Technische Universität München **Sommersemester 2007**
Theoretische Informatik **Übungsblatt 6**
Prof. J. Esparza / M. Luttenberger / S. Schwoon

Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt wird am besprochen.

Aufgabe 6.1

Bekanntlich findet sich in jedem siebten Überraschungsei eine Figur aus einem Figurenset. Wir nehmen an, dass ein Figurenset aus sieben Figuren besteht, dass die Figuren gleichverteilt sind und jedes Ei 50 Cent kostet.

Es sei dann Y die Zufallsvariable, die angibt, wie viele Überraschungseier gekauft werden müssen, bis man einen vollständigen Figurensatz hat. Dann ist $P = \frac{1}{2}Y$ der Preis in Euro, den man bezahlen muss.

Verwenden Sie nun sowohl die Markov-Ungleichung

$$\Pr[P \geq t] \leq \frac{\mathbb{E}[P]}{t}$$

als auch die Chebyshev-Ungleichung

$$\Pr[|P - \mathbb{E}[P]| \geq t] \leq \frac{\mathrm{Var}[P]}{t^2},$$

um jeweils ein Intervall anzugeben, in dem der Preis mit mindestens 95 % Wahrscheinlichkeit liegt.

Lösungsvorschlag Sei Y die Anzahl der benötigten Versuche, bis man schließlich das Set zusammen hat. Dann gilt $Y = X_1 + \ldots + X_7$, wobei die X_i unabhängig sind und jeweils $X_i \sim \mathrm{Geo}(\frac{8-i}{49})$ verteilt sind.

Der Preis, den man für ein vollständiges Set dann bezahlen muss, ist somit $P = \frac{1}{2}Y$.

Man erhält also

$$\mathbb{E}[P] = \frac{1}{2}\sum_{i=1}^{7}\mathbb{E}[X_i] = \frac{1}{2}\sum_{i=1}^{7}\frac{49}{8-i} = \frac{49}{2}\cdot\sum_{i=1}\frac{1}{i} = 63.525.$$

Entsprechend gilt wieder wegen der Unabhängigkeit der X_i:

$$\mathrm{Var}[P] = \frac{1}{4}\sum_{i=1}^{7}\mathrm{Var}[X_i] = \frac{1}{4}\sum_{i=1}^{7}\left(\frac{49^2}{(8-i)^2} - \frac{49}{8-i}\right) = -31.7625 + \frac{49^2}{4}\cdot\sum_{i=1}^{7}\frac{1}{i^2} \approx 875.70.$$

Mit der Chebyshev-Ungleichung

$$\Pr[|P - \mathbb{E}[P]| \geq t] \leq \frac{\mathrm{Var}[P]}{t^2}$$

folgt also

$$\Pr[|P - \mathbb{E}[P]| < t] \geq 1 - \frac{\mathrm{Var}[P]}{t^2} \stackrel{!}{=} 0.95$$

bzw.

$$t^2 = \frac{\mathrm{Var}[P]}{0.05} = 17514 \text{ oder } t = 132.34.$$

Nach Chebyshev befindet sich der Preis also in dem Intervall $\mathbb{E}[P] + [-132.34, 132.34] = [0, 195.5]$ mit W'keit 0.95.

Mit der Markov-Ungleichung erhält man

$$\Pr[P < t] \geq 1 - \frac{\mathbb{E}[P]}{t} \stackrel{!}{=} 0.95,$$

also

$$t = \frac{\mathbb{E}[P]}{0.05} = 1270.5.$$

Die Markov-Ungleichung liefert also eine viel schlechtere Vorhersage mit $[0, 1270]$.

Tatsächlich gilt $\Pr[P \leq t] > 0.95$ bereits für $t = 119.5$ (per Programm berechnet).

Aufgabe 6.2

Bei der Herstellung von Chips können Fehler auf Grund von Defekten im verwendeten Wafer auftreten. Der Chip sei in n gleich große Bereiche $R_1, \ldots, R_n$ von jeweils der Fläche $A_n = \frac{A}{n}$ unterteilt, wobei A die Gesamtfläche des Chips sei und jeder Bereich entweder defekt oder funktionsfähig sein kann.

Die Wahrscheinlichkeit, dass einer der Bereiche R_i defekt ist, sei direkt proportional zur Fläche des Bereichs: $\mu \cdot A_n \in (0,1)$.

(a) Geben Sie die Verteilung der Zufallsvariablen Y an, welche die Anzahl der Defekte auf dem Chip zählt für $n \to \infty$.

(b) Wie viele Defekte hat ein Chip im Durchschnitt, wenn 95% aller Chips keine Defekte aufweisen?

Lösungsvorschlag Für n sei Y_n die Zufallsvariable, welche die Anzahl der fehlerhaften Bereiche zählt. Dann ist Y_n nach Aufgabenstellung $\mathrm{Bin}(n, \mu \cdot A_n)$ verteilt.

