Technische Universität München
Fakultät für Informatik
Lehrstuhl Informatik XIV
Prof. Dr. Susanne Albers
Dr. Werner Meixner

Sommersemester 2014
Lösungsblatt 8
18. Juni 2014

Diskrete Wahrscheinlichkeitstheorie

Hausaufgabe 1 (5 Punkte)

Seien $X_1, X_2, \ldots, X_{100}$ unabhängige diskrete Zufallsvariable, die gleichverteilt auf $\{1, 2, \ldots, 20\}$ sind. Wir nehmen Zufallsvariablen $Y_i \sim \text{Bin}(1; \frac{8}{20})$ an, die genau dann den Wert 1 liefern, wenn $X_i > 12$ gilt.

Bestimmen Sie mit Hilfe einer geeigneten Chernoff-Schranke nach Vorlesung eine möglichst gute obere Schranke für die Wahrscheinlichkeit des Ereignisses

$$Y_1 + Y_2 + \ldots + Y_{100} \geq 50\,.$$

Lösung

Es gilt $p := \Pr[Y_i = 1] = \frac{8}{20}$.
Sei $Y = \sum_{i=1}^{100} Y_i$. Dann gilt $\mu := \mathbb{E}[Y] = 100 \cdot \frac{8}{20} = 40$.
Nach Chernoff gilt für alle $\delta \geq 0$

$$\Pr[X \geq (1+\delta)\mu] \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu .$$

Damit erhalten wir mit $\delta = \frac{1}{4} = 0,25$ und $50 = (1 + \frac{1}{4}) \cdot 40$

$$\Pr[Y \geq 50] \leq \left(\frac{e^{0,25}}{1,25^{1,25}}\right)^{40} \approx 0,31\,.$$

(5P)

Hausaufgabe 2 (5 Punkte)

Wir betrachten ein Münzwurfexperiment, das darin besteht, jede von drei unterschiedlichen Münzen A bzw. B bzw. C so lange zu werfen, bis Kopf erscheint. Dabei nehmen wir an, dass die Erfolgswahrscheinlichkeiten für einen einzigen Wurf mit A bzw. B bzw. C die Werte $p_1 = \frac{1}{3}$ bzw. $p_2 = \frac{1}{2}$ bzw. $p_3 = \frac{2}{3}$ sind. Die Münzen A und C sind also unfair.

X_A bzw. X_B bzw. X_C seien die entsprechenden unabhängigen Zufallsvariablen, die die Anzahl der Würfe mit A bzw. B bzw. C zählen. Die Gesamtzahl der Würfe sei gegeben durch die Zufallsvariable $Y = X_A + X_B + X_C$.

1. Sei $G_Y(s)$ die wahrscheinlichkeitserzeugende Funktion für Y. Bestimmen Sie $G'_Y(0)$.
2. Sei f_Y die Dichtefunktion von Y. Bestimmen Sie $f_Y(4)$.
3. Bestimmen Sie den Erwartungwert $\mathbb{E}[Y]$.
4. Zeigen Sie $\Pr[Y \geq 16{,}5] \leq \frac{1}{10}$.
   Hinweis: Benutzen Sie die Ungleichung von Chebyshev.

Lösung

1. $f_Y(1) = 0$. Daraus folgt $G'_Y(0) = 0$. (1P)

2. $f_Y(4) = p_1p_2p_3(1-p_1) + p_1p_2p_3(1-p_2) + p_1p_2p_3(1-p_3) = \frac{1}{6}$. (2P)

3. $\mathbb{E}[Y] = \mathbb{E}[X_A] + \mathbb{E}[X_B] + \mathbb{E}[X_C] = \frac{13}{2}$. (1P)

4. $\mathrm{Var}[Y] = \mathrm{Var}[X_A] + \mathrm{Var}[X_B] + \mathrm{Var}[X_C] = \frac{35}{4}$.
   $\Pr[Y \geq 16{,}5] = \Pr[|Y - 6{,}5| \geq 10] \leq \frac{35}{400} \leq \frac{1}{10}$. (1P)

Hausaufgabe 3 (5 Punkte)

Sei X eine binomialverteilte Zufallsvariable mit Parametern $n = 4$ und $p = \frac{1}{2}$, d.h. $X \sim \mathrm{Bin}(4, \frac{1}{2})$.

1. Geben Sie die erzeugende Funktion $G_X(s)$ in geschlossener Form an.

2. Berechnen Sie den Erwartungswert der bedingten Variablen $X|X \neq 2$.

3. Ein Experiment bestehe darin, dass die Zufallsvariable X wiederholt ausgewertet wird, und zwar so oft, bis bei der n-ten Wiederholung der Wert 2 erstmalig erscheint. Dann wird die Summe der aufgetretenen Werte $\neq 2$ gebildet.

