Übungen zu „Stochastische Modelle der Populationsgenetik"

16. *Die Länge des Kingman-Baums.*
a) Die Summe der Längen aller Zweige eines Kingman-Baums mit n Blättern lässt sich darstellen als

$$L_n := 2E_{\binom{2}{2}} + 3E_{\binom{3}{2}} + \cdots + nE_{\binom{n}{2}}$$

mit unabhängigen Exp $(\binom{j}{2})$-verteilten Zufallsvariablen $E_{\binom{j}{2}}$. Begründen Sie das.

b) Finden Sie die Verteilungsfunktion der Grenzverteilung von $\frac{1}{2}L_n - \ln(n-1)$ für $n \to \infty$.

17. *Warten auf den ersten Erfolg - einmal anders.*
a) In einer Pólya(1,1)-Urne sei die eine Ahnenkugel rot, die andere blau. Es sei L die Nummer des Zuges, bei der zum ersten mal eine blaue Kugel gezogen wird. Berechnen Sie die Gewichte der bedingten Verteilung von L, gegeben dass beim ersten Zug die rote Kugel gezogen wird.

b) Ein zufälliger Anteil U einer kontinuierlichen Population $\mathfrak{P}$ der Masse 1 sei rot und der Rest sei blau gefärbt; U sei uniform verteilt auf $[0,1]$. Die Zufallsvariablen U_1, U_2, ... seien unabhängige, uniform verteilte Züge aus $\mathfrak{P}$, sie seien auch unabhängig von U. Berechnen Sie die Verteilung der Nummer des Zuges, der erstmalig ins Blaue führt, gegeben, dass der allererste Zug ins Rote führt.

18. *de Finetti - il caso generale.*
$Y = (Y_1, Y_2, \ldots)$ sei eine zufällige Folge in einem Raum E, und $\Phi_n := \frac{1}{n}\sum_{i=1}^n \delta_{Y_i}$ sei die *n-te empirische Verteilung zu Y*. Der Satz von de Finetti lautet:

Sind die $Y_1, Y_2, \ldots$ austauschbar und ist E vollständig und separabel, dann konvergiert die Folge Φ_n fast sicher gegen eine zufällige Verteilung Φ, und gegeben Φ sind die $Y_1, Y_2, \ldots$ unabhängig und identisch verteilt mit Verteilung Φ.

(Mit der Konvergenz einer Folge (ρ_n) von Verteilungen auf E gegen eine Verteilung ρ auf E ist hier die schwache Konvergenz gemeint: $\int f(a)\rho_n(da) \to \int f(a)\rho(da)$ für alle stetigen, beschränkten $f : E \to \mathbb{R}$.)

Erweitern Sie den in der Vorlesung für den binären Fall diskutierten Beweis auf diesen allgemeinen Fall. Klären Sie dazu erst einmal, was es heißen soll, aus einer Verteilung der Form $\frac{1}{n}\sum_{i=1}^n \delta_{y_i}$ j-mal ohne Zurücklegen zu ziehen, und vergleichen Sie dieses mit einem j-maligen Ziehen mit Zurücklegen. Argumentieren Sie auf dieser Grundlage, dass für alle j und alle stetiegn beschränkten $f_1, \ldots, f_j : E \to \mathbb{R}$ gilt: $\mathbf{E}[f_1(Y_1)\cdots f_j(Y_j)\,|\,\Phi_n, Y_{n+1}, Y_{n+2}, \ldots]$ für große n nahe bei $\prod_{k=1}^j \mathbf{E}[f_j(Y_1)\,|\,\Phi_n, Y_{n+1}, Y_{n+2}, \ldots]$ ist, und zwar gleichmäßig in $\Phi_n, Y_{n+1}, Y_{n+2}, \ldots$. Folgern Sie, dass für ein geeignetes Teilfeld $\mathfrak{I}$ gilt: Bedingt unter $\mathfrak{I}$ sind $Y_1, Y_2, \ldots$ unabhängig und identisch verteilt. Weiter hilft dann das Starke Gesetz der Großen Zahlen in der Form von Glivenko-Cantelli...