Nach Aufgabenstellung gilt weiter, dass $n \cdot A_n = A$ für alle $n \in \mathbb{N}$, wobei A die gesamte Chipfläche ist.

Nach der Vorlesung folgt damit

$$\lim_{n\to\infty} \mathrm{b}(k; n, \mu \cdot A_n) = \mathrm{po}(k; \mu \cdot A) = e^{-\mu \cdot A} \frac{(\mu \cdot A)^k}{k!}.$$

Damit ist die Wahrscheinlichkeit, dass der Chip keinen Defekt hat, gegeben durch

$$\mathrm{po}(0; \mu \cdot A) = e^{-\mu \cdot A} \stackrel{!}{=} 0.95.$$

Damit folgt, dass die durchschnittliche Defektanzahl pro Chip gerade $\mu \cdot A = -\ln 0.95 \approx 0.051$ beträgt.

Aufgabe 6.3

Ein Angler sitzt am Fluss und wartet auf Fische. Die Anzahl der in einer Stunde vorbeikommenden Fische bezeichnen wir mit der Zufallsvariablen X, die $\mathrm{Po}(\lambda)$-verteilt sei. Jeder Fisch beißt mit Wahrscheinlichkeit p an, mit Wahrscheinlichkeit $1 - p$ nicht. Y bezeichne die Anzahl der Fische, die der Angler in einer Stunde fängt.

(a) Begründen Sie, dass Y unter der Bedingung $X = n$ die Verteilung $\mathrm{Bin}(n, p)$ hat, d.h. dass $\Pr[Y = k \mid X = n] = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$ gilt.

(b) Berechnen Sie die Verteilung von Y.

Lösungsvorschlag

(a) Wenn $X = n$ feststeht, kann das Anbeißen jedes einzelnen der n Fische als Bernoulli-Experiment mit Parameter p betrachtet werden; Y ergibt sich dann als die Summe der "Erfolge" bei den n Experimenten und ist demzufolge $\mathrm{Bin}(n, p)$-verteilt.

(b)

$$\begin{aligned}
\Pr[Y = k] &= \sum_{n=0}^{\infty} \Pr[Y = k \mid X = n] \cdot \Pr[X = n] \\
&= \sum_{n=k}^{\infty} \frac{n!}{k! \cdot (n-k)!} \cdot p^k \cdot (1-p)^{n-k} \cdot \mathrm{e}^{-\lambda} \cdot \frac{\lambda^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{1}{k! \cdot n!} \cdot p^k \cdot (1-p)^n \cdot \mathrm{e}^{-\lambda} \cdot \lambda^{n+k} \\
&= \frac{(\lambda p)^k}{k!} \cdot \mathrm{e}^{-\lambda} \cdot \sum_{n=0}^{\infty} \frac{(\lambda \cdot (1-p))^n}{n!} \\
&= \frac{(\lambda p)^k}{k!} \cdot \mathrm{e}^{-\lambda} \cdot \mathrm{e}^{\lambda \cdot (1-p)} \\
&= \mathrm{e}^{-\lambda p} \cdot \frac{(\lambda p)^k}{k!}
\end{aligned}$$

Also gilt $Y \sim \mathrm{Po}(\lambda p)$.

Aufgabe 6.4 (freiwillige Zusatzaufgabe - wird nicht besprochen)

Auf dem letzten Übungsblatt wurde folgendes einfaches Modell für den Kurs einer Aktie vorgestellt:

Es seien $t_0 < t_1 < \ldots < t_n$ Zeitpunkte, wobei der Wert der Aktie zum Zeitpunkt t_i mit S_i bezeichnet sei. Weiterhin seien $R_1, \ldots, R_n$ unabhängige, identisch verteilte Zufallsvariablen mit

$$\Pr[R_1 = u] = p \text{ und } \Pr[R_1 = d] = 1 - p =: q,$$

wobei $0 < p, d < 1$ und $u > 1$. Für den Aktienwert nehmen wir dann

$$S_k = S_0 \prod_{i=1}^{k} R_i$$

an. Damit folgt, dass der Wertebereich von S_n gerade

$$W_{S_n} = \{S_0 \cdot u^k \cdot d^{n-k} \mid 0 \le k \le n\}$$

ist.

Sie haben bereits gezeigt, dass

$$\Pr[S_n = S_0 \cdot u^k \cdot d^{n-k}] = \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

und

$$\mathbb{E}[S_n] = S_0 \cdot (u \cdot p + d \cdot q)^n$$

gilt.

Beim Handeln mit Aktien gibt es die Möglichkeit, sogenannte *Call-Optionen* (kurz: *Call*) zu erwerben. Eine Call-Option wird dabei zum Zeitpunkt t_0 zu einem gewissen Preis C erworben und sichert dem Käufer das Recht zu, die Aktie zum Zeitpunkt t_n zu einem Preis K zu kaufen. Dabei wird K beim Erwerb der Call-Option zum Zeitpunkt t_0 festgelegt.

