Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr
Dr. Werner Meixner

Sommersemester 2010
Lösungsblatt 8
28. Juni 2010

Diskrete Wahrscheinlichkeitstheorie

Hausaufgabe 1 (5 Punkte)

Wahr oder falsch? Beweisen Sie Ihre Antworten.

1. Jede stetige Zufallsvariable X mit $W_X = \mathbb{R}$, für deren Dichtefunktion f_X die Beziehung $f_X(x) = f_X(-x)$ für alle $x \in \mathbb{R}$ gilt und deren Erwartungswert existiert, besitzt den Erwartungswert 0.

2. Es gibt keine stetige Zufallsvariable X mit $W_X = \mathbb{R}$, für deren Verteilungsfunktion F_X die Beziehung $F_X(x) = F_X(-x)$ gilt für alle $x \in \mathbb{R}$.

3. Die gemeinsame Dichtefunktion $f_{X,Y}$ zweier stetiger Zufallsvariablen X und Y sei konstant gleich $\frac{2}{\pi}$ auf der Menge $B = \{(x,y)\,;\, x^2+y^2 \le r^2\}$ und sei 0 außerhalb von B. Dann gilt $r^2 = \frac{1}{4}$.

Lösungsvorschlag

1. Die Aussage ist wahr. Der Beweis erfordert eine sorgfältige Beachtung der Integrationsregeln, insbesondere der entsprechenden Substitutionsregel.

   Für alle $a, b \in \mathbb{R}$ und $x = \phi(t) = -t$ gilt $\phi'(t) = -1$ und $\phi(a) = -a, \phi(b) = -b$. Nach der Substitutionsregel erhalten wir

$$\begin{aligned}
\int_a^b x \cdot f_X(x)\,\mathrm{d}x &= \int_{-a}^{-b} \phi(t) \cdot f_X(\phi(t))\phi'(t)\,\mathrm{d}t \\
&= \int_{-a}^{-b} (-t) \cdot f_X(-t)(-1)\,\mathrm{d}t \\
&= \int_{-a}^{-b} t \cdot f_X(t)\,\mathrm{d}t \\
&= -\int_{-b}^{-a} t \cdot f_X(t)\,\mathrm{d}t\,.
\end{aligned}$$

   $\mathbb{E}[X] = 0$ folgt aus der Gleichung $\mathbb{E}[X] = -\mathbb{E}[X]$, die nun wie folgt gezeigt wird.

$$\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x \\
&= \lim_{a\to-\infty} \lim_{b\to\infty} \left(\int_a^b x \cdot f_X(x)\,\mathrm{d}x \right) \\
&= \lim_{a\to-\infty} \lim_{b\to\infty} \left(-\int_{-b}^{-a} t \cdot f_X(t)\,\mathrm{d}t \right) \\
&= -\int_{-\infty}^{\infty} t \cdot f_X(t)\,\mathrm{d}t \\
&= -\mathbb{E}[X]\,.
\end{aligned}$$

2. Die Aussage ist wahr. Der Beweis benutzt die folgenden Eigenschaften der Verteilungsfunktion F_X, die in der Vorlesung genannt wurden,

$$\lim_{x\to-\infty} F_X(x) = 0 \quad \text{und} \quad \lim_{x\to\infty} F_X(x) = 1\,.$$

Aus der Gleichung $F_X(x) = F_X(-x)$ folgt der Widerspruch

$$\begin{aligned} 0 = \lim_{x\to-\infty} F_X(x) &= \lim_{x\to-\infty} F_X(-x) \\ &= \lim_{x\to\infty} F_X(x) = 1\,. \end{aligned}$$

3. Die Aussage ist falsch. Beweis: Sei F_B die Fläche der Kreisscheibe B. Dann gilt $F_B = \pi r^2$ und wir erhalten den Widerspruch

$$\begin{aligned} 1 = \int_{\mathbb{R}^2} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_B f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_B \frac{2}{\pi}\,\mathrm{d}x\,\mathrm{d}y \\ &= \frac{2}{\pi}\cdot\pi r^2 \;=\; 2\cdot\frac{1}{4} \;=\; \frac{1}{2}\,. \end{aligned}$$

Hausaufgabe 2 (5 Punkte)

Auf einem Blatt Papier sind im Abstand von 4 cm horizontale Linien aufgemalt. Wir werfen eine 2 cm lange Nadel auf dieses Blatt Papier. Dabei treffen wir immer das Papier und werfen nicht daneben, so dass der Mittelpunkt der Nadel gleichverteilt wird auf dem Papier.

