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

Sommersemester 2009
Übungsblatt 8

LÖSUNG

Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt ist am 6. Juli bis 10:15 abzugeben und wird am 6./7. Juli besprochen.

Beachten Sie: Alle Ergebnisse sind zu begründen!

Aufgabe 8.1 — je 1P

Seien $X_1, X_2, \dots$ unabhängig und identisch verteilt mit $\mathbb{E}X_i = 0.8$ und Standardabweichung $\sigma = 0.4$.
Seien $Y_n := X_1 + \cdots + X_n$ und $A_n := \frac{Y_n}{n}$.
Berechnen Sie folgende Grenzwerte.

(a) $\lim_{n\to\infty} \Pr[A_n = 0.8]$

(b) $\lim_{n\to\infty} \Pr[0.7n < Y_n < 0.9n]$

(c) $\lim_{n\to\infty} \Pr[Y_n < 0.8n + 0.8\sqrt{n}]$

(d) $\lim_{n\to\infty} \Pr[0.79 < A_n < 0.81]$

Lösungsvorschlag: Der ZGWS besagt, dass die Verteilung von $\frac{Y_n - 0.8n}{0.4\sqrt{n}}$ gegen die Standardnormalverteilung konvergiert. Das wird jetzt immer wieder benutzt.

(a)
$$\Pr[A_n = 0.8] = \Pr[Y_n = 0.8n] = \Pr\left[\frac{Y_n - 0.8n}{0.4\sqrt{n}} = 0\right] = \Pr\left[\frac{Y_n - 0.8n}{0.4\sqrt{n}} \le 0\right] - \Pr\left[\frac{Y_n - 0.8n}{0.4\sqrt{n}} < 0\right]$$

Nach dem ZGWS konvergieren sowohl $\Pr\left[\frac{Y_n-0.8n}{0.4\sqrt{n}} \le 0\right]$ als auch $\Pr\left[\frac{Y_n-0.8n}{0.4\sqrt{n}} < 0\right]$ gegen $\Phi(0)$. Es folgt

$$\lim_{n\to\infty} \Pr[A_n = 0.8] = \Phi(0) - \Phi(0) = 0$$

(b)
$$\Pr[0.7n < Y_n < 0.9n] = \Pr\left[-\frac{0.1}{0.4}\sqrt{n} < \frac{Y_n - 0.8n}{0.4\sqrt{n}} < \frac{0.1}{0.4}\sqrt{n}\right]$$

Für $n \to \infty$ konvergiert diese W'keit gegen $\Phi(\infty) - \Phi(-\infty) = 1 - 0 = 1$.

(c)
$$\Pr\left[Y_n < 0.8n + 0.8\sqrt{n}\right] = \Pr\left[\frac{Y_n - 0.8n}{0.4\sqrt{n}} < 2\right]$$

Für $n \to \infty$ konvergiert diese W'keit gegen $\Phi(2) \approx 0.977 < 1$.

(d) Der Grenzwert ist 1, denn $\Pr[0.79 < A_n < 0.81] = \Pr[0.79n < Y_n < 0.81n]$ und dann weiter wie in (b).

Aufgabe 8.2 — 1P+2P+2P

Ein Restaurant hat 400 Gäste am Tag. Im Durchschnitt bestellen 20% einen Nachtisch. Sei X die Zahl der Gäste, die Nachtisch bestellen. Wir interessieren uns für ein Intervall, das X mit W'keit 0.95 enthält.
Hinweis: Benutzen Sie in dieser Aufgabe den Zentralen Grenzwertsatz und rechnen Sie mit $\Phi(2) = 0.975$.

(a) Das gesuchte Intervall soll möglichst klein sein. Begründen Sie (informell), warum das Intervall symmetrisch um 80, also von der Form $[80 - y, 80 + y]$ sein sollte.

(b) Berechnen Sie y.

(c) Wie viele Gäste müsste das Restaurant haben, damit mit 95-prozentiger W'keit zwischen 19 und 21 Prozent der Gäste Nachtisch bestellen?

Lösungsvorschlag: X kann als normalverteilt angenommen werden mit Erwartungswert $\mathbb{E}X = 80$ und Standardabweichung $\sigma = \sqrt{\mathrm{Var}X} = \sqrt{400 \cdot 0.8 \cdot 0.2} = 8$.

