Technische Universität München
Theoretische Informatik
Prof. J. Esparza / M. Luttenberger / S. Schwoon

Sommersemester 2007
Übungsblatt 7

Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt wird am besprochen.

Aufgabe 7.1

In dieser Aufgabe soll die erzeugende Funktion für eine negativ-binomial-verteilte Zufallsvariable berechnet werden.

(a) Zeigen Sie hierfür in einem ersten Schritt, dass für unabhängige Zufallsvariablen X, Y in $(\Omega, \Pr)$

$$G_{X+Y}(z) = G_X(z) \cdot G_Y(z)$$

gilt.

Hinweis: Da X und Y unabhängig sind, gilt dies auch für z^X und z^Y.

Bemerkung: Machen Sie sich klar, wie mit Hilfe dieses Resultats die erzeugende Funktion einer $\text{Bin}(n, p)$-verteilten Zufallsvariable sehr einfach aus der erzeugenden Funktion einer $\text{Bin}(1, p)$-verteilten Zufallsvarialbe gewonnen werden kann.

(b) Es sei nun Y negativ-binomial-verteilt, wobei Bernoulli-Experimente mit Erfolgswahrscheinlichkeit p solange wiederholt werden, bis der n.te Erfolg eintritt (kurz: $Y \sim \overline{\text{Bin}}(n, p)$).

Zeigen Sie, dass die erzeugende Funktion von Y gleich $\left(\frac{p \cdot z}{1-(1-p)\cdot z}\right)^n$ ist. Verwenden Sie (a) hierfür.

*(c) (Freiwillige Zusatzaufgabe - wird nicht in den Übungen besprochen.)

Falls Y $\overline{\text{Bin}}(n, p)$-verteilt ist, so sagt man, dass die Zufallsvariable $Z := Y - n$, welche nur die Anzahl der Misserfolge bis zum n.ten Erfolg zählt, $\overline{\text{Bin}}^+(n, p)$-verteilt ist.

Bestimmen Sie nun analog zu (b) zunächst die erzeugende Funktion $G_Z(z)$, und zeigen Sie dann, dass für $1 - p := \frac{\mu}{n}$ (mit $\mu > 0$)

$$\lim_{n\to\infty} G_Z(z) = e^{\mu\cdot(z-1)}$$

gilt.

Bemerkung: Wie auf den Folien vermerkt, bedeutet dies, dass die Verteilung $\overline{\text{Bin}}^+(n, 1 - \frac{\mu}{n})$ gegen die Verteilung $\text{Po}(\mu)$ konvergiert.

Lösungsvorschlag

(a) Da X und Y unabhängig sind, gilt dies auch für die Zufallsvariablen z^X und z^Y. Damit folgt

$$G_{X+Y}(z) = \mathbb{E}[z^{X+Y}] = \mathbb{E}[z^X \cdot z^Y] = \mathbb{E}[z^X] \cdot \mathbb{E}[z^Y] = G_X(z) \cdot G_Y(z).$$

(b) Y kann als die Summe von n geometrisch mit Parameter p verteilten Zufallsvariablen $X_1, \ldots, X_n$ geschrieben werden. Nach (a) folgt dann:

$$G_Y(z) = G_{X_1}(z)^n = \left(\frac{p \cdot z}{1-(1-p)\cdot z}\right)^n.$$

(c) Man kann Z wiederum als Summe von n Zufallsvariablen X_i mit $X_i \sim \overline{\text{Bin}}^+(1, p)$ schreiben. Dann erhält man zunächst

$$\Pr[X_i = k] = p \cdot q^k$$

und daher

$$G_{X_i}(z) = \sum_{k=0}^{\infty} z^k \cdot \Pr[X_i = k] = p \cdot \sum_{k=0}^{\infty} ((1-p)\cdot z)^k = \frac{p}{1-(1-p)\cdot z}.$$