   Sei X_i für $i \in \mathbb{N}$ die i-te Wiederholung von X, sei N die Zufallsvariable, die die Nummer n der letzten Wiederholung darstellt, und sei $S = \sum_{i=1}^{N-1} X_i$.

   Berechnen Sie den Erwartungswert $\mathbb{E}[S]$ von S.

Lösung

1. $G_X(s) = (1 - \frac{1}{2} + \frac{1}{2}s)^4 = \frac{1}{16}(1+s)^4 =$
   $= \frac{1}{16} + \frac{4}{16}s + \frac{6}{16}s^2 + \frac{4}{16}s^3 + \frac{1}{16}s^4$. (1P)

2.

$$\Pr[X = x | X \neq 2] = \begin{cases} \frac{\Pr[X=x]}{\Pr[X\neq 2]} & : \text{falls } x \neq 2 \\ 0 & : \text{sonst} \end{cases}$$
$$= \begin{cases} \frac{1}{10} & : \text{falls } x = 0 \ \vee \ x = 4 \\ \frac{2}{5} & : \text{falls } x = 1 \ \vee \ x = 3 \\ 0 & : \text{sonst} \end{cases}$$

   Es folgt: $\mathbb{E}[X|X \neq 2] = \frac{1}{10} \cdot 0 + \frac{2}{5} \cdot 1 + \frac{2}{5} \cdot 3 + \frac{1}{10} \cdot 4 = 2$. (2P)

3. N ist geometrisch verteilt mit $p = \frac{3}{8}$.
   Es folgt $\mathbb{E}[N] = \frac{8}{3}$, mithin $\mathbb{E}[N-1] = \frac{8}{3} - 1 = \frac{5}{3}$.
   Es folgt $\mathbb{E}[S] = \mathbb{E}[N-1] \cdot \mathbb{E}[X|X \neq 2] = \frac{5}{3} \cdot 2 = \frac{10}{3}$.
   (2P)

Hausaufgabe 4 (5 Punkte)

Es sei M_1 eine Maschine, die bei Aufruf zufällig mit gleicher Wahrscheinlichkeit eine der Zahlen 1, 2, 3 oder 4 ausgibt. Wir bezeichnen die entsprechende Zufallsvariable mit N. Eine Maschine M_2 werfe bei Aufruf eine faire Münze, die entweder „Kopf“ oder „Wappen“ zeigt.

Wir betrachten einen Algorithmus A, dessen Ausführung in 2 Schritten ein Ergebnis erzeugt. Im ersten Schritt wird M_1 veranlasst, eine Zahl k auszugeben. Im zweiten Schritt wird M_2 k mal aufgerufen. Das Ergebnis einer Ausführung von A definieren wir als diejenige Zahl, die angibt, wie oft im zweiten Schritt „Kopf“ geworfen wurde.
Es sei Y die Zufallsvariable, die die Ergebnisse des Algorithmus A beschreibt.

1. Geben Sie die wahrscheinlichkeitserzeugende Funktion $G_N(z)$ für N an.
2. Berechnen Sie den Erwartungswert $\mathbb{E}[Y]$.
3. Berechnen Sie die wahrscheinlichkeitserzeugende Funktion $G_Y(z)$ für Y.

Lösung

1. Ausgehend von der Gleichverteilung für 4 Ereignisse erhalten wir

$$G_N(z) = \sum_{k=1}^{4} \frac{1}{4} z^k \,.$$

(1P)

2. Die Indikatorvariable X gebe mit Wert 1 an, dass Kopf geworfen wurde. Die erzeugende Funktion für X ist

$$G_X(z) = \frac{1}{2} + \frac{1}{2} z.$$

Es gilt $G'_N(z) = \frac{1}{4} + \frac{2}{4} z + \frac{3}{4} z^2 + \frac{4}{4} z^3$

und $G'_X(z) = \frac{1}{2}\,.$

Mit $\mathbb{E}[X] = G'_X(1) = \frac{1}{2}$

und $\mathbb{E}[N] = G'_N(1) = \frac{1}{4}(1+2+3+4) = \frac{5}{2}$

erhalten wir

$$\mathbb{E}[Y] = \mathbb{E}[N] \cdot \mathbb{E}[X] = \frac{5}{2} \cdot \frac{1}{2} = \frac{5}{4}\,.$$ (2P)

3.

$$\begin{aligned} G_Y(z) &= G_N(G_X(z)) \\ &= \frac{1}{4}\left(\frac{1+z}{2} + \left(\frac{1+z}{2}\right)^2 + \left(\frac{1+z}{2}\right)^3 + \left(\frac{1+z}{2}\right)^4\right) \\ &= \frac{1}{64}(15 + 26z + 16z^2 + 6z^3 + z^4)\,. \end{aligned}$$

(2P)

__Hinweis:__ Die Vorbereitungsaufgaben bereiten die Tutoraufgaben vor und werden in der Zentralübung unterstützt. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Hausaufgaben sollen selbstständig bearbeitet und zur Korrektur und Bewertung abgegeben werden.

