Kapitel III
Größenverzerrung und Bäume mit Rückgrat: Ein probabilistischer Zugang zu GWP

Die vorhergehenden zwei Kapitel haben uns einen guten Einblick in die klassische Theorie von Galton-Watson-Prozessen vermittelt und gezeigt, daß diese trotz häufiger Verwendung von Martingalargumenten vornehmlich durch das Studium der auftretenden erzeugenden Funktionen und damit analytisch geprägt ist. Die Unabhängigkeitseigenschaften und die rekursive Struktur von GWP und GWPI lassen sich in brauchbare Eigenschaften dieser analytischen Transformierten übersetzen, und durch Taylor-Entwicklungen 1. oder 2. Ordnung gelangt man relativ schnell zu relevanten Aussagen, wobei sich der technische Anspruch in akzeptablen Grenzen hält. Zudem gestattet der Zugang, wie wir noch sehen werden, die Behandlung vieler Verallgemeinerungen von GWP, etwa solchen in variierenden Umgebungen (☞ Kapitel IV). Andererseits verhindert die Reduktion auf analytische und wenig intuitive Argumente ein tieferes Verständnis der betrachteten Prozesse: Man erkennt zwar, wie sich GWP asymptotisch verhalten, aber nicht unbedingt warum, da die geführten Beweise kaum eine probabilistische, auf das Pfadverhalten der Prozesse gerichtete Interpretation zulassen.

In diesem Kapitel wollen wir unseren Blickwinkel erweitern und einen eher genealogisch orientierten Standpunkt einnehmen. Dies bedeutet, daß wir unsere nachfolgenden Untersuchungen von GWP zu einem wesentlichen Teil auf die Betrachtung des assoziierten Stammbaums, also des Galton-Watson-Baums (GWB), stützen. GWB hatten wir zwar schon in Kapitel I vorgestellt, spielten aber dort mit Ausnahme von Abschnitt I.7 keine besondere Rolle. Zwei Zugänge sollen hiernach vorgestellt werden. Der erste stammt von Lyons, Pemantle und Peres [5] und basiert massiv auf dem Zusammenhang der Verteilung eines GWB's $\boldsymbol{GW}$ mit der Verteilung einer bestimmten assoziierten Größenverzerrung $\widehat{\boldsymbol{GW}}$ mit einem ausgezeichneten Pfad, genannt Rückgrat. Der zweite, etwas weniger bekannte Zugang geht auf Geiger [3] zurück, ist vor allem im subkritischen und kritischen Fall von Nutzen und basiert darauf, rekursiv eine Folge endlicher Bäume zu konstruieren, die zum einen eine geeignete Abhängigkeitsstruktur besitzen und deren Verteilungen zum anderen den bedingten Verteilungen $\mathbb{P}(\boldsymbol{GW} \in \cdot | z_n \circ \boldsymbol{GW} > 0)$, $n \geq 0$, entsprechen, wobei $z_n \circ \boldsymbol{GW}$ die Anzahl der Knoten in der n-ten Generation (auf Höhe n) angibt. Die Abschnitte 3–7 widmen sich dem ersten, die verbleibenden Abschnitte 8–? dem zweiten Zugang. Wir beginnen allerdings mit einigen notwendigen formalen Grundlagen zur Definition von (lokal endlichen) Bäumen sowie zur Metrisierung der Menge aller solchen Bäume, die es dann erlaubt, einen GWB als Zufallselement in dieser Menge aufzufassen.

1. Bäume

Im folgenden werden wir Bäume in kanonischer Weise als Teilmengen des Ulam-Harris-Baums $\mathbb{V} = \bigcup_{n\geq 0} \mathbb{N}^n$ definieren, wobei an die im Anschluß von Lemma I.1.2 eingeführten Bezeichnungen erinnert sei.

1.1. Definition. Eine Teilmenge $\tau \subset \mathbb{V}$ heißt *Baum*, falls sie den folgenden Bedingungen genügt:

(1) $\varnothing \in \tau$,

(2) $v_1...v_n \in \tau \;\Rightarrow\; v_1...v_k \in \tau$ für jedes $1 \leq k < n$ und $n \geq 1$,

(3) $v_1...v_{n-1}v_n \in \tau \;\Rightarrow\; v_1...v_{n-1}j \in \tau$ für jedes $j \in \{1, ..., v_n\}$ und $n \geq 1$.

Erfüllt τ ferner die Bedingung

(4) $z_n(\tau) \stackrel{\text{def}}{=} |\tau \cap \mathbb{N}^n| < \infty$ für alle $n \geq 0$,

so heißt τ *lokal endlich*. Die Elemente von τ nennen wir *Individuen*; das Individuum $\varnothing$ wird auch als *Wurzel* bezeichnet.

Im folgenden beschäftigen wir uns mit lokal endlichen Bäumen und definieren deshalb

$$\mathbb{T} \stackrel{\text{def}}{=} \big\{\tau \subset \mathbb{V} :\; \tau \text{ ist ein lokal endlicher Baum}\big\}.$$

Ferner sei für $\tau \in \mathbb{T}$ dessen *Höhe* durch

$$H(\tau) \stackrel{\text{def}}{=} \sup\{n \geq 0 :\; z_n(\tau) > 0\} \;\in\; \overline{\mathbb{N}}_0$$

definiert, wobei wir auf die Implikation

$$z_n(\tau) = 0 \quad\Rightarrow\quad z_{n+k}(\tau) = 0 \text{ für alle } k \geq 0$$

hinweisen.

Für $\tau \in \mathbb{T}$ und $n \geq 0$ setzen wir hiernach abkürzend

$$\begin{aligned}
\tau_n &\stackrel{\text{def}}{=} \tau \cap \mathbb{N}^n \;=\; \{v \in \tau : |v| = n\},\\
\tau_{|n} &\stackrel{\text{def}}{=} \bigcup_{k=0}^{n} \tau_k \;=\; \{v \in \tau : |v| \leq n\},\\
[\tau]_n &\stackrel{\text{def}}{=} \{\tau' \in \mathbb{T} : \tau'_{|n} = \tau_{|n}\}.
\end{aligned}$$

Die Menge $\mathbb{T}$ wird uns in den nachfolgenden Abschnitten als Wertebereich verschiedener Abbildungen begegnen. Um aus diesen Abbildungen Zufallsvariablen, d.h. meßbare Abildungen zu machen, verschaffen wir uns zunächst eine geeignete σ-Algebra auf $\mathbb{T}$, indem wir auf dieser Menge eine Metrik definieren, die eine Topologie vermittelt.

1.2. Lemma. *(a) Die Abbildung* $d : \mathbb{T} \times \mathbb{T} \to [0,1]$*, gegeben durch*

$$d(\tau,\tau') \stackrel{\text{def}}{=} \exp\Big(-\sup\{n \ge 0 : \tau_{|n} = \tau'_{|n}\}\Big), \quad \tau,\tau' \in \mathbb{T},$$

ist eine Metrik auf $\mathbb{T}$*, wobei natürlich* $e^{-\infty} \stackrel{\text{def}}{=} 0$ *vereinbart wird.*

(b) Der Raum $(\mathbb{T},d)$ *ist separabel. Genauer gilt:*

$$\mathbb{T}^e \stackrel{\text{def}}{=} \{\tau \in \mathbb{T} : H(\tau) < \infty\} = \{\tau \in \mathbb{T} : \tau \textit{ ist endlich}\}$$

ist abzählbar und liegt dicht in $\mathbb{T}$.

Beweis: (a) Es reicht offenbar der Nachweis der Dreiecksungleichung. Für beliebige $\tau,\tau',\tau'' \in \mathbb{T}$ gilt aber

$$\sup\{n : \tau_{|n} = \tau''_{|n}\} \ \ge\ \sup\{n : \tau_{|n} = \tau'_{|n}\} \wedge \sup\{n : \tau'_{|n} = \tau''_{|n}\}$$

und folglich

$$d(\tau,\tau'') \ \le\ d(\tau,\tau') \vee d(\tau',\tau'') \ \le\ d(\tau,\tau') + d(\tau',\tau'').$$

Eine Metrik, welche die erste obige Ungleichung erfüllt, heißt übrigens *Ultrametrik*.

(b) Mit $\mathbb{T}_n \stackrel{\text{def}}{=} \{\tau \in \mathbb{T} : H(\tau) = n\} = \{\tau \in \mathbb{T} : z_n(\tau) > 0, z_{n+1}(\tau) = 0\}$ gilt

$$\mathbb{T}^e = \bigcup_{n \ge 0} \mathbb{T}_n,$$

und da jedes $\mathbb{T}_n$ offenbar abzählbar ist, gilt dasselbe für $\mathbb{T}^e$. Sei nun $\tau \in \mathbb{T}$. Für $n \ge 0$ folgt dann

$$d(\tau,\tau_{|n}) = e^{-n},$$

und da $\tau_{|n} \in \mathbb{T}^e$ für jedes $n \ge 0$, zeigt dies die Dichtheit von $\mathbb{T}^e$ in $\mathbb{T}$. ◊

1.3. Bemerkungen. (a) Der metrische Raum $(\mathbb{T},d)$ ist vollständig. Da wir dieses Ergebnis jedoch nicht benötigen, verzichten wir auf den Beweis (☞ Übung 1.1).

(b) Für $\tau \in \mathbb{T}$ und $\varepsilon > 0$ bezeichne $\mathbb{B}(\tau,\varepsilon) \stackrel{\text{def}}{=} \{\tau' \in \mathbb{T} : d(\tau,\tau') < \varepsilon\}$ die offene ε-Kugel um τ bezüglich d. Da d nur Werte in der abzählbaren Menge $\{0\} \cup \{e^{-n} : n \in \mathbb{N}_0\}$ annimmt, gilt

$$\big\{\mathbb{B}(\tau,\varepsilon) : \tau \in \mathbb{T}, \varepsilon > 0\big\} = \big\{\mathbb{B}(\tau,e^{-n}) : \tau \in \mathbb{T}, n \ge 1\big\} = \big\{[\tau]_n : \tau \in \mathbb{T}, n \ge 1\big\},$$

wobei sich die zweite Gleichung vermöge der Äquivalenzkette

$$\tau' \in \mathbb{B}(\tau,e^{-k}) \quad\Leftrightarrow\quad \sup\{n : \tau_{|n} = \tau'_{|n}\} > n \quad\Leftrightarrow\quad \tau_{|n} = \tau'_{|n} \quad\Leftrightarrow\quad \tau' \in [\tau]_n$$

ergibt.

(c) Es seien $\tau, \tau' \in \mathbb{T}$ und $j \geq k \geq 1$. Dann gilt

$$[\tau]_j \cap [\tau']_k \;=\; \{\chi \in \mathbb{T} : \chi_{|j} = \tau_{|j},\, \chi_{|k} = \tau'_{|k}\} \;=\; \begin{cases} [\tau]_j, & \text{falls } \tau_{|k} = \tau'_{|k}, \\ \emptyset, & \text{falls } \tau_{|k} \neq \tau'_{|k}. \end{cases} \tag{1.1}$$

Folglich bildet

$$\mathcal{E} \overset{\text{def}}{=} \big\{\emptyset, \mathbb{T}\big\} \cup \big\{[\tau]_k : \tau \in \mathbb{T}, k \geq 1\big\} \tag{1.2}$$

ein $\cap$-stabiles System von Teilmengen von $\mathbb{T}$.

Wir definieren nun

$$\mathfrak{B}(\mathbb{T}) \overset{\text{def}}{=} \sigma(\mathcal{E}) \;=\; \sigma\Big(\big\{[\tau]_k : \tau \in \mathbb{T}, k \geq 1\big\}\Big) \tag{1.3}$$

und notieren:

1.4. Lemma. *Die gemäß (1.3) definierte σ-Algebra $\mathfrak{B}(\mathbb{T})$ über $\mathbb{T}$ entspricht der Borelschen σ-Algebra, also der von den (bzgl. der Metrik d) offenen Teilmengen erzeugten σ-Algebra.*

Beweis: Es genügt der Hinweis, daß sich jede nichtleere offene Teilmenge eines separablen Raums als abzählbare Vereinigung von ε-Kugeln schreiben läßt. $\diamondsuit$

Zum Ende dieses Abschnitts wollen wir noch kurz eine Ordnung auf $\mathbb{V}$ einführen, welche die Verwandtschaftsstruktur der als Individuen interpretierten Elemente von $\mathbb{V}$ widerspiegelt.

1.5. Definition. Es seien $v = v_1...v_m$ und $w = w_1...w_n$ Elemente von $\mathbb{V}$.

(a) Besitzt w die Darstellung vu für ein $u \in \mathbb{V}$, so heißt v *Vorfahre* von w sowie umgekehrt w ein *Nachkomme* oder *Nachfahre* von v, symbolisch durch $v \preceq w$ bzw. $w \succeq v$ ausgedrückt.

(b) Gilt in (a) speziell $u \in \mathbb{N}$, so heißt v *Mutter* von w sowie umgekehrt w *Kind* von v.

(c) Falls $\phi(v, w) \overset{\text{def}}{=} \inf\{k \geq 1 : v_k \neq w_k\}$, so heißt

$$v \wedge w \overset{\text{def}}{=} v_1...v_{\phi(u,v)-1}$$

der *erste gemeinsame Vorfahre* von v und w.

(d) Wir definieren im Fall $v \neq w$

$$v < w \quad \overset{\text{def}}{\Longleftrightarrow} \quad \begin{cases} |v| < |w|, & \text{falls } |v| \neq |w|, \\ v_{\phi(v,w)} < w_{\phi(v,w)}, & \text{falls } |v| = |w|, \end{cases}$$

und anschließend allgemein

$$v \leq w \quad \overset{\text{def}}{\Longleftrightarrow} \quad v = w \text{ oder } v < w.$$

Mit diesen Festsetzungen gilt $\varnothing \preceq v$ und $\varnothing \leq v$ für alle $v \in \mathbb{V}$.

Die in (d) eingeführte Relation "$\le$" bildet offenkundig eine Ordnung auf $\mathbb{V}$, die bei Einschränkung auf $\mathbb{N} \subset \mathbb{V}$ mit der geöhnlichen Ordnung übereinstimmt. Je zwei Elemente $v, w \in \mathbb{V}$ sind demnach stets vergleichbar, und jede endliche Teilmenge von $\mathbb{V}$ besitzt ein Minimum und ein Maximum. Für zweielementige Teilmengen $\{v, w\}$ entspricht das Minimum genau dem in (c) definierte ersten gemeinsamen Vorfahren $v \wedge w$, was die gewählte Notation "$\wedge$" erklärt. Abschließend erwähnen wir noch, daß man anschaulich "$v \le w$" als "v ist älter als w" interpretieren kann, sofern man eine gewisse Festlegung der Geburtsreihenfolge der Individuen von $\mathbb{V}$ unterstellt, die sich der Leser selbst überlegen mag (☞ hierzu auch Übung 1.3).

Übungen zu Abschnitt 1

Übung 1.1. Zeigen Sie, daß der metrische Raum $(\mathbb{T}, d)$ vollständig ist.

Übung 1.2. Zeigen Sie, daß neben d auch durch

$$d_2(\tau, \tau') \stackrel{\text{def}}{=} \frac{1}{1 + \sup\{n \ge 0 : \tau_{|n} = \tau'_{|n}\}}$$

eine Metrik auf $\mathbb{T}$ definiert wird, welche dieselbe Topologie und damit dieselbe Borelsche σ-Algebra induziert.

Übung 1.3. Es seien $v, w \in \mathbb{V}$ und $\tau \in \mathbb{T}^e$ ein endlicher Baum.

(a) Es bezeichne $v \vee w$ das Maximum von v, w bezüglich der in Definition 1.5(d) eingeführten Ordnung. Geben Sie eine anschauliche Charakterisierung von $v \vee w$.

(b) Geben Sie eine anschauliche Charakterisierung des minimalen und maximalen Elements von τ.

2. Der Galton-Watson-Baum: Formalitäten und Eigenschaften

Nach den Vorbereitungen im letzten Abschnitt sind wir nun in der Lage, den bereits in Abschnitt I.1.1 vorgestellten GWB $\boldsymbol{GW}$ zu einer gegebenen Reproduktionsverteilung $(p_j)_{j \ge 0}$ als Zufallselement in $(\mathbb{T}, \mathfrak{B}(\mathbb{T}))$ zu definieren. Dies läßt sich in naheliegender Weise vermöge des folgenden *Standardmodells* realisieren (vgl. die Ausführungen nach Lemma I.1.2):

Sei $\{X_v : v \in \mathbb{V}\}$ eine Familie unabhängiger, identisch nach $(p_j)_{j \ge 0}$ verteilter Zufallsgrößen auf einem W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$ und dann $\boldsymbol{GW} \stackrel{\text{def}}{=} \bigcup_{n \ge 0} \boldsymbol{GW}_n$ mit $\boldsymbol{GW}_0 \stackrel{\text{def}}{=} \{\varnothing\}$ sowie

$$\boldsymbol{GW}_n \stackrel{\text{def}}{=} \left\{v_1 ... v_n \in \mathbb{N}^n : v_1 ... v_{n-1} \in \boldsymbol{GW}_{n-1} \text{ und } 1 \le v_n \le X_{v_1 ... v_{n-1}}\right\} \tag{2.1}$$

für $n \ge 1$ (beachte die übliche Konvention $v_1 ... v_{n-1} \stackrel{\text{def}}{=} \varnothing$ im Fall $n = 1$). Dann ist $\boldsymbol{GW}$ offensichtlich eine $\mathbb{T}$-wertige Abbildung. Wir setzen ferner (mit $z_n : \mathbb{T} \to \mathbb{N}_0$ gemäß Definition 1.1)

$$Z_n \stackrel{\text{def}}{=} |\boldsymbol{GW}_n| = z_n \circ \boldsymbol{GW} \tag{2.2}$$

für $n \geq 0$.

2.1. Lemma. *Für die zuvor definierten Abbildungen gilt:*

(a) $\boldsymbol{GW} : \Omega \to \mathbb{T}$ *ist* $\mathfrak{A}$-$\mathfrak{B}(\mathbb{T})$*-meßbar und folglich ein Zufallselement auf* $(\Omega, \mathfrak{A}, \mathbb{P})$.

(b) Für jedes $n \geq 0$ *ist* $z_n : \mathbb{T} \to \mathbb{N}_0$ $\mathfrak{B}(\mathbb{T})$*-meßbar und somit* $Z_n = z_n \circ \boldsymbol{GW}$ *eine Zufallgröße auf* $(\Omega, \mathfrak{A}, \mathbb{P})$.

Beweis: Es ist lediglich $\boldsymbol{GW}^{-1}(\mathcal{E}) \subset \mathfrak{A}$ zu zeigen. Für $A = [\tau]_k$, $\tau \in \mathbb{T}$ und $k \geq 1$, gilt aber

$$\begin{aligned} \boldsymbol{GW}^{-1}(A) \;&=\; \{\omega \in \Omega : \boldsymbol{GW}_{|k}(\omega) = \tau_{|k}\} \\ &=\; \bigcap_{j=1}^{k} \{\omega \in \Omega : \boldsymbol{GW}_j(\omega) = \tau_j\} \\ &\in\; \sigma\big(\{X_v : |v| \leq k-1\}\big) \;\subset\; \mathfrak{A}. \end{aligned}$$

(b) Hier ergibt sich für $k \geq 1$

$$z_n^{-1}(\{k\}) \;=\; \{\tau \in \mathbb{T} : z_n(\tau) = k\} \;=\; \bigcup_{\tau \in \mathbb{T}_n : \tau_n = k} [\tau]_n \;\in\; \mathfrak{B}(\mathbb{T})$$

weil die auftretende Vereinigung abzählbar ist. Dies zeigt die Meßbarkeit von z_n. $\diamondsuit$

Nach diesen Vorbereitungen ist folgende Definition sinnvoll:

2.2. Definition. Wir setzen $\mathbf{Q} \overset{\text{def}}{=} \mathbb{P}(\boldsymbol{GW} \in \cdot)$ und bezeichnen dieses Bildmaß als das *Galton-Watson-Maß* auf $(\mathbb{T}, \mathfrak{B}(\mathbb{T}))$.

Es gilt nun für alle $n \geq 0$

$$\mathbb{P}(Z_n \in \cdot) \;=\; \mathbb{P}(z_n \circ \boldsymbol{GW} \in \cdot) \;=\; \mathbf{Q}(z_n \in \cdot) \tag{2.3}$$

sowie allgemeiner

$$\mathbb{P}((Z_n)_{n\geq 0} \in \cdot) \;=\; \mathbb{P}((z_n \circ \boldsymbol{GW})_{n\geq 0} \in \cdot) \;=\; \mathbf{Q}((z_n)_{n\geq 0} \in \cdot). \tag{2.4}$$

Die letzte Beziehung zeigt, daß wir einen GWP nunmehr auch als stochastische Folge auf dem W-Raum $(\mathbb{T}, \mathfrak{B}(\mathbb{T}), \mathbf{Q})$ realisieren können.

Als nächstes wollen wir das intuitiv auf der Hand liegende Resultat präzisieren und beweisen, daß Individuen derselben Generation einer Galton-Watson-Population vermöge ihres unabhängigen und identischen Reproduktionsverhaltens Teilbäume erzeugen, die unabhängig und wiederum nach $\mathbf{Q}$ verteilt sind. Hierfür bedarf es aber zunächst einiger weiterer Definitionen:

Gegeben $\tau \in \mathbb{T}$ und $u \in \tau$, heißt

$$\tau^u \;\overset{\text{def}}{=}\; \{v \in \tau : v \succeq u\}$$

der *in u verwurzelte Teilbaum von τ*. Er enthält alle Individuen in τ, die von u abstammen (einschließlich u selbst). Um aus dieser Teilmenge von τ wiederum einen Baum gemäß Definition 1.1 zu machen, definieren wir den *Shiftoperator* Θ_u zu u durch

$$\Theta_u(uv) \stackrel{\text{def}}{=} v, \quad v \in \mathbb{V},$$

speziell also $\Theta_u u = \varnothing$. Dann genügt $\Theta_u(\tau^u)$ den Bedingungen aus Definition 1.1. Im Prinzip wird der in u verwurzelte Teilbaum lediglich so verschoben, daß er $\varnothing$ als Wurzel besitzt. Mit anderen Worten, τ^u entspricht bis auf diese Verschiebung einem Baum im Sinne unserer Definition. Durch eine ähnliche Argumentation wie im Beweis von Lemma 2.1(a) erhält man im Fall $u \in \boldsymbol{GW}$ die $\mathfrak{A}$-$\mathfrak{B}(\mathbb{T})$-Meßbarkeit der Abbildung $\Theta_u \circ \boldsymbol{GW}^u$ (☞ Übung 2.1), und in der Tat läßt sich nun die oben gemachte Aussage hinsichtlich der Unabhängigkeit und identischen Verteilung der Teilbäume formal bestätigen:

2.3. Satz. *Für einen GWB $\boldsymbol{GW}$ mit Galton-Watson-Maß $\mathbf{Q}$, Reproduktionsverteilung $(p_j)_{j\ge 0}$ und assoziiertem GWP $(Z_n)_{n\ge 0}$ gilt:*

(a) *Ist $p_k > 0$ für ein $k \ge 1$, so sind, gegeben $Z_1 = k$, die von den Individuen der ersten Generation generierten (geshifteten) Teilbäume $\Theta_i \boldsymbol{GW}^i$, $1 \le i \le k$, bedingt stochastisch unabhängig mit Verteilung $\mathbf{Q}$.*

(b) *Für $n \ge 2$ gilt entsprechend: Ist $\mathbb{P}(Z_n = k) > 0$, so sind, gegeben $Z_n = k$, die von den Individuen der n-ten Generation generierten (geshifteten) Teilbäume $\Theta_u \boldsymbol{GW}^u$, $u \in \boldsymbol{GW}_n$, bedingt stochastisch unabhängig mit Verteilung $\mathbf{Q}$.*

Beweis: Wir beweisen nur Teil (a), da der Beweis von (b) ähnlich, wenn auch schreibaufwendiger ist. Seien $A_1, ..., A_k \in \mathcal{E}$ und o.B.d.A. alle nichtleer. Dann existiert eine Familie $\{B_v^{(i)} : v \in \mathbb{V}, 1 \le i \le k\}$ von Teilmengen von $\mathbb{N}_0$, so daß die Gleichungskette

$$\begin{aligned}
\prod_{i=1}^k \mathbf{Q}(A_i) &= \prod_{i=1}^k \prod_{v\in\mathbb{V}} \mathbb{P}(X_v \in B_v^{(i)}) \\
&= p_k^{-1} \prod_{i=1}^k \prod_{v\in\mathbb{V}} \mathbb{P}(X_{iv} \in B_v^{(i)}) \mathbb{P}(X_\varnothing = k) \\
&= p_k^{-1} \mathbb{P}\left(\{Z_1 = k\} \cap \bigcap_{i=1}^k \bigcap_{v\in\mathbb{V}} \{X_{iv} \in B_v^{(i)}\} \right) \\
&= \frac{\mathbb{P}(Z_1 = k, \Theta_i \boldsymbol{GW}^i \in A_i, 1 \le i \le k)}{\mathbb{P}(Z_1 = k)} \\
&= \mathbb{P}(\Theta_i \boldsymbol{GW}^i \in A_i, 1 \le i \le k)
\end{aligned}$$

gültig ist, wobei wir die Unabhängigkeit und identische Verteilung der Zufallsgrößen X_v, $v \in \mathbb{V}$, benutzt haben. Setzt man nun $A_i = \mathbb{T}$ für geeignete $i \in \{1, ..., k\}$, folgt zunächst die Verteilungsaussage und dann im Hinblick auf die obige Rechnung und unter Hinweis auf Satz 23.2 in [1] auch die Unabhängigkeit, da $\mathcal{E}$ $\cap$-stabil ist mit $\mathbb{T} \in \mathcal{E}$. ◊

Übungen zu Abschnitt 2

Übung 2.1. Zeigen Sie die $\mathfrak{A}$-$\mathfrak{B}(\mathbb{T})$-Meßbarkeit des geshifteten Teilbaums $\Theta_u \circ \boldsymbol{GW}^u$ für jedes $u \in \boldsymbol{GW}$.

