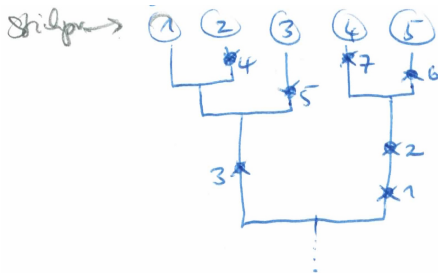


Erinnerungen: IMS-Modell

Datenmatrix

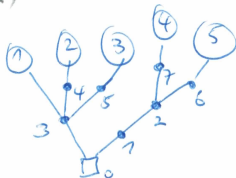


ein möglicher Koaleszenten-Baum



Äquivalente Kodierungen:

„Gen-Baum“:
(geumreht)



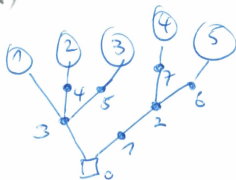
(innere
Knoten
≙ Mutationen,
Blätter ≙ Stücker)

Pfade zur Wurzel:

- ① : (3, 0)
- ② : (4, 0)
- ③ : (5, 0)
- ④ : (7, 2, 1, 0)
- ⑤ : (6, 2, 1, 0)

Erinnerung: Gewurzelte und ungewurzelte (Gen-)Bäume

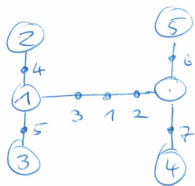
„Gen-Baum“:
(gewurzelt)



(innere Knoten
≙ Mutationen,
Blätter ≙ Stichproben)



Zug. ungewurzelter Baum



← eine „abgeleitete“ / „erschlossene“ / „gefolgerte“
(engl. „inferred“)
sequenz
(⊙ : (2, 1, 0))



(Wir betrachten im Folgenden nur den gewurzelten Fall)

$\theta > 0$, $n \in \mathbb{N}$, $D_n =$ Sequenzbeobachtungen an den Blättern eines n -Koaleszenten, längs dessen Kanten sich mit Rate $\theta/2$ Mutationen gemäß dem IMS-Modell ereignen

$$p_\theta(\mathbf{t}, \mathbf{n}) := \mathbb{P}_\theta(D_n \sim (\mathbf{t}, \mathbf{n}))$$

Erinnerung (Prop. 18):

$$\begin{aligned} p_\theta(\mathbf{t}, \mathbf{n}) &= \frac{\binom{n}{2}}{n\theta/2 + \binom{n}{2}} \sum_{i: n_i \geq 2} \frac{n_i - 1}{n - 1} p_\theta(\mathbf{t}, \mathbf{n} - \mathbf{e}_i) \\ &\quad + \frac{\theta/2}{n\theta/2 + \binom{n}{2}} \sum_{\substack{i: n_i=1, x_{j0} \text{ unique,} \\ s(\mathbf{t}_i) \neq \mathbf{t}_j \forall j}} p_\theta(s_i(\mathbf{t}), \mathbf{n}) \\ &\quad + \frac{\theta/2}{n\theta/2 + \binom{n}{2}} \frac{1}{d} \sum_{\substack{i: n_i=1, \\ x_{j0} \text{ unique}}} \sum_{j: s(\mathbf{t}_i) = \mathbf{t}_j} (n_j + 1) p_\theta(\tau_i(\mathbf{t}), \tau_i(\mathbf{n} + \mathbf{e}_j)). \end{aligned}$$

mit Randbedingung $p_\theta((1, 0), (1)) = 1$.

$(s(\mathbf{t}_i))$ entfernt die jüngste Mutation aus Typ i , $\tau_i(\mathbf{t})$ entfernt den i -ten Typ; ist strikt rekursiv in $C(\mathbf{t}, \mathbf{n}) := s + n$, der Komplexität der Stichprobe.

