

Übungen zu „Stochastische Modelle der Populationsgenetik“ Blatt 2

Abgabe: Mittwoch, 11. Mai 2016 vor der Vorlesung

4. Im Wright-Fisher-Modell mit Populationsgröße N sei in der Generation 0 ein Anteil p_N der Individuen vom Typ B , der Rest sei vom Typ b . Die Typen werden über die Generationen vererbt. In der Generation g werden zwei Individuen rein zufällig aus der Population gezogen.

a) Berechnen Sie die *erwartete Heterozygotie* h_g , d.h. die Wahrscheinlichkeit, dass die beiden Individuen nicht den gleichen Typ haben.

b) Es sei $t \geq 0$, und es existiere $\lim_{N \rightarrow \infty} p_N =: p$. Was ist dann $\lim_{N \rightarrow \infty} h_{\lfloor tN \rfloor}$?

5. Für $n \leq N \in \mathbb{N}$ sei $(\mathcal{C}_k^n)_{k \in \mathbb{N}_0}$ der austauschbare n -Coalescent zur Kinderzahlverteilung $\mathcal{L}(\mathcal{N})$ bei Populationsgröße N .

Zeigen Sie: $(\mathcal{C}_k^n)_{k \in \mathbb{N}_0}$ ist eine Markovkette mit Übergangswahrscheinlichkeiten

$$\mathbf{P}(\mathcal{C}_{k+1}^n = \eta | \mathcal{C}_k^n = \xi) = \frac{(N)_{a\downarrow}}{(N)_{b\downarrow}} \mathbf{E} \left[\prod_{\ell=1}^a (\mathcal{N}_\ell)_{b_\ell\downarrow} \right]$$

wobei ξ und η Partitionen von $[n]$ sind mit $a := \#\eta \leq \#\xi =: b$, und η so aus ξ hervorgeht, dass (für jedes $\ell = 1, \dots, a$) b_ℓ Partitionselemente von ξ zu einem Partitionselement von η vereinigt werden (mit $b_1 + \dots + b_a = b$). Beweisen Sie dazu Lemma 1 und 2 und deren Korollar (erst für $n = N$) im Ergänzungsblatt zur Vorlesung 3. Dabei kann sich ein Blick auf den Beweis von Lemma 1.5 im Manuskript von M. Birkner (siehe Link auf der LV-Seite) lohnen.

6. Für jede gerade natürliche Zahl N definieren wir eine Nachkommensverteilung in einem Cannings-Modell wie folgt: Sei $\mathcal{M}$ Multinomial($N, (\frac{1}{N}, \dots, \frac{1}{N})$) verteilt und $\mathcal{H} = (\mathcal{H}_1, \dots, \mathcal{H}_N)$ so, dass für ein rein zufällig ausgewähltes $I \in \{1, \dots, N\}$ die Komponente $\mathcal{H}_I$ den Wert $N/2$ hat und die restlichen Komponenten Multinomial($\frac{N}{2}, (\frac{1}{N-1}, \dots, \frac{1}{N-1})$) verteilt sind. Die Zufallsvariable $\mathcal{N}$ komme so zustande, dass mit W'keit $1 - 1/N$ die Zufallsvariable $\mathcal{M}$ realisiert wird und mit W'keit $1/N$ die Zufallsvariable $\mathcal{H}$.

a) Begründen Sie, warum die Bedingung von Möhle und Sagitov verletzt ist, d.h. Dreifachverschmelzungen nicht asymptotisch vernachlässigbar sind gegenüber Paarverschmelzungen.

b) Diskutieren Sie (ohne formalen Beweis), gegen welchen Prozess die Folge der zeitreskalierten Stichprobengenealogien für $N \rightarrow \infty$ konvergieren. (Mit "zeitreskaliert" ist hier gemeint: N Generationen pro Zeiteinheit.)

7. In einem Moran-Modell zur Populationsgröße N sind zur Zeit Null k Individuen vom Typ A . Die Anzahl der Typ A -Individuen zur Zeit t sei Z_t . Zu dieser Zeit t zieht man (ohne Zurücklegen) eine Stichprobe der Größe n . Begründen Sie aus der graphischen Konstruktion des Modells die Identität

$$\mathbf{E}_k \left[\frac{Z_t(Z_t - 1) \cdots (Z_t - n + 1)}{N(N - 1) \cdots (N - n + 1)} \right] = \mathbf{E}_n \left[\frac{k(k - 1) \cdots (k - D_t + 1)}{N(N - 1) \cdots (N - D_t + 1)} \right].$$

Dabei sei $D = (D_r)_{r \geq 0}$ der Anzahl-Prozess im Kingman-Coalescent.