3. Größenverzerrte Galton-Watson-Bäume mit Rückgrat

Ein fundamentales Werkzeug für unseren neuen Zugang zu einigen der klassischen Grenzwertsätze für GWP bilden die sogenannten *größenverzerrten GWB mit Rückgrat*, die wir im folgenden konstruieren. Dabei handelt es sich um zufällige unendliche, aber lokal endliche Bäume $\widehat{\boldsymbol{T}}$, in denen ein zufälliger aufsteigender Pfad $\boldsymbol{V}$, das sogenannte *Rückgrat* (engl. *spine*), besonders ausgezeichnet ist. Durch Vergleich der Verteilung dieses Hilfsbaums $\widehat{\boldsymbol{T}}$ mit dem Galton-Watson-Maß $\mathbf{Q}$ werden wir später neue Beweise für bekannte Grenzwertsätze geben können.

1. Größenverzerrte Verteilungen und Zufallsgrößen. Da größenverzerrte Zufallsgrößen einen wichtigen Baustein zur oben angekündigten Konstruktion bilden, werden diese zunächst formal eingeführt:

3.1. Definition. Es sei ν eine Verteilung auf $([0,\infty),\mathfrak{B}_{[0,\infty)})$ mit endlichem, positivem Erwartungswert $\gamma \stackrel{\text{def}}{=} \int_{[0,\infty)} x\,\nu(dx)$. Dann heißt $\widehat{\nu}$, definiert durch

$$\widehat{\nu}(B) \stackrel{\text{def}}{=} \frac{1}{\gamma}\int_B x\ \nu(dx), \quad B \in \mathfrak{B}_{[0,\infty)}, \tag{3.1}$$

die *zu ν gehörende größenverzerrte (engl. size-biased) Verteilung* auf $[0,\infty)$ oder auch *Größenverzerrung von ν*. Ist X eine Zufallsgröße mit Verteilung ν, so heißt die Zufallsgröße $\widehat{X}$ eine *Größenverzerrung von X*, falls $\widehat{X}$ die Verteilung $\widehat{\nu}$ besitzt.

3.2. Bemerkungen. (a) Wir werden hiernach hauptsächlich Größenverzerrungen von Verteilungen auf $\mathbb{N}_0$ betrachten. Ist $\nu = (\nu_k)_{k\ge 0}$ eine solche, so gilt offenkundig $\widehat{\nu} = (\widehat{\nu}_k)_{k\ge 0}$ mit

$$\widehat{\nu}_k = \gamma^{-1}k\nu_k, \quad k \ge 0,$$

wobei $\gamma = \sum_{k\ge 1} k\nu_k$ wie in 3.1 den Erwartungswert von ν angibt.

(b) Ist $\widehat{X}$ die Größenverzerrung einer Zufallsgröße X mit Verteilung ν, so besitzt $\widehat{\nu} = \mathbb{P}(\widehat{X} \in \cdot)$ die ν-Dichte

$$\frac{d\widehat{\nu}}{d\nu}(x) = \frac{x}{\gamma}\mathbf{1}_{[0,\infty)}(x) = \frac{x}{\mathbb{E}X}\mathbf{1}_{[0,\infty)}(x). \tag{3.3}$$

Deshalb ergibt sich für jede meßbare Funktion $\varphi : [0,\infty) \to [0,\infty)$ die für uns wichtige Rechenregel

$$\begin{aligned} \mathbb{E}\varphi(\widehat{X}) &= \int_{[0,\infty)} \varphi(x)\ \widehat{\nu}(dx) \\ &= \int_{[0,\infty)} \varphi(x)\,\frac{x}{\mathbb{E}X}\ \nu(dx) = \frac{\mathbb{E}X\varphi(X)}{\mathbb{E}X}. \end{aligned} \tag{3.4}$$

Ferner notieren wir, daß $\widehat{X}$ fast sicher positiv ist.

(c) Besitzt ν eine $\lambda\!\!\lambda$-Dichte g, so gilt natürlich dasselbe für $\widehat{\nu}$, und wir erhalten mit der Produktregel für Radon-Nikodym-Ableitungen

$$\widehat{g}(x) \stackrel{\text{def}}{=} \frac{d\widehat{\nu}}{d\lambda\!\!\lambda}(x) = \frac{d\widehat{\nu}}{d\nu}(x)\,\frac{d\nu}{d\lambda\!\!\lambda}(x) = \gamma^{-1}xg(x), \quad x \geq 0. \tag{3.5}$$

2. Konstruktion und Eigenschaften größenverzerrter GWB. Das Konzept der Größenverzerrung nichtnegativer Zufallsgrößen soll nun in geeigneter Weise auf einen GWB $\boldsymbol{GW}$ übertragen und eine Größenverzerrung $\widehat{\boldsymbol{GW}}$ von $\boldsymbol{GW}$ konstruiert werden. Später werden wir im Zusammenhang mit einem neuen Beweis des Satzes von Kesten und Stigum (Satz I.6.2) der Frage nachgehen, ob bzw. unter welchen Bedingungen – in Analogie zu (3.3) – die mit $\widehat{\mathbf{Q}}$ bezeichnete Verteilung von $\widehat{\boldsymbol{GW}}$ durch das assoziierte Galton-Watson-Maß $\mathbf{Q}$ dominiert wird. Ist dies nicht der Fall, sind die beiden Maße notwendig zueinander singulär (☞ Lemma 5.1 und Satz 5.2).

Zunächst schreiten wir jedoch zur angekündigten Konstruktion, wobei wir von einer Reproduktionverteilung $(p_j)_{j\geq 0}$ mit Reproduktionmittel $\mu \in (0,\infty)$ ausgehen.

3.3. Konstruktion von $\widehat{\boldsymbol{GW}}$. Wir erinnern daran, daß die Familie $\{X_v : v \in \mathbb{V}\}$ den GWB $\boldsymbol{GW}$ bestimmt, wobei X_v die Anzahl der Kinder des (potentiellen) Individuums v angibt. Auf $(\Omega, \mathfrak{A}, \mathbb{P})$ sei nun darüber hinaus eine Familie $\{(\widehat{X}_n, C_n) : n \geq 0\}$ von unabhängigen Zufallsvektoren gegeben, die von $\{X_v : v \in \mathbb{V}\}$ unabhängig ist. Für $n \geq 0$ gelte

$$\mathbb{P}(\widehat{X}_n \in \cdot) = (\widehat{p}_j)_{j\geq 0} = (\mu^{-1}jp_j)_{j\geq 0}$$

sowie im Fall $p_k > 0$ für $k \geq m \geq 1$

$$\mathbb{P}(C_n = m|\widehat{X}_n = k) = \frac{1}{k},$$

d.h. gegeben $\widehat{X}_n = k$ sei C_n Laplace-verteilt auf $\{1, ..., k\}$, und $\widehat{X}_n$ besitze die größenverzerrte Reproduktionsverteilung $(\widehat{p}_j)_{j\geq 0}$.

Der *größenverzerrte GWB* $\widehat{\boldsymbol{GW}} = \bigcup_{n\geq 0} \widehat{\boldsymbol{GW}}_n$ *mit Rückgrat* $\boldsymbol{V} = (V^n)_{n\geq 0}$ wird damit wie folgt definiert: Wir setzen $\widehat{\boldsymbol{GW}}_0 \stackrel{\text{def}}{=} \{\varnothing\}$, $V^0 \stackrel{\text{def}}{=} \varnothing$ sowie induktiv für $n \geq 1$

$$\widehat{\boldsymbol{GW}}_n \stackrel{\text{def}}{=} \boldsymbol{A}_n \cup \boldsymbol{B}_n$$

mit

$$\boldsymbol{A}_n \stackrel{\text{def}}{=} \left\{vi \in \mathbb{N}^n : v \in \widehat{\boldsymbol{GW}}_{n-1}\backslash\{V^{n-1}\},\, i \leq X_v\right\}$$

und

$$\boldsymbol{B}_n \stackrel{\text{def}}{=} \left\{V^{n-1}k :\, 1 \leq k \leq \widehat{X}_{n-1}\right\}$$

sowie

$$V^n \stackrel{\text{def}}{=} V^{n-1}C_{n-1}.$$

Beachte hierbei die Konvention, daß $V^{n-1}k$ ebenso wie $V^{n-1}C_{n-1}$ als Konkatenationen (und nicht als Produkte) zu lesen sind, also als (V^{n-1}, k) bzw. (V^{n-1}, C_{n-1}).

In Worten läßt sich die Konstruktion wie folgt beschreiben: Wir starten wieder mit $\varnothing$ als einzigem Individuum (Urahnen) in der 0-ten Generation, das zugleich die 0-te Komponente V^0 des Rückgrats bildet. Dieses Individuum bekommt eine zufällige Anzahl $\widehat{X}_0 \geq 1$ von Kindern. Aus diesen wird zufällig gemäß einer entsprechenden Laplace-Verteilung (vermöge C_0) die erste Komponente V^1 des Rückgrats ausgewählt. Dieses Individuum V^1 reproduziert sich gemäß der größenverzerrten Reproduktionsverteilung $(\widehat{p}_j)_{j\geq 0}$ und generiert $\widehat{X}_1$ Nachfahren, während die restlichen Individuen der ersten Generation $\widehat{\boldsymbol{GW}}_1$ voneinander unabhängig Kinder gemäß der Reproduktionsverteilung $(p_j)_{j\geq 0}$ hervorbringen. So fortfahrend, wird V^n zufällig aus den $\widehat{X}_{n-1}$ Kindern von V^{n-1} ausgewählt und produziert Nachfahren gemäß $(\widehat{p}_j)_{j\geq 0}$; alle Individuen, die nicht zu $\boldsymbol{V}$ gehören, erzeugen Kinder wie im gewöhnlichen GWB. Insgesamt erhalten wir somit einen zufälligen Baum $\widehat{\boldsymbol{GW}}$, in dem eine aufsteigende Folge von Knoten $(V_n)_{n\geq 0}$, das *Rückgrat*, besonders ausgezeichnet ist. Der Terminus "aufsteigend" bedeutet hier, daß in jeder Generation genau ein Individuum ausgewählt wird. Die untere Abbildung zeigt einen Ausschnitt aus einem größenverzerrten GWB, in dem die Rückgratkomponenten (Wirbel), nämlich $V^0 = \varnothing, V^1 = 3$, $V^2 = 31$ und $V^3 = 312$, als schwarze Kästchen (■) und die übrigen Knoten als graue Kreise mit anhängenden gewöhnlichen GWB dargestellt sind.

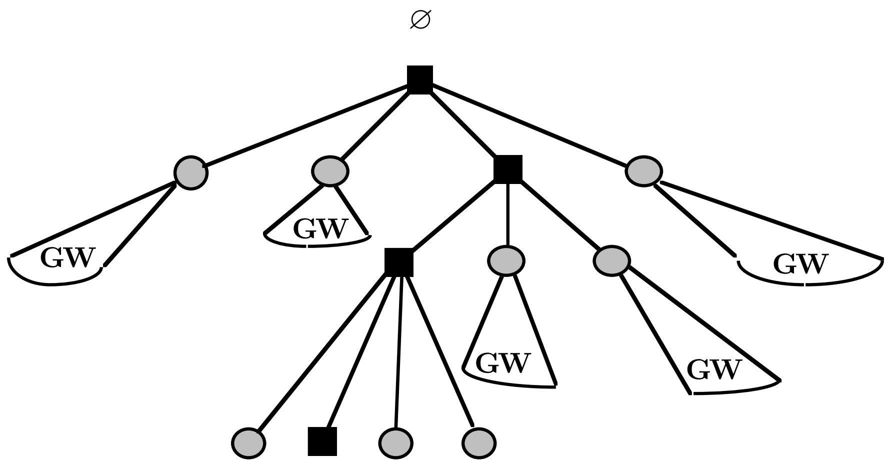

Bild 3.1. Ausschnitt aus einem größenverzerrten GWB

Wegen $\mathbb{P}(\widehat{X}_n = 0) = 0$ für $n \geq 0$ gilt offenbar $|\widehat{\boldsymbol{GW}}| = \infty$ f.s., die durch $\widehat{\boldsymbol{GW}}$ beschriebene Population kann also im Gegensatz zu der eines gewöhnlichen GWB's niemals aussterben.

Die obige Konstruktion liefert uns eine Abbildung

$$(\widehat{\boldsymbol{GW}}, \boldsymbol{V}) : \Omega \;\to\; \mathbb{T} \times \partial\mathbb{V}$$

mit

$$\partial\mathbb{V} \stackrel{\text{def}}{=} \big\{(v^j)_{j\geq 0} : v^j \in \mathbb{N}^j \text{ und } v_{j+1} \succeq v_j \text{ für } j \geq 0\big\}.$$

Offenbar können wir die Menge $\partial\mathbb{V}$ der unendlichen Pfade $(v^j)_{j\geq 0}$ in $\mathbb{V}$ mit der Menge $\mathbb{N}^\infty$ identifizieren vermöge der Zuordnung

$$(v^j)_{j\geq 0} \quad \longleftrightarrow \quad (v_j)_{j\geq 0},$$

wobei $v^j \stackrel{\text{def}}{=} v_1...v_j$ für jedes $j \geq 1$. Dies werden wir im folgenden, wann immer hilfreich, tun.

Wie schon für GWB bedarf es als nächstes der Angabe einer geeigneten σ-Algebra über $\mathbb{T}\times\partial\mathbb{V}$, so daß $(\widehat{\boldsymbol{GW}}, \boldsymbol{V})$ zu einer meßbaren Abbildung wird. Zu diesem Zweck betrachten wir den erweiterten Raum $\overline{\mathbb{V}} \stackrel{\text{def}}{=} \mathbb{V}\cup\partial\mathbb{V} = \bigcup_{n\in\overline{\mathbb{N}}_0}\mathbb{N}^n$ und geben zuerst das folgende Lemma, das sehr ähnlich zu Lemma 1.2 bewiesen werden kann. Wir weisen darauf hin, daß die in Definition 1.5 eingeführte Ordnung auf $\mathbb{V}$ und die resultierende Minimumsrelation $\wedge$ in offensichtlicher Weise auf $\overline{\mathbb{V}}$ ausgedehnt werden können.

3.4. Lemma. *Die Abbildung $d' : \overline{\mathbb{V}}\times\overline{\mathbb{V}} \to [0,1]$, gegeben durch*

$$d'(v,w) \stackrel{\text{def}}{=} e^{-|v\wedge w|}, \tag{3.6}$$

ist eine Metrik auf $\overline{\mathbb{V}}$, bezüglich der $\mathbb{V}$ eine abzählbar dichte und zugleich offene Teilmenge bildet. Der Raum $(\overline{\mathbb{V}}, d')$ ist ferner kompakt und $\partial\mathbb{V}$ der topologische Rand von $\mathbb{V}$ [1)]

Beweis: Übung 3.3. ◇

In Analogie zur Konstruktion von $\mathfrak{B}(\mathbb{T})$ setzen wir nun $\mathbb{B}'(v,\varepsilon) \stackrel{\text{def}}{=} \{w \in \overline{\mathbb{V}} : d'(v,w) < \varepsilon\}$ für $v \in \overline{\mathbb{V}}$, $\varepsilon > 0$ und dann

$$\mathfrak{B}(\overline{\mathbb{V}}) \stackrel{\text{def}}{=} \sigma\Big(\big\{\mathbb{B}'(v,\varepsilon) : v \in \overline{\mathbb{V}}, \varepsilon > 0\big\}\Big).$$

Wegen der Separabilität von $(\overline{\mathbb{V}}, d')$ ist dies wiederum die Borelsche σ-Algebra über $\overline{\mathbb{V}}$ mit dem abzählbaren Erzeuger (☞ Übung 3.4)

$$\mathcal{E}' \stackrel{\text{def}}{=} \big\{\mathbb{B}'(v,\varepsilon) : v \in \mathbb{V}, \varepsilon > 0\big\}.$$

Die gute Nachricht lautet dann:

3.5. Lemma. *Die Abbildung $(\widehat{\boldsymbol{GW}}, \boldsymbol{V})$ ist $\mathfrak{A}$-$\mathfrak{B}(\mathbb{T})\otimes\mathfrak{B}(\overline{\mathbb{V}})$-meßbar.*

Beweis: Unter Rückgriff auf Satz 19.3 in [1] reicht es zu zeigen, daß

(i) $\widehat{\boldsymbol{GW}}$ $\mathfrak{A}$-$\mathfrak{B}(\mathbb{T})$-meßbar und
(ii) $\boldsymbol{V}$ $\mathfrak{A}$-$\mathfrak{B}(\overline{\mathbb{V}})$-meßbar

ist, wobei der Nachweis von (i) wie der Beweis von Lemma 2.1(a) erfolgt: Gegeben $\tau\in\mathbb{T}$ und $k \geq 1$, erhalten wir

$$\widehat{\boldsymbol{GW}}^{-1}([\tau]_k) \;=\; \Big\{\omega\in\Omega : \widehat{\boldsymbol{GW}}_{|k}(\omega) = \tau_{|k}\Big\} \;\in\; \sigma\Big((\widehat{X}_n, C_n)_{0\leq n<k}, (X_v)_{|v|<k}\Big) \;\subset\; \mathfrak{A}$$

---

1) Dies rechtfertigt die Bezeichnung $\partial\mathbb{V}$.

und damit die behauptete Meßbarkeit von $\widehat{\boldsymbol{GW}}$.

Zum Nachweis von (ii) gehen wir ähnlich vor: Seien $v = v_1v_2... \in \mathbb{N}^\infty$, $v^0 \stackrel{\text{def}}{=} \varnothing$, $v^n \stackrel{\text{def}}{=} v_1...v_n$ für $n \geq 1$ und $0 < \varepsilon < 1$, o.E. $\varepsilon = e^{-m}$ für ein $m \geq 1$. Dann gilt

$$\boldsymbol{V}^{-1}(\mathbb{B}'(v,\varepsilon)) \;=\; \Big\{\omega \in \Omega : V^m(\omega) = v^m\Big\} \;\in\; \sigma\Big((V^k, C_k)_{0\leq k<m}\Big) \;\subset\; \mathfrak{A},$$

also wiederum das Verlangte. $\diamondsuit$

Wir setzen nun

$$\widehat{\mathbf{Q}}_* \;\stackrel{\text{def}}{=}\; \mathbb{P}\big((\widehat{\boldsymbol{GW}}, \boldsymbol{V}) \in \cdot\big), \quad \text{und} \quad \widehat{\mathbf{Q}} \;\stackrel{\text{def}}{=}\; \mathbb{P}\big(\widehat{\boldsymbol{GW}} \in \cdot\big)$$

und notieren das folgende Pendant zu Satz 2.3:

3.6. Satz. *Unter den Voraussetzungen von 3.3 und mit dessen Bezeichnungen gilt:*

(a) Ist $p_k > 0$ für ein $k \geq 1$, so sind, gegeben $(\widehat{X}_0, C_0) = (k, i)$ mit $i \in \{1, ..., k\}$, die Abbildungen $(\Theta_i \widehat{\boldsymbol{GW}}^i, \Theta_i \boldsymbol{V})$ und $\Theta_j \widehat{\boldsymbol{GW}}^j$, $1 \leq j \leq k, j \neq i$, meßbar und bedingt stochastisch unabhängig mit Verteilung $\widehat{\mathbf{Q}}_*$ bzw. $\mathbf{Q}$.

(b) Für $n \geq 2$ gilt entsprechend: Ist $\mathbb{P}(\widehat{\boldsymbol{GW}}_n = I_n, V^n = v) > 0$ für ein $I_n \subset \mathbb{N}^n$ und $v \in I_n$, so sind, gegeben $(\widehat{\boldsymbol{GW}}_n, V^n) = (I_n, v)$, die Abbildungen $(\Theta_v \widehat{\boldsymbol{GW}}^v, \Theta_v \boldsymbol{V})$ und $\Theta_u \widehat{\boldsymbol{GW}}^u$, $u \in I_n, u \neq v$, meßbar und bedingt stochastisch unabhängig mit Verteilung $\widehat{\mathbf{Q}}_*$ bzw. $\mathbf{Q}$.

Beweis: Übung 3.5. $\diamondsuit$

Der Zusammenhang zwischen den W-Maßen $\mathbf{Q}$ und $\widehat{\mathbf{Q}}$ bildet den Schlüssel für die im weiteren Verlauf vorgestellten Techniken und Inhalt des anschließenden Lemmas. Für $\tau \in \mathbb{T}$, $n \geq 0$ und $\varrho \in \tau_n$ definieren wir noch

$$[\tau; \varrho]_n \;\stackrel{\text{def}}{=}\; \big\{(\tau', (v^j)_{j\geq 0}) \in \mathbb{T} \times \partial\mathbb{V} : \tau' \in [\tau]_n, v^n = \varrho\big\}.$$

Diese Menge besteht aus allen Tupeln $(\tau', (v^j)_{j\geq 0}) \in \mathbb{T} \times \partial\mathbb{V}$, so daß der Baum τ' bis zur n-ten Generation mit τ übereinstimmt und der Pfad $(v^j)_{j\geq 0}$ durch ϱ verläuft.

3.7. Vergleichslemma. *Sei $w_n(\tau) \stackrel{\text{def}}{=} \mu^{-n} z_n(\tau)$ für $\tau \in \mathbb{T}$ und $n \geq 0$.*

(a) Für alle $\tau \in \mathbb{T}$ und $n \geq 0$ gilt

$$\mathbf{Q}\big([\tau]_{n+1}\big) \;=\; p_{z_1(\tau)} \prod_{j=1}^{z_1(\tau)} \mathbf{Q}\big([\Theta_j \tau^j]_n\big). \tag{3.7}$$

(b) Für alle $\tau \in \mathbb{T}$, $n \geq 0$ und $\varrho \in \tau_n$ gilt

$$\widehat{\mathbf{Q}}_*\big([\tau; \varrho]_n\big) \;=\; \mu^{-n} \mathbf{Q}\big([\tau]_n\big). \tag{3.8}$$

(c) *Für alle $\tau \in \mathbb{T}$ und $n \geq 0$ gilt*

$$\widehat{\mathbf{Q}}\big([\tau]_n\big) \;=\; w_n(\tau)\mathbf{Q}\big([\tau]_n\big). \tag{3.9}$$

(d) *Sei $\widehat{Z}_n \overset{\text{def}}{=} z_n \circ \widehat{\boldsymbol{GW}}$ für $n \geq 0$. Dann ist $\widehat{Z}_n$ eine Größenverzerrung von Z_n, d.h. es gilt*

$$\mathbb{P}(\widehat{Z}_n = k) \;=\; \frac{k\,\mathbb{P}(Z_n = k)}{\mu^n} \;=\; \frac{k\,\mathbb{P}(Z_n = k)}{\mathbb{E}Z_n} \tag{3.10}$$

für alle $k \geq 0$.