Interpretiere Genealogie mit Mutationen als „Historie“

(=Abfolge der Zustände in der Geschichte der Stichprobe zwischen jüngstem gemeinsamem Vorfahren und Gegenwart)

$$\mathcal{H} = (H_{-\tau+1} = ((1), (0)), H_{-\tau+2}, \dots, H_{-1}, H_0 = (\mathbf{t}, \mathbf{n}))$$

Bem.: Es können (u.U. sehr viele) verschiedene Historien zur selben Stichprobenkonfiguration (\mathbf{t}, \mathbf{n}) führen.

Interpretiere Genealogie mit Mutationen als „Historie“

(=Abfolge der Zustände in der Geschichte der Stichprobe zwischen jüngstem gemeinsamem Vorfahren und Gegenwart)

$$\mathcal{H} = (H_{-\tau+1} = ((1), (0)), H_{-\tau+2}, \dots, H_{-1}, H_0 = (\mathbf{t}, \mathbf{n}))$$

Bem.: Es können (u.U. sehr viele) verschiedene Historien zur selben Stichprobenkonfiguration (\mathbf{t}, \mathbf{n}) führen.

Simulation (auch) mit **Ethier-Griffiths-Urne**

- Beginne mit 2 Blättern
- Wenn aktuell k Blätter:

Füge einem Blatt eine neue Mutation hinzu m. Ws. $\frac{\theta}{\theta+(k-1)}$,

Verdopple ein Blatt m. Ws. $\frac{k-1}{\theta+(k-1)}$

(Blatt uniform aus den k Blättern gewählt).

- Stoppe sobald $n + 1$ Blätter, entferne das zuletzt erzeugte Blatt

Interpretiere Genealogie mit Mutationen als „Historie“

(=Abfolge der Zustände in der Geschichte der Stichprobe zwischen jüngstem gemeinsamem Vorfahren und Gegenwart)

$$\mathcal{H} = (H_{-\tau+1} = ((1), (0)), H_{-\tau+2}, \dots, H_{-1}, H_0 = (\mathbf{t}, \mathbf{n}))$$

Bem.: Es können (u.U. sehr viele) verschiedene Historien zur selben Stichprobenkonfiguration (\mathbf{t}, \mathbf{n}) führen.

Simulation (auch) mit **Ethier-Griffiths-Urne**

- Beginne mit 2 Blättern
- Wenn aktuell k Blätter:

Füge einem Blatt eine neue Mutation hinzu m. Ws. $\frac{\theta}{\theta+(k-1)}$,

Verdopple ein Blatt m. Ws. $\frac{k-1}{\theta+(k-1)}$

(Blatt uniform aus den k Blättern gewählt).

- Stoppe sobald $n + 1$ Blätter, entferne das zuletzt erzeugte Blatt

(Analogon zur Hoppe-Urne für das IMS-Modell; eine effiziente Alternative zum offensichtlichen 2-Schritt-Verfahren: 1. Simuliere n -Koaleszent, 2. simuliere IMS-Mutationen längs den Ästen)

(Nur der Vollständigkeit halber:)

Die Ethier-Griffiths-Urne ist eine Markovkette auf \mathcal{T}^* mit Übergängen

$(\mathbf{t}', \mathbf{n}') \rightarrow (\mathbf{t}'', \mathbf{n}'') =$

$$\begin{cases} \partial & \text{m. Ws. } \frac{n' - 1}{n' - 1 + \theta} & \text{falls } n' = n, \\ (\mathbf{t}', \mathbf{n}' + \mathbf{e}_i) & \text{m. Ws. } \frac{n' - 1}{n' - 1 + \theta} \frac{n'_i}{n'} & \text{falls } n' \leq n - 1, \\ (\alpha_i(\mathbf{t}'), \mathbf{n}') & \text{m. Ws. } \frac{\theta}{n' - 1 + \theta} \frac{1}{n'} & \text{falls } n'_i = 1, \\ (\epsilon_{i,j}(\mathbf{t}'), \epsilon_j(\mathbf{n}' - \mathbf{e}_j)) & \text{m. Ws. } \frac{\theta}{n' - 1 + \theta} \frac{n'_i}{n'} \frac{1}{d+1} & \text{falls } n'_i > 1 \ (j = 1, \dots, d + 1) \end{cases}$$

