

# 8. Übung zur Vorlesung “Algorithmische Massenspektrometrie”

Wintersemester 2018/2019

Sebastian Böcker, Martin Hoffmann

Ausgabe: 03. Januar 2019, Abgabe: 10. Januar 2019

1. **Berechnung von Isotopenmuster über die Binomialverteilung** Bei der Berechnung von Isotopenmustern von Molekülen aus dem CHNOPS Alphabet genügt es in der Regel sich auf Isotopenspezies zu beschränken, die maximal um 10 Dalton von der monoisotopischen Masse abweichen.
  - (a) Wie groß müsste die Masse eines Moleküls sein, damit ein +10 Peak noch über ein Prozent Intensität zeigt. Berechnen Sie dies für den *worst-case* – ein Molekül, das nur aus Schwefelatomen besteht und nur zwei Isotopen kennt:  $^{32}\text{S}$  mit Häufigkeit 94.99% und  $^{34}\text{S}$  mit Häufigkeit 5.01%. Beachten Sie, dass der nominale Massenunterschied zwischen den beiden Isotopen 2 Dalton beträgt.
  - (b) Organische Moleküle bestehen hauptsächlich aus den Elementen C und H. Wiederholen Sie die obige Aufgabe, diesmal jedoch für ein Molekül das nur aus C-Atomen besteht und die Isotopen  $^{12}\text{C}$  mit Häufigkeit 98.93% und  $^{13}\text{C}$  mit Häufigkeit 1.07% enthält.
  - (c) Warum sind Isotopenspezies, die um mehr als 10 Dalton von der monoisotopischen Masse abweichen für die Identifizierung von Metaboliten in Massenspektren nicht relevant?

(5 Punkte)

2. **Faltung** Gegeben seien drei (endliche) diskrete Zufallsvariablen  $X$ ,  $Y$  und  $Z$  mit  $Z = X + Y$ . Dann lässt sich die Verteilung von  $Z$  folgendermaßen aus den Verteilungen von  $X$  und  $Y$  berechnen:

$$\mathbb{P}(Z = i) = \sum_j \mathbb{P}(X = i - j) \cdot \mathbb{P}(Y = j) \quad (1)$$

Wenn wir die Wahrscheinlichkeiten der diskreten Zustände in einen Array speichern, ergibt sich folgender Algorithmus:

```
FOLD( $X = x_0, \dots, x_{m-1}$ ,  $Y = y_0, \dots, y_{n-1}$ )
   $Z = \text{new array with length } m + n \text{ filled with zeroes.}$ 
  FOR  $i \leftarrow 0$  UPTO  $m$ 
    FOR  $j \leftarrow 0$  UPTO  $n$ 
       $Z_{i+j} = Z_{i+j} + x_i \cdot y_j$ 
    END FOR;
  END FOR;
  RETURN  $Z$ 
END.
```

- (a) Gegeben sei ein gezinkter Würfel, der in  $\frac{1}{3}$  der Fälle die Zahl 1 würfelt, während die Wahrscheinlichkeit für die anderen Zahlen gleich ist. Nutzen Sie die Gleichung 1 um die Wahrscheinlichkeit zu berechnen, mit 2 Würfeln eine Augensumme von 4 zu würfeln.
- (b) Nutzen Sie den angegebenen Algorithmus um die Wahrscheinlichkeiten aller möglichen Augensummen beim Wurf zweier **fairer** Würfel (jede Zahl wird mit gleicher Wahrscheinlichkeit geworfen) zu berechnen.
- (c) Zurück zu den Summenformeln: Berechnen Sie die Wahrscheinlichkeiten der Isotopenspezies des Moleküls  $C_4O_2$ . Beschränken Sie sich auf die Isotopenspezies, deren Nominalmasse um maximal 2 Dalton von der monoisotopischen Nominalmasse abweichen. Die Isotopenverteilungen finden Sie in Tabelle 1. Betrachten Sie dabei jedes Atom wie einen Würfel. So entspricht die Isotopenverteilung von  $O_2$  gleich der Faltung der Isotopenverteilung von O mit sich selbst. Die Isotopenverteilung von  $C_4O_2$  ist dann die Faltung der Isotopenverteilung von  $C_4$  mit der Verteilung von  $O_2$ . Nutzen Sie den obigen Algorithmus um die Verteilungen zu berechnen.

(8 Punkte)

Element (Symbol)	Isotop	Masse	Häufigkeit (%)
Wasserstoff (H)	$^1H$	1.007825	99.985
	$^2H$	2.014102	0.015
Kohlenstoff (C)	$^{12}C$	12.0	98.890
	$^{13}C$	13.003355	1.110
Stickstoff (N)	$^{14}N$	14.003074	99.634
	$^{15}N$	15.000109	0.366
Sauerstoff (O)	$^{16}O$	15.994915	99.762
	$^{17}O$	16.999132	0.038
	$^{18}O$	17.999161	0.200
Phosphor (P)	$^{31}P$	30.973762	100
Schwefel (S)	$^{32}S$	31.972071	95.020
	$^{33}S$	32.971459	0.750
	$^{34}S$	33.967867	4.210
	$^{36}S$	35.967081	0.020

Tabelle 1: Relative Isotopenhäufigkeiten und Massen der sechs in Lebewesen häufigsten Elemente in Dalton, gerundet auf sechs Nachkommastellen.