Beweis: (a) Sei $\tau \in \mathbb{T}$, $n \geq 0$ und $k = z_1(\tau)$. O.E. werde $k \geq 1$ und $p_k > 0$ vorausgesetzt, denn im Fall $k = 0$ ist (3.7) wegen $[\tau]_{n+1} = \big\{\{\varnothing\}\big\}$ ebenso wie im Fall $p_k = 0$ trivial. Dann gilt unter Verwendung von Satz 2.3

$$\begin{aligned}
\mathbf{Q}\big([\tau]_{n+1}\big) &= \mathbb{P}\big(Z_1 = k, \Theta_i\,\boldsymbol{GW}^i \in [\Theta_i\tau^i]_n, 1 \leq i \leq k\big)\\
&= p_k\,\mathbb{P}\big(\Theta_i\,\boldsymbol{GW}^i \in [\Theta_i\tau^i]_n, 1 \leq i \leq k\big|Z_1 = k\big)\\
&= p_k \prod_{i=1}^{k} \mathbf{Q}\big([\Theta_i\tau^i]_n\big)
\end{aligned}$$

und somit das Gewünschte.

(b) Hier zeigen wir die Behauptung per Induktion über n. Falls $n = 0$, erhalten wir wegen $V^0 = \varrho = \varnothing$, $[\tau;\varnothing]_0 = \{(\tau', (v^j)_{j\geq 0}) : \tau \in \mathbb{T}, v^j \in \tau_j \text{ f.a. } j \geq 0\}$ und $[\tau]_0 = \mathbb{T}$ für alle $\tau \in \mathbb{T}$ offenkundig

$$\widehat{\mathbf{Q}}_*\big([\tau;\varrho]_0\big) \;=\; \widehat{\mathbf{Q}}_*\big([\tau;\varnothing]_0\big) \;=\; 1 \;=\; \mathbf{Q}\big([\tau]_0\big)$$

für alle $\tau \in \mathbb{T}$.

Für den Induktionsschritt nehmen wir an, die Behauptung sei bewiesen für ein $n \geq 0$, alle $\tau' \in \mathbb{T}$ und $\varrho' \in \tau'_n$. Wir benötigen zunächst die folgende Zwischenrechnung: Gegeben $\tau \in \mathbb{T}$ mit $z_1(\tau) = k \geq 1$ und $\varrho \in \tau_{n+1}$, existiert offenkundig ein eindeutiges $i \in \{1, ..., k\}$ derart, daß $\varrho \in \tau^i$. Daher gilt mit $A \overset{\text{def}}{=} \{\Theta_j\,\widehat{\boldsymbol{GW}}^j \in [\Theta_j\tau^j]_n, 1 \leq j \leq k, j \neq i\}$

$$\big\{\big(\widehat{\boldsymbol{GW}}, \boldsymbol{V}\big) \in [\tau;\varrho]_{n+1}\big\} \;=\; \big\{\widehat{X}_0 = k, C_0 = i, (\Theta_i\,\widehat{\boldsymbol{GW}}^i, \Theta_i\,\boldsymbol{V}) \in [\Theta_i\tau^i;\Theta_i\varrho]_n\big\} \cap A.$$

Sofern $p_k > 0$, was $\mathbb{P}(\widehat{X}_0 = k, C_0 = i) = \mu^{-1}kp_k\,k^{-1} = \mu^{-1}p_k > 0$ impliziert, ergibt sich unter Verwendung von Satz 3.6

$$\begin{aligned}
\widehat{\mathbf{Q}}_*\big([\tau;\varrho]_{n+1}\big) &= \mathbb{P}\big(\big(\widehat{\boldsymbol{GW}}, \boldsymbol{V}\big) \in [\tau;\varrho]_{n+1}\big)\\
&= \mu^{-1}p_k\,\mathbb{P}\big(\big\{(\Theta_i\,\widehat{\boldsymbol{GW}}^i, \Theta_i\,\boldsymbol{V}) \in [\Theta_i\tau^i;\Theta_i\varrho]_n\big\} \cap A\big|\widehat{X}_0 = k, C_0 = i\big)\\
&= \mu^{-1}p_k\,\widehat{\mathbf{Q}}_*\big([\Theta_i\tau^i;\Theta_i\varrho]_n\big) \prod_{j\neq i} \mathbf{Q}\big([\Theta_j\tau^j]_n\big).
\end{aligned}$$

Auf dieses Ergebnis wenden wir nun erst die Induktionsvoraussetzung und dann Teil (a) an und erhalten

$$\begin{aligned} \widehat{\mathbf{Q}}_*\big([\tau;\varrho]_{n+1}\big) &= \mu^{-1}p_k\mu^{-n}\mathbf{Q}\big([\Theta_i\tau^i]_n\big)\prod_{j\neq i}\mathbf{Q}\big([\Theta_j\tau^j]_n\big) \\ &= \mu^{-(n+1)}\mathbf{Q}\big([\tau]_{n+1}\big), \end{aligned}$$

was den Beweis von (3.8) beschließt.

(c) Hier ergibt sich mittels Teil (b) für alle $\tau \in \mathbb{T}$ und $n \geq 0$

$$\begin{aligned} \widehat{\mathbf{Q}}\big([\tau]_n\big) &= \mathbb{P}\big(\widehat{\boldsymbol{GW}} \in [\tau]_n\big) \\ &= \sum_{\varrho\in\tau_n}\mathbb{P}\big(\widehat{\boldsymbol{GW}} \in [\tau]_n, V^n = \varrho\big) \\ &= \sum_{\varrho\in\tau_n}\widehat{\mathbf{Q}}_*\big([\tau;\varrho]_n\big) \\ &= \sum_{\varrho\in\tau_n}\mu^{-n}\mathbf{Q}\big([\tau]_n\big) \\ &= w_n(\tau)\,\mathbf{Q}\big([\tau]_n\big). \end{aligned}$$

(d) Mit (c) folgt schließlich für $k \geq 1$

$$\begin{aligned} \mathbb{P}(\widehat{Z}_n = k) &= \sum_{\tau\in\mathbb{T}_n:z_n(\tau)=k}\widehat{\mathbf{Q}}\big([\tau]_n\big) \\ &= \sum_{\tau\in\mathbb{T}_n:z_n(\tau)=k}w_n(\tau)\,\mathbf{Q}\big([\tau]_n\big) \\ &= \sum_{\tau\in\mathbb{T}_n:z_n(\tau)=k}\frac{k}{\mu^n}\,\mathbf{Q}\big([\tau]_n\big) \\ &= \frac{k\,\mathbb{P}(Z_n = k)}{\mu^n}. \end{aligned}$$

◊

3.8. Bemerkungen. (a) Unter Verwendung der Teile (b) und (c) des Vergleichslemmas 3.7 läßt sich das anschaulich evidente Ergebnis zeigen, daß, gegeben $\widehat{\boldsymbol{GW}}_n = \{\varrho^1, ..., \varrho^p\}$ für ein $p \geq 1$, die Rückgratkomponente V^n eine Laplace-Verteilung auf dieser Menge besitzt (☞ Übung 3.6), was aber im weiteren Verlauf nicht ausdrücklich benötigt wird.

(b) Hinsichtlich der am Anfang dieses Unterabschnitts angesprochenen Frage, wann $\widehat{\mathbf{Q}}$ durch das Galton-Watson-Maß $\mathbf{Q}$ dominiert wird, gibt das Vergleichslemma noch keine vollständige Auskunft, zeigt aber immerhin, daß für alle $\tau \in \mathbb{T}$ und $n \geq 0$

$$\widehat{\mathbf{Q}}\big([\tau]_n\big) = w_n(\tau)\,\mathbf{Q}\big([\tau]_n\big) = \int_{[\tau]_n} w_n(\chi)\;\mathbf{Q}(d\chi)$$

gilt, weil w_n auf $[\tau]_n$ konstant ist. Wir belassen es hier bei dieser Feststellung und kehren im neuen Beweis des Satzes von Kesten, Stigum (☞ Satz 5.2) auf diesen Sachverhalt zurück.

Übungen zu Abschnitt 3

Übung 3.1. Sei $\nu = (\nu_k)_{k\ge 0}$ eine Verteilung auf $\mathbb{N}_0$ mit e.F. f und $0 < \sum_{k\ge 1} k\nu_k < \infty$. Bestimmen Sie die e.F. ihrer Größenverzerrung $\widehat{\nu}$.

Übung 3.2. Sei ν eine Verteilung auf $[0,\infty)$ mit $0 < \int x\,\nu(dx) < \infty$. Zeigen Sie, daß $\widehat{\nu}$ stets stochastisch größer als ν ist, d.h. $\widehat{\nu}(x,\infty) \ge \nu(x,\infty)$ für alle $x \ge 0$.

Übung 3.3. Beweisen Sie Lemma 3.4.

Übung 3.4. Zeigen Sie, daß $\mathcal{E}' = \{\mathbb{B}'(v,\varepsilon) : v \in \mathbb{V}, \varepsilon > 0\}$ abzählbar ist und die Borelsche σ-Algebra $\mathfrak{B}(\overline{\mathbb{V}})$ erzeugt.

Übung 3.5. Beweisen Sie Satz 3.6.

Übung 3.6. Seien $n, p \ge 1$ und $\varrho^1, ..., \varrho^p \in \mathbb{N}^n$ derart, daß $\mathbb{P}(\widehat{\boldsymbol{GW}}_n = \{\varrho^1, ..., \varrho^p\}) > 0$. Zeigen Sie

$$\mathbb{P}(V^n = \xi^k | \widehat{\boldsymbol{GW}}_n = \{\varrho^1, ..., \varrho^p\}) \;=\; \frac{1}{p}$$

für $k = 1, ..., p$, d.h. V^n ist bedingt unter $\widehat{\boldsymbol{GW}}_n = \{\varrho^1, ..., \varrho^p\}$ Laplace-verteilt auf dieser Menge.

4. Zum Grenzverhalten nichtkritischer GWPI: Ein probabilistischer Zugang

In diesem Abschnitt wenden wir uns nochmals den in Kapitel II behandelten GWPI zu und beweisen zwei Ergebnisse im nichtkritischen Fall, die im wesentlichen schon bekannt sind (☞ Korollar II.3.2 und Satz II.4.1), aber hier auf andere Weise bewiesen werden (ohne Verwendung e.F.). Der Grund für diese Rückkehr besteht aber vielmehr in der für uns wichtigen Tatsache, daß der Generationsgrößenprozeß eines größenverzerrten GWB's, nennen wir ihn größenverzerrten GWP, auch als GWPI aufgefaßt werden kann, wie wir als erstes zeigen wollen (☞ Satz 4.1 weiter unten).

Aus Zweckmäßigkeitsgründen variieren wir unsere in II.1.1 gegebene Definition eines GWPI $(Y_n)_{n\ge 0}$ dahingehend, daß wir $Y_0 \stackrel{\text{def}}{=} 0$ annehmen und $(Y_{n+1})_{n\ge 0}$ als GWPI im Sinne dieser Definition voraussetzen. Dies bedeutet, daß die betrachtete Population zum Zeitpunkt 0 keine Individuen besitzt und erst ab der ersten Generation eine zufällige Anzahl von Immigranten $\zeta_1, \zeta_2, ...$ aufnimmt, die dann jeweils u.i.v. GWP starten. Die Reproduktionsverteilung werde wie immer mit $(p_j)_{j\ge 0}$, die Immigrationsverteilung, also die Verteilung der unabhängigen Variablen ζ_n, wieder mit $(c_j)_{j\ge 0}$ bezeichnet. Es gilt demnach für $n \ge 1$

$$Y_n \;=\; \zeta_n \;+\; \sum_{k=1}^{Y_{n-1}} \xi_{nk} \tag{4.1}$$

mit unabhängigen, jeweils gemäß $(p_j)_{j\ge 0}$ verteilten Zufallsgrößen ξ_{nk}, welche die Anzahl der Nachkommen der Individuen angeben. Wir sprechen im folgenden dann auch von einem *GWPI* $(Y_n)_{n\ge 0}$ *mit Immigrationsfolge* $(\zeta_n)_{n\ge 1}$. Eine alternative Darstellung lautet (vgl. (II.1.1))

$$Y_n \;=\; \sum_{k=1}^{n}\sum_{j=1}^{\zeta_k} Z_{n-k}(k,j), \quad n\ge 1, \tag{4.2}$$

in der die $(Z_n(k,j))_{n\ge 0}$, $k,j\ge 1$, u.i.v. GWP mit einem Urahnen und Reproduktionsverteilung $(p_j)_{j\ge 0}$ bilden und auch unabhängig von $(\zeta_n)_{n\ge 1}$ sind. Wir interpretieren hierbei $(Z_n(k,j))_{n\ge 0}$ als den vom j-ten Immigranten der k-ten Generation initiierten Prozeß.

1. Verbindung zwischen GWPI und größenverzerrten GWB. Wir wollen nun zeigen, daß in einem größenverzerrten GWB mit Rückgrat die Nachkommen der Individuen des Rückgrats als Immigranten aufgefaßt werden können und so eine Verbindung zu GWPI entsteht. Die präzise Formulierung dieser Aussage gibt der folgende Satz, wobei wieder die Bezeichnungen aus Unterabschnitt 3.2 gelten:

4.1. Satz. *Es sei* $(Y_n)_{n\ge 0}$ *ein GWPI mit Immigrationsfolge* $\zeta_n \overset{\text{def}}{=} \widehat{X}_n - 1$ *für* $n\ge 1$. *Dann gilt*

$$\mathbb{P}\big((Y_n)_{n\ge 0}\in\cdot\big) \;=\; \widehat{\mathbf{Q}}\big((z_n-1)_{n\ge 0}\in\big) \;=\; \mathbb{P}\big((\widehat{Z}_n-1)_{n\ge 0}\in\cdot\big)$$

Beweis: Offenbar gilt $\widehat{\mathbf{Q}}(z_0-1\in\cdot)=\delta_0=\mathbb{P}(Y_0\in\cdot)$. Für $n\ge 0$ und $(k_1,...,k_{n+1})\in\mathbb{N}^{n+1}$ ergibt sich unter Beachtung der Unabhängigkeit von $\zeta_{n+1},(\xi_{n+1,j})_{j\ge 1}$ und $(Y_1,...,Y_n)$

$$\begin{aligned}
&\mathbb{P}(Y_1=k_1,...,Y_n=k_n,Y_{n+1}=k_{n+1})\\
&\;=\;\mathbb{P}\left(Y_1=k_1,...,Y_n=k_n,\zeta_{n+1}+\sum_{j=1}^{k_n}\xi_{n+1,j}=k_{n+1}\right)\\
&\;=\;\mathbb{P}(Y_1=k_1,...,Y_n=k_n)\,\mathbb{P}\left(\zeta_{n+1}+\sum_{j=1}^{k_n}\xi_{n+1,j}=k_{n+1}\right)\\
&\;=\;\mathbb{P}(Y_1=k_1,...,Y_n=k_n)\sum_{l=0}^{k_{n+1}}\mathbb{P}\left(\widehat{X}_{n+1}=l+1,\sum_{j=1}^{k_n}\xi_{n+1,j}=k_{n+1}-l\right)\\
&\;=\;\mathbb{P}(Y_1=k_1,...,Y_n=k_n)\sum_{l=0}^{k_{n+1}}\frac{(l+1)p_{l+1}}{\mu}\,p^{*(k_n)}_{k_{n+1}-l}
\end{aligned}$$

sowie andererseits mit $\Pi(k_1,...,k_n)\overset{\text{def}}{=}\{\tau\in\mathbb{T}_n : z_j(\tau)-1=k_j, 1\le j\le n\}$

$$\begin{aligned}
&\widehat{\mathbf{Q}}(z_1-1=k_1,...,z_n-1=k_n,z_{n+1}-1=k_{n+1})\\
&\;=\;\sum_{\tau\in\Pi(k_1,...,k_n)}\sum_{\varrho\in\tau_n}\mathbb{P}\left(\widehat{\boldsymbol{GW}}\in[\tau]_n,V^n=\varrho,\widehat{X}_{n+1}+\sum_{v\in\tau_n\setminus\{\varrho\}}X_v=k_{n+1}+1\right)
\end{aligned}$$

$$= \sum_{\tau\in\Pi(k_1,...,k_n)} \sum_{\varrho\in\tau_n} \widehat{\mathbf{Q}}_*\big([\tau;\varrho]_n\big)\,\mathbb{P}\Bigg(\widehat{X}_{n+1} + \sum_{v\in\tau_n\backslash\{\varrho\}} X_v = k_{n+1}+1\Bigg)$$

$$= \Bigg(\sum_{\tau\in\Pi(k_1,...,k_n)} \widehat{\mathbf{Q}}\big([\tau]_n\big)\Bigg)\Bigg(\sum_{l=0}^{k_{n+1}} \frac{(l+1)p_{l+1}}{\mu}\, p^{*(k_n)}_{k_{n+1}-l}\Bigg)$$

$$= \widehat{\mathbf{Q}}(z_1-1=k_1,...,z_n-1=k_n)\Bigg(\sum_{l=0}^{k_{n+1}} \frac{(l+1)p_{l+1}}{\mu}\, p^{*(k_n)}_{k_{n+1}-l}\Bigg).$$

Hierbei ging für die zweite Gleichheit die Unabhängigkeit von $\big(\widehat{\boldsymbol{GW}}_{|n}, V^n\big)$ und $\big(\widehat{X}_n, (X_v)_{v\in\mathbb{N}^n}\big)$ ein.

Aus den beiden obigen Rechnungen folgt induktiv die Identität

$$\mathbb{P}(Y_1=k_1,...,Y_n=k_n) \;=\; \widehat{\mathbf{Q}}(z_1-1=k_1,...,z_n-1=k_n)$$

für alle $n\geq 1$ und $(k_1,...,k_n)\in\mathbb{N}^n$ und damit die Behauptung. $\diamondsuit$

4.2. Bemerkung. Man kann sich das gerade gezeigte Ergebnis auch ohne formale Argumente recht gut klarmachen: Entfernt man aus dem zufälligen Baum $\widehat{\boldsymbol{GW}}$ das Rückgrat $\boldsymbol{V}$, so zerfällt jede Generation $n\geq 1$ in die Teilmenge der Nachkommen des Rückgratelements V^{n-1} der vorherigen Generation, die nicht als Wirbel gewählt wurden, d.h. $\boldsymbol{B}_n\backslash\{V^n\}$ (☞ 3.3), sowie der Menge aller anderen Nachkommen, deren Mutter irgendein Individuum in $\widehat{\boldsymbol{GW}}_{n-1}\backslash\{V^{n-1}\}$ ist, also $\boldsymbol{A}_n$. Die Elemente von $\boldsymbol{B}_n\backslash\{V^n\}$ interpretieren wir als Immigranten, sie reproduzieren gemäß $(\tilde{p}_j)_{j\geq 0}$ mit $\tilde{p}_j \overset{\text{def}}{=} \widehat{p}_{j+1}$, während $(p_j)_{j\geq 0}$ die Reproduktionsverteilung aller anderen Individuen bildet.

Für den Rest dieses Abschnitts beschäftigen wir uns mit der bereits angekündigten Herleitung von zwei Sätzen für nichtkritische GWPI, die uns später von Nutzen sein werden. Daß diese nur relativ geringfügige Erweiterungen von bekannten Resultaten aus Kapitel II darstellen, ist hier weniger von Belang als die Tatsache, daß hier keine e.F., sondern probabilistische Argumente verwendet werden.

2. Zum asymptotischen Wachstum superkritischer GWPI. Wir benötigen als erstes zwei Hilfsresultate, von denen das erste auf dem Borel-Cantelli-Lemma basiert.

4.3. Lemma. *Gegeben eine unabhängige Folge $(X_n)_{n\geq 0}$ identisch verteilter, nichtnegativer Zufallsgrößen, gelten die folgenden Aussagen:*

(a) $\lim_{n\to\infty} \frac{X_n}{n} = 0$ *f.s., falls* $\mathbb{E}X_0<\infty$.

(b) $\limsup_{n\to\infty} \frac{X_n}{n} = \infty$ *f.s., falls* $\mathbb{E}X_0=\infty$.

(c) $\sum_{n\geq 1} c^n e^{X_n} < \infty$ *f.s. für alle* $c\in(0,1)$, *falls* $\mathbb{E}X_0<\infty$.

(d) $\sum_{n\ge1} c^n e^{X_n} = \infty$ *f.s. für alle* $c \in (0,1)$*, falls* $\mathbb{E}X_0 = \infty$.

Beweis: Eine Anwendung von Ungleichung (A.12) auf S. 115 in [1] liefert

$$\begin{aligned}\sum_{n\ge1} \mathbb{P}\left(\frac{X_n}{n} > \varepsilon\right) \;&=\; \sum_{n\ge1} \mathbb{P}\left(\frac{X_0}{\varepsilon} > n\right)\\ &\le\; \frac{\mathbb{E}X_0}{\varepsilon} \;\le\; \sum_{n\ge0} \mathbb{P}\left(\frac{X_0}{\varepsilon} > \varepsilon\right) \;=\; \sum_{n\ge0} \mathbb{P}\left(\frac{X_n}{n} > \varepsilon\right)\end{aligned}$$

für alle $\varepsilon > 0$. Das Lemma von Borel-Cantelli impliziert daher $\mathbb{P}(\frac{X_n}{n} > \varepsilon \text{ u.o.}) = 0$ bzw. $= 1$ *für alle* $\varepsilon > 0$, falls $\mathbb{E}X_0 < \infty$ bzw. $= \infty$ gilt. Es folgen die Teile (a) und (b).

(c) Seien $\mathbb{E}X_0 < \infty$, $c \in (0,1)$, $0 < \varepsilon < -\log c$ und $A \overset{\text{def}}{=} \{\lim_{n\to\infty} \frac{X_n}{n} = 0\}$. Nach (a) ist $\mathbb{P}(A) = 1$. Für $\omega \in A$ und alle hinreichend großen n erhalten wir $X_n(\omega) \le n\varepsilon$, also

$$c^n e^{X_n(\omega)} \;\le\; e^{n(\varepsilon+\log c)}.$$

Da $e^{\varepsilon+\log c} < 1$ liefert dies die Konvergenz der Reihe $\sum_{n\ge1} c^n e^{X_n(\omega)}$.

(d) Falls $\mathbb{E}X_0 = \infty$, gilt $\mathbb{P}(B) = 1$ für $B \overset{\text{def}}{=} \{\limsup_{n\to\infty} \frac{X_n}{n} = \infty\}$. Ist nun $\omega \in B$, können wir nach eventuellem Übergang zu einer (von ω abhängigen) Teilfolge $\lim_{n\to\infty} \frac{X_n(\omega)}{n} = \infty$ annehmen. Für alle hinreichend großen n gilt daher $X_n(\omega) \ge -n\log c$, also $c^n e^{X_n(\omega)} \ge 1$, was die Divergenz der Reihe $\sum_{n\ge1} c^n e^{X_n(\omega)}$ beweist. ◊

Als weiteres Hilfsergebnis notieren wir eine Verallgemeinerung des Martingal-Konvergenzsatzes auf Folgen nichtnegativer Zufallsgrößen, die lediglich bedingt integrierbar sind und die Submartingaleigenschaft besitzen.