$(\mathbf{t}', \mathbf{n}')$ = aktueller Zustand (mit d Typen),

$\alpha_i(\mathbf{t}')$ fügt neue Mutation zu Typ i hinzu

$\epsilon_{i,j}(\mathbf{t}')$ kopiert Typ i , fügt eine neue Mutation hinzu und setzt den resultierenden

Typ an Position j in \mathbf{t} ein, $\epsilon_j(n''_1, \dots, n''_d) := (n''_1, \dots, n''_{j-1}, 1, n''_j, \dots, n''_d)$

(Nur der Vollständigkeit halber:)

Die Ethier-Griffiths-Urne ist eine Markovkette auf \mathcal{T}^* mit Übergängen

$(\mathbf{t}', \mathbf{n}') \rightarrow (\mathbf{t}'', \mathbf{n}'') =$

$$\begin{cases} \partial & \text{m. Ws. } \frac{n'-1}{n'-1+\theta} & \text{falls } n' = n, \\ (\mathbf{t}', \mathbf{n}' + \mathbf{e}_i) & \text{m. Ws. } \frac{n'-1}{n'-1+\theta} \frac{n'_i}{n'} & \text{falls } n' \leq n-1, \\ (\alpha_i(\mathbf{t}'), \mathbf{n}') & \text{m. Ws. } \frac{\theta}{n'-1+\theta} \frac{1}{n'} & \text{falls } n'_i = 1, \\ (\epsilon_{i,j}(\mathbf{t}'), \epsilon_j(\mathbf{n}' - \mathbf{e}_j)) & \text{m. Ws. } \frac{\theta}{n'-1+\theta} \frac{n'_i}{n'} \frac{1}{d+1} & \text{falls } n'_i > 1 \ (j = 1, \dots, d+1) \end{cases}$$

$(\mathbf{t}', \mathbf{n}')$ = aktueller Zustand (mit d Typen),

$\alpha_i(\mathbf{t}')$ fügt neue Mutation zu Typ i hinzu

$\epsilon_{i,j}(\mathbf{t}')$ kopiert Typ i , fügt eine neue Mutation hinzu und setzt den resultierenden

Typ an Position j in \mathbf{t} ein, $\epsilon_j(n''_1, \dots, n''_d) := (n''_1, \dots, n''_{j-1}, 1, n''_j, \dots, n''_d)$

Dies gestattet, die Übergangsw'keiten

$$\mathbb{P}_{\theta, n}(H_\ell = (\mathbf{t}'', \mathbf{n}'') \mid H_{\ell-1} = (\mathbf{t}', \mathbf{n}'))$$

der Historie abzulesen.

(vgl. z.B. B., Blath, Steinrücken, *Theor. Pop. Biology* 79 (2011) für kombinatorische Details)

$$p(\mathbf{t}, \mathbf{n}) = \sum_{\mathcal{H}: H_0 = (\mathbf{t}, \mathbf{n})} \mathbb{P}_{\theta, n}(\mathcal{H})$$

(und die Rekursion für $p(\mathbf{t}, \mathbf{n})$ aus Prop. 18 kann als eine Art der „Buchführung“ über mögliche Historien mit $H_0 = (\mathbf{t}, \mathbf{n})$ interpretiert werden)

$$p(\mathbf{t}, \mathbf{n}) = \sum_{\mathcal{H}: H_0 = (\mathbf{t}, \mathbf{n})} \mathbb{P}_{\theta, n}(\mathcal{H})$$

(und die Rekursion für $p(\mathbf{t}, \mathbf{n})$ aus Prop. 18 kann als eine Art der „Buchführung“ über mögliche Historien mit $H_0 = (\mathbf{t}, \mathbf{n})$ interpretiert werden)

Naiver Simulationsansatz:

$D^{(1)}, \dots, D^{(R)}$ u.a. simulierte n -Stichproben mit Parameter θ (z.B. via Ethier-Griffiths-Urne),

$$p_{\theta}(\mathbf{t}, \mathbf{n}) \approx \frac{1}{R} \sum_{r=1}^R \mathbf{1}(D^{(r)} \sim (\mathbf{t}, \mathbf{n})).$$

Korrekt im Limes $R \rightarrow \infty$,

Problem: Man braucht $R \gg 1/p_{\theta}(\mathbf{t}, \mathbf{n})$ (was leicht $> 10^{20}$ sein kann ...)

