Stochastik

Serie 5

1. Da es sich hierbei um eine hypergeometrische Verteilung handelt, benutze ich die Formel:

$$p_k = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \quad N = 10, M = 5, n = 4$$

a) k = 1:

$$p_1 = \frac{\binom{5}{1}\binom{5}{3}}{\binom{10}{4}} = \frac{5}{21} \approx 23,8\%$$

Die Wahrscheinlichkeit, daß genau eine Frau in der Gruppe ist, beträgt somit 23,8%.

b) k = 2:

$$p_2 = \frac{\binom{5}{2}\binom{5}{2}}{\binom{10}{4}} = \frac{10}{21} \approx 47,62\%$$

Die Wahrscheinlichkeit, daß genau zwei Frauen in der Gruppe sind, beträgt somit 47,62%.

2. Poissonverteilung: $p_k = \frac{\lambda^k}{k!} \cdot e^{-\lambda}$
Gegeben ist $\lambda = 2.5$.

$$\begin{aligned}
p_0 &= 2.5^0 \cdot e^{-2.5} \approx 8.2\% \\
p_1 &= \frac{2.5^1}{1!} \cdot e^{-2.5} \approx 20.5\% \\
p_2 &= \frac{2.5^2}{2!} \cdot e^{-2.5} \approx 25.7\% \\
p_3 &= \frac{2.5^3}{3!} \cdot e^{-2.5} \approx 21.4\% \\
p_4 &= \frac{2.5^4}{4!} \cdot e^{-2.5} \approx 13.4\% \\
p_5 &= \frac{2.5^5}{5!} \cdot e^{-2.5} \approx 6.7\%
\end{aligned}$$

a) $P(\{\text{kein Kunde}\}) = p_0 \approx 8.2\%$

b) $P(\{\text{genau ein Kunde}\}) = p_1 \approx 20.5\%$

c) $P(\{\text{mehr als drei Kunden}\}) = \overline{p_0 + p_1 + p_2 + p_3} \approx 24.2\%$

d) $P(\{\text{weniger als sechs Kunden}\}) = p_0 + \ldots + p_5 = \approx 95.9\%$

e) $P(\{\text{kein Kunde}\}) = p_0 \approx 8.2\%$

3. a) Die Wertetabelle mit Randwerten:

$Y \backslash X$	-1	0	1	Y
-1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$
X	$\frac{5}{12}$	$\frac{7}{24}$	$\frac{7}{24}$	

b) Ich berechne die Erwartungswerte, Varianz und den Korrelationskoeffizienten.

$$
\begin{aligned}
E(X) &= \sum_k (x_k p_k) = -1 \cdot \frac{5}{12} + 1 \cdot \frac{7}{24} \\
&= -\frac{1}{8} \\
V(X) &= (-1 + \frac{1}{8})^2 \cdot \frac{5}{12} + (\frac{1}{8})^2 \cdot \frac{7}{24} + (1 + \frac{1}{8})^2 \cdot \frac{7}{24} \\
&\approx 0.693 \\
\sigma(X) &= \sqrt{V(X)} = 0.8325 \\
E(Y) &= \sum_j (y_j p_j) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 0 \\
V(Y) &= 1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = 1 \\
\sigma(Y) &= 1 \\
C(X,Y) &= E((X - E(X)) \cdot (Y - E(Y))) = \sum_{\omega \in \Omega} (X_\omega - E(X)) \cdot (Y_\omega - E(Y)) \cdot P_\omega \\
&= -\frac{7}{8} \cdot (-1) \cdot \frac{1}{6} + \frac{1}{8} \cdot (-1) \cdot \frac{1}{6} + \frac{9}{8} \cdot (-1) \cdot \frac{1}{6} + -\frac{7}{8} \cdot \frac{1}{4} + \frac{1}{8} \cdot \frac{1}{8} + \frac{9}{8} \cdot \frac{1}{8} \\
&= -\frac{1}{8} \\
\rho(X,Y) &= \frac{C(X,Y)}{\rho(X) \cdot \rho(Y)} \\
&= \frac{-\frac{1}{8}}{0.8325 \cdot 1} \\
&\approx -0.150
\end{aligned}
$$

4. a) Hier die Tabelle mit den Randwerten:

$Y\backslash X$	1	2	3	4	5	6	Y
1	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{2}{36}$	0	0	0	0	$\frac{3}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{3}{36}$	0	0	0	$\frac{5}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{4}{36}$	0	0	$\frac{7}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{5}{36}$	0	$\frac{9}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$	$\frac{11}{36}$
X	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	

b) Ich berechne die Erwartungswerte und die Kovarianz:

$$\begin{aligned}
E(X) &= \sum_k (x_k p_k) = \frac{126}{36} \approx 3.5 \\
E(Y) &= \sum_j (y_j p_j) = \frac{161}{36} \approx 4.472 \\
C(X,Y) &= E((X-E(X))\cdot(Y-E(Y))) = \sum_{\omega\in\Omega}(X_\omega - E(X))\cdot(Y_\omega - E(Y))P(\omega) \\
&\approx 1.513^1
\end{aligned}$$