4.4. Lemma. *Sei* $M = (M_n)_{n\ge0}$ *eine Folge nichtnegativer Zufallsgrößen, die bezüglich einer Filtration* $(\mathcal{F}_n)_{n\ge0}$ *adaptiert ist. Außerdem gelte*

(1) $\mathbb{E}(M_{n+1}|\mathcal{F}_n) \ge M_n$ *f.s. für alle* $n \ge 0$,
(2) $\sup_{n\ge0} \mathbb{E}(M_n|\mathcal{F}_0) < \infty$ *f.s.*

Dann konvergiert M_n *f.s. gegen eine endliche Zufallsgröße* M_∞.

Es sei nochmals betont, daß die M_n nicht integrierbar sein müssen.

Beweis: Eine Kopie der Beweise von Satz 19.3 und Lemma 20.1 in [2] liefert für jede bezüglich $(\mathcal{F}_n)_{n\ge0}$ vorhersagbare Folge $H = (H_n)_{n\ge0}$ beschränkter nichtnegativer Zufallsgrößen und $M^{(a)} \overset{\text{def}}{=} ((M_n - a)^+)_{n\ge0}$, $a \in \mathbb{R}$, die Ungleichung

$$\mathbb{E}\big((H\cdot M^{(a)})_{n+1}|\mathcal{F}_n\big) \;\ge\; (H\cdot M^{(a)})_n \quad \mathbb{P}\text{-f.s.}$$

für alle $n \geq 0$. Daraufhin führt uns eine Kopie des Beweises der Überquerungs-Uungleichung 21.1 in [2] für beliebige $b > a$ zu der Ungleichung

$$(b-a)\,\mathbb{E}(U_n(a,b)|\mathcal{F}_0) \;\leq\; \mathbb{E}((M_n-a)^+|\mathcal{F}_0) - \mathbb{E}((M_0-a)^+|\mathcal{F}_0) \quad \mathbb{P}\text{-f.s.}, \tag{4.3}$$

wobei $U_n(a,b)$ die Anzahl der aufsteigenden Überquerungen des Intervalls (a,b) von M bis zum Zeitpunkt n angibt. Setzen wir $U_\infty(a,b) \stackrel{\text{def}}{=} \lim_{n\to\infty} U_n(a,b)$, so ergibt sich vermöge $\sup_{n\geq 0} \mathbb{E}(U_n(a,b)|\mathcal{F}_0) = \mathbb{E}(U_\infty(a,b)|\mathcal{F}_0)$ f.s., der Abschätzung (4.3) sowie Bedingung (2) des Lemmas

$$(b-a)\mathbb{E}(U_\infty(a,b)|\mathcal{F}_0) \;\leq\; \sup_{n\geq 0} \mathbb{E}((M_n-a)^+|\mathcal{F}_0) \;\leq\; |a| \;+\; \sup_{n\geq 0} \mathbb{E}(M_n|\mathcal{G}_0) \;<\; \infty \quad \mathbb{P}\text{-f.s.}$$

und somit $U_\infty(a,b) < \infty$ f.s. Wie im Beweis des Martingal-Konvergenzsatzes (☞ Satz 21.2 in [2]) gelangt man schließlich zu der Schlußfolgerung

$$\mathbb{P}\Big(\liminf_{n\to\infty} M_n < \limsup_{n\to\infty} M_n\Big) \;=\; 0,$$

also zur f.s. Konvergenz von M_n gegen ein M_∞ mit

$$\mathbb{E}(M_\infty|\mathcal{F}_0) \;\leq\; \sup_{n\geq 0} \mathbb{E}(M_n|\mathcal{G}_0) \quad \mathbb{P}\text{-f.s.}$$

unter Hinweis auf das Lemma für Fatou für bedingte Erwartungen. Folglich kann M_∞ als endliche Zufallsgröße gewählt werden. ♢

Unser angestrebtes Resultat in diesem Unterabschnitt lautet nun (vgl. Satz II.4.1):

4.5. Satz (Seneta). *Sei $(Y_n)_{n\geq 0}$ ein superkritischer GWPI mit endlichem Reproduktionsmittel μ und Immigrationsfolge $(\zeta_n)_{n\geq 1}$. Dann gilt:*

(a) Ist $\mathbb{E}\log^+\zeta_1 < \infty$, so konvergiert $\mu^{-n}Y_n$ f.s. gegen eine endliche Zufallsgröße Y_∞.

(b) Ist dagegen $\mathbb{E}\log^+\zeta_0 = \infty$, so gilt

$$\limsup_{n\to\infty} \frac{Y_n}{c^n} \;=\; \infty \quad \mathbb{P}\text{-}f.s.$$

für jedes $c \in (0,\infty)$.

Beweis: (a) Setzen wir $\mathcal{F}_0 \stackrel{\text{def}}{=} \sigma\big((\zeta_n)_{n\geq 1}\big)$, so ergibt sich vermöge (4.1)

$$\begin{aligned}
\mathbb{E}(Y_n|\mathcal{F}_0) \;&=\; \zeta_n \;+\; \sum_{k\geq 0} \mathbb{E}\left(\mathbf{1}_{\{Y_{n-1}=k\}} \sum_{j=1}^{k} \xi_{nj}\,\middle|\,\mathcal{F}_0\right) \\
&=\; \zeta_n \;+\; \sum_{k\geq 1} k\mu\,\mathbb{P}(Y_{n-1}=k|\mathcal{F}_0) \\
&=\; \zeta_n \;+\; \mu\,\mathbb{E}(Y_{n-1}|\mathcal{F}_0) \quad \mathbb{P}\text{-f.s.},
\end{aligned}$$

denn für jedes $k \geq 0$ ist $\sum_{j=1}^{k} \xi_{nj}$ unabhängig von $(Y_{n-1}, (\zeta_n)_{n\geq 1})$ [1)]. Unter Beachtung von $Y_0 = 0$ liefert dies induktiv

$$\mathbb{E}\left(\frac{Y_n}{\mu^n}\middle|\mathcal{F}_0\right) \;=\; \sum_{k=1}^{n} \frac{\zeta_k}{\mu^k} \quad \mathbb{P}\text{-f.s.}$$

für alle $n \geq 0$. Ein Anwendung von Lemma 4.3 auf die Folge $(\log^+ \zeta_n)_{n\geq 1}$ zeigt nun im Fall $\mathbb{E}\log^+ \zeta_1 < \infty$

$$\sup_{n\geq 0} \mathbb{E}\left(\frac{Y_n}{\mu^n}\middle|\mathcal{F}_0\right) \;=\; \sum_{k\geq 1} \frac{\zeta_k}{\mu^k} \;\leq\; \sum_{k\geq 1} \mu^{-k} e^{\log^+ \zeta_k} \;<\; \infty \quad \mathbb{P}\text{-f.s.} \tag{4.4}$$

Mittels dieser Vorüberlegung erhalten wir nun die f.s. Konvergenz von $(\mu^{-n} Y_n)_{n\geq 0}$, wenn wir noch verifizieren, daß diese Folge neben der mit (4.4) schon gezeigten Bedingung (2) auch die übrigen Voraussetzungen des vorherigen Lemmas erfüllt: Sei $\mathcal{F}_n \stackrel{\text{def}}{=} \sigma\big((\zeta_k)_{k\geq 1}, Y_0, ..., Y_n\big)$ für $n \geq 1$. Dann ist $(Y_n)_{n\geq 0}$ natürlich $(\mathcal{F}_n)_{n\geq 0}$-adaptiert, und wir erhalten für jedes $n \geq 0$

$$\mathbb{E}\left(\frac{Y_{n+1}}{\mu^{n+1}}\middle|\mathcal{F}_n\right) \;=\; \frac{\zeta_{n+1}}{\mu^{n+1}} \;+\; \frac{1}{\mu^{n+1}}\,\mathbb{E}\left(\sum_{j=1}^{Y_n} \xi_{n+1,j}\middle|\mathcal{F}_n\right) \;\geq\; \frac{Y_n}{\mu^n} \quad \mathbb{P}\text{-f.s.},$$

denn vermöge der Unabhängigkeit der Folge $(\xi_{n+1,j})_{j\geq 1}$ von $\mathcal{F}_n$ und der $\mathcal{F}_n$-Meßbarkeit von Y_n gilt offenbar $\mathbb{E}(\sum_{j=1}^{Y_n} \xi_{n+1,j}|\mathcal{F}_n) = Y_n\,\mathbb{E}\xi_{n+1,1} = \mu Y_n$ f.s. Also sind alle Voraussetzungen von Lemma 4.4 tatsächlich erfüllt für $(\mu^{-n} Y_n)_{n\geq 0}$, und wir erhalten das Gewünschte.

(b) Im Fall $\mathbb{E}\log^+ \zeta_1 = \infty$ liefert Lemma 4.3(b) unmittelbar

$$\limsup_{n\to\infty} \frac{\log^+ \zeta_n}{n} \;=\; \infty \quad \mathbb{P}\text{-f.s.}$$

und damit für jedes $c > 0$

$$\limsup_{n\to\infty} \frac{1}{c} \exp\left(\frac{\log \zeta_n}{n}\right) \;=\; \limsup_{n\to\infty} \left(\frac{\zeta_n^{1/n}}{c}\right)^n \;=\; \infty \quad \mathbb{P}\text{-f.s.},$$

also wegen $Y_n \geq \zeta_n$ für alle $n \geq 1$ die Behauptung. ◊

3. Zur Verteilungskonvergenz subkritischer GWPI. Auch hier benötigen wir zunächst das folgende Hilfsresultat, das eine Verallgemeinerung des Borel-Cantelli-Lemmas beinhaltet:

4.6. Lemma. *Es seien $(k_n)_{n\geq 1}$ eine Folge in $\mathbb{N}_0$, $(A_{nk})_{n,k\geq 1}$ eine Familie stochastisch unabhängiger Ereignisse sowie $B_n \stackrel{\text{def}}{=} \bigcup_{k=1}^{k_n} A_{nk}$ für $n \geq 1$ (die im Fall $k_n = 0$ als leere Menge interpretiert wird). Dann gilt die Implikation*

$$\mathbb{P}\Big(\limsup_{n\to\infty} B_n\Big) = 0 \quad \Rightarrow \quad \sum_{n\geq 1}\sum_{k=1}^{k_n} \mathbb{P}(A_{nk}) < \infty.$$

[1)] Benutze die Rechenregel $\mathbb{E}(XY|\mathcal{F}) = \mathbb{E}X\,\mathbb{E}(Y|\mathcal{F})$, falls X und $\sigma(Y,\mathcal{F})$ unabhängig sind

Beweis: Wäre die auftretende Reihe divergent, so folgte aus der Unabhängigkeit der A_{nk} der Widerspruch

$$\begin{aligned} 1 &= \mathbb{P}\Big(\liminf_{n\to\infty} B_n^c\Big) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{j\ge n}\bigcap_{k=1}^{k_j} A_{jk}^c\right) \\ &= \lim_{n\to\infty} \prod_{j\ge n}\prod_{k=1}^{k_j} \big(1-\mathbb{P}(A_{jk})\big) \\ &= \lim_{n\to\infty} \exp\left(-\sum_{j\ge n}\sum_{k=1}^{k_j} \log\big(1-\mathbb{P}(A_{jk})\big)\right) \\ &\le \lim_{n\to\infty} \exp\left(-\sum_{j\ge n}\sum_{k=1}^{k_j} \mathbb{P}(A_{jk})\right) \;=\; 0. \end{aligned}$$ ◊

Mithilfe dieses Lemmas zeigen wir nun (vgl. Korollar II.3.3)

4.7. Satz (Heathcote). *Für einen subkritischen GWPI $(Y_n)_{n\ge0}$ mit Immigrationsfolge $(\zeta_n)_{n\ge1}$ und $p_0<1$ gelten die folgenden Aussagen:*

(a) Ist $\mathbb{E}\log^+\zeta_1<\infty$, so konvergiert Y_n in Verteilung gegen eine Zufallsgröße Y_∞.

(b) Gilt dagegen $\mathbb{E}\log^+\zeta_1=\infty$, so konvergiert Y_n nach Wahrscheinlichkeit gegen ∞, d.h.

$$\lim_{n\to\infty}\mathbb{P}(Y_n>t) \;=\; 1$$

für alle $t>0$.

Beweis: Wir überlassen es dem Leser nachzuweisen, daß aus (4.2) die Verteilungsidentität

$$Y_n \overset{d}{=} \widehat{Y}_n \overset{\text{def}}{=} \sum_{k=1}^{n}\sum_{j=1}^{\zeta_k} Z_{k-1}(k,j), \quad n\ge1, \tag{4.5}$$

folgt, wobei die $(Z_n(k,j))_{n\ge0}$, $k,j\ge1$, u.i.v. GWP mit einem Urahnen und Reproduktionsverteilung $(p_j)_{j\ge0}$ bilden und auch unabhängig von $(\zeta_n)_{n\ge1}$ sind (☞ Übung 4.2). Im Unterschied zu Y_n hat $\widehat{Y}_n$ nichtnegative Zuwächse $\Lambda_k \overset{\text{def}}{=} \sum_{j=1}^{\zeta_k} Z_{k-1}(k,j)$ und konvergiert folglich f.s. gegen

$$\widehat{Y}_\infty \overset{\text{def}}{=} \sum_{k\ge1}\Lambda_k.$$

Da die Λ_k offenbar stochastisch unabhängig sind, erhalten wir mit dem Lemma von Borel-Cantelli unter Beachtung von $\mathbb{P}(\Lambda_k\in\mathbb{N}_0)=1$ für jedes k die Implikationen

$$\sum_{k\ge1}\mathbb{P}(\Lambda_k\ge1)\begin{cases}<\infty\\=\infty\end{cases} \Rightarrow \quad \mathbb{P}(\Lambda_k\ge1 \text{ u.o.})\begin{cases}=0\\=1\end{cases} \Rightarrow \quad \mathbb{P}(\widehat{Y}_\infty<\infty)\begin{cases}=1\\=0\end{cases}$$

und folglich

$$\mathbb{P}(\widehat{Y}_\infty < \infty) \;\in\; \{0,1\}. \tag{4.6}$$

(a) Es sei nun $\mathbb{E}\log^+\zeta_1 < \infty$ vorausgesetzt und weiterhin $\mathcal{F}_0 = \sigma\big((\zeta_n)_{n\ge 1}\big)$. Dann gilt unter Beachtung der Unabhängigkeit der $Z_{k-1}(k,j)$ von $\mathcal{F}_0$

$$\mathbb{E}(\widehat{Y}_\infty|\mathcal{F}_0) \;=\; \sum_{k\ge 1}\mathbb{E}(\Lambda_k|\mathcal{F}_0) \;=\; \sum_{k\ge 1}\mathbb{E}\left(\sum_{j=1}^{\zeta_k} Z_{k-1}(k,j)\middle|\mathcal{F}_0\right) \;=\; \sum_{k\ge 1}\mu^{k-1}\zeta_k \quad \mathbb{P}\text{-f.s.}$$

An dieser Stelle kommt erneut Lemma 4.3 ins Spiel, welches auf $(\log^+\zeta_n)_{n\ge 1}$ und $c = \mu \in (0,1)$ angewandt wegen $\mathbb{E}\log^+\zeta_1 < \infty$ die f.s. Konvergenz der Reihe $\sum_{k\ge 1}\mu^k e^{\log^+\zeta_k}$ garantiert, also

$$\sum_{k\ge 1}\mu^{k-1}\zeta_k \;=\; \mathbb{E}(\widehat{Y}_\infty|\mathcal{F}_0) \;<\; \infty \quad \mathbb{P}\text{-f.s.}$$

Daher gilt $\widehat{Y}_\infty < \infty$ f.s., und wir erhalten vermöge (4.5) das gewünschte Ergebnis

$$Y_n \;\xrightarrow{d}\; \widehat{Y}_\infty, \quad n\to\infty.$$

(b) Sei nun $\mathbb{E}\log^+\zeta_1 = \infty$ vorausgesetzt und $\mathbb{P}(\widehat{Y}_\infty = \infty) < 1$ angenommen, was im folgenden zum Widerspruch geführt wird. Gemäß (4.6) bedeutet die Annahme, daß

$$\mathbb{P}(\widehat{Y}_\infty = \infty) \;=\; \mathbb{P}\left(\sum_{k\ge 1}\Lambda_k = \infty\right) \;=\; 0,$$

so daß für $D_k \stackrel{\text{def}}{=} \{\Lambda_k \ge 1\} = \big\{\sum_{j=1}^{\zeta_k} Z_{k-1}(k,j)\ge 1\big\}$

$$\mathbb{P}\Big(\limsup_{k\to\infty} D_k\Big) \;=\; \mathbb{P}\Big(\limsup_{k\to\infty} D_k\Big|\mathcal{F}_0\Big) \;=\; 0 \quad \mathbb{P}\text{-f.s.}$$

gilt. Wegen der Unabhängigkeit von $\mathcal{F}_0$ und $\big((Z_n(k,j))_{n\ge 0}\big)_{j,k\ge 1}$ erhalten wir weiter

$$\begin{aligned} 0 \;&=\; \mathbb{P}\Big(\limsup_{k\to\infty} D_k\Big|\zeta_k = y_k, k\ge 1\Big) \\ &=\; \mathbb{P}\left(\limsup_{k\to\infty}\left\{\sum_{j=1}^{y_k} Z_{k-1}(k,j)\ge 1\right\}\right) \;=\; \mathbb{P}\Big(\limsup_{k\to\infty} B_k\Big) \end{aligned}$$

für $\mathbb{P}((\zeta_k)_{k\ge 1}\in\cdot)$-fast alle $(y_k)_{k\ge 1}$, wobei wir abkürzend $B_k \stackrel{\text{def}}{=} \bigcup_{j=1}^{y_k}\{Z_{k-1}(k,j)\ge 1\}$ gesetzt haben. Die Familie $(A_{nk})_{n,k\ge 1}$ mit $A_{kj} \stackrel{\text{def}}{=} \{Z_{k-1}(k,j)\ge 1\}$ erfüllt offenbar die Voraussetzungen von Lemma 4.6 (mit $k_n = y_n$ für $n\ge 1$), woraus wir mit diesem

$$\sum_{k\ge 1}\sum_{j=1}^{y_k}\mathbb{P}(A_{kj}) \;=\; \sum_{k\ge 1}\sum_{j=1}^{y_k}\mathbb{P}(Z_{k-1}(k,j)\ge 1) \;=\; \sum_{k\ge 1} y_k\mathbb{P}(Z_{k-1}(1,1)\ge 1) \;<\; \infty$$

für $\mathbb{P}((\zeta_k)_{k\ge 1} \in \cdot)$-fast alle $(y_k)_{k\ge 1}$, also

$$\sum_{k\ge 1} \zeta_k \mathbb{P}(Z_{k-1}(1,1) \ge 1) \;<\; \infty \quad \mathbb{P}\text{-f.s.}$$

schließen können. Beachten wir für $k \ge 1$ noch die einfache Ungleichung $e^{\log^+ \zeta_k} \le 1+\zeta_k$ sowie $\mathbb{P}(Z_{k-1}(1,1) \ge 1) \le (1-p_0)^{k-1}$, so ergibt sich wegen $0 < p_0 < 1$

$$\sum_{k\ge 1} (1-p_0)^k e^{\log^+ \zeta_k} \;<\; \infty$$

und folglich unter Hinweis auf Lemma 4.3(d) (mit $c = 1 - p_0$) der Widerspruch

$$\mathbb{E} \log^+ \zeta_1 < \infty.$$

Es gilt demnach $\widehat{Y}_\infty = \infty$ f.s., also unter nochmaliger Benutzung von (4.5)

$$\lim_{n\to\infty} \mathbb{P}(Y_n > t) \;=\; \lim_{n\to\infty} \mathbb{P}(\widehat{Y}_n > t) \;=\; \mathbb{P}(\widehat{Y}_\infty > t) \;=\; 1$$

für alle $t > 0$. $\diamondsuit$

Übungen zu Abschnitt 4

Übung 4.1. Führen Sie die Details im Beweis von Lemma 4.3 (Erweiterung des Martingal-Konvergenzsatzes) genauer aus.

Übung 4.2. Geben Sie einen Beweis von (4.5) (etwa mithilfe e.F.)

5. Superkritische GWP: Nochmals der Satz von Kesten, Stigum

Endlich sind wir in der Lage, ein erstes klassisches Resultat für GWP, nämlich den Satz von Kesten, Stigum, mittels der zuvor entwickelten Theorie für GWB und ihre Größenverzerrung auf alternative Weise zu beweisen. Wie schon zu Beginn des Kapitels bemerkt, stammt die vorgestellte Methode von Lyons, Peres und Pemantle [5], und stützt sich im wesentlichen auf das Vergleichslemma 3.7, die Sätze 4.1 und 4.5 sowie das nachfolgende, sehr allgemein formulierte Lemma über den Zusammenhang von zwei W-Maßen auf einem filtrierten W-Raum:

5.1. Lemma. *Es seien $\mathbf{P}, \mathbf{Q}$ W-Maße auf einem meßbaren Raum $(\mathfrak{X}, \mathcal{A})$ sowie $(\mathcal{A}_n)_{n\ge 0}$ eine Filtration über $\mathfrak{X}$ mit $\sigma(\bigcup_{n\ge 0} \mathcal{A}_n) = \mathcal{A}$. Für $\mathbf{P}_n \stackrel{\text{def}}{=} \mathbf{P}_{|\mathcal{A}_n}$, $\mathbf{Q}_n \stackrel{\text{def}}{=} \mathbf{Q}_{|\mathcal{A}_n}$ gelte $\mathbf{P}_n \ll \mathbf{Q}_n$, und es sei $X_n \stackrel{\text{def}}{=} \frac{d\mathbf{P}_n}{d\mathbf{Q}_n}$ für alle $n \ge 0$. Dann folgt mit $X \stackrel{\text{def}}{=} \limsup_{n\to\infty} X_n$:*

(a) $\mathbf{P}$ besitzt die folgende Zerlegung in einen $\mathbf{Q}$-stetigen und einen $\mathbf{Q}$-singulären Anteil:

$$\mathbf{P}(A) \;=\; \int_A X \, d\mathbf{Q} \;+\; \mathbf{P}(A \cap \{X = \infty\}), \quad A \in \mathcal{A}. \tag{5.1}$$

(b) Die folgenden Aussagen sind äquivalent:

(b1) $\mathbf{P} \ll \mathbf{Q}$,

(b2) $X < \infty$ $\mathbf{P}$-*f.s.*,

(b3) $\int_{\mathfrak{X}} X\, d\mathbf{Q} = \mathbf{P}(\mathfrak{X}) = 1$.

(c) Dual zu diesen Aussagen sind äquivalent:

(c1) $\mathbf{P} \perp \mathbf{Q}$,

(c2) $X = \infty$ $\mathbf{P}$-*f.s.*,

(c3) $\int_{\mathfrak{X}} X\, d\mathbf{Q} = 0$.

Beweis: Wir beschränken uns auf den Nachweis von Teil (a) und überlassen die dann relativ kanonischen Argumente zum Beweis von (b) und (c) dem Leser (☞ Übung 5.1).

Wir zeigen zuerst, daß $(X_n)_{n\ge 0}$ ein $\mathbf{Q}$-Martingal bezüglich $(\mathcal{A}_n)_{n\ge 0}$ bildet: Jedes X_n ist offenkundig $\mathcal{A}_n$-meßbar und wegen $\int X_n\, d\mathbf{Q} = \mathbf{P}(\mathfrak{X}) = 1$ auch $\mathbf{Q}$-integrierbar. Für jedes $A \in \mathcal{A}_n$ ergibt sich ferner unter Beachtung von $\mathcal{A}_n \subset \mathcal{A}_{n+1}$

$$\int_A X_{n+1}\, d\mathbf{Q} \;=\; \mathbf{P}_{n+1}(A) \;=\; \mathbf{P}_n(A) \;=\; \int_A X_n\, d\mathbf{Q}$$

und somit $\mathbb{E}_{\mathbf{Q}}(X_{n+1}|\mathcal{A}_n) = X_n$ $\mathbf{Q}$-f.s., was die Martingaleigenschaft beweist. Als nichtnegatives $\mathbf{Q}$-Martingal konvergiert X_n nach dem Martingal-Konvergenzsatz $\mathbf{Q}$-f.s. gegen X, und X ist $\mathbf{Q}$-integrierbar.

