Stochastische Modelle der Populationsbiologie SS 2015

Blatt 1

Aufgabe 1.1 Sei $(c_N)_{N\in\mathbb{N}} \subset (0,1)$ eine Nullfolge, $\lambda > 0$, $Z_i^{(N)}$, $i \in \mathbb{N}$ u.i.v. $\sim \mathrm{Ber}(\lambda c_N)$ und $T_\ell^{(N)}$ der Zeitpunkt des ℓ-ten Erfolgs in der Münzwurffolge $(Z_i^{(N)})_{i\in\mathbb{N}}$. Zeigen Sie, dass

$$c_N T_\ell^{(N)} \xrightarrow[N\to\infty]{d} \Gamma(\ell,\lambda), \quad \text{d.h. für } t \geqslant 0 \text{ gilt} \quad \mathbb{P}\big(c_N T_\ell^{(N)} \leqslant t\big) \xrightarrow[N\to\infty]{} \int_0^t \frac{\lambda^\ell}{\Gamma(\ell)} u^{\ell-1} e^{-\lambda u}\, du.$$

[*Hinweis.* Verwenden Sie beispielsweise die in der Vorlesung gezeigte Konvergenz von $\big(M^{(N)}_{\lfloor c_N^{-1} t\rfloor}\big)_{t\geqslant 0}$ gegen einen Poissonprozess oder nutzen Sie die Tatsache, dass $T_\ell^{(N)} - \ell$ negativ binomial verteilt ist, $\mathbb{P}(T_\ell^{(N)} - \ell = k) = \binom{k+\ell-1}{k}(\lambda c_N)^\ell (1-\lambda c_N)^k$ (warum?).]

Aufgabe 1.2 a) Im sog. Moran-Modell (mit Populationsgröße N) ist die Verteilung des Nachkommensvektors $\nu = (\nu_1, \dots, \nu_N)$ die einer rein zufällig gewählten Permutation von

$$\big(2, 0, \underbrace{1, 1, \dots, 1}_{(N-2)\text{-mal}}\big),$$

d.h. ein zufällig gewähltes Individuum hat 2 Nachkommen, ein anderes, zufällig gewähltes Individuum hat 0 Nachkommen und alle anderen genau einen (man kann dies auch so interpretieren, dass ein Individuum einen Nachkommen hat, ein anderes stirbt und alle übrigen weiterleben). Berechnen Sie c_N und d_N wie in Satz 1.8. Sind die Voraussetzungen von Satz 1.8 erfüllt?

b) Sei nun $\psi \in (0,1)$, $\gamma > 0$ und im N-ten Populationsmodell sei $\nu = (\nu_1, \dots, \nu_N)$ eine rein zufällige Permutation eines Vektors, der die Werte

$$\big(2, 0, \underbrace{1, 1, \dots, 1}_{(N-2)\text{-mal}}\big) \quad \text{mit W'keit } 1 - N^{-\gamma} \quad \text{und}$$

$$\big(\lfloor \psi N \rfloor, \underbrace{0, 0, 0, \dots, 0}_{(\lfloor \psi N\rfloor -1)\text{-mal}}, \underbrace{1, 1, 1, 1, 1, \dots, 1}_{(N-\lfloor \psi N\rfloor)\text{-mal}}\big) \quad \text{mit W'keit } N^{-\gamma}$$

annimmt. Berechnen Sie c_N und d_N wie in Satz 1.8. Unter welcher Bedingung an γ sind die Voraussetzungen von Satz 1.8 erfüllt?