5. Für die Verteilung der Zufallsgrösse $Z = X + Y$ ergibt sich folgendes: $P(Z = z) = P(X = x) \cdot P(Y = y)$, wobei $z = y + z$ gelten muss. Daraus ergibt sich folgender Term:

$$P(Z=z) = \sum_{k=0}^{z} (P(X=k) \cdot P(Y=z-k))$$

Wenn man die Formeln für die Poisson Verteilung einsetzt erhält man folgendes:

$$\begin{aligned}
P(Z=z) &= \sum_{k=0}^{z} \left(\frac{\lambda^k \cdot e^{-\lambda}}{k!} \cdot \frac{\mu^{z-k}\cdot e^{-\mu}}{(z-k)!} \right) \\
&= e^{-(\lambda+\mu)} \cdot \sum_{k=0}^{z} \left(\frac{\lambda^k}{k!} \cdot \frac{\mu^{z-k}}{(z-k)!} \right)
\end{aligned}$$

Die Summe erweitere ich jetzt mit $\frac{z!}{z!}$:

$$\begin{aligned}
P(Z=z) &= e^{-(\lambda+\mu)} \cdot \sum_{k=0}^{z} \left(\frac{\lambda^k}{k!} \cdot \frac{\mu^{z-k}}{(z-k)!} \cdot \frac{z!}{z!} \right) \\
&= \frac{e^{-(\lambda+\mu)}}{z!} \cdot \sum_{k=0}^{z} \left(\frac{\lambda^k \cdot \mu^{z-k} \cdot z!}{k! \cdot (z-k)!} \right)
\end{aligned}$$

[1] für die Berechnung der Kovarianz habe ich ein kleines C-Programm verwendet, welches auf Nachfrage nachgeliefert werden kann.

Die Summe kann man nach dem binomischen Satz dann wie folgt schreiben:

$$P(Z = z) \quad = \quad \frac{e^{-(\lambda+\mu)}}{z!} \cdot (\lambda + \mu)^z$$

Dies ist eine Poisson Verteilung mit dem Faktor $(\lambda + \mu)$. Demnach ist die aus der Summe von X und Y zusammengesetzte Zufallsgrösse wiederum Poissonverteilt.

6. a) Hier unser Pseudocode:

- Seien p_j die Wahrscheinlichkeiten
- $r = random([0, 1))$
- Solange $j < r$
  - $j = j + p_i$
  - $i = i + 1$
- i enthält das Ereignis

b) Der zugehörige Quellcode:

```
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <math.h>

#define BINN 10

/**********************************************************
 * Liefert eine Zufallszahl zwischen [0,1)
 */
double myrandom() {
  return rand()/(double)RAND_MAX;
}

/**********************************************************
 * Berechnet die Fakultaet
 */
double fak(int f) {
  double ret = 1.0;

  while (f > 0)
    ret *= (double) f--;

  return ret;
}

/**********************************************************
```

```
 * Simuliert das Ereignis
 */
int simulate(double* p) {
  int i = 0;
  double j = p[0];
  double r = myrandom();

  // addiere die Wahrscheinlichkeiten
  while(j < r)
    j += p[++i];

  return i;
}


/***********************************************************
 * Berechnet die Binominalverteilung
 */
double binomial(int k) {
  const int n = BINN ;
  const double p = 0.2;
  double ret = fak(n)/(fak(k) * fak(n-k));

  ret *= pow(p,(double)k);
  ret *= pow(1.0 - p,n-k);

  return ret;
}

/***********************************************************
 * main
 */
int main() {
  double p[BINN];
  int i;
  int sim_erg[BINN];

  // Berechne alle genauen Werte der binomial Verteilung
  srand(time(0));
  for(i = 0; i < BINN; i++) {
    p[i] = binomial(i);
    sim_erg[i] = 0;
  }

  // simuliere das Ereignis
  for(i = 0; i < 1000; i++) {
    sim_erg[simulate(p)]++;
  }
```

```
  printf("Genau | Sim | Diff\n");
  for(i = 0; i < BINN; i++)
    printf("          %f & %f & %f \\\\\n",
      p[i],
      sim_erg[i]/1000.0,
      p[i] - (sim_erg[i]/1000.0));

  return EXIT_SUCCESS;
}
```

. . . und die Ergebnisse:

Wahrscheinlichkeit	Simulation	Differenz
0.107374	0.099000	0.008374
0.268435	0.255000	0.013435
0.301990	0.317000	-0.015010
0.201327	0.196000	0.005327
0.088080	0.097000	-0.008920
0.026424	0.029000	-0.002576
0.005505	0.007000	-0.001495
0.000786	0.000000	0.000786
0.000074	0.000000	0.000074
0.000004	0.000000	0.000004

Da nur tausend Versuche durchgeführt wurden, liegt die maximale Genauigkeit der simulierten Werte bei der dritten Nachkommastelle. Dies erklärt die letzten drei Simulationswerte. Bei den hohen Wahrscheinlichkeiten ist die Simulation mit einem Fehler ¡ 5% hinreichend genau.