Zum Nachweis von (5.1) sei zunächst $\mathbf{P} \ll \mathbf{Q}$ angenommen und $Y \stackrel{\text{def}}{=} \frac{d\mathbf{P}}{d\mathbf{Q}}$. Für jedes $n \ge 0$ und $A \in \mathcal{A}_n$ ergibt sich dann

$$\int_A Y\, d\mathbf{Q} \;=\; \mathbf{P}(A) \;=\; \mathbf{P}_n(A) \;=\; \int_A X_n\, d\mathbf{Q}_n \;=\; \int_A X_n\, d\mathbf{Q}$$

und somit $X_n = \mathbb{E}_{\mathbf{Q}}(Y|\mathcal{A}_n)$ $\mathbf{Q}$-f.s. Vermöge des Struktursatzes für g.i. Martingale (☞ Satz 22.3 in [2]) folgt $X_n \to Y$ $\mathbf{Q}$-f.s. und in $\mathfrak{L}_1(\mathbf{Q})$ und dann $\frac{d\mathbf{P}}{d\mathbf{Q}} = Y = X$ $\mathbf{Q}$-f.s. (wegen der Eindeutigkeit des Grenzwerts) sowie

$$\mathbf{P}(X = \infty) \;=\; \int_{\{X=\infty\}} X\, d\mathbf{Q} \;=\; 0.$$

Die Beziehung (5.1) ist nun offensichtlich.

Für den allgemeinen Fall setzen wir $\nu \stackrel{\text{def}}{=} \frac{\mathbf{P}+\mathbf{Q}}{2}$ und $\nu_n \stackrel{\text{def}}{=} \nu_{|\mathcal{A}_n}$. Dann gilt $\mathbf{P} \ll \nu$, $\mathbf{Q} \ll \nu$, und wir können das zuvor Bewiesene auf $U_n \stackrel{\text{def}}{=} \frac{d\mathbf{P}_n}{d\nu_n}$ und $V_n \stackrel{\text{def}}{=} \frac{d\mathbf{Q}_n}{d\nu_n}$ anwenden. Dazu setzen wir $U \stackrel{\text{def}}{=} \limsup_{n\to\infty} U_n$ sowie $V_n \stackrel{\text{def}}{=} \limsup_{n\to\infty} V_n$. Für jedes $n \ge 1$ und $A \in \mathcal{A}_n$ gilt

$$\int_A (U_n + V_n)\, d\nu \;=\; \mathbf{P}(A) + \mathbf{Q}(A) \;=\; 2\nu(A) \;=\; \int_A 2\, d\nu$$

und somit

$$\nu(U_n + V_n = 2) \;=\; 1. \tag{5.2}$$

Dieselbe Argumentation wie oben liefert $U_n \to U = \frac{d\mathbf{P}}{d\nu}$ und $V_n \to V = \frac{d\mathbf{Q}}{d\nu}$ ν-f.s., was in Verbindung mit (5.2) außerdem $\nu(U = V = 0) = 0$ zeigt. Deshalb ist $\frac{U}{V}$ ν-f.s. wohldefiniert, und es folgt unter Hinweis auf Korollar 13.5 in [1]

$$\frac{U}{V} \;=\; \lim_{n\to\infty} \frac{U_n}{V_n} \;=\; \lim_{n\to\infty} \frac{d\mathbf{P}_n/d\nu_n}{d\mathbf{Q}_n/d\nu_n} \;=\; \lim_{n\to\infty} \frac{d\mathbf{P}_n}{d\mathbf{Q}_n} \;=\; \lim_{n\to\infty} X_n \;=\; X \quad \nu\text{-f.s.},$$

insbesondere $\{V = 0\} = \{X = \infty\}$ ν-f.s. Schließlich erhalten wir

$$\begin{aligned}\mathbf{P}(A) \;=\; \int_A U\, d\nu \;&=\; \int_{A\cap\{V>0\}} U\, d\nu \;+\; \int_{A\cap\{V=0\}} U\, d\nu \\ &=\; \int_A XV\, d\nu \;+\; \int_{A\cap\{X=\infty\}} U\, d\nu \\ &=\; \int_A X\, d\mathbf{Q} \;+\; \mathbf{P}(A\cap\{X=\infty\}),\end{aligned}$$

also (5.1). ♢

Hier ist noch einmal der aus I.6 bekannte Satz von Kesten, Stigum in der auf die äquivalenten Kernaussagen reduzierten Form (vgl. Satz I.6.2):

5.2. Satz (Kesten, Stigum). *Es seien $(Z_n)_{n\ge 0}$ ein superkritischer GWP mit einem Urahnen, endlichem Reproduktionsmittel $\mu \in (1,\infty)$, Aussterbewahrscheinlichkeit q und $W_n \stackrel{\text{def}}{=} \mu^{-n} Z_n$ für $n \ge 0$. Für $W \stackrel{\text{def}}{=} \lim_{n\to\infty} W_n$ sind dann folgende Aussagen äquivalent:*

$$\mathbb{P}(W = 0) = q, \tag{5.3}$$

$$\mathbb{E}W = 1, \tag{5.4}$$

$$\mathbb{E}Z_1 \log Z_1 \;=\; \sum_{k\ge 1} p_k\, k \log k \;<\; \infty. \tag{ZlogZ}$$

Beweis: Für $n \ge 0$ setzen wir $\mathcal{E}_n \stackrel{\text{def}}{=} \{\emptyset, \mathbb{T}\} \cup \{[\tau]_n : \tau \in \mathbb{T}\}$ und $\mathcal{A}_n \stackrel{\text{def}}{=} \sigma(\mathcal{E}_n)$. Zum Beweis des Satzes wenden wir Lemma 5.1 auf $\mathfrak{X} = \mathbb{T}$, $\mathcal{A} = \mathfrak{B}(\mathbb{T})$, $\mathbf{P} = \widehat{\mathbf{Q}}$ und $\mathbf{Q}$ an. Wir weisen darauf hin, daß offenbar $\mathcal{E} = \bigcup_{n\ge 0} \mathcal{E}_n$ für $\mathcal{E}$ gemäß (1.2) gilt und somit die im Lemma geforderte Eigenschaft $\mathcal{A} = \sigma\big(\bigcup_{n\ge 0} \mathcal{A}_n\big) = \sigma(\mathcal{E})$ gemäß (1.3) wirklich erfüllt ist.

Betrachten wir nun für $n \ge 1$ das durch

$$\Gamma_n(B) \;\stackrel{\text{def}}{=}\; \int_B w_n(\tau)\; \mathbf{Q}(d\tau), \quad B \in \mathfrak{B}(\mathbb{T})$$

definierte W-Maß auf $(\mathbb{T}, \mathfrak{B}(\mathbb{T}))$, das wegen

$$\Gamma_n(\mathbb{T}) \;=\; \int_{\mathbb{T}} w_n(\tau)\; \mathbf{Q}(d\tau) \;=\; \mathbb{E}(w_n \circ \boldsymbol{GW}) \;=\; \mathbb{E}W_n \;=\; 1$$

tatsächlich normiert ist. Teil (c) des Vergleichslemmas 3.7 liefert für $A = [\tau]_n \in \mathcal{E}_n$

$$\widehat{\mathbf{Q}}(A) \;=\; w_n(\tau)\mathbf{Q}([\tau]_n) \;=\; \int_{[\tau]_n} w_n(\tau)\;\mathbf{Q}(d\chi) \;=\; \int_A w_n(\chi)\;\mathbf{Q}(d\chi) \;=\; \Gamma_n(A),$$

denn für $\chi \in [\tau]_n$ gilt offenbar $w_n(\chi) = w_n(\tau)$. Somit stimmen die W-Maße Γ_n und $\widehat{\mathbf{Q}}$ auf $\mathcal{E}_n$ überein. Da jedoch $\mathcal{E}_n$ ein $\mathbb{T}$ enthaltendes, $\cap$-stabiles Mengensystem bildet, wie man mit (1.1) sofort überprüft, gilt bereits $\Gamma_n = \widehat{\mathbf{Q}}$ auf ganz $\mathcal{A}_n = \sigma(\mathcal{E}_n)$ (☞ Satz 3.8 in [1]).

Insgesamt erhalten wir

$$\widehat{\mathbf{Q}}_{|\mathcal{A}_n} \ll \mathbf{Q}_{|\mathcal{A}_n}$$

mit

$$\frac{d\widehat{\mathbf{Q}}_{|\mathcal{A}_n}}{d\mathbf{Q}_{|\mathcal{A}_n}} = w_n.$$

Dabei ist noch anzumerken, daß für $k \geq 1$

$$z_n^{-1}(\{k\}) \;=\; \{\tau \in \mathbb{T} : z_n(\tau) = k\} \;=\; \bigcup_{\tau \in \mathbb{T}_n : |\tau_n| = k} [\tau]_n \;\in\; \mathcal{A}_n$$

gilt (abzählbare Vereinigung) und deshalb z_n ebenso wie w_n $\mathcal{A}_n$-meßbar ist.

Für den restlichen Beweis setzen wir $w \stackrel{\text{def}}{=} \limsup_{n\to\infty} w_n$, so daß $W = w \circ \boldsymbol{GW}$ $\mathbb{P}$-f.s. gilt. Nunmehr sind alle Voraussetzungen von Lemma 5.1 erfüllt, aus dessen Teilen (b) und (c) sich deshalb die entscheidende Dichotomie ergibt: Wir haben einerseits

$$\int_{\mathbb{T}} w\, d\mathbf{Q} = 1 \quad \Leftrightarrow \quad \widehat{\mathbf{Q}} \ll \mathbf{Q} \quad \Leftrightarrow \quad w < \infty\; \widehat{\mathbf{Q}}\text{-f.s.} \tag{5.5}$$

und andererseits

$$w = 0\; \mathbf{Q}\text{-f.s.} \quad \Leftrightarrow \quad \widehat{\mathbf{Q}} \perp \mathbf{Q} \quad \Leftrightarrow \quad w = \infty\; \widehat{\mathbf{Q}}\text{-f.s.} \tag{5.6}$$

Diese Dichotomie wird uns helfen, weil sie die Verteilung von w unter $\mathbf{Q}$ mit der unter $\widehat{\mathbf{Q}}$ in Verbindung bringt, zu deren Untersuchung sich im folgenden Satz 4.5 heranziehen läßt:

"$(ZlogZ) \Rightarrow (5.4)$" Die Integrierbarkeit von $Z_1 \log^+ Z_1$ impliziert offensichtlich (unter Hinweis auf (3.4))

$$\mathbb{E}\log^+(\widehat{X}_1 - 1) \;=\; \mu^{-1}\mathbb{E}Z_1\log^+(Z_1 - 1) \;<\; \infty.$$

Bezeichnet nun $(Y_n)_{n\geq 0}$ einen GWPI mit Reproduktionsverteilung $(p_j)_{j\geq 0}$ und Immigrationsfolge $\zeta_n = \widehat{X}_n - 1$, $n \geq 1$, so zeigt Satz 4.5

$$Y_\infty \stackrel{\text{def}}{=} \lim_{n\to\infty} \mu^{-n} Y_n \;<\; \infty \quad \mathbb{P}\text{-f.s.}$$

Da die Verteilungen $\mathbb{P}((\mu^{-n}Y_n)_{n\geq 0} \in \cdot)$ und $\widehat{\mathbf{Q}}((w_n - \mu^{-n})_{n\geq 0} \in \cdot)$ gemäß Satz 4.1 übereinstimmen, gilt unter Beachtung von $\mu^{-n} \to 0$ offenbar auch

$$\widehat{\mathbf{Q}}(w < \infty) \;=\; \widehat{\mathbf{Q}}\Big(\limsup_{n\to\infty}(w_n - \mu^{-n}\Big) \;=\; 1,$$

was uns mittels (5.5) schließlich zu

$$1 \;=\; \int_{\mathbb{T}} w \; d\mathbf{Q} \;=\; \int_{\Omega} w \circ \boldsymbol{GW} \; d\mathbb{P} \;=\; \mathbb{E}W$$

und somit (5.4) bringt.

"(5.4) $\Rightarrow$ (5.3)" Im Fall $\mathbb{E}W = 1$ folgt notwendigerweise $\mathbb{P}(W = 0) < 1$, also unter Rückgriff auf Lemma I.5.2 wie behauptet $\mathbb{P}(W = 0) = q$.

"(5.3) $\Rightarrow$ ($ZlogZ$)" Dies zeigen wir durch Kontraposition, nehmen also $\mathbb{E}Z_1 \log^+ Z_1 = \infty$ an. Eine analoge Argumentation wie im Beweisteil "($ZlogZ$) $\Rightarrow$ (5.4)" liefert dann unter Benutzung der Sätze 4.1 und 4.5

$$\widehat{\mathbf{Q}}(w = \infty) \;=\; \mathbb{P}\Big(\limsup_{n\to\infty} \mu^{-n} Y_n = \infty\Big) \;=\; 1,$$

also unter Berücksichtigung von (5.6) die gewünschte Aussage $\mathbf{Q}(w = 0) = 1$. ◇

Übungen zu Abschnitt 5

Übung 5.1. Beweisen Sie die Teile (b) und (c) von Lemma 5.1 [natürlich unter Benutzung von Teil (a)].

6. Zum Grenzverhalten subkritischer GWP

Im subkritischen Fall ($\mu < 1$) gibt die Abschätzung

$$\mathbb{P}(Z_n > 0) \;\le\; \mathbb{E}Z_n \;=\; \mu^n \;\to\; 0, \quad n \to \infty$$

einen ersten Anhaltspunkt für die Konvergenzgeschwindigkeit der Überlebenswahrscheinlichkeit $\mathbb{P}(Z_n > 0)$ gegen 0. Wie schon in I.8, aber mit anderen Methoden, gehen wir nun der Frage nach, unter welchen Bedingungen μ^n die richtige Konvergenzrate von $\mathbb{P}(Z_n > 0)$ bildet. Dies legt nahe, die Folgen

$$c_n \;\stackrel{\text{def}}{=}\; \mu^{-n}\mathbb{P}(Z_n > 0), \quad n \ge 0$$

und

$$\mu_n^+ \;\stackrel{\text{def}}{=}\; \mathbb{E}(Z_n | Z_n > 0) \;=\; c_n^{-1}, \quad n \ge 0$$

zu untersuchen. Die bereits bekannte Antwort, Inhalt des Satzes I.8.1 (von Kolmogorov), geben wir in Satz 6.2 nach dem folgenden einfachen Lemma über größenverzerrte Verteilungen.

6.1. Lemma. *Sei $(X_n)_{n\ge 0}$ eine Folge von integrierbaren $\mathbb{N}$-wertigen Zufallsgrößen und $(\widehat{X}_n)_{n\ge 0}$ eine assoziierte Folge von Größenverzerrungen. Dann gilt mit $\mathbf{P}_n \stackrel{\text{def}}{=} \mathbb{P}(X_n \in \cdot)$, folglich $\widehat{\mathbf{P}}_n = \mathbb{P}(\widehat{X}_n \in \cdot)$:*

(a) Ist die Folge $(\widehat{\mathbf{P}}_n)_{n\ge 0}$ straff, folgt $\sup_{n\ge 0} \mathbb{E}X_n < \infty$.

(b) *Gilt dagegen* $\widehat{X}_n \xrightarrow{\mathbb{P}} \infty$, *also* $\lim_{n\to\infty} \mathbb{P}(\widehat{X}_n \le t) = 0$ *für alle* $t > 0$, *so ist* $\lim_{n\to\infty} \mathbb{E}X_n = \infty$.

Beweis: (a) Die Straffheit der $\widehat{\mathbf{P}}_n$ garantiert die Existenz eines $N \in \mathbb{N}$, so daß

$$\inf_{n\ge 0} \mathbb{P}(\widehat{X}_n \in \{1,...,N\}) \;=\; \inf_{n\ge 0} \frac{1}{\mathbb{E}X_n} \sum_{k=1}^{n} k\mathbf{P}_n(\{k\}) \;\ge\; \frac{1}{2}$$

und damit

$$\frac{1}{N} \sup_{n\ge 0} \mathbb{E}X_n \;\le\; \sup_{n\ge 0} \left(\mathbb{E}X_n \Big/ \sum_{k=1}^{N} k\mathbf{P}_n(\{k\}) \right) \;\le\; 2.$$

(b) Nehmen wir an, daß $\liminf_{n\to\infty} \mathbb{E}X_n < \infty$, also o.B.d.A. (nach Übergang zu einer Teilfolge) $M \stackrel{\text{def}}{=} \sup_{n\ge 0} \mathbb{E}X_n < \infty$. Da dann für jedes $N \ge 1$

$$\lim_{n\to\infty} \sum_{k=1}^{N} k\mathbf{P}_n(\{k\}) \;\le\; \lim_{n\to\infty} \frac{M}{\mathbb{E}X_n} \sum_{k=1}^{N} k\mathbf{P}_n(\{k\}) \;=\; M \lim_{n\to\infty} \mathbb{P}(\widehat{X}_n \le N) \;=\; 0$$

gilt, erhalten wir [1)]

$$\lim_{n\to\infty} \mathbb{P}(X_n \le N) \;\le\; \lim_{N\to\infty} \sum_{k=1}^{N} k\mathbf{P}_n(\{k\}) \;=\; 0.$$

Folglich existiert zu jedem $N \ge 1$ ein $k_N \ge 1$ derart, daß $\mathbb{P}(X_{k_N} > N) \ge \frac{1}{2}$ und somit auch $\mathbb{E}X_{k_N} \ge N/2$ gilt, was unserer Annahme $M < \infty$ offenkundig widerspricht. ♢

6.2. Satz. *Für jeden GWP* $(Z_n)_{n\ge 0}$ *mit Reproduktionsmittel* $\mu \in (0,\infty)$ *ist die Folge* $c_n = \mu^{-n}\mathbb{P}(Z_n > 0)$, $n \ge 0$, *monoton fallend und folglich konvergent. Im subkritischen Fall* $\mu < 1$ *sind außerdem die folgenden Aussagen äquivalent zur* $(Z\log Z)$*-Bedingung:*

$$c \;\stackrel{\text{def}}{=}\; \lim_{n\to\infty} c_n \;>\; 0, \tag{6.1}$$

$$\sup_{n\ge 0} \mu_n^+ \;<\; \infty. \tag{6.2}$$

Beweis: Die Äquivalenz von (6.1) und (6.2) ist wegen $\mu_n^+ = \mathbb{E}(Z_n|Z_n > 0) = c_n^{-1}$ klar, sobald wir die Monotonie der c_n (oder der μ_n^+) nachgewiesen haben.

Wir nehmen wieder an, daß Z_n in der Form $z_n \circ \boldsymbol{GW}$ mit einem GWB $\boldsymbol{GW}$ vorliegt, und schreiben abkürzend $\mathbf{P}_n \stackrel{\text{def}}{=} \mathbb{P}(Z_n \in \cdot|Z_n > 0)$ und $Z_n(v) \stackrel{\text{def}}{=} z_n \circ \Theta_v \circ \boldsymbol{GW}^v$ für $n \ge 0$ und $v \in \boldsymbol{GW}$. $Z_n(v)$ gibt demnach die Größe der n-ten Generation des in $v \in \boldsymbol{GW}$ verwurzelten (geshifteten) Teilbaums $\Theta_v \circ \boldsymbol{GW}^v$ (eine Kopie von $\boldsymbol{GW}$) an. Auf $\{Z_n > 0\}$ setzen wir für $n \ge 1$ weiter

$$\varrho_n \;\stackrel{\text{def}}{=}\; \inf \underbrace{\{v \in \boldsymbol{GW}_1 : Z_{n-1}(v) > 0\}}_{\neq\emptyset \text{ auf } \{Z_n>0\}}$$

[1)] Dies läßt sich übrigens auch direkt aus Übung 3.2 folgern.

und dann

$$H_n \stackrel{\text{def}}{=} Z_{n-1}(\varrho_n)\,\mathbf{1}_{\{Z_n>0\}}.$$

Folglich ist ϱ_n das kleinste Individuum der ersten Generation mit Nachfahren in der n-ten Generation, während H_n die Anzahl der Nachfahren desselben angibt. Eine einfache Rechnung liefert für $k,n\ge 1$

$$\begin{aligned}
\mathbb{P}(H_n=k) &= \sum_{i\ge 1}\mathbb{P}(\varrho_n=i, Z_n>0, H_n=k)\\
&= \sum_{i\ge 1}\mathbb{P}(Z_1\ge i, Z_{n-1}(l)=0 \text{ für } l<i, Z_{n-1}(i)=k)\\
&= \sum_{i\ge 1}\sum_{j\ge i:p_j>0} p_j\,\mathbb{P}(Z_{n-1}(l)=0 \text{ für } l<i, Z_{n-1}(i)=k|Z_1=j)\\
&= \sum_{i\ge 1}\sum_{j\ge i:p_j>0} p_j\,\mathbb{P}(Z_{n-1}=k)\,\mathbb{P}(Z_{n-1}=0)^{i-1}\\
&= \mathbb{P}(Z_{n-1}=k)\sum_{i\ge 1}\mathbb{P}(Z_1\ge i)\,\mathbb{P}(Z_{n-1}=0)^{i-1},
\end{aligned}$$

wobei für die vorletzte Zeile einmal mehr Satz 2.3 eingegangen ist. Es folgt nun

$$\mathbb{P}(Z_n>0) = \sum_{k\ge 1}\mathbb{P}(H_n=k) = \mathbb{P}(Z_{n-1}>0)\sum_{i\ge 1}\mathbb{P}(Z_1\ge i)\,\mathbb{P}(Z_{n-1}=0)^{i-1} \tag{6.3}$$

und schließlich per Division beider Ergebnisse für $k\ge 1$

$$\mathbb{P}(H_n=k|Z_n>0) = \frac{\mathbb{P}(H_n=k)}{\mathbb{P}(Z_n>0)} = \frac{\mathbb{P}(Z_{n-1}=k)}{\mathbb{P}(Z_{n-1}>0)} = \mathbb{P}(Z_{n-1}=k|Z_{n-1}>0),$$

d.h. $\mathbb{P}(H_n\in\cdot|Z_n>0)=\mathbf{P}_{n-1}$ für $n\ge 1$.

Wegen $H_n\le Z_n$ erhalten wir nun für $k,n\ge 1$

$$\begin{aligned}
\mathbf{P}_n([k,\infty)) &= \mathbb{P}(Z_n\ge k|Z_n>0)\\
&\ge \mathbb{P}(H_n\ge k|Z_n>0)\\
&= \mathbb{P}(Z_{n-1}\ge k|Z_{n-1}>0)\\
&= \mathbf{P}_{n-1}([k,\infty))
\end{aligned}$$

und daraus weiter

$$\mathbb{E}(Z_n|Z_n>0) = \sum_{k\ge 1}\mathbf{P}_n([k,\infty)) \ge \sum_{k\ge 1}\mathbf{P}_{n-1}([k,\infty)) = \mathbb{E}(Z_{n-1}|Z_{n-1}>0),$$

also die Monotonie der Folge $(\mu_n^+)_{n\ge 0}$.

Für den restlichen Beweis setzen wir $\mu<1$ voraus und bringen den größenverzerrten GWP $\big(\widehat{Z}_n\big)_{n\ge 0}$ ins Spiel, wobei an $\widehat{Z}_n = z_n\circ\widehat{\boldsymbol{GW}}$ erinnert sei. Bezeichnet wie üblich $\widehat{\mathbf{P}}_n$ die

Größenverzerrung von $\mathbf{P}_n = \mathbb{P}(Z_n \in \cdot | Z_n > 0)$, so folgt $\widehat{Z}_n \stackrel{d}{=} \widehat{\mathbf{P}}_n$ gemäß Teil (d) des Vergleichslemmas 3.7 und Übung 6.2, also gemäß Satz 4.1 auch

$$\widehat{\mathbf{P}}_n \;=\; \mathbb{P}(1 + Y_n \in \cdot) \tag{6.4}$$

für alle $n \geq 0$, sofern $(Y_n)_{n\geq 0}$ weiterhin einen GWPI mit Immigrationsfolge $\zeta_n = \widehat{X}_n - 1$ für $n \geq 1$ bezeichnet.