Analog zu (b) folgt somit

$$G_Z(z) = \left(\frac{p}{1-(1-p)\cdot z}\right)^n.$$

Setzt man nun $p = 1 - \frac{\mu}{n}$ ein, so erhält man

$$G_Z(z) = \left(\frac{1-\frac{\mu}{n}}{1-\frac{\mu}{n}\cdot z}\right)^n = \left(1-\frac{\mu}{n}\right)^n \cdot \left(1-\frac{\mu}{n}\cdot z\right)^{-n} = \left(1+\frac{-\mu}{n}\right)^n \cdot \left(1+\frac{\mu\cdot z}{n}\right)^n \xrightarrow{n\to\infty} e^{-\mu}\cdot e^{\mu\cdot z} = e^{\mu\cdot(z-1)}.$$

Aufgabe 7.2

In Aufgabe 6.3 haben wir einen Angler betrachtet, an dessen Leine stündlich $X \sim \mathrm{Po}(\lambda)$ Fische vorbeikommen, wobei jeder Fisch F_i nur mit einer gewissen Wahrscheinlichkeit p anbeißt, d.h. $F_i \sim \mathrm{Bin}(1,p)$. Mit Y ist dann die Anzahl gefangener Fische bezeichnet wurden.

Es besteht also der folgende Zusammenhang zwischen den Zufallsvariablen Y, X und den F_i:

$$Y(\omega) = \sum_{i=1}^{X(\omega)} F_i(\omega).$$

Erklärung: Ein konkretes Elementarereignis ω legt bereits vollständig fest, dass $X(\omega)$ Fische am Angler vorbeikommen, und ob Fisch i anbeisst ($F_i(\omega) = 1$) oder nicht ($F_i(\omega) = 0$).

Wir berechnen erneut die Verteilung von Y, diesmal mit Hilfe von erzeugenden Funktionen.

(a) Da Summen mit zufälliger Summationsgrenze nicht nur speziell für Angler interessant sind, betrachten wir zunächst die etwas allgemeinere Situation, in welcher wir die Verteilung von X und der F_i nicht kennen.

Wie in Aufgabe 6.3 nehmen wir jedoch an, dass X und die F_i unabhänig von einander sind, die F_i identisch verteilt sind, und alle Zufallsvariablen nur Werte in $\mathbb{N}$ annehmen.

Insbesondere soll $\Pr[X = k] > 0$ für alle $k \in \mathbb{N}$ gilt. Zeigen Sie nun für die Zufallsvariable Y, dass Folgendes gilt:

$$G_Y(z) = G_X(G_{F_1}(z))$$

Hinweise:

- Nach Definition gilt $G_Y(z) = \mathbb{E}[z^Y]$. Da $\Pr[X = k] > 0$ für alle $k \in \mathbb{N}$ angenommen wird, kann Satz 12 angewendet werden.
- Beachten Sie, dass für unabhängige Zufallsvariablen U, V mit $\Pr[U \in A] > 0$ stets

$$\mathbb{E}[V \mid U \in A] = \sum_{v\in W_V} v \cdot \Pr[V = v \mid U \in A] = \sum_{v\in W_V} v \cdot \Pr[V = v] = \mathbb{E}[W]$$

gilt.

(b) Berechnen Sie nun mit Hilfe von (a) die Verteilung von Y im Fall unseres Anglers, d.h. für $X \sim \mathrm{Po}(\lambda)$ und $F_i \sim \mathrm{Bin}(1,p)$.

(c) In manchen Situationen mag die Verteilung der Zufallsvariablen X und F_i nicht bekannt sein, man weiss jedoch, dass die Annahmen aus (a) zutreffen, und kennt weiterhin auf Grund von Statistiken (ungefähr) $\mathbb{E}[X], \mathrm{Var}[X]$ und $\mathbb{E}[F_1] = \mathbb{E}[F_i], \mathrm{Var}[F_1] = \mathrm{Var}[F_i]$.