Importance sampling

$$\begin{aligned}
 p_{\theta}(\mathbf{t}, \mathbf{n}) &= \mathbb{P}_{\theta, n}(H_0 = (\mathbf{t}, \mathbf{n})) = \sum_{\mathcal{H}: H_0 = (\mathbf{t}, \mathbf{n})} \mathbb{P}_{\theta, n}(\mathcal{H}) \\
 &= \sum_{\mathcal{H}: H_0 = (\mathbf{t}, \mathbf{n})} \underbrace{\frac{\mathbb{P}_{\theta, n}(\mathcal{H})}{Q(\mathcal{H})}}_{=: w_Q(\mathcal{H})} Q(\mathcal{H}), \\
 &\qquad\qquad\qquad \text{importance weight}
 \end{aligned}$$

für jede Wahrscheinlichkeitsverteilung Q auf Historien, die

$\mathbb{P}_{\theta, n} \Big|_{\{H_0 = (\mathbf{t}, \mathbf{n})\}} \ll Q$ erfüllt.

Importance sampling

$$\begin{aligned}
 p_{\theta}(\mathbf{t}, \mathbf{n}) &= \mathbb{P}_{\theta, n}(H_0 = (\mathbf{t}, \mathbf{n})) = \sum_{\mathcal{H}: H_0=(\mathbf{t}, \mathbf{n})} \mathbb{P}_{\theta, n}(\mathcal{H}) \\
 &= \sum_{\mathcal{H}: H_0=(\mathbf{t}, \mathbf{n})} \underbrace{\frac{\mathbb{P}_{\theta, n}(\mathcal{H})}{Q(\mathcal{H})}}_{=: w_Q(\mathcal{H})} Q(\mathcal{H}), \\
 &\hspace{10em} \text{importance weight}
 \end{aligned}$$

für jede Wahrscheinlichkeitsverteilung Q auf Historien, die $\mathbb{P}_{\theta, n} \Big|_{\{H_0=(\mathbf{t}, \mathbf{n})\}} \ll Q$ erfüllt. Demnach

$$p_{\theta}(\mathbf{t}, \mathbf{n}) \approx \hat{p}_{\theta, Q, R}(\mathbf{t}, \mathbf{n}) := \frac{1}{R} \sum_{i=1}^R w_Q(\mathcal{H}^{(i)}),$$

mit $\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(R)}$ u.a., $\sim Q$, und

$$\mathbb{E}_{\theta}[\hat{p}_{\theta, Q, R}(\mathbf{t}, \mathbf{n})] = p_{\theta}(\mathbf{t}, \mathbf{n}), \quad \text{Var}_{\theta}(\hat{p}_{\theta, Q, R}(\mathbf{t}, \mathbf{n})) = \frac{1}{R} \text{Var}_Q(w_Q(\mathcal{H}))$$

(und die Kunst ist, diese Varianz möglichst klein zu machen...).

Importance sampling, 2: Für gegebenes (\mathbf{t}, \mathbf{n})

$$p_{\theta}(\mathbf{t}, \mathbf{n}) \approx \hat{p}_{\theta, Q, R}(\mathbf{t}, \mathbf{n}) := \frac{1}{R} \sum_{i=1}^R w_Q(\mathcal{H}^{(i)}),$$

mit $\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(R)}$ u.a., $\sim Q$, $w_Q(\mathcal{H}) = \mathbb{P}_{\theta, n}(\mathcal{H}) / Q(\mathcal{H})$, benötigt (nur)
 $\mathbb{P}_{\theta, n} |_{\{H_0 = (\mathbf{t}, \mathbf{n})\}} \ll Q$

