (1-\epsilon)^{n-1})} - 1 \\ &= \sum_{n=N}^{\infty} \frac{\exp(\theta (1-\epsilon)^{n-1}) - 1}{1-2\epsilon \exp(\theta (1-\epsilon)^{n-1})} \\ &\stackrel{(ii)+(a)}{\leq} \sum_{n=N}^{\infty} \frac{\frac{1}{1-\theta(1-\epsilon)^{n-1}} - 1}{1 - \frac{2\epsilon}{1-\theta(1-\epsilon)^{n-1}}} \end{aligned}$$

$$\begin{aligned}
&= \sum_{n=N}^{\infty} \frac{\theta(1-\epsilon)^{n-1}}{1-2\epsilon-\theta(1-\epsilon)^{n-1}} \\
&\stackrel{(b)}{\leq} \sum_{n=N}^{\infty} \frac{-(1-\epsilon)^{n-1} \log(2\epsilon)}{1-2\epsilon+(1-\epsilon)^{n-1} \log(2\epsilon)} \\
&\stackrel{(iii)+(iv)}{\leq} \sum_{n=N}^{\infty} \frac{\left(\frac{1}{2\epsilon}-1\right)(1-\epsilon)^{n-1}}{1-2\epsilon+\left(1-\frac{1}{2\epsilon}\right)(1-\epsilon)^{n-1}} \\
&= \sum_{n=N}^{\infty} \frac{(1-2\epsilon)(1-\epsilon)^{n-1}}{2\epsilon(1-2\epsilon)+(2\epsilon-1)(1-\epsilon)^{n-1}} \\
&= \sum_{n=N}^{\infty} \frac{(1-\epsilon)^{n-1}}{-(1-\epsilon)^{n-1}+2\epsilon} \\
&\stackrel{(c)}{\leq} \sum_{n=N}^{\infty} \frac{1}{n^2} < \infty.
\end{aligned}$$

Da aber auch die ersten $(N-1)$ Summanden der Reihe endlich sind, folgt die Behauptung.

□

Literaturverzeichnis

- [1] G. Alsmeyer. *Verzweigungsprozesse (Skript zur Vorlesung)*.
- [2] G. Alsmeyer. *Wahrscheinlichkeitstheorie (3.Auflage)*. Skripten zur Mathematischen Statistik, Nr. 30, Universität Münster (2003).
- [3] G. Alsmeyer. *Stochastische Prozesse Teil 1 (3.Auflage)*. Skripten zur Mathematischen Statistik, Nr. 33, Universität Münster (2005).
- [4] K. B. Athreya and P. E. Ney. *Branching Processes*. Springer, New York, 1972.
- [5] P. Bickel and D. A. Freedman. Some asymptotic theory for the bootstrap. *Annals of Statistics*, 9:1196–1217, 1981.
- [6] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.
- [7] M. Cramer. *Stochastische Analyse rekursiver Algorithmen mit idealen Metriken, Doktorarbeit, Universität Freiburg*, 1995.
- [8] M. Cramer and L. Rüschendorf. Analysis of recursive algorithms by the contraction method. *Lecture Notes in Statistics*, 114:18–33, 1995.
- [9] L. Devroye. Exponential bounds for the running time of a selection algorithm. *Journal of Computer and System Sciences*, 29:1–7, 1984.
- [10] L. Devroye. On the probabilistic worst-case time of "find". *Algorithmica*, 31:291–303, 2001.
- [11] C. M. Goldie and R. Grübel. Perpetuities with thin tails. *Advances in Applied Probability*, 28:463–480, 1996.
- [12] R. Grübel. Hoare's selection theorem: A markov chain approach. *Journal of Applied Probability*, 35:36–45, 1998.
- [13] R. Grübel and U. Rösler. Asymptotic distribution theory for hoare's selection algorithm. *Advances in Applied Probability*, 28:252–269, 1996.
- [14] C. A. R. Hoare. Algorithm 65, find. *Communication of the Association for Computing Machinery*, 4:321–322, 1961.

- [15] P. Kirschenhofer, C. Martinez, and H. Prodinger. Analysis of hoare's find algorithm with median-of-three partition. *Random Structures and Algorithms*, 10(1):143–156, 1997.
- [16] D. Knof. *Konvergenz von Find zu einem Maß auf $D[0,1]$* , Diplomarbeit, Universität Kiel, 2004.
- [17] D. Knof. *Struktur von Fixpunkten aus Prozessgleichungen*, Doktorarbeit, Universität Kiel, 2007.
- [18] D. Knof and U. Rösler. *The Analysis of Find or Perpetuities on Cadlag Functions*, <http://www.math.uni-kiel.de/stochastik/roesler/research/KnoRoe.pdf>, besucht: 07.10.2009, 2007.
- [19] D. E. Knuth. Mathematical analysis of algorithms. *Information Processing*, 71:19–27 (Proceedings of IFIP Congress, Ljubljana, 1971), 1972.
- [20] B. Kodaj and T. F. Mori. On the number of comparisons in hoare's algorithm "find". *Studia Scientiarum Mathematicarum Hungarica*, 33:185–207, 1997.
- [21] C. Martinez, D. Panario, and A. Viola. Optimal sampling strategies in quicksort and quickselect. *SIAM Journal Computing*, 31 (3):683–705, 2001.
- [22] V. Paulsen. The moments of find. *Journal of Applied Probability*, 34:1079–1082, 1997.
- [23] S. T. Rachev and L. Rüschendorf. Probability metrics and recursive algorithms. *Advances in Applied Probability*, 27:770–799, 1995.
- [24] U. Rösler. A limit theorem for "quicksort". *RAIRO, Theoretical Informatics and Applications*, 25:85–100, 1991.
- [25] U. Rösler. A fixed point theorem for distributions. *Stochastic Processes and their Applications*, 42:195–214, 1992.
- [26] U. Rösler. On the analysis of stochastic divide and conquer algorithms. *Algorithmica*, 29:238–261, 2001.
- [27] U. Rösler. *The Backward View on the Algorithm Find*, *Berichtsreihe des Mathematischen Seminars der Universität Kiel*, <http://www.numerik.uni-kiel.de/reports/2003/03-6.ps.gz>, besucht: 07.10.2009, 2003.
- [28] U. Rösler. Quickselect revisited. *Journal of the Iranian Statistical Institute*, 3:271–296, 2004.
- [29] U. Rösler and L. Rüschendorf. The contraction method for recursive algorithms. *Algorithmica*, 29:3–33, 2001.