Mit dem Kauf einer Call-Option spekuliert man also darauf, dass sich der Aktienkurs zum Zeitpunkt t_n überhalb von K befinden wird. Der Wert des Calls zum Zeitpunkt t_n ist daher einfach $\max\{S_n - K, 0\}$: ist die Aktie zum Zeitpunkt t_n mehr als K wert, so kann man den Gewinn $S_n - K$ erhalten, indem man die Call-Option wahrnimmt, die Aktie zum Preis K kauft und dann direkt zum Preis S_n wieder verkauft; gilt andererseits $S_n \le K$, so ist die Call-Option wertlos.

Es stellt sich nun die Frage, wie man den Kaufpreis C der Call-Option zum Zeitpunkt t_0 in Abhängigkeit von S_n und K festlegen soll.

Wie Black, Scholes und Merton gezeigt haben, sollte der erwartete Wert der Call-Option zum Zeitpunkt t_n gerade mit dem Wert übereinstimmen, den man erhalten hätte, wenn man zum Zeitpunkt t_0 den Wert C auf ein Bankkonto eingezahlt hätte.

Wir nehmen an, dass das Konto mit einem Faktor $r > 1$ verzinst ist. Das heißt, befindet sich zu t_0 der Wert C auf dem Konto, so befindet sich dann $C \cdot r^n$ zum Zeitpunkt t_n auf dem Konto.

(a) Zeigen Sie, dass unter obigen Annahmen und Forderungen dann

$$C = S_0 \cdot \left(\frac{u \cdot p + d \cdot q}{r}\right)^n \cdot (1 - F_{n, \frac{u \cdot p}{u \cdot p + d \cdot q}}(\kappa)) - K \cdot r^{-n} \cdot (1 - F_{n,p}(\kappa)),$$

wobei

$$\kappa := \max\{k \in \mathbb{N} \mid S_0 \cdot u^k \cdot d^{n-k} \le K\}$$

und $F_{n,p}$ die Verteilungsfunktion der Verteilung $\mathrm{Bin}(n, p)$ sei.

Anmerkung: Im Fall, dass die R_i normalverteilt sind, lässt sich die Rechnung zu obiger Formel verallgemeinern. Dies führt dann auf die sogenannte „Black-Scholes“-Formel (`http://en.wikipedia.org/wiki/Black-Scholes_formula`), welche eine so wichtige Rolle in der Finanzmathematik spielt, dass Merton und Scholes u.a. für jene Formel 1997 den Nobelpreis erhalten haben.

Lösungsvorschlag Es gilt

$$
\begin{aligned}
& \mathbb{E}[\max\{S_n - K, 0\}] \\
= \; & \sum_{\omega \in [S_n > K]} (S_n(\omega) - K)\Pr[\omega] \\
= \; & -K \cdot \Pr[S_n > K] + \sum_{\omega \in [S_n > K]} S_n(\omega)\Pr[\omega] \\
= \; & -K \cdot \Pr[S_n > K] + \sum_{k=\kappa+1}^{n} S_0 \cdot u^k \cdot d^{n-k} \cdot \binom{n}{k} \cdot p^k \cdot q^{n-k} \\
= \; & -K \cdot \Pr[S_n > K] + S_0 \cdot \mathbb{E}[X_1]^n \cdot \sum_{k=\kappa+1}^{n} \binom{n}{k} \cdot \frac{(u \cdot p)^k \cdot (d \cdot q)^{n-k}}{(u \cdot p + d \cdot q)^n} \\
= \; & -K \cdot \Pr[S_n > K] + S_0 \cdot \mathbb{E}[X_1]^n \cdot \sum_{k=\kappa+1}^{n} \binom{n}{k} \cdot \left(\frac{u \cdot p}{u \cdot p + d \cdot q} \right)^k \cdot \left(\frac{d \cdot q}{u \cdot p + d \cdot q} \right)^{n-k} \\
= \; & S_0 \cdot \mathbb{E}[X_1]^n \cdot (1 - F_{n, \frac{u \cdot p}{u \cdot p + d \cdot q}}(\kappa)) - K \cdot (1 - F_{n,p}(\kappa)),
\end{aligned}
$$

wobei $F_{n,p}$ die Verteilungsfunktion der Verteilung $\mathrm{Bin}(n, p)$ sei. Insgesamt also

$$
C = S_0 \cdot \left(\frac{u \cdot p + d \cdot q}{r} \right)^n \cdot (1 - F_{n, \frac{u \cdot p}{u \cdot p + d \cdot q}}(\kappa)) - K \cdot r^{-n} \cdot (1 - F_{n,p}(\kappa)).
$$