(a) Das wird von der Form der Glockenkurve nahegelegt: Gesucht ist ein möglichst kleines Intervall mit gegebener Fläche (0.95) unter der Glockenkurve. Da die Glockenkurve bei 80 das Maximum hat und nach beiden Seiten symmetrisch abfällt, sollte das Intervall symmetrisch um 80 herum liegen.

(b)

$$\begin{aligned} 0.95 &\overset{!}{=} \Pr[80 - y \le X \le 80 + y] \\ &= \Pr\left[-\frac{y}{8} \le \frac{X - \mathbb{E}X}{\sigma} \le \frac{y}{8}\right] \\ &= \Phi(\frac{y}{8}) - \Phi(-\frac{y}{8}) \\ &= 2\Phi(\frac{y}{8}) - 1 \end{aligned}$$

Auflösen nach $\Phi(\frac{y}{8})$ ergibt $\Phi(\frac{y}{8}) = 0.975$. Mit dem Hinweis folgt $\frac{y}{8} = 2$, also $y = 16$.

(c) Sei n die Zahl der Gäste.

$$\begin{aligned} 0.95 &\overset{!}{=} \Pr[0.19n \le X \le 0.21n] \\ &= \Pr\left[\frac{-0.01n}{\sqrt{0.2 \cdot 0.8 \cdot n}} \le \frac{X - 0.2n}{\sqrt{0.2 \cdot 0.8 \cdot n}} \le \frac{0.01n}{\sqrt{0.2 \cdot 0.8 \cdot n}}\right] \\ &= 2\Phi\left(\frac{1}{40}\sqrt{n}\right) - 1 \end{aligned}$$

Ähnlich wie in (b) folgt $\frac{1}{40}\sqrt{n} = 2$, also $n = 6400$.

Aufgabe 8.3 **1P+1P+2P+1P+2P**

90% der Computerchips eines bestimmten Herstellers haben eine erwartete Lebensdauer von 150 Monaten, mit einer Standardabweichung von 50 Monaten. 10% der Chips weisen aber einen Materialfehler auf. Die fehlerhaften Chips halten im Mittel nur 20 Monate, mit einer Standardabweichung von 10 Monaten. Sie haben einen Chip von diesem Hersteller gekauft. Sei die Zufallsvariable X seine Lebensdauer in Monaten und seien „schlecht" bzw. „gut" die Ereignisse, dass der Chip den Materialfehler aufweist bzw. nicht aufweist.

(a) Zeigen Sie: $\Pr[X \ge 50] \ge 0.9 \cdot \Pr[X \ge 50 \mid \text{„gut"}]$.

(b) Benutzen Sie (a) und die Chebyshev-Ungleichung, um eine untere Schranke für $\Pr[X \ge 50]$ zu berechnen.

Setzen Sie im folgenden sowohl für schlechte als auch für gute Chips eine normalverteilte Lebensdauer voraus.

(c) Berechnen Sie $\Pr[X \ge 50]$.

(d) Nach genau 50 Monaten geht Ihr Chip kaputt. Sie interessieren sich für die W'keit, dass Sie einen fehlerhaften Chip erwischt haben. Ihr Übungspartner schlägt vor, dafür die bedingte W'keit $\Pr[\text{„schlecht"} \mid X = 50]$ zu berechnen. Wo liegt das Problem?

(e) Berechnen Sie $\Pr[\text{„schlecht"} \mid 50 - \varepsilon \le X \le 50 + \varepsilon]$ für kleine ε.

Hinweis: Benutzen Sie $\int_{50-\varepsilon}^{50+\varepsilon} f(x)\, dx \approx 2\varepsilon \cdot f(50)$ für kleine ε.

Lösungsvorschlag:

(a) Mit dem Satz von der totalen W'keit gilt:

$$\Pr[X \ge 50] = \Pr[X \ge 50 \mid \text{„gut"}] \cdot \underbrace{\Pr[\text{„gut"}]}_{0.9} + \Pr[X \ge 50 \mid \text{„schlecht"}] \cdot \Pr[\text{„schlecht"}] \ge 0.9 \cdot \Pr[X \ge 50 \mid \text{„gut"}]$$

(b) Sei Y die Lebensdauer eines guten Chips. Es gilt:

$$\Pr[Y \le 50] = \Pr[Y - \mathbb{E}Y \le 50 - 150] \le \Pr[|Y - \mathbb{E}Y| \ge 100] \overset{\text{Chebyshev}}{\le} \frac{\mathrm{Var}[Y]}{100^2} = \frac{50^2}{100^2} = \frac{1}{4}$$

