Institut für angewandte Mathematik
Sommersemester 2009
Andreas Eberle, Martin Slowik

3. Übungsblatt „Algorithmische Mathematik II"

Abgabe bis Dienstag, 14 Uhr, in der Mathematikbibliothek (LWK)

---

1. (Zufallsstichproben mit und ohne Zurücklegen)

a) Geben Sie die Massenfunktionen der Binomialverteilung mit Parametern n und p, und der hypergeometrischen Verteilung mit Parametern n, m und r an (Bezeichnungen wie in der Vorlesung). Versuchen Sie, zunächst nicht in den Unterlagen nachzusehen. Überlegen Sie sich, was die Verteilungen beschreiben, und wie die Massenfunktionen daher aussehen.

b) Eine Prüfung besteht aus 12 Fragen, die mit *ja* oder *nein* zu beantworten sind. Sie gilt bei mindestens 8 richtigen Antworten als bestanden.

   i) Ein Student kreuzt auf gut Glück die Antworten an. Mit welcher Wahrscheinlichkeit besteht er die Prüfung?

   ii) Wie ändert sich die Wahrscheinlichkeit, wenn er 2 Fragen mit Sicherheit beantworten kann und nur den Rest zufällig ankreuzt?

   iii) Falls er gar nichts weiß, wäre es dann für ihn günstiger, zufällig 6-mal *ja* und 6-mal *nein* anzukreuzen, vorausgesetzt, dass für 6 Fragen die richtige Antwort *ja* lautet?

c) Zeigen Sie: Für $m \to \infty$ und $r \to \infty$ mit $p = \frac{r}{m}$ konstant konvergiert die hypergeometrische Verteilung gegen die Binomialverteilung mit Parametern n und p. Interpretieren Sie die Aussage anschaulich.

2. (Erwartungswert von positiven Zufallsvariablen mit ganzzahligen Werten)

a) Sei T eine Zufallsvariable mit Werten in $\mathbb{Z}_+$. Zeigen Sie

$$E[T] = \sum_{k=1}^{\infty} P[T \geq k].$$

b) Berechnen Sie, wieviele Würfe im Schnitt nötig sind, bis beim Würfeln zum ersten Mal eine "6" fällt.

3. (Unabhängige Ereignisse) . Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum. Zwei Ereignisse $A_1, A_2 \in \mathcal{A}$ heißen unabhängig, falls

$$P[A_1 \cap A_2] = P[A_1]P[A_2]$$

gilt. Allgemeiner heißen Ereignisse $A_1, \ldots, A_n \in \mathcal{A}$ unabhängig, falls

$$P[A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}] \;=\; P[A_{i_1}] \cdot P[A_{i_2}] \cdot \ldots \cdot P[A_{i_k}]$$

für alle $1 \leq k \leq n$ und $1 \leq i_1 < i_2 \ldots < i_k \leq n$ gilt.

Seien nun $A_1, \ldots, A_n \in \mathcal{A}$ unabhängige Ereignisse, $n \geq 2$. Zeigen Sie:

a) Auch die Ereignisse A_1^c und A_2, sowie A_1^c und A_2^c sind jeweils unabhängig.

b) Allgemeiner gilt für alle $k = 0, 1, \ldots, n$:

$$P[A_1^c \cap \ldots \cap A_k^c \cap A_{k+1} \cap \ldots \cap A_n] \;=\; P[A_1^c] \cdot \ldots \cdot P[A_k^c] \cdot P[A_{k+1}] \cdot \ldots \cdot P[A_n].$$

c) Für $B_i \in \{A_i, A_i^c\}$, $1 \leq i \leq n$, sind die Ereignisse $B_1, \ldots, B_n$ ebenfalls unabhängig.

d) Gilt $P[A_i] = p \in [0, 1]$ für alle $1 \leq i \leq n$, dann ist die Anzahl

$$S_n(\omega) \;:=\; |\{1 \leq i \leq n \,:\, \omega \in A_i\}|$$

der eingetretenen Ereignisse binomialverteilt mit Parametern n und p.

4. (Zufallspermutationen)

a) Eine aufsteigende Teilfolge einer Permutation $\omega \in S_n$ ist eine Sequenz $\omega(i_1) < \omega(i_2) < \cdots < \omega(i_k)$ mit $1 \leq i_1 < i_2 < \cdots < i_k \leq n$. Sei $N(\omega)$ die Anzahl der aufsteigenden Teilfolgen der Permutation ω. Zeigen Sie, daß für eine (gleichverteilte) Zufallspermutation aus S_n gilt:

$$E[N] \;=\; \sum_{k=1}^{n} \frac{1}{k!}\binom{n}{k}.$$

b) Sei

$$\Omega_n \;:=\; \{1, 2, \ldots, n\} \times \{2, 3, \ldots, n\} \times \cdots \times \{n-1, n\}.$$

Zeigen Sie, daß die Abbildung $X : \Omega_n \to S_n$,

$$X(\omega) \;=\; \tau_{n-1,\omega(n-1)} \circ \tau_{n-2,\omega(n-2)} \circ \cdots \circ \tau_{2,\omega(2)} \circ \tau_{1,\omega(1)}$$

eine Bijektion ist. Hierbei bezeichnet $\tau_{i,j}$ die Transposition von i und j. Folgern Sie, daß der Algorithmus aus der Vorlesung tatsächlich eine gleichverteilte Zufallspermutation erzeugen würde, wenn man die Pseudozufallszahlen durch echte unabhängige gleichverteilte Zufallszahlen ersetzen würde.

P. (Zufällige Teilmengen und Monte Carlo Simulation)

a) Implementieren Sie einen Algorithmus, der eine (pseudo)zufällige n-elementige Teilmenge aus der Menge $\{1, 2, \ldots, m\}$ auswählt. Die Teilmenge kann in Mathematica als Liste gespeichert werden, und soll über eine Funktion **rsubset[m,n]** abrufbar sein.

b) Verwenden Sie die Funktion **rsubset[m,n]** um eine Funktion **rhypgeom[m,r,n]** zu definieren, die eine Stichprobe von der hypergeometrischen Verteilung mit Parametern m, r und n erzeugt.

c) Obwohl jede n-elementige Teilmenge mit derselben Wahrscheinlichkeit auftritt, empfindet man Mengen $\omega \subseteq \{1, 2, \ldots, m\}$, die viele aufeinanderfolgende Zahlen enthalten, als weniger "zufällig". Diese Eigenschaft quantifizieren wir nun durch die Zahl

$$X(\omega) \ := \ \max \{k \ : \ k \text{ aufeinanderfolgende Zahlen gehören zu } \omega \}.$$

Explizite Formeln für die Wahrscheinlichkeiten

$$p_X(k) \ = \ P[\{\omega : \ X(\omega) = k\}]$$

unter der Gleichverteilung auf den n-elementigen Teilmengen von $\{1, 2, \ldots, m\}$ sind schwierig zu finden. Ein möglicher Ausweg sind *Monte Carlo Schätzer*: Man simuliert s zufällige n-elementige Teilmengen $\omega_1, \ldots, \omega_s$, und berechnet die relativen Häufigkeiten

$$\hat{p}_X(k) \ := \ \frac{|\{1 \leq i \leq s \ : \ X(\omega_i) = k\}|}{s}$$

als Schätzwert für die Wahrscheinlichkeiten $p_X(k)$.

Berechnen Sie Monte-Carlo Schätzer für $n = 15$, $m = 30$, und verschiedene Werte für s zwischen 1 und 10000. Stellen Sie $\hat{p}_X$ jeweils graphisch dar.