Musterlösung für die Klausur zur Vorlesung

Stochastik I

Wintersemester 2011/2012

1. Aufgabe (6 Punkte)
Im Wohnort von Herrn Müller herrscht an 25% aller Tage Sonnenschein, an 50% aller Tage ist der Himmel bewölkt und an 25% aller Tage regnet es unaufhörlich. Jeden Morgen schaut Herr Müller, bevor er das Haus verlässt, nach dem Wetter. Wenn es regent, nimmt er seinen Regenschirm mit Wahrscheinlichkeit 0.9 mit (er ist offensichtlich vergesslich), bei bewölktem Himmel mit Wahrscheinlichkeit 0.5 (er ist unentschlossen) und bei Sonnenschein mit Wahrscheinlichkeit 0.2 (er ist Pessimist).

a) Wie groß ist die Wahrscheinlichkeit, dass Herr Müller den Schirm zuhause lässt?

b) Wenn Herr Müller morgens mit dem Schirm das Haus verlässt, wie groß ist dann die Wahrscheinlichkeit, dass die Sonne scheint?

Lösung:
a)(3 Punkte) Mit dem Satz von der totalen Wahrscheinlichkeit gilt

$$\begin{aligned}
\mathbb{P}(\text{kein Schirm}) &= \mathbb{P}(\text{kein Schirm}|\text{Sonne}) \cdot \mathbb{P}(\text{Sonne}) \\
&\quad +\mathbb{P}(\text{kein Schirm}|\text{bewölkt}) \cdot \mathbb{P}(\text{bewölkt}) \\
&\quad +\mathbb{P}(\text{kein Schirm}|\text{Regen}) \cdot \mathbb{P}(\text{Regen}) \\
&= 0.8 \cdot 0.25 + 0.5 \cdot 0.5 + 0.1 \cdot 0.25 \\
&= 0.475 \quad \left(= \frac{19}{40}\right)
\end{aligned}$$

b)(3 Punkte) Mit dem Satz von Bayes gilt

$$\begin{aligned}
\mathbb{P}(\text{Sonne}|\text{Schirm}) &= \frac{\mathbb{P}(\text{Schirm}|\text{Sonne}) \cdot \mathbb{P}(\text{Sonne})}{\mathbb{P}(\text{Schirm})} \\
&= \frac{\mathbb{P}(\text{Schirm}|\text{Sonne}) \cdot \mathbb{P}(\text{Sonne})}{1 - \mathbb{P}(\text{kein Schirm})} \\
&= \frac{0.2 \cdot 0.25}{1 - 0.475} \\
&= 0.0952 \quad \left(= \frac{2}{21}\right)
\end{aligned}$$

2. Aufgabe (10 Punkte)
Auf $\Omega = \{1, 2, 3, 4\}$ mit der Gleichverteilung betrachte man die Zufallsvariablen

$$X(\omega) = \begin{cases} 1 & \text{falls } \omega \text{ gerade} \\ 0 & \text{sonst} \end{cases} \qquad \text{und} \qquad Y(\omega) = \begin{cases} 1 & \text{falls } \omega \in \{2\} \\ 0 & \text{sonst.} \end{cases}$$

a) Prüfen Sie, ob X und Y unabhängig sind.

b) Berechnen Sie $\mathbb{E}(X + Y)$ und $\mathrm{Var}(X + Y)$.

c) Sei $\mathcal{E} = \mathcal{P}(\Omega)$ die Potenzmenge von Ω. Finden Sie drei verschiedene Ereignisse $A, B, C \in \mathcal{E}$, für die gilt: A und B sind unabhängig, B und C sind unabhängig, aber A und C sind nicht unabhängig. D.h., die Eigenschaft der Unabhängigkeit ist nicht transitiv.