Aufgabe 1.3 Sei Q die Sprungratenmatrix einer zeitkontinuierlichen Markovkette auf der endlichen Menge E. Überzeugen Sie sich, dass

$$\frac{\partial}{\partial t} \exp(tQ) = Q \exp(tQ) = \exp(tQ) Q, \tag{1}$$

gilt, demnach für $t \geqslant 0$ und $P_t = e^{tQ} = (p_t(x,y))_{x,y\in E}$

$$\frac{\partial}{\partial t} P_t = Q P_t, \text{ d.h. } \forall\, x, y \in E: \quad \frac{\partial}{\partial t} p_t(x,y) = \sum_z q_{x,z} p_t(z,y) = \sum_z q_{x,z}\big(p_t(z,y) - p_t(x,y)\big), \tag{2}$$

$$\frac{\partial}{\partial t} P_t = P_t Q, \text{ d.h. } \forall\, x, y \in E: \quad \frac{\partial}{\partial t} p_t(x,y) = \sum_z p_t(x,z) q_{z,y} = p_t(x,y) q_{y,y} + \sum_{z\neq y} p_t(x,z) q_{z,y}. \tag{3}$$

[*Hinweis.* Warum ist gliedweise Differentiation der Exponentialreihe hier erlaubt?

Bericht: Die Gleichungen (2) und (3) haben eine stochastische Interpretation. (2) heißt Kolmogorovs *Rückwärtsgleichung*, denn es gilt (wir schreiben $\mathbb{P}_x(\cdot)$ für $\mathbb{P}(\,\cdot\,|X_0 = x)$)

$$\begin{aligned}
\frac{p_{t+h}(x,y) - p_t(x,y)}{h} &= \frac{1}{h}\big(\mathbb{P}_x(X_{t+h} = y) - \mathbb{P}_x(X_t = y)\big) \\
&= \frac{1}{h}\Big(\textstyle\sum_z \mathbb{P}_x(X_{t+h} = y | X_h = z)\mathbb{P}_x(X_h = z) - \mathbb{P}_x(X_t = y)\Big) \\
&= \frac{1}{h}\Big(\textstyle\sum_z p_t(z,y)\big(1_{\{x=z\}} + h q_{x,z} + o(h)\big) - p_t(x,y)\Big) = \displaystyle\sum_z q_{x,z} p_t(z,y) + o(1),
\end{aligned}$$

man leitet sie also aus her durch „Rückwärtszerlegung“ des Prozesses X im Intervall $[0, t+h]$ gemäß dem Verhalten am Anfang des Intervalls. Analog heißt (3) Kolmogorovs *Vorwärtsgleichung*, sie entsteht durch Zerlegung gemäß dem Wert bei t:

$$\begin{aligned} \frac{p_{t+h}(x,y)-p_t(x,y)}{h} &= \frac{1}{h}\Big(\sum\nolimits_z \mathbb{P}_x(X_{t+h}=y|X_t=z)\mathbb{P}_x(X_t=z)-\mathbb{P}_x(X_t=y)\Big) \\ &= \frac{1}{h}\Big(\sum\nolimits_z p_t(x,z)\big(1_{\{z=y\}}+hq_{y,z}+o(h)\big)-p_t(x,y)\Big)=\sum_z q_{x,z}p_t(z,y)+o(1). \end{aligned}$$

Sowohl (2) als auch (3) sind (im Fall $|E|<\infty$) eindeutig lösbar und beide bestimmen die Halbgruppe von Übergangsmatrizen $(P_t)_{t\geqslant 0}$ – es sind beides endliche Systeme linearer Differentialgleichungen mit konstanten Koeffizienten.]

b) Sei $E=\{0,1\}$, $Q=\left(\begin{smallmatrix}-a & a\\ b & -b\end{smallmatrix}\right)$ mit $a,b>0$.

Zeigen Sie: Für $x,y\in\{0,1\}$ gilt $p_t(x,y)=\delta_{x,y}e^{-(a+b)t}+(1-e^{-(a+b)t})\mu(y)$ mit $\mu(0)=b/(a+b)$, $\mu(1)=a/(a+b)$.

c) Sei $E=\{0,1,2,3\}$ (oder auch $E=\{\mathsf{A},\mathsf{G},\mathsf{C},\mathsf{T}\}$) und

$$Q=\begin{pmatrix} -3 & 1 & 1 & 1\\ 1 & -3 & 1 & 1\\ 1 & 1 & -3 & 1\\ 1 & 1 & 1 & -3 \end{pmatrix}.$$

Zeigen Sie: Für $x,y\in E$ gilt $p_t(x,y)=\delta_{x,y}e^{-4t}+\frac{1}{4}(1-e^{-4t})$.