Mit welcher Wahrscheinlichkeit berührt die Nadel eine Linie?

Lösungsvorschlag

Die Gleichverteilung von Punkten auf einer Fläche bedeutet, dass gleich große Flächenstücke im Mittel gleich oft von erzeugten Punkten berührt werden, hier z. B. durch den Mittelpunkt der geworfenen Nadel.

Sei der Abstand des Mittelpunkts der Nadel zu seiner nächsten Linie gleichverteilt auf $[0,2]$ (in cm). Sei ferner Θ der kleinste Winkel zwischen der Linie und der Nadel (im Bogenmaß). Dann gehen wir davon aus, dass Θ gleichverteilt ist auf $[0,\frac{\pi}{2}]$. Wenn y die Distanz zwischen dem Mittelpunkt der Nadel und der nächsten Linie ist, dann schneidet oder berührt sie die Linie genau dann, wenn

$$y \le \frac{2\ (\mathrm{cm})}{2}\sin\Theta.$$

Die angenommene Unabhängigkeit von Abstand und Winkel zusammen mit den Gleichverteilungsannahmen bei Abstand und Winkel hat zur Folge, dass im Koordinatensystem von Abstand und Winkel innerhalb der Fläche $\{(d,\alpha)\,;\,0\le d\le 2\wedge 0\le\alpha\le\frac{\pi}{2}\}$ eine Gleichverteilung der erzeugten Punkte vorliegt. Die Wahrscheinlichkeit, dass die Nadel die nächste Linie schneidet oder berührt, ist also

$$\frac{\int_0^{\pi/2} 1\ \mathrm{cm}\cdot\sin\Theta\ \mathrm{d}\Theta}{\int_{0\ \mathrm{cm}}^{2\ \mathrm{cm}}\int_0^{\pi/2} 1\ \mathrm{d}\Theta\ \mathrm{d}y} = \frac{1\ \mathrm{cm}}{\pi\ \mathrm{cm}} = \frac{1}{\pi}.$$

Hausaufgabe 3 (5 Punkte)

Sei X eine beliebige kontinuierliche Zufallsvariable.
Zeigen Sie, dass die Funktion $\psi(a) = \mathbb{E}[-(X-a)^2]$ ihr Maximum bei $a = \mathbb{E}[X]$ hat.

Lösungsvorschlag

$$\begin{aligned}
\psi'(a) &= \frac{\mathrm{d}}{\mathrm{d}a} \int_{-\infty}^{\infty} -(x-a)^2 f_X(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}a} \left[-(x-a)^2 f_X(x)\right] \mathrm{d}x \\
&= \int_{-\infty}^{\infty} 2(x-a) f_X(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} 2x f_X(x)\,\mathrm{d}x - 2a \int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x \\
&= 2\mathbb{E}[X] - 2a\,.
\end{aligned}$$

Offenbar gilt $\psi'(\mathbb{E}[X]) = 0$ und $\psi''(\mathbb{E}[X]) = -2 < 0$. Bekanntlich folgt daraus, dass $\psi(\mathbb{E}[X])$ ein Maximum ist.

Hausaufgabe 4 (5 Punkte)

Die Lebensdauer T eines Monitors habe die folgende (mit $\lambda > 0$ und a parametrisierte) Dichte:

$$f_T(t) = \begin{cases} a\lambda^2 t e^{-\lambda t} & : t \geq 0 \\ 0 & : t < 0 \end{cases}$$

1. Welchen Wert muss a besitzen, so dass für alle $\lambda > 0$ die Funktion f_T tatsächlich eine Dichte ist, d. h., dass $\int_{-\infty}^{\infty} f_T(t)\,dt = 1$ gilt. Beweis!
2. Berechnen Sie $\Pr[T > \frac{2}{\lambda}]$.
3. Berechnen Sie $\mathbb{E}[T]$ in Abhängigkeit von λ.