Für Praktikabilität wünschenswert:

- Q sollte konzentriert sein auf $\{\mathcal{H} : H_0 = (\mathbf{t}, \mathbf{n})\}$
- $H_0, H_{-1}, H_{-2}, \dots$ Markovkette unter Q

Importance sampling, 2: Für gegebenes (\mathbf{t}, \mathbf{n})

$$p_{\theta}(\mathbf{t}, \mathbf{n}) \approx \hat{p}_{\theta, Q, R}(\mathbf{t}, \mathbf{n}) := \frac{1}{R} \sum_{i=1}^R w_Q(\mathcal{H}^{(i)}),$$

mit $\mathcal{H}^{(1)}, \dots, \mathcal{H}^{(R)}$ u.a., $\sim Q$, $w_Q(\mathcal{H}) = \mathbb{P}_{\theta, n}(\mathcal{H}) / Q(\mathcal{H})$, benötigt (nur) $\mathbb{P}_{\theta, n} |_{\{H_0 = (\mathbf{t}, \mathbf{n})\}} \ll Q$

Für Praktikabilität wünschenswert:

- Q sollte konzentriert sein auf $\{\mathcal{H} : H_0 = (\mathbf{t}, \mathbf{n})\}$
- $H_0, H_{-1}, H_{-2}, \dots$ Markovkette unter Q

Der Vorschlag von Griffiths & Tavaré

Q_{GT} (Verteilung der) Markovkette startend von $H_0 = (\mathbf{t}, \mathbf{n})$ mit

$$Q_{\text{GT}}(H_{i-1} | H_i) \propto \mathbb{P}_{n, \theta}(H_i | H_{i-1})$$

R.C. Griffiths, S. Tavaré, Ancestral inference in population genetics. *Statist. Sci.* 9:307–319, (1994).

(vgl. Rekursion in Prop. 18; dies kann als allgemeiner Ansatz zur Lösung linearer Rekursionsgleichungen aufgefasst werden)

Optimale Vorschlagsverteilung

$Q_{\text{opt}}(\cdot) := \mathbb{P}_{n,\theta}(\cdot \mid H_0 = (\mathbf{t}, \mathbf{n}))$ ist optimal:

- Varianz des Schätzers=0 denn $w_{\text{opt}}(\mathcal{H}^{(i)}) \equiv p_{\theta}(\mathbf{t}, \mathbf{n})$.
- (Leider i.A. nur hypothetische Lösung): Q_{opt} zu konstruieren ist genauso schwer wie die Berechnung von $p_{\theta}(\mathbf{t}, \mathbf{n})$
- H_0, H_{-1}, \dots ist Markovkette unter Q_{opt} . (Dies ist eine allgemeine Eigenschaft der Zeitumkehr von Markovketten, vgl. Nagasawas Formel)

Optimale Vorschlagsverteilung

$Q_{\text{opt}}(\cdot) := \mathbb{P}_{n,\theta}(\cdot | H_0 = (\mathbf{t}, \mathbf{n}))$ ist optimal:

- Varianz des Schätzers=0 denn $w_{\text{opt}}(\mathcal{H}^{(i)}) \equiv p_{\theta}(\mathbf{t}, \mathbf{n})$.
- (Leider i.A. nur hypothetische Lösung): Q_{opt} zu konstruieren ist genauso schwer wie die Berechnung von $p_{\theta}(\mathbf{t}, \mathbf{n})$
- H_0, H_{-1}, \dots ist Markovkette unter Q_{opt} . (Dies ist eine allgemeine Eigenschaft der Zeitumkehr von Markovketten, vgl. Nagasawas Formel)

Der Vorschlag von Stephens & Donnelly

Q_{SD} (Verteilung der) Markovkette startend von $H_0 = (\mathbf{t}, \mathbf{n})$ mit folgenden Übergängen: Wähle eine Linie uniform,

- wenn gezogener Typ und dessen „äusserste“ Mutation nur einmal vorhanden: entferne diese Mutation
- wenn mindestens zwei Linien diesen Typ haben: verschmelze ein solches Paar

M. Stephens, P. Donnelly, Inference in Molecular Population Genetics, Journal of the Royal Statistical Society, Series B, 62, 605–655, (2000)

(Dies wäre tatsächlich optimal im Fall des IMA-Modells, vgl. Ewens'sche Stichprobenformel.)