Aufgabe 1.4* a) Sei E ein polnischer Raum, $D([0,\infty),E)$ die Menge aller Funktionen $f:[0,\infty)\to E$, die rechtsstetig sind und an jeder Stelle einen linken Limes besitzen (oft mit dem französischen Akronym càdlàg bezeichnet).

Sei $\Lambda:=\{\lambda:[0,\infty)\to[0,\infty):\lambda$ bijektiv und stetig$\}$, die Lipschitz-Konstante von $\lambda(\cdot)$ bezeichnen wir mit $\gamma(\lambda):=\sup_{0\leqslant s<t}\frac{\lambda(t)-\lambda(s)}{t-s}$ $(\leqslant\infty)$. Für $f,g\in D([0,\infty),E)$ definieren wir die Skorokhod-Metrik als

$$d(f,g):=\inf_{\lambda\in\Lambda}\Big\{\gamma(\lambda)\vee\int_0^\infty e^{-t}\sup_{s\geqslant 0}d_E\big(f(s\wedge t),g(\lambda(s)\wedge t)\big)\,dt\Big\}$$

Zeigen Sie: Dies *ist* eine Metrik [d.h. $d(\cdot,\cdot)$ ist symmetrisch, $d(f,g)=0 \iff f=g$, die Dreiecksungleichung gilt], damit ausgestattet ist $D([0,\infty),E)$ ein vollständiger und separabler metrischer Raum.

[*Hinweis.* Vgl. auch Kap. 3.5 in S.N. Ethier, T.G. Kurtz, Markov processes: characterization and convergence, Wiley, 1986]

b) Sei nun $|E|<\infty$ und E mit der diskreten Metrik $d_E(x,y)=\mathbf{1}(x\neq y)$ ausgestattet. Für $f\in D([0,\infty),E)$ sei $\tau_0^{(f)}:=0$, sofern $\tau_{i-1}^{(f)}<\infty$ setzen wir

$$\tau_i^{(f)}:=\inf\{t>\tau_0^{(f)}:f(t)\neq f(\tau_{i-1}^{(f)})\},\quad i\in\mathbb{N}$$

(ansonsten sei auch $\tau_i^{(f)}=\infty$) und

$$\xi_j^{(f)}:=f(\tau_j^{(f)})\quad\text{für } j\in\mathbb{N}_0 \text{ mit } j\leqslant M^{(f)}:=\inf\{k:\tau_k^{(f)}=\infty\}.$$

Seien $f_n,f\in D([0,\infty),E)$. Zeigen Sie: Die Folge $\tau_{i-1}^{(f)}$, $i\in\mathbb{N}$ besitzt keine Häufungspunkte im Endlichen. Es gilt $d(f_n,f)\to 0$ für $n\to\infty$ genau dann, wenn gilt

1. $M^{(f_n)}\to M^{(f)}$ für $n\to\infty$,
2. für jedes $k\in\mathbb{N}$ ist $\xi_{k\wedge M^{(f)}}^{(f_n)}=\xi_{k\wedge M^{(f)}}^{(f)}$ für alle genügend großen n und
3. für jedes $k\in\mathbb{N}$ gilt $\tau_{k\wedge M^{(f)}}^{(f_n)}\xrightarrow[n\to\infty]{}\tau_{k\wedge M^{(f)}}^{(f)}$.

c) Zeigen Sie: Unter den Voraussetzungen von Lemma 1.6 konvergieren die Prozesse $(X_{\lfloor t/c_N\rfloor})_{t\geqslant 0}$, aufgefasst als $D([0,\infty),E)$-wertige Zufallsvariablen, gegen die dort angegebene zeitkontinuierliche Markovkette X.