Lösungsvorschlag

1. Wir wenden partielle Integration an wie folgt.

$$\begin{aligned}
\int_{-\infty}^{\infty} f_T(t)\,\mathrm{d}t &= a\lambda^2 \int_0^{\infty} t e^{-\lambda t}\,\mathrm{d}t \\
&= a\lambda^2 \cdot \left(\left[-\frac{1}{\lambda} t e^{-\lambda t} \right]_0^{\infty} - \int_0^{\infty} -\frac{1}{\lambda} e^{-\lambda t}\,\mathrm{d}t \right) \\
&= a\lambda^2 \cdot \left(0 + \left[-\frac{1}{\lambda^2} e^{-\lambda t} \right]_0^{\infty} \right) \\
&= a\,.
\end{aligned}$$

   Daraus folgt $a = 1$.

2. Natürlich müssen wir nun $a = 1$ setzen. Wir berechnen zunächst $\Pr[T \leq \frac{2}{\lambda}]$ analog wie oben.

$$\begin{aligned}
\Pr\left[T \leq \frac{2}{\lambda}\right] &= \int_0^{2/\lambda} f_T(t)\,\mathrm{d}t \\
&= \lambda^2 \cdot \left(\left[-\frac{1}{\lambda} t e^{-\lambda t}\right]_0^{2/\lambda} + \left[-\frac{1}{\lambda} e^{-\lambda t}\right]_0^{2/\lambda}\right) \\
&= \lambda^2 \cdot \left(-\frac{1}{\lambda} \cdot \frac{2}{\lambda} e^{-2} - \frac{1}{\lambda^2} e^{-2} + \frac{1}{\lambda^2}\right) \\
&\approx 0{,}59399415\,.
\end{aligned}$$

   Wir erhalten $\Pr[T > \frac{2}{\lambda}] \approx 0{,}40600585\,.$

3. Teilweise wie vorausgehend rechnen wir

$$\begin{aligned}
\mathbb{E}[T] &= \int_0^{\infty} t \cdot f_T(t)\,\mathrm{d}t \\
&= \int_0^{\infty} \lambda^2 t^2 e^{-\lambda t}\,\mathrm{d}t \\
&= \left[-\lambda t^2 e^{-\lambda t}\right]_0^{\infty} - \int_0^{\infty} -2\lambda t e^{-\lambda t}\,\mathrm{d}t \\
&= 2\lambda \int_0^{\infty} t e^{-\lambda t}\,\mathrm{d}t \\
&= 2\lambda \cdot \frac{1}{\lambda^2} \quad \text{(siehe oben)} \\
&= \frac{2}{\lambda}\,.
\end{aligned}$$

Tutoraufgabe 1

1. Seien X und Y unabhängige und positivwertige kontinuierliche Zufallsvariablen. Drücken Sie die Dichtefunktion f_Z von $Z = X/Y$ durch die Dichtefunktionen f_X und f_Y von X bzw. Y aus.

2. Die Zufallsvariablen X und Y seien gegeben durch die Koordinaten eines (gleichverteilt) zufällig gewählten Punktes $P \in \{(s,t); s^2 + t^2 \leq 4\}$ der x,y-Ebene.

   Berechnen Sie die gemeinsame Dichte $f_{X,Y}(x,y)$!
   Sind X und Y unabhängig?
   Berechnen Sie die Randdichten f_X und f_Y!

Lösungsvorschlag

1. Wir berechnen die Dichte f_Z als Ableitung der Verteilung F_Z. Da X und Y positivwertig sind, gilt zunächst $f_X(x) = 0$ und $f_Y(y) = 0$ für alle $x \leq 0$ und $y \leq 0$. Daraus folgt ohne Rechnung $f_Z(z) = 0$ für alle $z \leq 0$, mithin $F_Z(z) = 0$ für alle $z \leq 0$.