Leiten Sie mittels (a) Formeln für den Erwartungswert und die Varianz von Y her, welche allein auf $\mathbb{E}[X], \mathrm{Var}[X]$ und $\mathbb{E}[F_1], \mathrm{Var}[F_1]$ beruhen.

Hinweis: Beachten Sie, dass nach Vorlesung für eine Zufallsvariable Z mit erzeugender Funktion G_Z gilt:

$$\begin{array}{lcl} \mathbb{E}[Z] & = & G'_Z(1) \\ \mathrm{Var}[Z] & = & G''_Z(1) + G'_Z(1) - (G'_Z(1))^2 \\ 1 & = & G_Z(0). \end{array}$$

Lösungsvorschlag

(a)

$$\begin{aligned}
G_Y(z) = \mathbb{E}[z^Y] &= \textstyle\sum_{k\in\mathbb{N}} \mathbb{E}[z^Y \mid X = k] \cdot \Pr[X = k] \\
&= \textstyle\sum_{k\in\mathbb{N}} \mathbb{E}[z^{F_1+\ldots+F_k} \mid X = k] \cdot \Pr[X = k] \\
&\overset{\text{Unabhängigkeit von } X, F_1, \ldots, F_k}{=} \textstyle\sum_{k\in\mathbb{N}} \mathbb{E}[z^{F_1+\ldots+F_k}] \cdot \Pr[X = k] \\
&= \textstyle\sum_{k\in\mathbb{N}} \mathbb{E}[z^{F_1}]^k \cdot \Pr[X = k] \\
&= \mathbb{E}\left[\mathbb{E}[z^{F_1}]^X\right] \\
&= G_X(G_{F_1}(z)).
\end{aligned}$$

(b) Nach der Vorlesung gilt $G_X(z) = e^{\lambda\cdot(z-1)}$ und $G_{F_1}(z) = 1 - p + p \cdot z$. Mit (a) folgt daher

$$G_Y(z) = e^{\lambda\cdot(1-p+p\cdot z-1)} = e^{\lambda\cdot p\cdot(z-1)}$$

Wegen der Eindeutigkeit der erzeugenden Funktion ist Y also $\mathrm{Po}(\lambda \cdot p)$ verteilt.

(c) Es sei $\mu_X := \mathbb{E}[X]$, $\sigma_X^2 := \mathrm{Var}[X]$, $\mu_F := \mathbb{E}[F_1]$, $\sigma_F^2 := \mathrm{Var}[F_1]$.

$$\begin{aligned}
\mathbb{E}[Y] &\overset{\text{nach Vorlesung}}{=} G'_Y(1) \\
&= G'_X(G_{F_1}(1)) \cdot G'_{F_1}(1) \\
&\overset{\text{G(1) = 1}}{=} G'_X(1) \cdot G'_{F_1}(1) \\
&= \mu_X \cdot \mu_F.
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}[Y] &\overset{\text{nach Vorlesung}}{=} G''_Y(1) + G'_Y(1) - G'_Y(1)^2 \\
&= G''_X(G_{F_1}(1)) \cdot G'_{F_1}(1)^2 + G'_X(G_{F_1}(1)) \cdot G''_{F_1}(1) + \mu_X \cdot \mu_F - \mu_X^2 \cdot \mu_F^2 \\
&= G''_X(1) \cdot \mu_F^2 + \mu_X \cdot G''_{F_1}(1) + \mu_X \cdot \mu_F - \mu_X^2 \cdot \mu_F^2 \\
&= (\sigma_X^2 - \mu_X + \mu_X^2) \cdot \mu_F^2 + \mu_X \cdot (\sigma_F^2 - \mu_F + \mu_F^2) + \mu_X \cdot \mu_F - \mu_X^2 \cdot \mu_F^2 \\
&= \sigma_X^2 \cdot \mu_F^2 + \mu_X \cdot \sigma_F^2
\end{aligned}$$