Lösung:
a)(2 Punkte) Damit X und Y unabhängig sind, müsste $\mathbb{P}(X = x \wedge Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y)$ für alle möglichen Werte $x, y \in \{0, 1\}$ gelten. Es ist aber beispielsweise

$$\mathbb{P}(X = 1 \wedge Y = 1) = \mathbb{P}(\{2\}) = \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{4} = \mathbb{P}(\{2, 4\}) \cdot \mathbb{P}(\{2\}) = \mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1),$$

d.h., X und Y sind nicht unabhängig.
b)(2+3=5 Punkte) Es ist $\mathbb{E}(X) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 0 = \frac{1}{2}$ und $\mathbb{E}(Y) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot 0 = \frac{1}{4}$, sowie

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$$

Da X und Y nicht unabhängig sind, gilt im Allgemeinen <u>nicht</u> $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. Stattdessen rechenen wir direkt mit der Definition der Varianz

$$\begin{aligned} \mathrm{Var}(X+Y) &= \sum_{\omega} (X(\omega) + Y(\omega) - \mathbb{E}(X+Y))^2 \cdot \mathbb{P}(\omega) \\ &= \left(0 - \frac{3}{4}\right)^2 \cdot \frac{1}{4} + \left(2 - \frac{3}{4}\right)^2 \cdot \frac{1}{4} + \left(0 - \frac{3}{4}\right)^2 \cdot \frac{1}{4} + \left(1 - \frac{3}{4}\right)^2 \cdot \frac{1}{4} \\ &= \frac{11}{16} = 0.6875 \end{aligned}$$

oder mit der Formel $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$:
Es ist $\mathrm{Var}(X) = \frac{1}{2}\left(1 - \frac{1}{2}\right)^2 + \frac{1}{2}\left(0 - \frac{1}{2}\right)^2 = \frac{1}{4}$ und $\mathrm{Var}(Y) = \frac{1}{4}\left(1 - \frac{1}{4}\right)^2 + \frac{3}{4}\left(0 - \frac{1}{4}\right)^2 = \frac{3}{16}$, sowie $\mathbb{E}(XY) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot 0 = \frac{1}{4}$ und $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{4} - \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, also

$$\mathrm{Var}(X+Y) = \frac{1}{4} + \frac{3}{16} + 2 \cdot \frac{1}{8} = \frac{11}{16}.$$

c)(3 Punkte) Wähle zum Beispiel $A = \{1,2\}$, $B = \{2,4\}$ und $C = \{3,4\}$. Dann gilt

$$\mathbb{P}(A \cap B) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

und

$$\mathbb{P}(B \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(B) \cdot \mathbb{P}(C)$$

aber

$$\mathbb{P}(A \cap C) = 0 \neq \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A) \cdot \mathbb{P}(C).$$

3. Aufgabe (6 Punkte)
Die zufällige Wartezeit X bei einer Telefonhotline sei durch die Verteilungsfunktion F charakterisiert mit

$$F(x) = \begin{cases} 0 & : \ x < 0 \\ a - be^{-\lambda x} & : \ x \geq 0 \end{cases}$$

für gewisse Parameter $a, b, \lambda \in \mathbb{R}$, wobei $\lambda > 0$. Weiter sei bekannt, dass $\mathbb{P}(X = 0) = \frac{1}{2}$ (d.h. mit Wahrscheinlichkeit $\frac{1}{2}$ kommt man sofort dran), sowie $\mathbb{P}(X > 1) = \frac{1}{4}$ (d.h. mit Wahrscheinlichkeit $\frac{1}{4}$ muss man länger als eine Minute warten).

a) Bestimmen Sie die Paramter a, b, λ mithilfe der gegebenen Informationen so, dass F tatsächlich eine Verteilungsfunktion ist.

b) Lässt sich die Verteilung auch durch eine Dichtefunktion beschreiben? Begründen Sie Ihre Antwort und geben Sie die Dichte gegebenenfalls an.