   Für die Verteilungsfunktion $F_Z(z)$ gilt

$$
\begin{aligned}
F_Z(z) &= \Pr[Z \leq z] \\
&= \Pr[X/Y \leq z] \\
&= \Pr[(X,Y) \in \{(x,y)\,;\, x \leq zy\}] \\
&= \int_{\{(x,y)\,;\, x \leq zy\}} f_{X,Y}(x,y)\ \mathrm{d}x\,\mathrm{d}y \\
&= \int_{\{(x,y)\,;\, 0<x, 0<y,\, x \leq zy\}} f_X(x) \cdot f_Y(y)\ \mathrm{d}x\,\mathrm{d}y \\
&= \int_0^\infty f_Y(y) \cdot \left(\int_0^{z\cdot y} f_X(x)\ \mathrm{d}x \right)\ \mathrm{d}y \\
&= \int_0^\infty f_Y(y) \cdot F_X(zy)\ \mathrm{d}y\,.
\end{aligned}
$$

   Damit errechnet sich die Dichte zu

$$
\begin{aligned}
f_Z(z) &= \frac{\mathrm{d}}{\mathrm{d}z} \int_0^\infty f_Y(y) \cdot F_X(zy)\ \mathrm{d}y \\
&= \int_0^\infty f_Y(y) \cdot f_X(zy) \cdot y\ \mathrm{d}y\,.
\end{aligned}
$$

2. (a) Die gegebene Punktmenge $K = \{(s,t); s^2 + t^2 \leq 4\}$ ist eine Kreisscheibe mit Radius 2 und Mittelpunkt $(0,0)$. Gleichverteilung auf der Kreisscheibe bedeutet für die Dichte die Gleichung

$$
f_{X,Y}(x,y) = c
$$

   für alle $(x,y) \in K$ mit einer konstanten Zahl c. Außerhalb von K ist die Dichte gleich 0. Das Integral über $\mathbb{R}^2$ muss 1 ergeben. Da für die Kreisfläche $F_K = 4\pi$ gilt, folgt

$$
c = \frac{1}{4\pi}\,.
$$

(b) Wären X und Y unabhängig, dann müsste für die Verteilungsfunktionen $F_{X,Y}$, F_X und F_Y für alle $x, y \in \mathbb{R}$ gelten

$$F_{X,Y}(x,y) \;=\; F_X(x) \cdot F_Y(y)\,.$$

Dies führt man zum Widerspruch, in dem man für $(x,y) = (-\sqrt{2}, -\sqrt{2})$ zeigt

$$F_{X,Y}(-\sqrt{2}, -\sqrt{2}) \;\neq\; F_X(-\sqrt{2}) \cdot F_Y(-\sqrt{2})\,.$$

Auf der Menge $B = \{(x,y)\,;\,(x \leq -\sqrt{2}) \wedge (y \leq -\sqrt{2})\}$ hat die gemeinsame Dichte $f_{X,Y}$ den Wert 0. Damit gilt

$$\begin{aligned}
F_{X,Y}(-\sqrt{2}, -\sqrt{2}) &= \Pr[X \leq -\sqrt{2}\ ,\ Y \leq -\sqrt{2}] \\
&= \int_{-\infty}^{-\sqrt{2}} \int_{-\infty}^{-\sqrt{2}} f_{X,Y}(x,y)\ \mathrm{d}x\,\mathrm{d}y \\
&= \int_B f_{X,Y}(x,y)\ \mathrm{d}x\,\mathrm{d}y \\
&= 0
\end{aligned}$$

Die Randverteilungen $F_X(-\sqrt{2})$ und $F_Y(-\sqrt{2})$ dagegen sind beide nicht gleich Null.

$$\begin{aligned}
F_X(-\sqrt{2}) &= \Pr[X \leq -\sqrt{2}] \\
&= \int_{-\infty}^{-\sqrt{2}} \left[\int_{-\infty}^{\infty} f_{X,Y}(u,v)\ \mathrm{d}v\right]\ \mathrm{d}u \\
&= \int_{-1}^{-\sqrt{2}} \frac{\sqrt{4-u^2}}{2\pi}\ \mathrm{d}u \\
&\neq 0
\end{aligned}$$

Analog zeigt man $F_Y(-\sqrt{2}) \neq 0$.

