# Übungen zur Vorlesung „Stochastik für die Informatik“

Abgabe der Lösungen zu den S-Aufgaben: Freitag, 27. Oktober 2017, vor der Vorlesung (12:25-12:30 im Magnus HS)

**1. S.** Wie in der Vorlesung betrachten wir eine Gesamtfläche $G$ bestehend aus $g$ Pixeln und eine Teilfläche $F$ bestehend aus $f$ Pixeln. Jetzt geht es um eine viermal wiederholte rein zufällige Wahl eines Pixels aus $G$ (anders gesagt: um ein Ziehen aus $G$ mit Zurücklegen), beschrieben durch eine auf dem Wertebereich $\{1,\ldots,g\}^4$ uniform verteilte Zufallsvariable $X = (X_1, X_2, X_3, X_4)$.
a) Wieviele verschiedene mögliche Ausgänge von $X$ gibt es, bei denen eines der $X_i$ auf die Menge $\{1,\ldots,f\}$ und drei auf die Menge $\{f+1,\ldots,g\}$ fallen? Drücken Sie das Ergebnis durch $g$ und durch $p := f/g$ aus.
b) Wie wahrscheinlich ist es, dass (in der Situation von a) von den vier zufällig aus $G$ gewählten Pixeln genau eines aus $F$ gewählt wird?
c) Es sei $M$ die zufällige Trefferquote von $F$ . Illustrieren Sie für $p = 0.195$ und $n = 4$ das Ergebnis aus b) mittels des über den Link auf der StofI-Web-Seite zur Verfügung gestellten R-Programms “Monte Carlo Simulation”.[1] Betrachten Sie dazu ein Histogramm der Schätzwerte aus (z.B.) 1000 Wiederholungen des Zufallsexperiments.

**2.** Erkunden Sie in der in Aufgabe 1 beschriebenen Situation (wieder für $p = 0.195$) mittels des R-Programms “Monte Carlo Simulation”, wie sich die Genauigkeit der Schätzung verändert, wenn (i) $n = 100$ (ii) $n = 400$ (iii) $n = 1600$ Punkte in die Menge $G$ geworfen werden: Um welchen Faktor (circa) wird jeweils das Histogramm der Schätzwerte schmäler?

**3.** Es sei $n \in \mathbb{N}$. Wir nennen $a = (a_0, a_1, \ldots, a_n)$ einen *Standardpfad der Länge* $n$, falls $a_0 = 0$ und $a_i - a_{i-1} \in \{-1, +1\}$, $i = 1, \ldots, n$. (Wir interpretieren $a_i$ als Position des Pfades zum Zeitpunkt $i$.) Es sei $S$ die Menge der Standardpfade der Länge 20 und $S_1 := \{a \in S : \text{ es existiert ein } i \text{ mit } a_i \geq 1\}$.
a) Begründen Sie die folgende Aussage:

$$\#\{a \in S_1 : a_{20} > 1\} = \#\{a \in S_1 : a_{20} < 1\}.$$

*Betrachten Sie dazu den letzten Zeitpunkt $\ell = \ell(a)$, zu dem der Pfad $a \in S_1$ die Position 1 besucht, und denken Sie sich Sie das Pfadstück $(a_\ell, a_{\ell+1}, \ldots, a_{20})$ an der Höhe 1 gespiegelt.*
b) Gilt folgende Aussage:

$$2\#\{a \in S_1 : a_{20} > 1\} = 2\#\{a \in S : a_{20} > 1\} = \#\{a \in S : a_{20} \neq 0\} \quad ?$$

*(Hinweis: Nach einer geraden Anzahl von Schritten landet man bei einer geraden Zahl.)*
c) Es sei $X = (X_0, X_1, \ldots, X_{20})$ ein rein zufälliger Standardpfad der Länge 20. Drücken Sie die Wahrscheinlichkeit des Ereignisses $\{X \text{ erreicht jemals einen Wert } \geq 1\}$ durch die Wahrscheinlichkeit des Ereignisses $\{X_{20} = 0\}$ aus.

**4 S.** Es sei $n \in \mathbb{N}$. Wir nennen $a = (a_1, \ldots, a_n)$ eine 01-*Folge der Länge* $n$, falls $a_i \in \{0, 1\}$, $i = 1, \ldots, n$. Wir sagen, dass $a$ das Muster 0110 enthält, falls $a_i a_{i+1} a_{i+2} a_{i+3} = 0110$ für ein $i \in \{1, \ldots, n-3\}$. Für $k \in \mathbb{N}$ sei nun $X = (X_1, \ldots, X_{4k})$ eine rein zufällige 01-Folge der Länge $4k$. Warum gilt folgende Aussage:

$$\mathbf{P}(X \text{ enthält das Muster } 0110) \geq 1 - (1 - 2^{-4})^k?$$

[1] Das frei verfügbare statistische Programmpaket R bekommen Sie über www.r-project.org, zu finden auch über google → R, auf Ihren Rechner.