



## Blatt 3

**Aufgabe 3.1** Wir betrachten eine DNA-Sequenz der Länge  $N$ . Wir nehmen an, dass alle Positionen statistisch unabhängig sind und die Nukleotide an jeder einzelnen Position gleichverteilt auftreten.

- (a) Wir betrachten wie in der Vorlesung die (überlappend) gezählte Anzahl  $k$  des Auftretens des Muster *gcgg*. Bestimmen Sie Erwartungswert und Varianz.
- (b) Wir nehmen nun (fälschlicherweise) an, daß die Anzahl  $k$  der Auftreten des Musters binomial verteilt ist.
  - (i) Warum ist diese Annahme falsch?
  - (ii) Geben Sie den resultierenden Erwartungswert und Varianz an.
- (c) Erweitern Sie die Überlegungen aus Teilaufgabe (a) für ein beliebiges Muster der Länge 5.

**Aufgabe 3.2** Betrachten Sie  $N$  statistisch unabhängige Zufallsvariablen  $X_n$ , die alle einer Normalverteilung  $\mathcal{N}(x|\mu, \sigma^2)$  folgen.

- (a) Bestimmen Sie die Maximum Likelihood Schätzer für  $\mu$  und  $\sigma$ .
- (b) Vergleichen Sie dies mit den erwartungstreuen Schätzern.

**Aufgabe 3.3** Generieren Sie 10 normalverteilte Zufallszahlen  $x_n$  mit der Dichtefunktion

$$P(X = x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

- (a) Berechnen Sie nun aus dieser "Stichprobe" die MLEs  $\hat{\mu}$  und  $\hat{\sigma}^2$ .
- (b) Wiederholen Sie das Ganze  $10^6$  mal und erstellen Sie Histogramme für  $\hat{\mu}$  und  $\hat{\sigma}^2$ .
- (c) Welchen Zusammenhang gibt es zwischen den Histogrammen und den Erwartungswerten  $E(\hat{\mu})$  und  $E(\hat{\sigma}^2)$ . Welchen Zusammenhang gibt zwischen den Histogrammen und den Schätzern, wenn wir diese als Zufallsvariable betrachten?
- (d) Wiederholen Sie alles. Ändert sich etwas?

Hinweis (für octave): `help normal_rnd`