

7. Übung zur Vorlesung “Einführung in die Bioinformatik I, 2. Teil”

Sommersemester 2014

Sebastian Böcker, Martin Engler

Ausgabe: 5. Juni 2014,
Abgabe: 12. Juni 2014 nach der Übung

Aufgabe 1 (5 Punkte): Gegeben sei die Distanzmatrix:

A	B	C		
0	6	8		A
	0	5		B
		0		C

Benutzen Sie Neighbor Joining, um einen additiven Baum zu rekonstruieren. Berechnen Sie die zum additiven Baum gehörige Baum-Metrik, und vergleichen Sie sie mit der Eingabe-Matrix.

Aufgabe 2 (5 Punkte): Gegeben seien zwei faire sechsseitige Würfel. Diese werden nacheinander geworfen. Folgende Ereignisse sollen betrachtet werden:

- A = „Die Augenzahl des ersten Würfels ist 3, die des zweiten ist 5“
- B = „Die Summe der Augen ist 8“
- C = „Die Augenzahl des ersten Würfels ist ungleich der des zweiten Würfels“

Bestimmen Sie folgende Wahrscheinlichkeiten: $\mathbb{P}(A)$, $\mathbb{P}(B)$, $\mathbb{P}(C)$, $\mathbb{P}(B \cap C)$, $\mathbb{P}(B|C)$, $\mathbb{P}(C|B)$. Sind B und C unabhängig?

Aufgabe 3 (8 Punkte): Der HIV-Test ELISA erkennt eine vorhandene HIV-Infektion mit einer Wahrscheinlichkeit von 99,9 % und liefert bei einer nicht infizierten Person mit einer Wahrscheinlichkeit von 99,8 % ein negatives Testergebnis. Angenommen, es wird mit diesem Test eine Reihenuntersuchung in einer Bevölkerung, in der 0,1 % der Personen HIV-positiv sind, vorgenommen. Wieviel Prozent der Personen, deren Testergebnis positiv ist, sind infiziert? Verwenden Sie für Ihre Berechnung Bayes'sche Statistik.

Aufgabe 4 (7 Punkte): Sie spielen das „Zonk-Spiel“: In einer Spielshow befindet sich der Hauptgewinn hinter einer von drei Türen, hinter den beiden anderen befinden sich „Zonks“ (Nieten). Sie wählen eine der drei Türen, die aber noch nicht geöffnet wird. Jetzt öffnet der Spielleiter eine der beiden anderen Türen, hinter der sich *immer* ein Zonk befindet (der Spielleiter weiß, hinter welcher Tür der Gewinn und hinter welchen Türen die beiden Zonks sind). Schließlich bietet Ihnen der Spielleiter an, Ihre Wahl zu überdenken und sich vor die andere noch geschlossene Tür zu stellen. Was ist statistisch sinnvoller: Bleiben Sie vor Ihrer Tür stehen, oder wechseln Sie zu der dritten, noch geschlossenen Tür? Erläutern Sie Ihre Wahl mit Hilfe von Bayes'scher Statistik, Modell ist „Preis ist hinter meiner zuerst gewählten Tür“.

Bonusaufgabe (10 Punkte): Gegeben sei folgende Distanzmatrix:

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	
0	2	8	12	18	18	<i>A</i>
	0	4	8	18	18	<i>B</i>
		0	6	18	18	<i>C</i>
			0	8	12	<i>D</i>
				0	10	<i>E</i>
					0	<i>F</i>

Benutzen Sie die folgenden agglomerativen Cluster-Verfahren, um phylogenetische Wurzelbäume aus der Matrix zu rekonstruieren: (a) UPGMA, (b) Complete linkage clustering. Vergleichen Sie die Bäume.