Wir zeigen als nächstes "(6.2) $\Rightarrow$ $(ZlogZ)$": Ist die Folge $(\mu_n^+)_{n\geq 0}$ beschränkt, so zeigt Lemma 6.1(b) unmittelbar, daß $\widehat{Z}_n$ nicht nach Wahrscheinlichkeit gegen ∞ streben kann. Vermöge (6.4) folgt dann dasselbe auch für Y_n, so daß Satz 4.7(b) in Kombination mit Bemerkung 3.2(b)

$$\mathbb{E} \log^+ \zeta_1 \;=\; \mathbb{E} \log^+(\widehat{X}_1 - 1) \;=\; \mu^{-1} \mathbb{E} Z_1 \log^+(Z_1 - 1) \;<\; \infty \tag{6.5}$$

und damit $(ZlogZ)$ liefert.

Für "$(ZlogZ)$ $\Rightarrow$ (6.2)" argumentieren wir folgendermaßen: Mit $Z_1 \log^+ Z_1$ ist gemäß (6.5) auch $\log^+ \zeta_1$ integrierbar. Deshalb konvergiert Y_n (und damit natürlich auch $Y_n + 1$) gemäß Satz 4.7(a) in Verteilung gegen eine endliche Zufallsgröße, was vermöge (6.4) dasselbe für $\widehat{Z}_n$ garantiert. Dies sichert insbesondere die Straffheit der zugehörigen Verteilungsfolge, also von $(\widehat{\mathbf{P}}_n)_{n\geq 0}$. Unter nochmaliger Anwendung von Lemma 6.1 folgern wir schließlich

$$\sup_{n\geq 0} \int x \, \mathbf{P}_n(dx) \;=\; \sup_{n\geq 0} \mathbb{E}(Z_n | Z_n > 0) \;<\; \infty,$$

also das Gewünschte. ♢

Im folgenden bezeichne $\|\mathbf{P} - \mathbf{Q}\|$ den Variationsabstand zweier Verteilungen $\mathbf{P}$ und $\mathbf{Q}$ auf $\mathbb{N}_0$ (☞ Anhang für eine Zusammenstellung der wichtigsten Fakten über $\| \cdot \|$). Wir zeigen als nächstes die Konvergenz in Totalvariation von $\mathbf{P}_n = \mathbb{P}(Z_n \in \cdot | Z_n > 0)$ und damit eine Verschärfung des Satzes I.8.2 (von Yaglom), der lediglich schwache Konvergenz aussagt.

6.3. Satz. *Gegeben einen subkritischen GWP $(Z_n)_{n\geq 0}$ mit $\mu \in (0,1)$, gilt für die (bedingten) Verteilungen $\mathbf{P}_n = \mathbb{P}(Z_n \in \cdot | Z_n > 0)$, $n \geq 0$*

$$\sum_{n\geq 1} \|\mathbf{P}_n - \mathbf{P}_{n-1}\| \;<\; \infty. \tag{6.6}$$

Insbesondere konvergiert $\mathbf{P}_n$ in Totalvariation gegen eine Verteilung π auf $\mathbb{N}_0$.

Beweis: Im Beweis von Satz 6.2, dessen Bezeichnungen hier natürlich Bestand haben, haben wir gezeigt, daß die bedingte Verteilung der dort definierten Zufallsvariablen H_n gegeben $Z_n > 0$ durch $\mathbf{P}_{n-1}$ gegeben ist. Folglich bildet (H_n, Z_n) unter $\mathbb{P}(\cdot | Z_n > 0)$ eine Kopplung von $(\mathbf{P}_{n-1}, \mathbf{P}_n)$, und wir erhalten vermöge der Kopplungsungleichung A.3 im Anhang

$$\|\mathbf{P}_n - \mathbf{P}_{n-1}\| \;\leq\; \mathbb{P}(H_n \neq Z_n | Z_n > 0).$$

Da das Ereignis $\{H_n \neq Z_n\} = \sum_{i\ge 1}\{\varrho_n = i, H_n \neq Z_n\}$ in der Form

$$\sum_{i\ge 1}\sum_{j\ge i}\left\{Z_1 = j, \sum_{u=1}^{i-1} Z_{n-1}(u) = 0, Z_{n-1}(i) > 0, \sum_{u=i+1}^{j} Z_{n-1}(u) > 0\right\}$$

geschrieben werden kann, zeigt eine erneute Anwendung von Satz 2.3

$$\begin{aligned}\|\mathbf{P}_n - \mathbf{P}_{n-1}\| &\le \mathbb{P}(H_n \neq Z_n | Z_n > 0)\\ &= \frac{\mathbb{P}(Z_{n-1} > 0)}{\mathbb{P}(Z_n > 0)}\sum_{i\ge 1}\sum_{j\ge i} p_j\, \mathbb{P}(Z_{n-1} = 0)^{i-1}\big(1 - \mathbb{P}(Z_{n-1} = 0)^{j-i}\big).\end{aligned}$$

Setzen wir

$$\alpha(k) \stackrel{\text{def}}{=} \min\{n \ge 1 : \mathbb{P}(Z_n > 0) < 1/k\}, \quad k \ge 1,$$

und beachten, daß $\delta \stackrel{\text{def}}{=} \sup_{n\ge 1}\frac{\mathbb{P}(Z_{n-1}>0)}{\mathbb{P}(Z_n>0)} < \infty$ wegen

$$\frac{\mathbb{P}(Z_n > 0)}{\mathbb{P}(Z_{n-1} > 0)} = \mathbb{E}(1 - p_0^{Z_{n-1}} | Z_{n-1} > 0) \ge 1 - p_0$$

für alle $n \ge 1$ gilt, so gelangen wir durch geeignete Umordnung zu der Abschätzung

$$\sum_{n\ge 1}\|\mathbf{P}_n - \mathbf{P}_{n-1}\| \le \delta\sum_{n\ge 0}\sum_{i\ge 1}\sum_{j\ge i} p_j\, \mathbb{P}(Z_n = 0)^{i-1}\big(1 - \mathbb{P}(Z_n = 0)^{j-i}\big) = \delta(I_1 + I_2),$$

wobei

$$\begin{aligned}I_1 &\stackrel{\text{def}}{=} \sum_{j\ge 1} p_j \sum_{n=0}^{\alpha(j)-1}\sum_{i=1}^{j}\mathbb{P}(Z_n = 0)^{i-1}\big(1 - \mathbb{P}(Z_n = 0)^{j-i}\big) \quad \text{und}\\ I_2 &\stackrel{\text{def}}{=} \sum_{j\ge 1} p_j \sum_{n\ge\alpha(j)}\sum_{i=1}^{j}\mathbb{P}(Z_n = 0)^{i-1}\big(1 - \mathbb{P}(Z_n = 0)^{j-i}\big).\end{aligned}$$

Die Terme I_1 und I_2 werden nun getrennt abgeschätzt. Unter Benutzung der aus der Monotonie der c_n resultierenden Ungleichung

$$\mathbb{P}(Z_{\alpha(k)-1} > 0) \le \mu^{\alpha(k)-1-n}\, \mathbb{P}(Z_n > 0)$$

für $0 \le n < \alpha(k)$ ergibt sich

$$\begin{aligned}I_1 &\le \sum_{j\ge 1} p_j \sum_{n=0}^{\alpha(j)-1}\sum_{i=1}^{j}\mathbb{P}(Z_n = 0)^{i-1}\\ &\le \sum_{j\ge 1} p_j \sum_{n=0}^{\alpha(j)-1}\frac{1}{\mathbb{P}(Z_n > 0)}\end{aligned}$$

$$\begin{aligned}
&\le \sum_{j\ge 1} p_j \sum_{n=1}^{\alpha(j)-1} \mu^{\alpha(j)-1-n} \underbrace{\frac{1}{\mathbb{P}(Z_{\alpha(j)-1}>0)}}_{\le j} \\
&\le \sum_{j\ge 1} j p_j \sum_{n\ge 0} \mu^n \\
&= \frac{\mu}{1-\mu} \; < \; \infty.
\end{aligned}$$

Eine ähnliche Rechnung unter Benutzung von

$$\mathbb{P}(Z_n > 0) \;\le\; \mu^{n-\alpha(k)}\, \mathbb{P}(Z_{\alpha(k)} > 0)$$

für $n \ge \alpha(k)$, die sich natürlich erneut aus der Monotonie der c_n ergibt, sowie der für $x \in [0,1]$ und $j \ge 0$ gültigen Abschätzung $1-(1-x)^j \le jx$ liefert

$$\begin{aligned}
I_2 &\le \sum_{j\ge 1} p_j \sum_{n\ge \alpha(j)} \sum_{i=1}^{j} \big(1-\mathbb{P}(Z_n=0)^{j-i}\big) \\
&\le \sum_{j\ge 1} p_j \sum_{n\ge \alpha(j)} \mathbb{P}(Z_n>0) \sum_{i=1}^{j} (j-i) \\
&\le \sum_{j\ge 1} j^2 p_j \sum_{n\ge 0} \mu^n \underbrace{\mathbb{P}(Z_{\alpha(j)}>0)}_{\le 1/j} \\
&\le \frac{\mu}{1-\mu} \; < \; \infty.
\end{aligned}$$

Insgesamt haben wir also

$$\sum_{n\ge 1} \|\mathbf{P}_n - \mathbf{P}_{n-1}\| \;\le\; \delta(I_1+I_2) \;<\; \infty,$$

also (6.6) verifiziert. Insbesondere bildet $(\mathbf{P}_n)_{n\ge 0}$ damit eine Cauchy-Folge bezüglich $\|\cdot\|$ und konvergiert folglich gegen in Totalvariation einen Limes π (☞ Satz A.4). $\diamondsuit$

Für Aussagen zur Konvergenzgeschwindigigkeit von $\mathbf{P}_n$ gegen π verweisen wir auf Übung 6.4., während die Eigenschaft der *Quasi-Invarianz* von π in Übung 6.5 behandelt wird.

Übungen zu Abschnitt 6

Übung 6.1. Beweisen Sie die folgende Verschärfung von Lemma 6.1(a): Genau dann ist die Familie $(\widehat{\mathbf{P}}_n)_{n\ge 0}$ straff, wenn die Folge $(X_n)_{n\ge 0}$ gleichgradig integrierbar ist. [Hinweis: Zeigen Sie zuerst, daß die Straffheit von $(\widehat{\mathbf{P}}_n)_{n\ge 0}$ äquivalent ist zur Existenz einer monoton wachsenden, unbeschränkten Funktion $\varphi : [0,\infty) \to [0,\infty)$ mit

$$\sup_{n\ge 0} \mathbb{E}\varphi(\widehat{X}_n) \;<\; \infty.$$

Benutzen sie anschließend (3.4) und Satz 50.2 in [1].]

Übung 6.2. Sei X eine nichtnegative, integrierbare Zufallsgröße mit Verteilung $\mathbf{P}$ und $\mathbf{Q} \stackrel{\text{def}}{=} \mathbb{P}(X \in \cdot | X > 0)$. Zeigen Sie, daß die zugehörigen Größenverzerrungen übereinstimmen, also $\widehat{\mathbf{P}} = \widehat{\mathbf{Q}}$ gilt.

Übung 6.3. Gegeben sei die Situation und Notation von Satz 6.2 und dessen Beweis. Sei ferner $\widetilde{Z}_1$ eine Zufallsgröße mit $\mathbb{P}(\widetilde{Z}_1 = i) = \mu^{-1}\mathbb{P}(Z_1 \geq i)$ für $i \in \mathbb{N}$. Zeigen Sie mittels (6.3), daß

$$c_n - c_{n+1} \;=\; c_n\Big(1 - \mathbb{E}(1-\mu^n c_n)^{\widetilde{Z}_1 - 1}\Big) \;\leq\; \mu^n\, \mathbb{E}(\widetilde{Z}_1 - 1)^2$$

gilt (was insbesondere die Monotonie der c_n zeigt) und folgern Sie daraus weiter

$$c_n \;=\; c + O(\mu^n), \quad n \to \infty,$$

falls $(Z_n)_{n\geq 0}$ endliche Reproduktionsvarianz besitzt.

Übung 6.4. Gegeben sei die Situation und Notation von Satz 6.3 und dessen Beweis. Zeigen Sie:

(a) Aus $(ZlogZ)$ folgt $\alpha(k) \simeq \log_{1/\mu} k = \log k / |\log \mu|$ für $k \to \infty$.

(b) Für jedes $p \geq 1$ gilt

$$\sum_{j\geq 1} j \log^p j\, p_j < \infty \quad \Rightarrow \quad \sum_{n\geq 1} n^p \|\mathbf{P}_n - \mathbf{P}_{n-1}\| < \infty \quad \Rightarrow \quad \lim_{n\to\infty} n^p \|\mathbf{P}_n - \pi\| = 0.$$

(c) Für jedes $1 < \gamma < \mu$ gilt

$$\sum_{j\geq 1} j^{\gamma+1} p_j < \infty \quad \Rightarrow \quad \sum_{n\geq 1} \gamma^n \|\mathbf{P}_n - \mathbf{P}_{n-1}\| < \infty \quad \Rightarrow \quad \lim_{n\to\infty} \gamma^n \|\mathbf{P}_n - \pi\| = 0.$$

Übung 6.5. Gegeben sei weiterhin die Situation von Satz 6.3, und es bezeichne

$$p_{ij}^{(k)} = \mathbb{P}(Z_{n+1} = j | Z_n = i) \;=\; \mathbb{P}_i(Z_1 = j)$$

wieder die k-Schritt Übergangswahrscheinlichkeiten von $(Z_n)_{n\geq 0}$ für $i, j, k \in \mathbb{N}_0$ (☞ I.1). Eine Verteilung $\pi = (\pi_i)_{i\geq 0}$ heißt

— *β-invariant* für $(Z_n)_{n\geq 0}$ für ein $0 < \beta < 1$, falls $\beta\pi_j = \sum_{i\geq 1} \pi_i p_{ij}$ für alle $j \in \mathbb{N}_0$ gilt (wegen $p_{00} = 1$ folgt daraus zwingend $\pi_0 = 1$).

— *quasi-invariant*, falls $\pi = \mathbb{P}_\pi(Z_n \in \cdot | Z_n > 0)$ für alle $n \in \mathbb{N}$, d.h.

$$\pi_j \;=\; \frac{\mathbb{P}_\pi(Z_n = j)}{\mathbb{P}_\pi(Z_n > 0)} \;=\; \frac{\sum_{i\geq 1} \pi_i p_{ij}^{(n)}}{\sum_{k\geq 1}\sum_{i\geq 1} \pi_i p_{ik}^{(n)}}$$

für alle $j, n \in \mathbb{N}$ gilt.

— *quasi-stationär*, falls $\mathbb{P}_\lambda(Z_n \in \cdot|Z_n > 0) \xrightarrow{w} \pi$ für eine Anfangsverteilung λ gilt, d.h.

$$\lim_{n\to\infty} \mathbb{P}_\lambda(Z_n = j|Z_n > 0) \; = \; \pi_j \tag{6.7}$$

für alle $j \in \mathbb{N}$ gilt.

Zeigen Sie für die Limesverteilung π aus Satz 6.3, die offensichtlich quasi-stationär ist (mit $\lambda = \delta_1$):

(a) π ist μ-invariant.

(b) π ist quasi-invariant.

(c) π erfüllt (6.7) unter jeder Anfangsverteilung λ mit $\lambda_0 = 0$. [Hinweis: Zeigen Sie die Behauptung für jedes $\lambda = \delta_k$, $k \in \mathbb{N}$, per Induktion über k.]

(d) π ist die einzige quasi-invariante Verteilung für $(Z_n)_{n\ge 0}$.

7. Und schließlich kritische GWP

Wir kommen schließlich zur Untersuchung kritischer GWP unter Benutzung größenverzerrter Bäume mit Rückgrat, wobei die folgende Zerlegung von $\widehat{Z}_n = z_n \circ \widehat{\boldsymbol{GW}}$ grundlegend ist: Unter Beibehaltung der bisherigen Notation sei

$$\widehat{Z}_n(v) \;\stackrel{\text{def}}{=}\; z_n \circ \Theta_v \circ \widehat{\boldsymbol{GW}}^{\,v}, \quad n \ge 0$$

für $v \in \widehat{\boldsymbol{GW}}$ und dann für $(j,n) \in \mathbb{N}^2_{\le} \stackrel{\text{def}}{=} \{(k,l) \in \mathbb{N}^2 : k \le l\}$

$$S_{n,j} \;\stackrel{\text{def}}{=}\; \sum_{\varrho \in \widehat{\boldsymbol{GW}}_j : V^{j-1} \preceq \varrho \ne V^j} \widehat{Z}_{n-j}(\varrho) \;=\; \widehat{Z}_{n-j+1}(V^{j-1}) - \widehat{Z}_{n-j}(V^j)$$

sowie

$$R_{n,j} \;\stackrel{\text{def}}{=}\; \sum_{\varrho \in \widehat{\boldsymbol{GW}}_j : V^{j-1} \preceq \varrho, V^j < \varrho} \widehat{Z}_{n-j}(\varrho).$$

Damit gibt $S_{n,j}$ die Anzahl der Nachfahren von V^{j-1} in $\widehat{\boldsymbol{GW}}_n$ an, die keine Nachfahren von V^j sind, während $R_{n,j}$ die Anzahl der Nachfahren von V^{j-1} in $\widehat{\boldsymbol{GW}}_n$ beschreibt, deren (eindeutig bestimmter) Vorfahre in Generation j des größenverzerrten GWB's sich rechts von V^j befindet, also größer als V^j ist.

$\widehat{Z}_n$ besitzt nun offenbar die Darstellung (Teleskop-Summe)

$$\widehat{Z}_n \;=\; 1 + \widehat{Z}_n(V^0) - \widehat{Z}_0(V^n) \;=\; 1 + \sum_{j=1}^{n} S_{n,j}, \tag{7.1}$$

und es gilt $R_{n,j} \le S_{n,j}$ für $(j,n) \in \mathbb{N}^2_{\le}$. Setzen wir noch

$$A_{n,j} \;\stackrel{\text{def}}{=}\; \{R_{n,j} = S_{n,j}\}$$

für $(j,n) \in \mathbb{N}^2_\leq$, so hält das folgende Lemma weitere wichtige Eigenschaften der zuvor eingeführten Zufallsgrößen fest.

7.1. Lemma. *Sei $(Z_n)_{n\geq 0}$ ein kritischer GWP mit Reproduktionsvarianz $\sigma^2 > 0$. Dann gilt in den zuvor eingeführten Bezeichnungen:*

(a) Für jedes $n \geq 1$ sind die Zufallsvektoren $(R_{n,1}, S_{n,1}), ..., (R_{n,n}, S_{n,n})$ und damit ebenso die Ereignisse $A_{n,1}, ..., A_{n,n}$ stochastisch unabhängig.

(b) Für jedes $(j,n) \in \mathbb{N}^2_\leq$ ist $\mathbb{E}R_{n,j} = \sigma^2/2$.

(c) Für jedes $(j,n) \in \mathbb{N}^2_\leq$ ist $\mathbb{P}(A_{n,j}) = \frac{\mathbb{P}(Z_{n-j+1}>0)}{\mathbb{P}(Z_{n-j}>0)}$ und konvergiert gegen 1, falls $n-j \to \infty$.

(d) Für jedes $j \geq 1$ gilt

$$\lim_{n\to\infty} \mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}} \;=\; \frac{\sigma^2}{2}$$

und im Fall $\sigma^2 < \infty$ außerdem

$$\lim_{n\to\infty} \mathbb{E}R_{n,j}\mathbf{1}_{A^c_{n,j}} \;=\; 0.$$

Beweis: (a) Für $1 \leq j \leq n$ seien $(r_j, s_j) \in \mathbb{N}^2_\leq$,

$$M_j \stackrel{\text{def}}{=} \{\varrho \in \widehat{\boldsymbol{GW}}_j : \varrho \succeq V^{j-1},\ \varrho < V^j\}$$

und

$$N_j \stackrel{\text{def}}{=} \{\varrho \in \widehat{\boldsymbol{GW}}_j : \varrho \succeq V^{j-1},\ \varrho > V^j\}.$$

Folglich enthält M_j bzw. N_j alle Kinder des Rückgratknotens V^{j-1}, die kleiner bzw. größer als V^j sind, sich also links bzw. rechts von V^j in $\widehat{\boldsymbol{GW}}$ befinden.

Auf dem Ereignis $D_n \stackrel{\text{def}}{=} \bigcap_{i=0}^{n-1}\{\widehat{X}_i = k_i, C_i = v_i\}$ gilt einerseits

$$V^j \;=\; v_0...v_{j-1} \;\in\; \mathbb{N}^j,$$

andererseits sind im Fall $\mathbb{P}(D_n) = \prod_{i=0}^{n-1} p_{k_i} > 0$ die Zufallsgrößen

$$\widehat{Z}_{n-j}(v), \quad 1 \leq j \leq n,\ v \in M_j \cup N_j$$

bedingt unter D_n stochastisch unabhängig und für festes j identisch nach $\Gamma_{n-j} \stackrel{\text{def}}{=} \mathbb{P}(Z_{n-j} \in \cdot)$ verteilt, wie sich aus der Konstruktion 3.3 größenverzerrter GWB in Analogie zu Satz 2.3 ergibt. Da für jedes j offenbar

$$\{R_{n,j} = r_j, S_{n,j} = s_j\} \;=\; \left\{\sum_{v\in M_j} \widehat{Z}_{n-j}(v) = s_j - r_j,\ \sum_{v\in N_j} \widehat{Z}_{n-j}(v) = r_j\right\}$$

gilt, folgt

$$\begin{aligned}
&\mathbb{P}\Big(R_{n,j} = r_j, S_{n,j} = s_j,\, 1 \le j \le n\Big) \\
&\qquad = \sum_{*} \prod_{i=0}^{n-1} p_{k_i} \prod_{j=1}^{n} \Gamma_{n-j}^{*(v_{j-1}-1)}(\{s_j - r_j\})\Gamma_{n-j}^{*(k_{j-1}-v_{j-1})}(\{r_j\}) \\
&\qquad = \prod_{j=1}^{n} \sum_{1\le v_{j-1}\le k_{j-1}} p_{k_{j-1}} \Gamma_{n-j}^{*(v_{j-1}-1)}(\{s_j - r_j\})\Gamma_{n-j}^{*(k_{j-1}-v_{j-1})}(\{r_j\}),
\end{aligned}$$

wobei sich die Summation $\sum_*$ über alle Tupel $\big((v_0, k_0), ..., (v_{n-1}, k_{n-1})\big) \in (\mathbb{N}^2_{\le})^n$ erstreckt und die letzte Zeile nach Umordnung dieser Summe ergibt. Dies zeigt die behauptete Unabhängigkeit der $(R_{n,j}, S_{n,j})$ mit

$$\mathbb{P}(R_{n,j} = r, S_{n,j} = s) \;=\; \sum_{1\le m\le k} p_k \Gamma_{n-j}^{*(m-1)}(\{s-r\})\Gamma_{n-j}^{*(k-m)}(\{r\}) \tag{7.2}$$

für $(r, s) \in \mathbb{N}^2_{\le}$ und $1 \le j \le n$.

