

6.6.2022

Übung Bioinformatik 5

1. Identische Zufallssequenzen: Wahrscheinlichkeit bestimmen, dass 2 Sequenzen der Länge N mit einem Alphabet aus A Buchstaben identisch ist.

(0/1)

- pro Stelle 2 Möglichkeiten $\rightarrow N$ -mal A Möglichkeiten
 \rightarrow wie bei Würfeln (3. geworfen: 6^3 Möglichkeiten)
- Anzahl möglicher Sequenzen: A^N
- Wahrscheinlichkeit für auftreten einer bestimmten Sequenz: $p = \frac{1}{A^N}$
- 2 Sequenzen genau diese eine Sequenz: $p_2 = p^2 = \left(\frac{1}{A^N}\right)^2$

2. Wahrscheinlichkeit von Zufallssequenzen

Bestimmen Sie die Wahrscheinlichkeit $\Pr(N=100, A=2, \text{Pattern}="01", t=1)$

■ über Gegenwahrscheinlichkeit

$$p = 1 - p_{\text{geg}}$$

- dass es nicht zutrifft: 100 Möglichkeiten
- Möglichkeiten insg.: $A^N = 2^{100}$

$$p_{\text{geg}} = \frac{100}{2^{100}} = 7,9 \cdot 10^{-29}$$

$$p = 1 - 7,9 \cdot 10^{-29} = 0,999 \Rightarrow \text{zu } 99,9\% \text{ kommt die Sequenz } 01 \text{ in einer } 100\text{-stelligen Zufallssequenz vor}$$

0000...
1111...
1ser ab 1. Stelle bis vorletzter Stelle: 98 Möglichkeiten