



Prof. Hanspeter Herzel
Prof. Nils Blüthgen
Dr. Manuela Benary
Dr. Christoph Schmal

www.sys-bio.net/teaching
h.herzel@biologie.hu-berlin.de
nils.bluthgen@charite.de
manuela.benary@biologie.hu-berlin.de
christoph.schmal@charite.de

BIOINFORMATIK SS 2017 – ÜBUNGSBLATT 6

Gib deine Lösung bitte am 19.06.2017 in der Vorlesung ab. Alternativ kannst du die Lösung auch per E-Mail an christoph.schmal@charite.de schicken.

1. Wiederholung Statistik

A) Man betrachte den Wurf zweier sechseitiger idealer bzw. „fairer“ Würfel. Gib die Wahrscheinlichkeit folgender Ereignisse an:

- | | |
|------------------|------------------|
| (a) Doppelsechs | (c) Augensumme 3 |
| (b) Augensumme 2 | (d) Augensumme 7 |

B) Wie viel verschiedene Trinukleotide und Hexanukleotide treten in DNA-Sequenzen auf?

C) Berechne oder vereinfache folgende Ausdrücke:

- | | |
|--------------------|-----------------------|
| (a) $5!$ | (c) $\log_2 4^6$ |
| (b) $\binom{6}{2}$ | (d) $\ln x^2 - \ln x$ |

2. Sequenzstatistik I

Gegeben ist ein Plasmid mit 10.000 Basenpaaren und unabhängigen Nukleotiden mit folgenden Wahrscheinlichkeiten: $p_T = p_A = 0,3$ und $p_C = p_G = 0,2$. Wieviele Erkennungssequenzen CGTACG für ein Restriktionsenzym können im Durchschnitt erwartet werden? Wie groß ist die Standardabweichung? Berechnen Sie die Wahrscheinlichkeiten, dass die Sequenz CGTACG NICHT gefunden wird, genau einmal auftritt und mindestens zweimal gefunden wird.

3. Sequenzstatistik II

In einem Genom sei der C+G-Gehalt 40% und die Nukleotide seien statistisch unabhängig. Berechnen Sie die Wahrscheinlichkeiten für die drei Stopkodons. Geben Sie den mittleren Abstand zufälliger Stopkodons an. Wiederholen Sie die Rechnungen für einen C+G-Gehalt von 70%.