Damit und mit (a) gilt:

$$\Pr[X \ge 50] \ge 0.9 \cdot \Pr[X \ge 50 \mid \text{„gut"}] = 0.9 \cdot \Pr[Y \ge 50] = 0.9 \cdot (1 - \Pr[Y \le 50]) \ge 0.9 \cdot (1 - \frac{1}{4}) = 0.675$$

(c) Sei Y die Lebensdauer eines guten Chips. Es gilt

$$\Pr[Y \le 50] = \Pr\left[\frac{Y - 150}{50} \le \frac{50 - 150}{50}\right] = \Phi(-2) = 1 - \Phi(2)$$

und folglich

$$\Pr[Y \ge 50] = 1 - \Pr[Y \le 50] = \Phi(2) \approx 0.9772\,.$$

Sei Z die Lebensdauer eines schlechten Chips. Es gilt

$$\Pr[Z \le 50] = \Pr\left[\frac{Z - 20}{10} \le \frac{50 - 20}{10}\right] = \Phi(3)$$

und folglich

$$\Pr[Z \ge 50] = 1 - \Pr[Z \le 50] = 1 - \Phi(3) \approx 0.0013\,.$$

Mit dem Satz von der totalen W'keit folgt ähnlich wie in (a)

$$\Pr[X \ge 50] \approx 0.9 \cdot 0.9772 + 0.1 \cdot 0.0013 \approx 0.880\,.$$

(d) Da X eine stetige Zufallsvariable ist, gilt $\Pr[X = 50] = 0$. Deswegen ist $\Pr[\text{„schlecht"} \mid X = 50]$ nicht definiert. (Man würde 0 durch 0 teilen.)

(e) Sei $A_\varepsilon :=$ „$50 - \varepsilon \le X \le 50 + \varepsilon$". Dann gilt

$$\begin{aligned}\Pr[A_\varepsilon \mid \text{„schlecht"}] &= \frac{1}{\sqrt{2\pi} \cdot 10} \int_{50-\varepsilon}^{50+\varepsilon} \exp\left(-\frac{(x-20)^2}{2 \cdot 10^2}\right)\, dx \\ &\overset{\varepsilon \text{ klein}}{=} \frac{1}{\sqrt{2\pi} \cdot 10} \cdot 2\varepsilon \cdot \exp\left(-\frac{(30)^2}{2 \cdot 10^2}\right)\end{aligned}$$

und ähnlich

$$\begin{aligned}\Pr[A_\varepsilon \mid \text{„gut"}] &= \frac{1}{\sqrt{2\pi} \cdot 50} \int_{50-\varepsilon}^{50+\varepsilon} \exp\left(-\frac{(x-150)^2}{2 \cdot 50^2}\right)\, dx \\ &\overset{\varepsilon \text{ klein}}{=} \frac{1}{\sqrt{2\pi} \cdot 50} \cdot 2\varepsilon \cdot \exp\left(-\frac{(-100)^2}{2 \cdot 50^2}\right)\end{aligned}$$

Mit Bayes

$$\Pr[\text{„schlecht"} \mid A_\varepsilon] = \frac{\Pr[A_\varepsilon \mid \text{„schlecht"}] \cdot \Pr[\text{„schlecht"}]}{\Pr[A_\varepsilon \mid \text{„schlecht"}] \cdot \Pr[\text{„schlecht"}] + \Pr[A_\varepsilon \mid \text{„gut"}] \cdot \Pr[\text{„gut"}]}$$

und einigem Kürzen (vor allem ε kürzt sich raus) bekommen wir

$$\begin{aligned} &= \frac{0.1 \cdot \frac{1}{10} e^{-9/2}}{0.1 \cdot \frac{1}{10} e^{-9/2} + 0.9 \cdot \frac{1}{50} e^{-2}} \\ &= \frac{e^{-9/2}}{e^{-9/2} + \frac{9}{5} e^{-2}} \\ &\approx 0.044 \end{aligned}$$

Aufgabe 8.4 **2P+2P**

Es seien $T_1, T_2, T_3, \ldots$ unabhängige und identisch verteilte ZV mit $T_i \sim \exp(\lambda)$ und $\lambda > 0$.
Für $n \geq 1$ bezeichne S_n die Summe $\sum_{i=1}^n T_i$.