<u>Lösung:</u>
a)(4 Punkte) Damit F eine Verteilungsfunktion ist, muss gelten: $\lim_{x\to\infty} F(x) = 1$. Hier ist, wegen $\lim_{x\to\infty} e^{-\lambda x} = 0$,

$$\lim_{x\to\infty} F(x) = a,$$

also folgt $a = 1$.
Damit $\mathbb{P}(X = 0) = \frac{1}{2}$ gilt, muss die Verteilungsfunktion bei 0 um den Wert $\frac{1}{2}$ springen. Wegen $F(x) = 0$ für $x < 0$ folgt $F(0) = \frac{1}{2}$. Hier ist

$$F(0) = a - b = 1 - b,$$

und es folgt $b = \frac{1}{2}$.
Zuletzt folgt wegen $F(1) = \mathbb{P}(X \leq 1) = 1 - \mathbb{P}(X > 1) = \frac{3}{4}$ und

$$F(1) = a - b \cdot e^{-\lambda} = 1 - \frac{1}{2} e^{-\lambda}$$

dass $1 - \frac{1}{2} e^{-\lambda} = \frac{3}{4}$, also nach Umformen $\lambda = \ln(2)$.
Insgesamt ergibt sich

$$F(x) = 1 - \frac{1}{2} e^{-\ln(2)x}.$$

b)(2 Punkte) Da die Verteilungsfunktion nicht stetig (und damit erst recht nicht differenzierbar) ist, siehe Sprung bei $x = 0$, kann die Verteilung nicht durch eine Dichtefunktion beschrieben werden.

4. Aufgabe (7 Punkte)
Man betrachte das Intervall $[0, 1]$ mit der Gleichverteilung und definiere die Zufallsvariablen $X_n : [0, 1] \to \mathbb{R}$ durch

$$X_n(x) = \begin{cases} n & \text{falls } 0 \leq x < \frac{1}{n} \\ 0 & \text{sonst.} \end{cases}$$

a) Prüfen Sie nach, ob $\mathbb{E}(X_n) \stackrel{n\to\infty}{\longrightarrow} 0$ gilt.

b) Prüfen Sie nach, ob $X_n \stackrel{\text{i.W.}}{\longrightarrow} 0$ gilt. (Verwenden Sie dazu bitte nur die Definition von "Konvergenz in Wahrscheinlichkeit".)

c) Prüfen Sie nach, ob $X_n \stackrel{\text{f.s.}}{\longrightarrow} 0$ gilt. (Verwenden Sie dazu bitte nur die Definition von "Konvergenz fast sicher".)

Lösung:
a)(2 Punkte) Es ist

$$\mathbb{E}(X_n) = \frac{1}{n} \cdot n + \left(1 - \frac{1}{n}\right) \cdot 0 = 1$$

für alle n, d.h. $\mathbb{E}(X_n)$ konvergiert nicht gegen 0.
b)(3 Punkte) Für $\varepsilon > 0$ ist

$$\mathbb{P}(|X_n - 0| > \varepsilon) = \mathbb{P}(|X_n| > \varepsilon) \leq \mathbb{P}(X_n = n) = \frac{1}{n} \stackrel{n\to\infty}{\longrightarrow} 0,$$

also konvergiert X_n in Wahrscheinlichkeit gegen 0. Dabei wurde verwendet, dass X_n nur die Werte 0 und n annehmen kann, und somit $|X_n| > \varepsilon$ nur für $X_n = n$ gelten kann, woraus $\mathbb{P}(|X_n| > \varepsilon) \leq \mathbb{P}(X_n = n)$ folgt.
c)(4 Punkte) Für $x \in (0, 1]$ gilt $X_n(x) = 0$, falls $x \geq \frac{1}{n}$, d.h. falls $n \geq \frac{1}{x}$. D.h., für jedes dieser x gibt es einen Index $n_0 \geq \frac{1}{x}$ so dass $X_n(x) = 0$ für alle $n > n_0$, und somit konvergiert $X_n(x)$ gegen 0. Nur für $x = 0$ gilt $X_n(x) = n > 0$ für alle n, so dass $X_n(x)$ nicht gegen 0 konvergiert. Es ist aber $\mathbb{P}(0) = 0$ bzw. $\mathbb{P}((0, 1]) = 1$, da es sich um eine stetige Verteilung handelt, und somit gilt

$$\mathbb{P}(x \in [0, 1] : X_n(x) \to 0) = \mathbb{P}((0, 1]) = 1,$$

und X_n konvergiert fast sicher gegen 0.