Aufgabe 7.3

Nach dem Gesetz der großen Zahlen gilt für jede Folge von identisch verteilten, unabhängigen Zufallsvariablen $X_1, \ldots, X_n$

$$\Pr\left[\left|\mathbb{E}[X_1] - \frac{1}{n}\sum_{i=1}^{n} X_i\right| \geq \delta\right] \leq \frac{\mathrm{Var}[X_1]}{n \cdot \delta^2}.$$

Sind die X_i insbesondere $\mathrm{Bin}(1, p)$ verteilt, so gilt

$$\Pr\left[\left|p - \frac{1}{n}\sum_{i=1}^{n} X_i\right| \geq \delta\right] \leq \frac{p \cdot (1-p)}{n \cdot \delta^2}.$$

Dieser Sachverhalt kann benutzt werden, um die Güte eines Zufallszahlengenerators zu testen.

(a) Wie oben beschrieben, soll eine endliche Anzahl n von Bernoulli-Experimenten $X_i \sim \mathrm{Bin}(1, p)$ verwendet werden, um den Zufallsgenerator zu testen.

Verwenden Sie das Gesetz der großen Zahl dafür, um n in Abhängigkeit von p so zu bestimmen, dass der relative Fehler $\left|\frac{p - \frac{1}{n}\sum_{i=1}^{n} X_i}{p}\right|$ nur mit Wahrscheinlichkeit $\leq \varepsilon$ größer gleich θ ist.

(b) Schreiben Sie nun ein Programm in der Programmiersprache Ihrer Wahl, welches als Parameter p, die Schranke für den relativen Fehler θ und die Schranke für den Wahrscheinlichkeit ε übergeben bekommt, hieraus die Anzahl n der benötigten Bernoulli-Experimente berechnet, dann n Bernoulli-Experimente mit Erfolgswahrscheinlichkeit p durchführt, und schließlich sowohl die Approximation für p als auch den relativen Fehler ausgibt.

Um wieviel weicht die relative Erfolgshäufigkeit von p ab für $p = 0.1, \theta = 0.01, \varepsilon = 0.01$? Was sagt das über die Güte des verwendeten Zufallsgenerators aus?

Lösungsvorschlag

(a)

$$\Pr\left[\frac{\left|p - \frac{1}{n}\sum_{i=1}^{n} X_i\right|}{p} \geq \theta\right] = \Pr\left[\left|p - \frac{1}{n}\sum_{i=1}^{n} X_i\right| \geq \theta \cdot p\right] \leq \frac{p \cdot (1-p)}{n \cdot \theta^2 \cdot p^2} \leq \frac{1-p}{p} \cdot \frac{1}{n \cdot \theta^2} \leq \varepsilon.$$

Somit folgt:

$$n \geq \frac{1-p}{p} \cdot \frac{1}{\theta^2 \cdot \varepsilon}.$$

Speziell für $p = 0.1$, $\theta = 0.01$ und $\varepsilon = 0.01$ folgt also

$$n \geq \frac{0.9}{0.1} \cdot 10^6 = 9 \cdot 10^6.$$

(b)

```
#include <iostream>
#include <cstdlib>
#include <cmath>

using namespace std;

int
main( int argc, char* argv[] ) {
  if( argc < 3 )
     return -1;

  double p = atof( argv[1] );
  const double theta = atof( argv[2] );
  const double eps   = atof( argv[3] );

  const double P = ceil(RAND_MAX * p);

  p = P / RAND_MAX;

  const unsigned long N = (unsigned int) ceil( (1-p)/( p * theta * theta * eps));



  cout << " N = " << N << " P = " << P <<  " p = " << ((float)P) / RAND_MAX << endl;

  unsigned long S = 0;
  srand( time(NULL) );
  for( unsigned long i = 0; i < N; ++i ) {
    if( rand() < P ) ++S;
  }

  double pApprox = S / (double) N;

  cout << " approx. p = " << pApprox << " relative error " << fabs( p - pApprox) / p  << endl;
  return 0;
}
```