Der Vorschlag von Hobolt, Uyenoyama & Wiuf

A. Hobolt A, M.K. Uyenoyama, C. Wiuf, Importance sampling for the infinite sites model, Stat. Appl. Genet. Mol. Biol. 7 (2008)

Beob.: Für Stichproben, die nur eine Mutation enthalten,
 $(\mathbf{t}, \mathbf{n}) = (((1, 0), (0)), (m, n - m))$ kann man die Dynamik unter
 $\mathcal{Q}_{\text{opt}} = \mathbb{P}_{n,\theta}(\cdot \mid H_0 = (\mathbf{t}, \mathbf{n}))$ bestimmen:

$$\begin{aligned}
 p_{\theta}^{(1)}(n, m) &:= \mathbb{P}_{n,\theta} \left(H_{-1} = (((1, 0), (0)), (m - 1, n - 1 - m)) \right. \\
 &\qquad \qquad \qquad \left. \mid H_0 = (((1, 0), (0)), (m, n - m)) \right) \\
 &= \frac{\sum_{k=2}^{n-m+1} \frac{m-1}{n-k} \frac{1}{k-1+\theta} \binom{n-m-1}{k-2} \binom{n-1}{k-1}^{-1}}{\sum_{k'=2}^{n-m+1} \frac{1}{k'-1+\theta} \binom{n-m-1}{k'-2} \binom{n-1}{k'-1}^{-1}}
 \end{aligned}$$

(vgl. Kor. 2.7 zur Skelettkette des Kingman-Koaleszenten)

Der Vorschlag von Hobolt, Uyenoyama & Wiuf: Q_{HUW}

Sei

$$u_{i,m} = \begin{cases} p_{\theta}^{(1)}(n, d_m) \cdot \frac{n_i}{d_m} & \text{falls Typ } i \text{ Mutation } m \text{ hat,} \\ (1 - p_{\theta}^{(1)}(n, d_m)) \cdot \frac{n_i}{n - d_m} & \text{sonst,} \end{cases}$$

wobei $d_m = \text{Anz. Ind. in der aktuellen Konfiguration, die Mutation } m \text{ tragen.}$

Q_{HUW} schlägt Typ i vor mit W 'keit

$$q_{\text{HUW}}(i | (\mathbf{t}, \mathbf{n})) \propto \begin{cases} \sum_m u_{i,m} & \text{wenn Typ } i \text{ „erlaubt“} \\ 0 & \text{sonst.} \end{cases}$$

(Ein Typ ist nicht erlaubt, wenn er Vielfachheit 1 hat, aber seine äusserste Mutation in einem weiteren Typ vorkommt.)

Übergang, wenn vorgeschlagener Typ i ist

- Vielfachheit 1 hat : entferne äusserste Mutation
- Vielfachheit ≥ 2 hat : Verschmelzung innerhalb Typ i

Aspekte der Genealogie, gegeben die Daten

Zur Verteilung von T_{MRCA} gegeben $D_n \sim (\mathbf{t}, \mathbf{n})$:

Gegeben $\mathcal{H} = (H_0, H_1, \dots, H_\tau)$ ist $T_{\text{MRCA}} \sim \sum_{i=1}^\tau W_i$ mit W_i u.a.,
 $W_i \sim \text{Exp}(\ell_i \frac{\theta}{2} + \binom{\ell_i}{2})$ wenn in H_i ℓ_i Linien.