5. Aufgabe (9 Punkte)
Die Zufallsvariable X sei poissonverteilt zum Parameter $\lambda > 0$, d.h. es ist $\mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}$ für $n \in \mathbb{N}_0 = \{0, 1, 2, 3, ...\}$. Dabei ist λ unbekannt und soll durch k unabhängige Messungen $X_1, ..., X_k$ geschätzt werden. Wir wählen den Schätzer

$$T_k = T(X_1, ..., X_k) = \frac{X_1 + ... + X_k}{k}.$$

a) Untersuchen Sie den Schätzer T_k auf Erwartungstreue und Konsistenz.

b) Angenommen der wahre Wert ist $\lambda = 4$. Wie groß ist die Wahrscheinlichkeit, bei 10.000 Messungen diesen Wert bis auf einen Fehler von $\varepsilon = 0.05$ genau zu schätzen? Berechnen Sie diese Wahrscheinlichkeit entweder approximativ mithilfe des zentralen Grenzwertsatzes oder schätzen Sie sie mithilfe der Tschebyschev-Ungleichung ab. (D.h., Sie können sich einen der beiden Lösungswege aussuchen!)

Lösung:
a) (2+3=5 Punkte) Für die Poissonverteilte Zufallsvariable X gilt $\mathbb{E}(X) = \lambda$, d.h. es gilt ebenso $\mathbb{E}(X_i) = \lambda$ für alle unabhängigen Kopien X_i, $i = 1, ..., k$, von X. Damit folgt

$$\mathbb{E}(T_k) = \mathbb{E}\left(\frac{X_1 + ... + X_k}{k}\right) = \frac{1}{k}(\mathbb{E}(X_1) + ... + \mathbb{E}(X_k)) = \mathbb{E}(X) = \lambda,$$

d.h., T_k ist erwartungstreu.
Nach dem schwachen Gesetz der großen Zahl gilt

$$\mathbb{P}(|\bar{X} - \mathbb{E}(X)| > \varepsilon) \to 0 \quad \text{für alle } \varepsilon > 0,$$

d.h. $\bar{X} \xrightarrow{\text{i.W.}} \mathbb{E}(X)$. Wegen $T_k = \bar{X}$ und $\lambda = \mathbb{E}(X)$ folgt sofort $T_k \xrightarrow{\text{i.W.}} \lambda$ und T_k ist konsistent.
b) (4 Punkte) Verwende den zentralen Grenzwertsatz: Danach ist $\bar{X}$ annähernd normalverteilt, d.h. es gilt

$$\mathbb{P}(|\bar{X} - 4| < 0.05) = \mathbb{P}\left(\left|\frac{\bar{X} - 4}{\sqrt{1/10.000 \cdot 4}}\right| < \frac{0.05}{\sqrt{1/10.000 \cdot 4}}\right) \approx \mathbb{P}(|Y| < 2.5),$$

wobei Y standardnormalverteilt ist. Umformen und Ablesen aus der Tabelle zur Standardnormalvertilung ergibt

$$\mathbb{P}(|Y| < 2.5) = \mathbb{P}(Y < 2.5) - \mathbb{P}(Y \leq -2.5) = 2 \cdot \mathbb{P}(Y < 2.5) - 1 = 0.9876.$$

Die gesuchte Wahrscheinlichkeit beträgt also annähernd 99%.
Verwende die Tschebyschev-Ungleichung: Danach gilt

$$\mathbb{P}(|\bar{X} - 4| < 0.05) = 1 - \mathbb{P}(|\bar{X} - 4| \geq 0.05) \geq 1 - \frac{\mathrm{Var}(\bar{X})}{0.05^2} = 1 - \frac{1/10.000 \cdot 4}{0.05^2} = 0.84.$$

Die gesuchte Wahrscheinlichkeit beträgt also mindestens 84%.