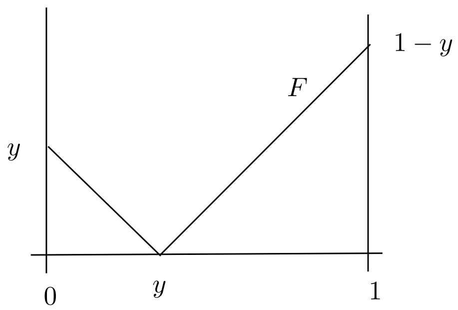
Übungen zu „Stochastische Modelle der Populationsgenetik“

21. *Doch noch nicht fixiert!* Wir betrachten das Infinite-Sites-Modell mit Mutationsrate $\theta/2$. Sei $n \geq 1$. Finden sie den Erwartungswert der Anzahl der Mutationen, die von jedem Individuum in einer n-Stichprobe getragen werden, aber nicht in der (als unendlich groß angenommenen) Gesamtpopulation fixiert sind. (Hinweis: Betrachten Sie die Differenz aus der Tiefe eines ∞-Coalescent und der Tiefe eines n-Coalescent.)

22. *Variation is the spice of life.* a) Wieviele segregierende Sites, die jeweils zwischen 1% und 99% der Population betreffen, sind in einer großen Population im Infinite Sites Modell mit Mutationsrate $\theta/2$ zu erwarten?(Hinweis: Die in der Vorlesung hergeleitete Formel $f_i := \theta/i$ ist hilfreich.)
b) Überschlagen Sie das Ergebnis für das menschliche Genom mit einer geschätzten Mutationsrate (auf der evolutionären Zeitskala) von $\approx 1/1330$ pro Nukleotid und rund 3 Milliarden Nukleotiden pro Genom. (Quelle: R. Durrett, Probability Models for DNA sequence evolution, 2nd ed., Springer 2008 (p. 51); L. Kruglyak, D. A. Nickerson, Variation is the spice of life, *Nature Genetics* 156 (2001) 234-236.)

23. *Der Kern des Müllers Green.*[1] Es sei $W = (W_t)$ ein Standard-Wienerprozess auf $[0,1]$, gestoppt beim ersten Treffen des Randes $\{0,1\}$. Für die erwartete Zeit, die der in x gestartete zufällige Pfad W in dy verbringt, schreiben wir $G(x,y)\,dy$, $x \in [0,1], y \in (0,1)$. Unser Ziel ist es, $G(x,y)$ zu berechnen. Gehen Sie dazu folgenden Weg:
a) Finden Sie ein Argument, warum für jedes feste y die Abbildung $x \to G(x,y)$ jeweils auf $[0,y]$ und $[y,1]$ affin linear ist. (Hinweis: Erinnern Sie sich an Aufgabe 1.)
b) Es geht jetzt um $G(y,y) = \mathbf{E}_y \left[\int_0^T \delta_y(W_t)\,dt\right]$. Dabei ist T die erste Treffzeit von $\{0,1\}$, und $z \to \delta_y(z)$ die durch $\int_0^1 h(z)\delta_y(z)dz := h(y)$ definierte “verallgemeinerte Funktion”. (Definiert man die Ableitung F' über die partielle Integrationsformel $\int_0^1 h(z)F'(z)dz := -\int_0^1 h'(z)F(z)\,dz$ für alle am Rand verschwindenden glatten h, dann lässt sich δ_y als Ableitung von $\mathbf{1}_{[y,1]}$ auffassen, und dieses wiederum als $\frac{1}{2}F''$, mit der skizzierten Funktion F. Wenden Sie (couragiert) die Itô-Formel auf $F(W_T) - F(W_0)$ an. (Eigentlich braucht es hier ja noch ein Approximationsargument, aber der Kern der Sache ist schon das Gesagte...)



[1]Zur wunderbaren Biographie von George Green (1793-1841) siehe http://www.gap-system.org/ history/Biographies/Green.html