Simuliere $\mathcal{H}^{(r)}$, $r = 1, \dots, R$ u.a. gemäß \mathcal{Q} , dazu jeweils

$$T^{(r)} = \sum_{i=1}^\tau W_i^{(r)} :$$

$$\mathbb{P}_\theta \left(T_{\text{MRCA}} \leq t \mid D_n \sim (\mathbf{t}, \mathbf{n}) \right) \approx \frac{\frac{1}{R} \sum_{r=1}^R \frac{\mathbb{P}_\theta(\mathcal{H}^{(r)})}{\mathcal{Q}(\mathcal{H}^{(r)})} \mathbf{1}(T^{(r)} \leq t)}{\hat{p}_{\theta, \mathcal{Q}, R}(\mathbf{t}, \mathbf{n})}$$

Sequenzinformation aus der Kontrollregion der mitochondrialen DNS einer Stichprobe von $n = 55$ weiblichen Nuh-Chah-Nulth ($s = 18$ segregierende Stellen)

Daten aus R.H. Ward, B.L. Frazier, K. Dew-Jager, and S. Pääbo, Extensive Mitochondrial Diversity Within a Single Amerindian Tribe, *Proc. Nat. Acad. Sci. USA*, 88:8720–872, (1991)

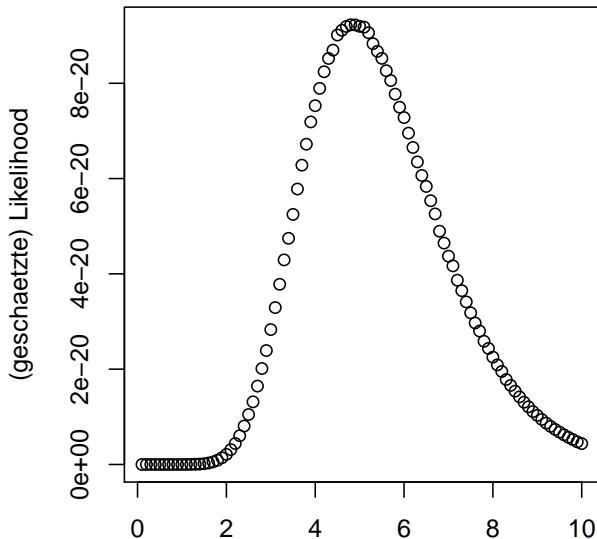
wie editiert in R.C. Griffiths, S. Tavaré, Ancestral inference in population genetics, *Statist. Sci.* 9: 307–319, (1994)

```

 2 : A G G A A T C C T T T C T C T T C
 2 : A G G A A T C C T T T C T C T T C
 1 : G A G G A C C C T C T T C C C T T T
 3 : G G A G A C C C C C T T C C C T T C
19 : G G G A A T C C T C T T C T C T T C
 1 : G G G A G T C C T C T T C T C T T C
 1 : G G G G A C C C T C C C C C T T T
 1 : G G G G A C C C T C C C T C C T T T
 4 : G G G G A C C C T C T T C C C C C T
 8 : G G G G A C C C T C T T C C C C T T
 5 : G G G G A C C C T C T T C C C T T C
 4 : G G G G A C C C T C T T C C C T T T
 3 : G G G G A C C T T C T T C C C T T C
 1 : G G G G A C T C T C T T C C T T T C

```


Likelihood-Funktion

(basierend auf $R = 1.2 \cdot 10^7$ Replikaten pro Punkt)Mini-Studie zu den
Importance Sampling-Verfahren

G+T:

R	10^5	$5 \cdot 10^5$	10^6
---	--------	----------------	--------

rel.F.	0.53	0.11	0.08
--------	------	------	------

S+D:

R	10^5	$5 \cdot 10^5$	10^6
---	--------	----------------	--------

rel.F.	0.049	0.037	0.027
--------	-------	-------	-------

H+U+W:

R	10^5	$5 \cdot 10^5$	10^6
---	--------	----------------	--------

rel.F.	0.072	0.024	0.015
--------	-------	-------	-------

Verteilungsfunktion der Zeit bis zum MRCA, gegeben die Daten

