

Übung 6

6.1) $P_0(X_{n+1}=1 | X_1 + \dots + X_n = s) = (s+1) / (n+2)$

$n=7$
 $s=4$

$\rightarrow P_0(X_{n+1}=1 | X_1 + \dots + X_n = s) = \frac{8}{9}$

6.2)

G A T A C
G A T C A
A T T G A
G C T C G

A: .1 .2 0 .1 .2
C: 0 .1 0 .2 .1
G: .3 0 0 .1 .1
T: 0 .1 .4 0 0

\rightarrow

A	1/4	2/4	0	1/4	2/4
C	0	1/4	0	2/4	1/4
T	3/4	0	0	1/4	1/4
G	0	1/4	1	0	0

mit Pseudocounts: + $\frac{1}{4}$

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

\rightarrow

A	2/8	3/8	1/8	2/8	3/8
C	1/8	2/8	1/8	3/8	2/8
T	4/8	1/8	1/8	2/8	2/8
G	1/8	2/8	5/8	1/8	1/8

↳ Stellen mit ursprünglich 0 werden durch Pseudocounts durch geringe Wahrscheinlichkeit von 1/8 ersetzt

↳ dadurch werden die anderen Wahrscheinlichkeiten leicht verringert

6.3)

A	1/2	1/4	0	1
C	1/2	1/4	0	0
T	0	1/4	1/2	0
G	0	1/4	1/2	0

1. $P_0(ACTA | \text{Profile}) = 1/2 \cdot 1/4 \cdot 1/2 \cdot 1 = \frac{1}{16}$

2. $P_0(CTG | \text{Profile}) = 1/2 \cdot 1/4 \cdot 1/2 \cdot 0 = 0$

3. $P_0(CCTAA | \text{Profile}) = 1/2 \cdot 1/4 \cdot 1/2 \cdot 1 \cdot 1/4 = \frac{1}{64}$

6.4)

- y -Achse : bits zwischen 0-2
↳ gibt an, wie viel Information an der jeweiligen Stelle gegeben ist
- x -Achse zeigt Positionen in 5'-3'-Richtung
- relative Größe der Buchstaben gibt deren Häufigkeit in der Sequenz wieder
- Informationsgehalt spiegelt wieder, wie konserviert die Position ist
- RBS ist reich an A und G
- erstes translatiertes Codon rechts von y -Achse → Startcodon
- Vermutlich Shine-Dalgarno-Sequenz bei -12 bis -7