(b) Vermöge (7.2) erhalten wir

$$\begin{aligned}
\mathbb{E}R_{n,j} \;&=\; \sum_{r\ge1} r\, \mathbb{P}(R_{n,j} = r) \\
&=\; \sum_{r\ge1} r \sum_{k\ge1} \sum_{m=1}^{k} p_k \Gamma_{n-j}^{*(k-m)}(\{r\}) \\
&=\; \sum_{k\ge1} p_k \sum_{m=1}^{k} \sum_{r\ge1} r \Gamma_{n-j}^{*(k-m)}(\{r\}) \\
&=\; \sum_{k\ge1} p_k \sum_{m=1}^{k} (k-m) \underbrace{\mathbb{E}Z_{n-j}}_{=1} \\
&=\; \sum_{k\ge1} p_k \frac{k(k-1)}{2} \\
&=\; \frac{\sigma^2}{2}.
\end{aligned}$$

(c) Auch hier benutzen wir (7.2) und erhalten

$$\begin{aligned}
\mathbb{P}(A_{n,j}) \;&=\; \mathbb{P}(R_{n,j} = S_{n,j}) \\
&=\; \sum_{k\ge1} \sum_{m=1}^{k} p_k \Gamma_{n-j}^{*(m-1)}(\{0\}) \\
&=\; \sum_{k\ge1} p_k \sum_{m=1}^{k} \mathbb{P}(Z_{n-j} = 0)^{m-1}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\mathbb{P}(Z_{n-j} > 0)} \sum_{k\ge 1} p_k\big(1 - \mathbb{P}(Z_{n-j} = 0)^k\big)\\
&= \frac{1}{\mathbb{P}(Z_{n-j} > 0)} \sum_{k\ge 1} p_k\,\mathbb{P}(Z_{n-j+1} > 0 | Z_1 = k)\\
&= \frac{\mathbb{P}(Z_{n-j+1} > 0)}{\mathbb{P}(Z_{n-j} > 0)}.
\end{aligned}
$$

Die Behauptung $\mathbb{P}(A_{n,j}) \to 1$, falls $n - j \to \infty$, ergibt sich wegen $\mathbb{P}(Z_{n-j} = 0) \uparrow 1$ aufgrund monotoner Konvergenz des dritten Ausdrucks der obigen Gleichungskette gegen $\sum_{k\ge 1} kp_k = 1$.

(d) Der besagte Erwartungswert besitzt gemäß (7.2) die Darstellung

$$
\begin{aligned}
\mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}} &= \sum_{r\ge 1} r\,\mathbb{P}(R_{n,j} = r, R_{n,j} = S_{n,j})\\
&= \sum_{r\ge 1} r \sum_{k\ge 1}\sum_{m=1}^{k} p_k \Gamma_{n-j}^{*(m-1)}(\{0\})\Gamma_{n-j}^{*(k-m)}(\{r\})\\
&= \sum_{k\ge 1}\sum_{m=1}^{k} p_k\,\mathbb{P}(Z_{n-j} = 0)^{m-1} \sum_{r\ge 1} r\Gamma_{n-j}^{*(k-m)}(\{r\})\\
&= \sum_{k\ge 1}\sum_{m=1}^{k} p_k\,\mathbb{P}(Z_{n-j} = 0)^{m-1}(k-m).
\end{aligned}
$$

Eine erneute Anwendung des Satzes von der monotonen Konvergenz in Verbindung mit der Rechnung in Teil (b)

$$
\lim_{n\to\infty} \mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}} = \sum_{k\ge 1}\sum_{m=1}^{k} p_k(k-m) = \frac{\sigma^2}{2}.
$$

Im Fall $\sigma^2 < \infty$ folgt nun mit (b) außerdem sofort

$$
\mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}^c} = \frac{\sigma^2}{2} - \mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}} \to 0, \quad n \to \infty,
$$

was den Beweis komplettiert. ◊

7.2. Bemerkungen. (a) Ein ähnliches Vorgehen wie im Beweis von Lemma 7.1(b) liefert $\mathbb{E}S_{n,j} = \sigma^2$, ein Ergebnis, das wir allerdings nicht benötigen und dehalb auch nicht herleiten, sondern dem Leser als Übung ans Herz legen (☞ Übung 7.1).

(b) Der obige Beweis hat gezeigt, daß die Ausdrücke $\mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}}$, $\mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}^c}$ und $\mathbb{P}(A_{n,j})$ von (j,n) nur über $n - j$ abhängen. Deshalb verwenden wir im folgenden die Schreibweisen

$$
\alpha_{n-j} \stackrel{\text{def}}{=} \mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}}, \quad \beta_{n-j} \stackrel{\text{def}}{=} \mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}^c} \quad \text{und} \quad \gamma_{n-j} \stackrel{\text{def}}{=} \mathbb{P}(A_{n,j}),
$$

und Lemma 7.1 sichert dann

$$
\lim_{n\to\infty} \alpha_n = \sigma^2/2, \quad \lim_{n\to\infty} \gamma_n = 1
$$

sowie im Fall $\sigma^2 < \infty$

$$\lim_{n\to\infty} \beta_n \;=\; 0.$$

Wir sind nun in der Lage, einen probabilistischen Beweis des folgenden, bereits in Satz I.9.1 enthaltenen Resultats zu geben, welches erstmals von Kolmogorov unter der Voraussetzung $\mathbb{E}Z_1^3 < \infty$ gezeigt wurde, in dieser Allgemeinheit aber von Kesten, Ney und Spitzer stammt:

7.3. Satz. *Für einen kritischen GWP $(Z_n)_{n\ge 0}$ mit Reproduktionsvarianz $\sigma^2 \in (0,\infty]$ gilt*

$$\lim_{n\to\infty} n\,\mathbb{P}(Z_n > 0) \;=\; \lim_{n\to\infty} \Big(\mathbb{E}\big(Z_n/n|Z_n > 0\big)\Big)^{-1} \;=\; \frac{2}{\sigma^2}$$

mit der üblichen Vereinbarung $\frac{1}{\infty} \stackrel{\text{def}}{=} 0$.

Beweis: Für $n \ge 1$ setzen wir

$$R_n \;\stackrel{\text{def}}{=}\; 1 + \sum_{j=1}^{n} R_{n,j}.$$

Da $R_{n,j}$ die Anzahl der Individuen in $\widehat{\boldsymbol{GW}}_n$ angibt, die größer als V^n sind und von V^{j-1}, nicht aber von V^j abstammen, bezeichnet R_n gerade die Anzahl der Individuen in $\widehat{\boldsymbol{GW}}_n$, die größer oder gleich V^n sind. [1)]

Zum besseren Verständnis des weiteren Vorgehens, das wiederum auf einer Kopplung basiert, seien folgende Hinweise gegeben: Mittels Lemma 7.1(b) sieht man schnell, daß $\mathbb{E}R_n/n$ denselben Limes $\sigma^2/2$ besitzt wie im Satz für $\mathbb{E}(Z_n/n|Z_n > 0)$ behauptet. Deshalb werden wir im Anschluß Zufallsgrößen R_n^*, $n \ge 1$, konstruieren, die einerseits

(i) $R_n^* \stackrel{d}{=} \mathbf{P}_n = \mathbb{P}(Z_n \in \cdot|Z_n > 0)$ für alle $n \ge 1$ (☞ (7.9))

und andererseits (im Fall $\sigma^2 < \infty$)

(ii) $n^{-1}\mathbb{E}|R_n - R_n^*| \to 0$ für $n \to \infty$ (☞ (7.10))

erfüllen. Dies liefert dann wie gewünscht $\mathbb{E}(Z_n/n|Z_n > 0) \to \sigma^2/2$.

Die Konstruktion der R_n^* ist allerdings keineswegs offensichtlich und erfordert als erstes die Einsicht, daß $\widehat{Z}_n$ bedingt unter dem Ereignis

$$A_n \;\stackrel{\text{def}}{=}\; \{V^n = \min \widehat{\boldsymbol{GW}}_n\},$$

(alle Individuen in $\widehat{\boldsymbol{GW}}_n \backslash \{V^n\}$ befinden sich rechts von V^n) ebenfalls nach $\mathbf{P}_n$ verteilt ist (☞ (7.5)): Wegen $R_{n,j} \le S_{n,j}$ für jedes $j \le n$ folgt offenkundig

$$A_n \;=\; \{R_n = \widehat{Z}_n\} \;=\; \bigcap_{j=1}^{n} \{R_{n,j} = S_{n,j}\} \;=\; \bigcap_{j=1}^{n} A_{n,j} \tag{7.3}$$

1) $R_n - 1$ gibt also die Anzahl der Individuen in $\widehat{\boldsymbol{GW}}_n$ an, die sich rechts von V^n befinden.

und via Lemma 7.1

$$\mathbb{P}(A_n) \;=\; \prod_{j=1}^{n} \mathbb{P}(A_{n,j}) \;=\; \prod_{j=1}^{n} \frac{\mathbb{P}(Z_{n-j+1}>0)}{\mathbb{P}(Z_{n-j}>0)} \;=\; \frac{\mathbb{P}(Z_n>0)}{\mathbb{P}(Z_0>0)} \;=\; \mathbb{P}(Z_n>0). \tag{7.4}$$

Die für $k \geq 1$ gültige Rechnung

$$\begin{aligned}
\mathbb{P}(\{\widehat{Z}_n = k\} \cap A_n) \;&=\; \mathbb{P}(\widehat{Z}_n = k, V^n = \min \widehat{\boldsymbol{GW}}_n) \\
&=\; \sum_{\tau \in \mathbb{T}_n : z_n(\tau) = k} \widehat{\mathbf{Q}}_*\big([\tau; \min \tau_n]_n\big) \\
&=\; \sum_{\tau \in \mathbb{T}_n : z_n(\tau) = k} \mathbf{Q}\big([\tau]_n\big) \\
&=\; \mathbb{P}(Z_n = k),
\end{aligned}$$

in die für die vorletzte Gleichheit einmal mehr das Vergleichslemma 3.7 eingeflossen ist, impliziert somit die angekündigte Verteilungsidentität

$$\mathbb{P}(\widehat{Z}_n \in \cdot | A_n) \;=\; \mathbb{P}(Z_n \in \cdot | Z_n > 0) \;=\; \mathbf{P}_n. \tag{7.5}$$

Damit können wir uns der Konstruktion der Zufallsgrößen R_n^* mit Verteilung $\mathbf{P}_n$ in Angriff nehmen: Bei festem $n \geq 1$ seien hierzu $R'_{n,1}, ..., R'_{n,n}$ stochastisch unabhängige Zufallsgrößen auf $(\Omega, \mathfrak{A}, \mathbb{P})$ mit

$$\mathbb{P}(R'_{n,j} \in \cdot) \;=\; \mathbb{P}(R_{n,j} \in \cdot | A_{n,j}) \;=\; \mathbb{P}(S_{n,j} \in \cdot | A_{n,j}) \tag{7.6}$$

für $1 \leq j \leq n$, wobei auf $\mathbb{P}(A_{n,j}) = \gamma_{n-j} > 0$ hingewiesen sei. Der Vektor $(R'_{n,1}, ..., R'_{n,n})$ sei außerdem unabhängig von $(R_{n,1}, S_{n,1}), ..., (R_{n,n}, S_{n,n})$. Definieren wir nun

$$R_{n,j}^* \;\stackrel{\text{def}}{=}\; R_{n,j} \mathbf{1}_{A_{n,j}} + R'_{n,j} \mathbf{1}_{A_{n,j}^c} \tag{7.7}$$

sowie

$$R_n^* \;\stackrel{\text{def}}{=}\; 1 + \sum_{j=1}^{n} R_{n,j}^*, \tag{7.8}$$

so sind vermöge Lemma 7.1 auch die Zufallsgrößen $R_{n,1}^*, ..., R_{n,n}^*$ stochastisch unabhängig, und es gilt für $k \geq 1$ und $j \leq n$ aufgrund der getroffenen Unabhängigskeitsannahmen

$$\begin{aligned}
\mathbb{P}(R_{n,j}^* = k) \;&=\; \mathbb{P}(\{R_{n,j} = k\} \cap A_{n,j}) \;+\; \mathbb{P}(\{R'_{n,j} = k\} \cap A_{n,j}^c) \\
&=\; \mathbb{P}(\{S_{n,j} = k\} \cap A_{n,j}) \;+\; \mathbb{P}(R'_{n,j} = k)\, \mathbb{P}(A_{n,j}^c) \\
&=\; \mathbb{P}(S_{n,j} = k | A_{n,j}) \big(\mathbb{P}(A_{n,j}) + \mathbb{P}(A_{n,j}^c)\big) \\
&=\; \mathbb{P}(S_{n,j} = k | A_{n,j})
\end{aligned}$$

und damit im Hinblick auf (7.1) und (7.3)

$$\begin{aligned}
\mathbb{P}(R_n^* = k) &= \sum_{\substack{k_1,\ldots,k_n\ge 0,\\ k_1+\ldots+k_n=k-1}} \mathbb{P}(R_{n,j}^* = k_j, 1\le j\le n) \\
&= \sum_{\substack{k_1,\ldots,k_n\ge 0,\\ k_1+\ldots+k_n=k-1}} \prod_{j=1}^n \mathbb{P}(R_{n,j}^* = k_j) \\
&= \sum_{\substack{k_1,\ldots,k_n\ge 0,\\ k_1+\ldots+k_n=k-1}} \prod_{j=1}^n \mathbb{P}(S_{n,j} = k_j | A_{n,j}) \\
&= \frac{1}{\mathbb{P}(A_n)} \sum_{\substack{k_1,\ldots,k_n\ge 0,\\ k_1+\ldots+k_n=k-1}} \mathbb{P}(\{S_{n,j} = k_j, 1\le j\le n\}\cap A_n) \\
&= \frac{1}{\mathbb{P}(A_n)} \mathbb{P}(\{\widehat{Z}_n = k\}\cap A_n) \\
&= \mathbb{P}(\widehat{Z}_n = k | A_n).
\end{aligned}$$

Dies garantiert mit Blick auf (7.5) wie angekündigt

$$\mathbb{P}(R_n^* \in \cdot) \;=\; \mathbf{P}_n \;=\; \mathbb{P}(Z_n \in \cdot | Z_n > 0). \tag{7.9}$$

Sei nun zunächst $\sigma^2 < \infty$ vorausgesetzt. Dann gilt nach Definition der $R_{n,j}^*$ sowie unter Hinweis auf Lemma 7.1

$$\begin{aligned}
\frac{1}{n}|\mathbb{E}R_n - \mathbb{E}R_n^*| &\le \frac{1}{n}\mathbb{E}|R_n - R_n^*| \\
&\le \frac{1}{n}\sum_{j=1}^n \mathbb{E}|R_{n,j} - R_{n,j}^*| \\
&\le \frac{1}{n}\sum_{j=1}^n \int_{A_{n,j}^c} (R_{n,j} + R_{n,j}^*)\, d\mathbb{P} \\
&= \frac{1}{n}\sum_{j=1}^n \int_{A_{n,j}^c} (R_{n,j} + R_{n,j}')\, d\mathbb{P} \\
&= \frac{1}{n}\sum_{j=1}^n \left(\beta_{n-j} + \mathbb{P}(A_{n,j}^c)\,\mathbb{E}R_{n,j}'\right) \\
&= \frac{1}{n}\sum_{j=1}^n \left(\beta_{n-j} + \frac{1-\gamma_{n-j}}{\gamma_{n-j}}\mathbb{E}R_{n,j}\right) \\
&= \frac{1}{n}\sum_{j=1}^n \left(\beta_j + \frac{1-\gamma_j}{\gamma_j}\frac{\sigma^2}{2}\right).
\end{aligned} \tag{7.10}$$

Der letzte Ausdruck ist jedoch das n-te Césaro-Mittel der Nullfolge $(\beta_j + \frac{1-\gamma_j}{\gamma_j}\frac{\sigma^2}{2})_{j\ge 1}$ (☞ Bemerkung 7.2(b)) und konvergiert folglich ebenfalls gegen 0. Damit erhalten wir

$$\begin{aligned}\left|\frac{\mathbb{E}R_n^*}{n}-\frac{\sigma^2}{2}\right| &\le \frac{|\mathbb{E}R_n^*-\mathbb{E}R_n|}{n} + \left|\frac{\mathbb{E}R_n}{n}-\frac{\sigma^2}{2}\right| \\ &= o(1) + \left|\frac{1}{n}+\underbrace{\frac{1}{n}\sum_{j=1}^n \mathbb{E}R_{n,j}}_{=\sigma^2/2}-\frac{\sigma^2}{2}\right| = o(1).\end{aligned}$$

Daraus folgt schließlich vermöge (7.9)

$$\frac{1}{n\,\mathbb{P}(Z_n>0)} = \frac{\mathbb{E}(Z_n|Z_n>0)}{n} = \frac{\mathbb{E}R_n^*}{n} \to \frac{\sigma^2}{2}, \quad n\to\infty.$$

Für den noch unerledigten Fall $\sigma^2=\infty$ notieren wir unter Hinweis auf (7.7) und (7.8)

$$R_n^* \ge \sum_{j=1}^n R_{n,j}\mathbf{1}_{A_{n,j}}$$

und folgern

$$\frac{1}{n\,\mathbb{P}(Z_n>0)} = \frac{\mathbb{E}R_n^*}{n} \ge \frac{1}{n}\sum_{j=1}^n \mathbb{E}R_{n,j}\mathbf{1}_{A_{n,j}} = \frac{1}{n}\sum_{j=0}^{n-1}\alpha_j \to \frac{\sigma^2}{2} = \infty, \quad n\to\infty,$$

denn der Satz von Césaro bleibt richtig, wenn das Césaro-Mittel über eine Folge mit Limes ∞ (hier $(\alpha_j)_{j\ge 0}$) gebildet wird. ♢

7.4. Bemerkung. Die im obigen Beweis als (ii) formulierte und mit (7.10) bewiesene Aussage im Fall $\sigma^2<\infty$ besagt gerade die $\mathfrak{L}_1$-Konvergenz von $n^{-1}(R_n-R_n^*)$ gegen 0, was insbesondere

$$\frac{R_n}{n}-\frac{R_n^*}{n} \xrightarrow{\mathbb{P}} 0, \quad n\to\infty$$

garantiert. Diese Tatsache werden wir im Beweis von Satz 7.8 verwenden, denn mit dem Satz von Slutsky (☞ Satz 36.12 in [1]) garantiert sie uns, daß die Verteilungskonvergenz von R_n/n und die von R_n^*/n äquivalente Bedingungen sind (bei notwendig gleichen Grenzverteilungen).

Wir kommen nun zur Konvergenz der bedingten Verteilung von Z_n/n gegeben $Z_n>0$ gegen eine Exponentialverteilung mit Parameter $2/\sigma^2$ (die dritte Aussage in Satz I.9.1). Die drei nachfolgenden Lemmata bilden die notwendige Vorarbeit, wobei das erste Lemma die Verträglichkeit von Verteilungskonvergenz und Größenverzerrung beleuchtet. Den einfachen Beweis überlassen wir dem Leser zur Übung (☞ Übung 7.2).

7.5. Lemma. *Es sei $(X_n)_{n\ge 0}$ eine Folge nichtnegativer Zufallsgrößen mit endlichen, positiven Erwartungswerten sowie $(\widehat{X}_n)_{n\ge 0}$ eine Folge assoziierter Größenverzerrungen. Es gelte $X_n \xrightarrow{d} X_0$. Dann folgt aus der Verteilungskonvergenz der $\widehat{X}_n$ bereits $\widehat{X}_n \xrightarrow{d} \widehat{X}_0$.*

Unser zweites Lemma stellt einen Zusammenhang zwischen R_n und $\widehat{Z}_n$ her: Lax gesprochen, erhält man R_n durch Stauchung von $\widehat{Z}_n$ mit einem auf $(0,1)$ gleichverteilten Faktor.

7.6. Lemma. *Sei U eine von $(\widehat{Z}_n)_{n\ge 0}$ unabhängige, $R(0,1)$-verteilte Zufallsgröße. Dann gilt für jedes $n \ge 1$ die Verteilungsidentität*

$$R_n \stackrel{d}{=} \lceil U\widehat{Z}_n \rceil,$$

wobei für $x \in \mathbb{R}$ wie üblich $\lceil x \rceil \stackrel{\text{def}}{=} \min\{n \in \mathbb{Z} : n \ge x\}$ gesetzt sei.

Beweis: Wie im Beweis von Satz 7.3 gesehen, gibt $R_n - 1$ die Anzahl der Individuen in $\widehat{\boldsymbol{GW}}_n$ an, die sich rechts von V^n befinden; insbesondere gilt $R_n \le \widehat{Z}_n$. Ordnen wir für $\tau \in \mathbb{T}$ mit $z_n(\tau) = m$, $m \ge 1$, die Individuen in τ_n in der Form

$$\varrho_n^{(1)}(\tau) < \ldots < \varrho_n^{(m)}(\tau)$$

und ziehen erneut das Vergleichslemma 3.7 zu Rate, so folgt für $k \ge 1$

$$\begin{aligned}
\mathbb{P}(R_n = k) &= \sum_{l\ge k} \mathbb{P}(\widehat{Z}_n = l, R_n = k) \\
&= \sum_{l\ge k} \sum_{\tau\in\mathbb{T}_n : z_n(\tau)=l} \widehat{\mathbf{Q}}_*\big([\tau; \varrho_n^{(l-k+1)}(\tau)]_n\big) \\
&= \sum_{l\ge k} \sum_{\tau\in\mathbb{T}_n : z_n(\tau)=l} \mathbf{Q}\big([\tau]_n\big) \\
&= \sum_{l\ge k} \mathbb{P}(Z_n = l) \\
&= \sum_{l\ge k} \mathbb{P}(\widehat{Z}_n = l)\, \underbrace{\mathbb{P}\left(\left(\frac{k-1}{l} < U \le \frac{k}{l}\right)\right.}_{=1/l} \\
&= \mathbb{P}(k-1 < U\widehat{Z}_n \le k) \\
&= \mathbb{P}(\lceil U\widehat{Z}_n \rceil = k).
\end{aligned}$$

Dies beweist offenbar die Behauptung. ◊

Wir kommen schließlich zu zwei interessanten Charakterisierungen von Exponentialverteilungen, die auch von eigenständigem Interesse sind. Bei der Charakterisierung (7.13) handelt es sich einmal mehr um eine stochastische Fixpunktgleichung (vgl. Übungen I.6.2 und I.8.5), die in Übung 7.3 noch eingehender studiert wird. Für weitere Informationen einschließlich analoger Aussagen für mehrdimensionale Verteilungen sei auch auf [4] und [6] hingewiesen.

7.7. Lemma. *Es seien X, X_1, X_2 u.i.v. nichtnegative Zufallsgrößen mit Erwartungswert $\nu \in (0,\infty)$, $\widehat{X}$ eine Größenverzerrung von X und U eine auf $(0,1)$ gleichverteilte, von allen* *vorherigen Variablen unabhängige Zufallsgröße. Dann sind äquivalent:*

$$X \stackrel{d}{=} Exp(1/\nu), \tag{7.11}$$

$$X \stackrel{d}{=} U\widehat{X}, \tag{7.12}$$

$$X \stackrel{d}{=} U(X_1 + X_2). \tag{7.13}$$

BEWEIS: Wir zeigen als erstes "(7.11)⇔(7.12)". Bezeichnet φ die L.T. von X, so bedeutet (7.12) vermöge des Eindeutigkeitssatzes für L.T. (benutze (3.4) und den Satz von Fubini)

$$\begin{aligned} \varphi(t) &= \mathbb{E}e^{-tU\widehat{X}} = \int_0^1 \mathbb{E}e^{-tu\widehat{X}}\,du \\ &= \frac{1}{\nu}\int_0^1 \mathbb{E}Xe^{-tuX}\,du = \frac{1}{\nu}\mathbb{E}\left(X\int_0^1 e^{-tuX}\,du\right) \\ &= \frac{1-\varphi(t)}{\nu t} \end{aligned}$$

für alle $t > 0$ und somit

$$\varphi(t) = \frac{1/\nu}{t+1/\nu}, \quad t > 0,$$

was, wiederum mit dem Eindeutigkeitssatz für L.T., zu (7.11) äquivalent ist.

