Institut für angewandte Mathematik
Sommersemester 2013
Andreas Eberle, Matthias Erbar / Jan Maas

3. Übungsblatt „AlMa 2" und „Stochastik (Lehramt)"

Abgabe bis Dienstag 13 Uhr in den Postfächern gegenüber der Bibliothek

---

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). Sehen Sie dabei nicht in den Unterlagen nach, sondern 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 Gück 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 diese Aussage.

2. (Erwartungswert von positiven Zufallsvariablen mit ganzzahligen Werten)

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

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

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

3. (Random Walk)
Sei P die Gleichverteilung auf $\Omega = \{-1, +1\}^N$ und $X_i(\omega) = \omega_i$. Wir interpretieren

$$S_0 \ := \ 0 \quad \text{und} \quad S_n \ := \ X_1 + \ldots + X_n, \qquad n = 1, \ldots, N,$$

als zufällige Irrfahrt eines Teilchens auf $\mathbb{Z}$ mit Start in 0. Für $\lambda \in \mathbb{N}$ sei

$$T_\lambda \ = \ \min\{n > 0 \mid S_n = \lambda\}$$

der Zeitpunkt des ersten Besuchs in λ. Zeigen Sie, dass

a) die Verteilung von S_n gegeben ist durch

$$P[S_n = k] \ = \ \begin{cases} 0, & \text{falls } n + k \text{ ungerade oder } |k| > n, \\ 2^{-n}\binom{n}{(n+k)/2} & \text{sonst.} \end{cases}$$

b) das Reflexionsprinzip gilt, d.h. für jedes $c > 0$ erhält man:

$$P[S_n = \lambda - c, \, T_\lambda \leq n] \ = \ P[S_n = \lambda + c].$$

c) die Verteilung von T_λ gegeben ist durch:

$$P[T_\lambda \leq n] \ = \ P[S_n \geq \lambda] + P[S_n > \lambda].$$

4. (Bandrika 1)
Du hast dich im Nationalpark von Bandrika verlaufen. Von den Besuchern im Park sind zwei Drittel Touristen. Fragen nach der Richtung werden von diesen mit Wahrscheinlichkeit $\frac{3}{4}$ richtig beantwortet. Dabei sind Antworten auf wiederholte Fragen unabhängig, auch wenn die Frage und die Person dieselben sind. Wenn man hingegen einen Bandrikaner fragt, ist die Antwort immer falsch.

a) Du fragst eine Person, ob der Ausgang sich in Richtung Osten oder Westen befindet. Als Antwort erhälst du Osten. Wie groß ist die Wahrscheinlichkeit, dass das richtig ist?

b) Du fragst dieselbe Person nochmals und bekommst dieselbe Antwort. Zeige, dass die Wahrscheinlichkeit, nun die richtige Antwort erhalten zu haben, $\frac{1}{2}$ beträgt.

c) Du richtest dieselbe Frage ein drittes Mal an dieselbe Person und erhälst wieder die Antwort Osten. Wie hoch ist jetzt die Wahrscheinlichkeit, dass die Antwort stimmt?

d) Ein viertes Mal wird der geduldige Passant von dir gefragt, doch die Antwort ist wieder Osten. Zeige, dass die Antwort mit Wahrscheinlichkeit $\frac{27}{70}$ richtig ist.

e) Zeige für den Fall, dass die vierte Antwort Westen wäre, dass die Richtung Osten mit Wahrscheinlichkeit $\frac{9}{10}$ zutrifft.

5. (Datenanalyse; nur Stochastik für Lehramt)
Betrachten Sie die Statistiken auf der nächsten Seite. Stellen Sie den Bezug zur Vorlesung her.

P. (Zufällige Teilmengen und Monte Carlo Simulation; nur AlMa 2)

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

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

Explizite Formeln für die Wahrscheinlichkeiten

$$p_X(k) \;=\; P\big[\{\omega \mid X(\omega) = k\}\big]$$

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

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

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

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

Hinweis: Sie können in Mathematica ähnliche Befehle wie in C++ verwenden, es stehen aber auch diverse Befehle anderer Programmiersprachen zur Verfügung. Hier können Sie zum Beispiel die Mathematica-Funktionen **Module***,* **Do***,* **If***,* **Take** *und* **Table** *verwenden.*

Table 2. Total Income and Total Tax (in thousands of dollars), and Tax Rate for Taxable Income Tax Returns, by Income Category and Year

Adjusted Gross Income	1974 Income	1974 Tax	1974 Tax Rate	1978 Income	1978 Tax	1978 Tax Rate
under $ 5,000	41,651,643	2,244,467	.054	19,879,622	689,318	.035
$ 5,000 to $ 9,999	146,400,740	13,646,348	.093	122,853,315	8,819,461	.072
$ 10,000 to $14,999	192,688,922	21,449,597	.111	171,858,024	17,155,758	.100
$ 15,000 to $99,999	470,010,790	75,038,230	.160	865,037,814	137,860,951	.159
$ 100,000 or more	29,427,152	11,311,672	.384	62,806,159	24,051,698	.383
Total	880,179,247	123,690,314		1,242,434,934	188,577,186	
Overall Tax Rate			.141			.152