

**Aufgabe 4.1 (Konvergenz der Genealogien von Cannings-Modellen gegen den Kingman-Koaleszenten)** Für  $N \in \mathbb{N}$  sei  $\nu^{(N)} = (\nu_1^{(N)}, \dots, \nu_N^{(N)})$  ein austauschbarer Nachkommensanzahlvektor mit  $\nu_1^{(N)} + \dots + \nu_N^{(N)} = N$ , sei weiter

$$c_N = \frac{\mathbb{E}[\nu_1^{(N)}(\nu_1^{(N)} - 1)]}{N - 1} \quad \text{und} \quad d_N = \frac{\mathbb{E}[\nu_1^{(N)}(\nu_1^{(N)} - 1)(\nu_1^{(N)} - 2)]}{(N - 1)(N - 2)}$$

Es gelte  $c_N \rightarrow 0$  und  $d_N/c_N \rightarrow 0$  für  $N \rightarrow \infty$ .

In Satz 1.8 des Skripts *Stochastische Modelle der Populationsbiologie*, WS 15/16, <https://www.staff.uni-mainz.de/birkner/SMPB1516/smpb1516.pdf> wird gezeigt, dass die Ahnenverhältnisse einer  $n$ -Stichprobe in einer Schar von Cannings-Modellen basierend auf den (Verteilungen der) Nachkommensvektoren  $\nu^{(N)}$  in Verteilung gegen den  $n$ -Koaleszenten konvergieren, wenn die Zeit in Vielfachen von  $1/c_N$  gemessen wird.

Lesen und verstehen Sie diesen Satz und seinen Beweis. Was sind die entscheidenden Beweisschritte und -gedanken?

**Aufgabe 4.2 (Moran-Modell und Eldon-Wakeley-Modell)** 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

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

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 aus dem Skript *Stochastische Modelle der Populationsbiologie*, WS 15/16, <https://www.staff.uni-mainz.de/birkner/SMPB1516/smpb1516.pdf>. Sind die Voraussetzungen dieses Satzes 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

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

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

annimmt. Berechnen Sie  $c_N$  und  $d_N$  wie in Satz 1.8 aus dem genannten Skript. Unter welcher Bedingung an  $\gamma$  sind die Voraussetzungen von Satz 1.8 erfüllt?

c) Falls in b) die Voraussetzungen von Satz 1.8 nicht erfüllt sind, wie verhält sich für  $N \rightarrow \infty$  die Wahrscheinlichkeit, dass in einer Stichprobe der Größe  $n$  innerhalb einer Generation  $k$  Individuen zugleich einen gemeinsamen Ahnen haben, für  $2 < k \leq n$ ?

b.w.

**Aufgabe 4.3 (Stichprobenformel im 2-Typ-Fall)** Sei  $n \in \mathbb{N}$ ,  $0 < \theta_0, \theta_1 < \infty$ ,  $\theta := \theta_0 + \theta_1$ . Längs den Ästen eines  $n$ -Koaleszenten (aufgefasst als Baum mit kontinuierlichen Astlängen) legen wir gemäß den Sprüngen eines Poissonprozesses mit Rate  $\theta/2$  Mutationsereignisse, zudem werden diese Mutationsereignisse unabhängig mit einem Typ aus  $\{0, 1\}$  markiert: Mit Wahrscheinlichkeit  $\theta_i/\theta$  ist eine Mutation eine Typ- $i$ -Mutation,  $i = 0, 1$ . Weiter markieren wir die Wurzel des Baums unabhängig mit einer „künstlichen“ Mutation, die ebenfalls (und abhängig von allem anderen) mit Wahrscheinlichkeit  $\theta_i/\theta$  Typ  $i$  trägt,  $i = 0, 1$ . Der Typ eines Blatts sei definiert als der Typ der ersten Mutation, die man auf dem eindeutigen Pfad von diesem Blatt zur Wurzel trifft. Für  $n_0, n_1 \in \mathbb{N}_0$  mit  $n_0 + n_1 = n$  sei

$$p(n_0, n_1) := \mathbb{P}(n_0 \text{ Blätter haben Typ } 0 \text{ und } n_1 \text{ Blätter haben Typ } 1)$$

Zeigen Sie:

$$p(n_0, n_1) = \frac{(n_0 + n_1)!}{n_1! n_0!} \frac{(\theta_0)_{n_0 \uparrow} (\theta_1)_{n_1 \uparrow}}{(\theta_0 + \theta_1)_{(n_1 + n_0) \uparrow}}$$

wobei  $(x)k \uparrow = x(x+1) \cdots (x+k-1)$ .

[*Hinweis.* Ein Ansatz besteht darin, gemäß dem am nächsten zu den Blättern gelegenen Mutationsereignis zu zerlegen. Es kann dabei kombinatorisch angenehmer sein, zunächst

$$\tilde{p}(n_0, n_1) = \mathbb{P}(n_0 \text{ vorgegebene Blätter haben Typ } 0, \text{ die übrigen } n_1 \text{ Blätter Typ } 1)$$

zu betrachten. (Wie verhalten sich  $p(n_1, n_0)$  und  $\tilde{p}(n_0, n_1)$  zueinander?)