Zur Einsicht von "(7.11),(7.12)⇒(7.13)" genügt der Hinweis, daß ein Dichtevergleich (☞ (3.5) für die Dichte von $\widehat{X}$) sofort $\widehat{X} \stackrel{d}{=} X_1 + X_2 \stackrel{d}{=} \Gamma(2, 1/\nu)$ und damit auch (7.13) liefert. Es bleibt also nur noch "(7.13)⇒(7.11)" zu zeigen. Da $X_1 + X_2$ die L.T. φ^2 besitzt, ergibt sich die Äquivalenz von (7.13) mit

$$\varphi(t) = \int_0^1 \varphi(ut)^2\,du = \frac{1}{t}\int_0^t \varphi(u)^2\,du, \quad t > 0, \tag{7.14}$$

was nach Differentiation

$$\varphi'(t) = -\frac{1}{t^2}\int_0^t \varphi(u)^2\,du + \frac{\varphi(t)^2}{t}, \quad t > 0 \tag{7.15}$$

liefert und per Kombination von (7.14) und (7.15) zu der Differentialgleichung

$$t\varphi'(t) + \varphi(t) = \varphi(t)^2, \quad t > 0 \tag{7.16}$$

führt. Betrachten wir nun die auf $(0, \infty)$ positive Funktion $\Psi \stackrel{\text{def}}{=} \frac{1-\varphi}{\varphi}$ mit Ableitung $\Psi' = -\frac{\varphi'}{\varphi^2}$ und $\Psi(0) = 0$, so folgt aus (7.16) weiter

$$\frac{\Psi'(t)}{\Psi(t)} = \frac{\varphi'(t)}{\varphi(t)^2 - \varphi(t)} = \frac{1}{t}, \quad t > 0$$

und dann bei Integration über beliebige $(s, t) \subset (0, \infty)$

$$\log\Big(\Psi(t)/\Psi(s)\Big) = \int_s^t \frac{\Psi'(u)}{\Psi(u)}\,du = \int_s^t \frac{1}{u}\,du = \log(t/s),$$

was schließlich

$$\Psi(t) \;=\; \frac{\Psi(s)}{s}t, \quad t > s$$

für alle $s > 0$ ergibt. Damit gilt $\Psi(t) = at$ für ein $a > 0$ auf ganz $[0, \infty)$, also

$$\varphi(t) \;=\; \frac{1}{1+\Psi(t)} \;=\; \frac{1}{1+at} \;=\; \frac{1/a}{t+1/a}, \quad t > 0$$

und somit $X \stackrel{d}{=} \operatorname{Exp}(1/a)$, wobei dann nach Voraussetzung $a = \mathbb{E}X = \nu$ gelten muß. ♢

Nach diesen Vorbereitungen sind wir in der Lage, das auf Yaglom zurückgehende Konvergenzresultat (I.9.4) probabilistisch zu beweisen.

7.8. Satz (Yaglom). *Gegeben einen kritischen GWP $(Z_n)_{n\ge 0}$ mit Reproduktionsvarianz $\sigma^2 \in (0,\infty)$, gilt für $t \ge 0$*

$$\mathbb{P}(Z_n/n \le t | Z_n > 0) \;=\; 1 - \exp(-2t/\sigma^2),$$

d.h. die bedingte Verteilung $\mathbb{P}(Z_n/n \in \cdot | Z_n > 0)$ konvergiert schwach gegen die Exponentialverteilung mit Parameter $2/\sigma^2$.

Beweis: Die Bezeichnungen von Satz 7.3 und dessen Beweis seien weiter gültig. Da $R_n^* \stackrel{d}{=} \mathbb{P}(Z_n \in \cdot | Z_n > 0)$ für alle $n \ge 1$ gemäß (7.9), müssen wir $R_n^*/n \stackrel{d}{\to} \operatorname{Exp}(2/\sigma^2)$ zeigen.

Die Voraussetzung $\sigma^2 < \infty$ sichert wegen $\mathbb{E}R_n^*/n \to \sigma^2/2$ (☞ Beweis von Satz 7.3) und $\mathbb{E}\widehat{Z}_n = \mathbb{E}Z_n^2/\mathbb{E}Z_n = \operatorname{Var}Z_n + 1 = n\sigma^2 + 1$ (☞ Satz I.2.2)

$$\kappa \stackrel{\text{def}}{=} \sup_{n\ge 1} \frac{\mathbb{E}R_n^*}{n} \;<\; \infty, \quad \sup_{n\ge 1} \frac{\mathbb{E}R_n}{n} \;=\; \frac{\sigma^2}{2} + 1 \;<\; \infty \quad \text{und} \quad \sup_{n\ge 1} \frac{\mathbb{E}\widehat{Z}_n}{n} \;=\; \sigma^2 + 1 \;<\; \infty$$

und damit die Straffheit der Folgen

$$\big(\mathbb{P}(R_n^*/n) \in \cdot)\big)_{n\ge 1}, \quad \big(\mathbb{P}(R_n/n) \in \cdot)\big)_{n\ge 1} \quad \text{und} \quad \big(\mathbb{P}(\widehat{Z}_n/n) \in \cdot)\big)_{n\ge 1},$$

wie sich leicht mit der Markov-Ungleichung zeigen läßt. Nach dem Auswahlsatz von Helly-Bray (☞ Satz 44.1 in [1]) und unter Beachtung von $(R_n - R_n^*)/n \stackrel{\mathbb{P}}{\to} 0$ (☞ Bemerkung 7.4) existieren deshalb (o.B.d.A. auf $(\Omega, \mathfrak{A}, \mathbb{P})$ definierte) Zufallsgrößen S, T mit

$$R_{n_k}^*/n_k \stackrel{d}{\to} S, \quad R_{n_k}/n_k \stackrel{d}{\to} S \quad \text{und} \quad \widehat{Z}_{n_k}/n_k \stackrel{d}{\to} T \tag{7.17}$$

für eine geeignete aufsteigende Teilfolge $(n_k)_{k\ge 1}$ in $\mathbb{N}$. Ferner liefert (7.9) für $t > 0$ und $n \ge 1$

$$\begin{aligned}\mathbb{P}(\widehat{Z}_n/n \le t) \;&=\; \frac{\mathbb{E}Z_n \mathbf{1}_{\{Z_n/n \le t\}}}{\mathbb{E}Z_n} \\ &=\; \frac{\mathbb{E}(Z_n \mathbf{1}_{\{Z_n/n\le t\}} | Z_n > 0)}{\mathbb{E}(Z_n | Z_n > 0)} \\ &=\; \frac{\mathbb{E}\big(\mathbf{1}_{\{R_n^*/n \le t\}} R_n^*/n\big)}{\mathbb{E}(R_n^*/n)},\end{aligned}$$

so daß $\widehat{Z}_n/n$ eine Größenverzerrung von R_n^*/n bildet. Dies liefert für $t > 0$ und $n \geq 1$ insbesondere

$$\mathbb{E}\big(\mathbf{1}_{\{R_n^*/n>t\}}R_n^*/n\big) \;\leq\; \kappa\,\mathbb{P}(\widehat{Z}_n/n > t),$$

also im Hinblick auf die Straffheit von $(\mathbb{P}(\widehat{Z}_n/n \in \cdot))_{n\geq 1}$ die gleichgradige Integrierbarkeit der Folge $(R_{n_k}^*/n_k)_{k\geq 1}$, welche vermöge Satz 50.5 in [1]

$$\mathbb{E}S \;=\; \lim_{k\to\infty} \frac{\mathbb{E}R_{n_k}^*}{n_k} \;=\; \frac{\sigma^2}{2} \;\in\; (0,\infty) \tag{7.18}$$

impliziert. Eine Anwendung von Lemma 7.5 (mit $X_0 = S, X_k = R_{n_k}^*/n_k$ und $\widehat{X}_k = \widehat{Z}_{n_k}/n_k$ für $k \geq 1$) liefert nun $T \stackrel{d}{=} \widehat{S}$ für jede auf $(\Omega, \mathfrak{A}, \mathbb{P})$ definierte Größenverzerrung $\widehat{S}$ von S und somit

$$\widehat{Z}_{n_k}/n_k \;\stackrel{d}{\to}\; \widehat{S}, \quad k \to \infty.$$

Bezeichnet im folgenden U eine von $\widehat{S}$ und $(\widehat{Z}_n)_{n\geq 1}$ unabhängige, $R(0,1)$-verteilte Zufallsgröße, so folgt unter Hinweis auf Satz 36.11 in [1]

$$U\widehat{Z}_{n_k}/n_k \;\stackrel{d}{\to}\; U\widehat{S}, \quad k \to \infty.$$

Wegen

$$0 \;\leq\; \Delta_n \;\stackrel{\text{def}}{=}\; \frac{\lceil U\widehat{Z}_n\rceil}{n} - \frac{U\widehat{Z}_n}{n} \;\leq\; \frac{1}{n} \;\to\; n, \quad n \to \infty$$

zeigt Satz 36.12 in [1] in Verbindung mit Lemma 7.6

$$\frac{R_{n_k}}{n_k} \;\stackrel{d}{=}\; \frac{\lceil U\widehat{Z}_{n_k}\rceil}{n_k} \;=\; \frac{U\widehat{Z}_{n_k}}{n_k} + \Delta_{n_k} \;\stackrel{d}{\to}\; U\widehat{S}, \quad k \to \infty.$$

Folglich sind im Hinblick auf die erste Aussage in (7.17) die Zufallsgrößen S und $U\widehat{S}$ identisch verteilt, was $S \stackrel{d}{=} \mathrm{Exp}(2/\sigma^2)$ vermöge Lemma 7.7 (und $\mathbb{E}S = \sigma^2/2$ gemäß (7.18)) beweist.

Ist nun $(n_k')_{k\geq 1}$ irgendeine andere aufsteigende Teilfolge derart, daß $R_{n_k'}^*/n_k'$ in Verteilung konvergiert und o.E. gleiches auch für $R_{n_k'}/n_k'$ und $\widehat{Z}_{n_k'}/n_k'$ gilt (sonst dünne $(n_k')_{k\geq 1}$ nochmals geeignet aus), so liefert dieselbe Argumentation wie zuvor, daß auch $R_{n_k'}^*/n_k'$ die Grenzverteilung $\mathrm{Exp}(2/\sigma^2)$ besitzt. Jede verteilungskonvergente Teilfolge von $(R_n^*/n)_{n\geq 1}$ hat demnach denselben Limes, was gemäß Korollar 44.4 in [1] schließlich

$$R_n^*/n \;\stackrel{d}{\to}\; \mathrm{Exp}(2/\sigma^2), \quad n \to \infty$$

und so die Behauptung impliziert. ◇

7.9. Bemerkung. Zur probabilistischen Interpretation des obigen Beweises läßt sich folgendes anmerken: Die mit S bezeichnete Grenzvariable genügt gemäß Lemma 7.7 der Verteilungsidentität

$$S \;\stackrel{d}{=}\; U\widehat{S} \;\stackrel{d}{=}\; U(S_1 + S_2),$$

wobei S_1, S_2 voneinander und von U unabhängige Kopien von S bilden. Schreiben wir $\widehat{Z}_n$ in der Form

$$\widehat{Z}_n \;=\; 1 + L_n + R_n,$$

wobei L_n die links von V^n befindlichen Individuen in $\widehat{\boldsymbol{GW}}_n$ angibt, so lassen sich S_1 und S_2 als (Verteilungs-)Grenzwerte von L_n/n bzw. R_n/n auffassen.

Übungen zu Abschnitt 7

Übung 7.1. Beweisen Sie die in Bemerkung 7.2(a) aufgestellte Behauptung $\mathbb{E}S_{n,j} = \sigma^2$ für alle $(j,n) \in \mathbb{N}^2_{\le}$.

Übung 7.2. Beweisen Sie Lemma 7.5. [Hinweis: Überlegen Sie sich, daß es reicht, die vage Konvergenz der $\mathbb{P}(\widehat{X}_n \in \cdot)$ gegen $\mathbb{P}(\widehat{X}_0 \in \cdot)$ zu zeigen, und benutzen Sie Übung 6.1.]

Übung 7.3. Betrachten Sie die Situation von Lemma 7.7, und sei $\{U_v : v \in \mathbb{V}_{2,\infty}\}$ eine unabhängige Familie identisch $R(0,1)$-verteilter Zufallsgrößen, wobei $\mathbb{V}_{2,\infty} = \bigcup_{n\ge 0}\{1,2\}^n$ wieder den homogenen Baum der Ordnung 2 und Höhe ∞ bezeichnet (☞ Abschnitt I.12). Für $v = v_1 ... v_n \in \mathbb{V}_{2,\infty}$ setzen wir

$$T(v) \;=\; (T_1(v), T_2(v)) \;\stackrel{\text{def}}{=}\; (U_v, U_v)$$

und dann

$$L(v) \;\stackrel{\text{def}}{=}\; T_{v_1}(\varnothing)T_{v_2}(v_1)\cdot ... \cdot T_{v_n}(v_1...v_{n-1}) \;=\; U_\varnothing U_{v_1}\cdot ... \cdot U_{v_1...v_{n-1}}.$$

Dies läßt sich so interpretieren, daß wir zunächst jeder Kante $(v, vi) \in \mathbb{V}^2_{2,\infty}$ das Gewicht $T_i(v)$ zuordnen. Anschließend erhält jeder Pfad $\varnothing \to v$ das Gewicht, welches sich durch Multiplikation der Kantengewichte entlang dieses Pfades ergibt. Sei schließlich $\{X_v : v \in \mathbb{V}_{2,\infty}\}$ eine weitere unabhängige Familie identisch verteilter, nichtnegativer Zufallsgrößen mit positivem und endlichem Erwartungswert ν, die auch unabhängig von $\{U_v : v \in \mathbb{V}_{2,\infty}\}$ ist. Die Verteilung der X_v werde mit F bezeichnet. Zeigen Sie:

(a) Aus $F = \mathrm{Exp}(1/\nu)$ folgt

$$X_\varnothing \;\stackrel{d}{=}\; \sum_{v\in\mathbb{V}_{2,\infty},|v|=n} L(v)X_v$$

für alle $n \ge 1$.

(b) Für allgemeines F mit Erwartungswert ν und endlicher Varianz gilt

$$\sum_{v\in\mathbb{V}_{2,\infty},|v|=n} L(v)X_v \;\stackrel{d}{\to}\; \mathrm{Exp}(1/\nu), \quad n\to\infty.$$

[Hinweis: Betrachten Sie hierzu $\Delta_n \stackrel{\text{def}}{=} \sum_{v\in\mathbb{V}_{2,\infty},|v|=n} L(v)(X_v - Y_v)$ für eine von allen vorher eingeführten Zufallsgrößen unabhängige Familie $\{Y_v : v \in \mathbb{V}_{2,\infty}\}$ u.i.v. Zufallsgrößen mit Verteilung $\mathrm{Exp}(1/\nu)$. Zeigen Sie $\Delta_n \xrightarrow{\mathbb{P}} 0$ durch Berechnung von $Var\Delta_n$ und benutzen Sie (a).]

(c) Im Fall $F = \delta_1$ definiert

$$W_n \stackrel{\text{def}}{=} \sum_{v \in \mathbb{V}_{2,\infty}, |v|=n} L(v) X_v = \sum_{v \in \mathbb{V}_{2,\infty}, |v|=n} L(v)$$

ein Martingal und konvergiert f.s. gegen ein $W \stackrel{d}{=} \mathrm{Exp}(1)$.

Übung 7.4. Gegeben sei die Situation von Lemma 7.7, jedoch mit *zwei* unabhängigen $R(0,1)$-verteilten Zufallsgrößen U_1, U_2, die auch von X, X_1, X_2 unabhängig sind. Zeigen Sie die Äquivalenz von

$$X \stackrel{d}{=} \Gamma(2, 1/\nu),$$

und

$$X \stackrel{d}{=} U_1 X_1 + U_2 X_2.$$

Anhang: Der Variationsabstand

Im folgenden geben wir eine kurze Zusammenstellung der wichtigsten Fakten über den Variationsabstand. Beginnen wir mit der Definition:

A.1. Definition. Sei $\mathcal{P} = \mathcal{P}(\mathfrak{X}, \mathcal{A})$ die Menge der W-Maße auf einem meßbaren Raum $(\mathfrak{X}, \mathcal{A})$. Dann definiert

$$d_V(P,Q) \stackrel{\text{def}}{=} \sup_{A \in \mathcal{A}} |P(A) - Q(A)|, \quad P, Q \in \mathcal{P} \tag{A.1}$$

eine Metrik auf $\mathcal{P}$, genannt *Variations-* oder *Supremumsabstand*. Anstelle von $d_V(P,Q)$ schreiben wir auch $\|P-Q\|$. Konvergenz bezüglich d_V wird als *Konvergenz in Totalvariation* bezeichnet.

Die Schreibweise $\|P-Q\|$ deutet darauf hin, daß d_V durch eine Norm induziert wird. In der Tat bildet

$$\|\Lambda\| \stackrel{\text{def}}{=} \sup_{A \in \mathcal{A}} |\Lambda(A)|$$

eine Norm auf dem Vektorraum der *endlichen signierten Maße* auf $(\mathfrak{X}, \mathcal{A})$, gegeben durch

$$\mathcal{S} = \mathcal{S}(\mathfrak{X}, \mathcal{A}) \stackrel{\text{def}}{=} \{\Lambda = \lambda P - \mu Q : \lambda, \mu \in [0, \infty), P, Q \in \mathcal{P}\}.$$

Dies spielt für unsere Zwecke aber keine weitere Rolle.

Das anschließende Lemma zeigt, daß das Supremum in (A.1) stets angenommen wird.

A.2. Lemma. *Gegeben $P, Q \in \mathcal{P}$ mit Dichten f_P und f_Q bezüglich eines σ-endlichen Maßes μ (z.B. $\mu = P + Q$), gilt mit $A^+ \stackrel{\text{def}}{=} \{f_P > f_Q\}$ und $A^- \stackrel{\text{def}}{=} \{f_P < f_Q\}$*

$$d_V(P, Q) \;=\; |P(A^\pm) - Q(A^\pm)| \;=\; \frac{1}{2}\int |f_P - f_Q|\; d\mu. \tag{A.2}$$

Beweis: Wir bemerken zuerst, daß aus $\int (f_P - f_Q)\, d\mu = 0$

$$\int_{\{f_P > f_Q\}} (f_P - f_Q)\; d\mu \;=\; \int_{\{f_Q > f_P\}} (f_Q - f_P)\; d\mu \tag{A.3}$$

folgt, und dies dann weiter

$$\begin{aligned} \int |f_P - f_Q|\; d\mu \;&=\; \int_{\{f_P > f_Q\}} (f_P - f_Q)\; d\mu \;+\; \int_{\{f_Q > f_P\}} (f_Q - f_P)\; d\mu \\ &=\; 2\int_{\{f_P > f_Q\}} (f_P - f_Q)\; d\mu \\ &=\; 2\int_{\{f_Q > f_P\}} (f_Q - f_P)\; d\mu. \end{aligned} \tag{A.4}$$

liefert. Nun gilt für alle $A \in \mathcal{A}$ unter Verwendung von (A.3) und (A.4)

$$\begin{aligned} |P(A) - Q(A)| \;&=\; \left|\int_A (f_P - f_Q)\; d\mu\right| \\ &=\; \left|\int_{A\cap\{f_P > f_Q\}} (f_P - f_Q)\; d\mu \;-\; \int_{A\cap\{f_Q > f_P\}} (f_Q - f_P)\; d\mu\right| \\ &\le\; \max\left\{\int_{\{f_P > f_Q\}} (f_P - f_Q)\; d\mu, \int_{\{f_Q > f_P\}} (f_Q - f_P)\; d\mu\right\} \\ &=\; \frac{1}{2}\int |f_P - f_Q|\; d\mu, \end{aligned}$$

und es ergibt sich die Behauptung, da diese Ungleichung bei Wahl von $A = A^+$ oder $= A^-$ eine Gleichung wird. ♢

Ist $\mathfrak{X}$ abzählbar und μ das Zählmaß auf $\mathfrak{X}$, so liefert (A.2) offenbar

$$\|P - Q\| \;=\; \frac{1}{2}\sum_{i\in\mathfrak{X}} |p_i - q_i|, \tag{A.5}$$

wobei $p_i \stackrel{\text{def}}{=} P(\{i\})$ und $q_i \stackrel{\text{def}}{=} Q(\{i\})$.

Gegeben $P, Q \in \mathcal{P}$, nennt man ein Paar (X, Y) von $\mathfrak{X}$-wertigen Zufallsvariablen auf einem W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$ eine *Kopplung von* (P, Q), wenn $X \stackrel{d}{=} P$ und $Y \stackrel{d}{=} Q$. Eine häufig effiziente Methode zur Abschätzung des Variationsabstandes besteht in der Verwendung solcher Kopplungen. Grundlage hierfür bildet die folgende *Kopplungsungleichung*.

A.3. Lemma (Kopplungsungleichung). *Zu W-Maßen $P, Q \in \mathcal{P}$ sei eine Kopplung (X, Y) gegeben. Dann gilt die Kopplungsungleichung*

$$\|P - Q\| \ \le\ \mathbb{P}(X \ne Y).$$

Beweis: Die Behauptung ergibt sich leicht vermöge

$$\begin{aligned} |P(A) - Q(A)| &= |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)| \\ &= |\mathbb{P}(X \in A, X \ne Y) - \mathbb{P}(Y \in A, X \ne Y)| \\ &\le \mathbb{P}(X \ne Y) \end{aligned}$$

für alle $A \in \mathcal{A}$. ♢

Als letztes notieren wir noch:

A.4. Satz. *Der Raum $(\mathcal{P}, d_V)$ ist vollständig.*

Beweis: Sei $(P_n)_{n\ge1}$ eine Cauchyfolge bezüglich d_V. Sei weiter μ ein σ-endliches dominierendes Maß für diese Folge (z.B. $\mu = \sum_{n\ge1} 2^{-n} P_n \in \mathcal{P}$) und $f_n \stackrel{\text{def}}{=} \frac{dP_n}{d\mu}$. Vermöge (A.2) gilt dann

$$\|P_n - P_m\| \ =\ \frac{1}{2}\int |f_n - f_m|\, d\mu \ =\ \|f_n - f_m\|_1$$

für alle $m, n \ge 1$, wobei $\|\cdot\|_1$ die übliche Norm auf $\mathfrak{L}_1(\mu)$ bezeichnet. Da $(\mathfrak{L}_1(\mu), \|\cdot\|_1)$ (modulo Gleichheit f.ü.) bekanntlich einen Banachraum bildet, folgt die Existenz eines $f \in \mathfrak{L}_1(\mu)$ mit $\|f_n - f\|_1 \to 0$ und $\|f\|_1 = 1$. Mit $P \stackrel{\text{def}}{=} f\mu \in \mathcal{P}$ erhält man schließlich

$$\|P_n - P\| \ =\ \frac{1}{2}\|f_n - f\|_1 \ \to\ 0, \quad n \to \infty$$

und damit die Behauptung. ♢

Literatur

[1] ALSMEYER, G. *Wahrscheinlichkeitstheorie (5. Auflage).* Skripten zur Mathematischen Statistik, Nr. 30, Universität Münster (2007).

[2] ALSMEYER, G. *Stochastische Prozesse, Teil I (3. Auflage).* Skripten zur Mathematischen Statistik, Nr. 33, Universität Münster (2005).

[3] GEIGER, J. Elementary new proofs of classical limit theorems for Galton-Watson processes. *J. Appl. Probab.* **36**, 301-309 (1999).

GEIGER, J. A new proof of Yaglom's exponential limit law. *Mathematics and computer science (Versailles, 2000), 245-249, Trends Math.*, Birkhäuser, Basel (2000).

KESTEN, H. und STIGUM, B. A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Stat.* **37**, 1211-1223 (1966).

KESTEN, H., NEY, P. und SPITZER, F. The Galton-Watson process with mean one and finite variance. *Theory Probab. Appl.* **11**, 513-540 (1966).

[4] KOTZ, S. und STEUTEL, F.W. Note on a characterization of exponential distributions. *Statist. Probab. Letters* **6**, 201-203 (1988)

[5] LYONS, R., PEMANTLE, R. und PERES, Y. Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23**, 1125-1138 (1995).

[6] PAKES, A.G. und KHATTREE, R. Length-biasing, characterization of laws and the moment problem. *Austral. J. Statist.* **34**, 307-322 (1992).

SENETA, E. On recent theorems concerning the supercritical Galton-Watson process. *Ann. Math. Stat.* **39**, 2098-2102 (1968).

YAGLOM, A.M. Certain limit theorems of the theory of branching processes. *Dokl. Acad. Nauk. SSSR* **56**, 795-798 (1947).