(a) Zeigen Sie mittels Induktion, dass S_n die folgende Dichte besitzt:

$$f_{S_n}(z) = \begin{cases} \frac{\lambda^n z^{n-1}}{(n-1)!} e^{-\lambda z} & \text{für } z \geq 0 \\ 0 & \text{sonst.} \end{cases}$$

(b) Zeigen Sie nun, dass für die Verteilungsfunktion von S_n für $z \geq 0$ und $n \geq 2$ gilt:

$$F_{S_n}(z) = -\frac{\lambda^{n-1} z^{n-1}}{(n-1)!} e^{-\lambda z} + F_{S_{n-1}}(z).$$

Geben Sie dann $F_{S_n}(z)$ in geschlossener Form an.

Hinweis: Sie dürfen verwenden, dass S_n und T_{n+1} unabhängig sind.

Lösungsvorschlag: Mit f_T sei die Dichte der Exponentialverteilung mit Parameter λ bezeichnet.

(a) $(n = 1)$: Es gilt $S_1 = T_1$, also

$$f_{S_1}(z) = I_{[0,\infty)}(z) \cdot \lambda e^{-\lambda z} = I_{[0,\infty)}(z) \cdot \frac{\lambda^1 z^0}{0!} e^{-\lambda z}.$$

Hierbei bezeichnet $I_{[0,\infty)}$ die charakteristische Funktion des Intervalls $[0, \infty)$.

$(n \to n+1)$: Da $T_1, \ldots, T_{n+1}$ unabhängig sind, sind auch S_n und T_{n+1} unabhängig.
Damit ist die Dichte von $f_{S_{n+1}}$ durch die Faltung von $f_{T_n} = f_T$ und f_{S_n} gegeben:

$$f_{S_{n+1}}(z) = \int_{-\infty}^{\infty} f_T(z-t) f_{S_n}(t) \mathrm{d}t$$

Man überlegt sich zunächst, dass der Integrand 0 ist für $t < 0$ und $z < t$:

- Für $t < 0$ folgt dies, da f_{S_n} diese Eigenschaft nach Induktionsannahme besitzt.
- Für $z < t$ folgt dies aus der Definition der Exponentialverteilung.

Wir müssen somit nur über $[0, z]$ integrieren:

$$\begin{aligned} f_{S_{n+1}}(z) = \int_0^z f_T(z-t) f_{S_n}(t) \mathrm{d}t &= \int_0^z \lambda e^{-\lambda(z-t)} \cdot \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t} \mathrm{d}t \\ &= \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda z} \int_0^z t^{n-1} \mathrm{d}t = \frac{\lambda^{n+1}}{(n-1)!} e^{-\lambda z} \frac{z^n}{n} = \frac{\lambda^{n+1} z^n}{n!} e^{-\lambda z}. \end{aligned}$$

(b) Sei $z \geq 0$ und $n \geq 2$:

$$F_{S_{n+1}}(z) = \int_{-\infty}^z f_{S_{n+1}}(t) \mathrm{d}t = \int_0^z \frac{\lambda^{n+1} t^n}{n!} e^{-\lambda t} \mathrm{d}t = \int_0^z -\frac{\lambda^n t^n}{n!} \cdot -\lambda e^{-\lambda t} \mathrm{d}t$$

(Partielle Intergration: $U = -\frac{\lambda^n t^n}{n!}$, $v = -\lambda e^{-\lambda t}$, $u = -\frac{\lambda^n t^{n-1}}{(n-1)!}$, $V = e^{-\lambda t}$)

$$\begin{aligned} &= \left[-\frac{\lambda^n t^n}{n!} e^{-\lambda t} \right]_0^z - \int_0^z -\frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t} \mathrm{d}t \\ &= -\frac{\lambda^n z^n}{n!} e^{-\lambda z} + \int_0^z f_{S_{n-1}}(t) \mathrm{d}t \\ &= -\frac{\lambda^n z^n}{n!} e^{-\lambda z} + F_{S_n}(z). \end{aligned}$$

Da $F_{S_1}(z) = F_T(z) = 1 - e^{-\lambda z}$ folgt mittels Induktion nach n:

$$F_{S_n}(z) = 1 - e^{-\lambda z} \sum_{k=0}^{n-1} \frac{\lambda^k z^k}{k!}.$$