(c) Die Randdichten berechnen sich wie folgt.

$$\begin{aligned}
f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\ \mathrm{d}y \\
&= \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \frac{1}{4\pi}\ \mathrm{d}y \\
&= \frac{\sqrt{4-x^2}}{2\pi}\,,
\end{aligned}$$

und analog folgt

$$f_Y(y) \;=\; \frac{\sqrt{4-y^2}}{2\pi}\,.$$

Tutoraufgabe 2

Seien $X_1, \ldots, X_n$ unabhängige, mit Parameter $\lambda = 1$ exponentialverteilte Zufallsvariablen. Zeigen Sie, dass für alle reellen x gilt

$$\Pr\left[\max_{1 \le i \le n} X_i \le \log n + x\right] = \left(1 - \frac{e^{-x}}{n}\right)^n .$$

Geben Sie den Grenzwert an für $n \to \infty$!

Lösungsvorschlag

Damit das Maximum die geforderte Bedingung erfüllt, muss jedes X_i diese Bedingung erfüllen. Wir haben also

$$\Pr\left[\max_{1 \le i \le n} X_i \le \log n + x\right] = \Pr\left[\bigcap_i (X_i \le \log n + x)\right].$$

Wegen der Unabhängigkeit der Zufallsvariablen folgt

$$\begin{aligned}
\Pr\left[\bigcap_i (X_i \le \log n + x)\right] &= \prod_{1 \le i \le n} \Pr[X_i \le \log n + x] \\
&= (1 - e^{-\log n - x})^n = \left(1 - \frac{e^{-x}}{n}\right)^n .
\end{aligned}$$

Für $n \to \infty$ gilt

$$\lim_{n \to \infty} \Pr\left[\max_{1 \le i \le n} X_i \le \log n + x\right] = \lim_{n \to \infty} \left(1 - \frac{e^{-x}}{n}\right)^n = e^{-e^{-x}} .$$

Tutoraufgabe 3

Wir modellieren die tägliche Preisänderung am Aktienmarkt durch eine Zufallsvariable mit Erwartungswert 0 und Varianz σ^2. Ausgehend vom „heutigen“ Preis Y_0 ergibt sich dann der Preis Y_n einer Aktie am folgenden n-ten Tag durch

$$Y_n = Y_{n-1} + X_n \quad (n \ge 1) ,$$

wobei die Zufallsvariablen $X_1, X_2, \ldots$ unabhängig und identisch verteilt sind mit Erwartungswert 0 und Varianz σ^2. Wir nehmen an, dass die Verteilung der Y_n durch den zentralen Grenzwertsatz ausreichend genau beschrieben werden kann.

Wie groß ist die Wahrscheinlichkeit, dass eine Aktie in 30 Tagen mehr als 112€ wert ist, wenn heute die Aktie 100€ kostet und $\sigma^2 = 1$ gilt?

Lösungsvorschlag

Nach dem zentralen Grenzwertsatz ist

$$\sqrt{30}\left(\frac{\sum_{i=1}^{30} X_i}{30}\right) = \frac{\sum_{i=1}^{30} X_i}{\sqrt{30}}$$

annähernd standardnormalverteilt. Damit ist

$$\Pr[Y_{30} > 112 \mid Y_0 = 100] = \Pr\left[\sum_{i=1}^{30} X_i > 12\right] =$$

$$= \Pr\left[\frac{\sum_{i=1}^{30} X_i}{\sqrt{30}} > \frac{12}{\sqrt{30}}\right] \approx 1 - \Phi(2.19) =\approx 1.4\% \,.$$