Vorbereitung 1

Seien X_1 und X_2 unabhängige exponentialverteilte Zufallsvariablen mit Parametern λ_1 bzw. λ_2.

1. Berechnen Sie die Dichtefunktion von $Y = X_1 + X_2$ durch Anwendung der Faltungsformel

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X_1}(x) \cdot f_{X_2}(y-x)\,\mathrm{d}x$$

und vereinfachen Sie das Ergebnis im Fall $\lambda_1 = \lambda_2$ so weit wie möglich.

2. Seien X_1, X_2, X_3 unabhängig exponentialverteilt mit gleichem Parameter λ und $Y = X_1+X_2+X_3$. Berechnen Sie die Verteilungsfunktion F_Y in geschlossener Form.

Lösung

1. Es gilt $f_{X_1}(x) = \lambda_1 e^{-\lambda_1 x}$ und $f_{X_2}(x) = \lambda_2 e^{-\lambda_2 x}$. Damit folgt, wenn $\lambda_1 \neq \lambda_2$ gilt, für $y \geq 0$

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 (y-x)}\,\mathrm{d}x \\
&= \lambda_1\lambda_2 e^{-\lambda_2 y} \int_0^y e^{(\lambda_2-\lambda_1)x}\,\mathrm{d}x \\
&= \lambda_1\lambda_2 e^{-\lambda_2 y} \left[\frac{e^{(\lambda_2-\lambda_1)x}}{\lambda_2-\lambda_1}\right]_{x=0}^{x=y} \\
&= \lambda_1\lambda_2 e^{-\lambda_2 y} \cdot \frac{e^{(\lambda_2-\lambda_1)y}-1}{\lambda_2-\lambda_1} \\
&= \lambda_1\lambda_2 \cdot \frac{e^{-\lambda_1 y}-e^{-\lambda_2 y}}{\lambda_2-\lambda_1}\,.
\end{aligned}$$

Im Fall $\lambda_1 = \lambda_2 =: \lambda$ gilt

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} \lambda e^{-\lambda x} \cdot \lambda e^{-\lambda(y-x)}\,\mathrm{d}x \\
&= \lambda^2 e^{-\lambda y} \int_0^y \mathrm{d}x \\
&= \lambda^2 y e^{-\lambda y}\,.
\end{aligned}$$

Für $y \leq 0$ folgt in allen Fällen direkt $f_Y(y) = 0$.

2. Wir wenden die Faltungsformel noch einmal an, und zwar auf $f_{X_1+X_2}$ und f_{X_3} wie folgt.

$$\begin{aligned}
f_Y(y) &= \int_{-\infty}^{\infty} f_{X_1+X_2}(x) \cdot f_{X_3}(y-x)\,\mathrm{d}x \\
&= \int_0^y \lambda^2 x e^{-\lambda x} \cdot \lambda e^{-\lambda(y-x)}\,\mathrm{d}x \\
&= \lambda^3 e^{-\lambda y} \cdot \int_0^y x\,\mathrm{d}x \\
&= \frac{\lambda^3 y^2}{2} e^{-\lambda y}\,.
\end{aligned}$$

Die Verteilungsfunktion F_Y kann nun durch Integration der Dichtefunktion f_Y berechnet werden, wie es im Folgenden ausführlich dokumentiert wird. Wir wenden insbesondere partielle Integration an.

$$
\begin{aligned}
F_Y(y) &= \int_{-\infty}^{y} f_Y(x)\,\mathrm{d}x \\
&= \int_0^y \frac{\lambda^3 x^2}{2} e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2) \int_0^y \frac{x^2}{2} \cdot (-\lambda) e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2) \left[\frac{x^2}{2} \cdot e^{-\lambda x}\right]_{x=0}^{x=y} \;-\; (-\lambda^2) \int_0^y x \cdot e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2) \frac{y^2}{2} e^{-\lambda y} \;-\; \lambda \int_0^y x \cdot (-\lambda) e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2) \frac{y^2}{2} e^{-\lambda y} \;-\; \lambda \left(\left[x \cdot e^{-\lambda x}\right]_{x=0}^{x=y} - \int_0^y e^{-\lambda x}\,\mathrm{d}x \right) \\
&= (-\lambda^2) \frac{y^2}{2} e^{-\lambda y} \;-\; \lambda y e^{-\lambda y} - \int_0^y (-\lambda) e^{-\lambda x}\,\mathrm{d}x \\
&= (-\lambda^2) \frac{y^2}{2} e^{-\lambda y} \;-\; \lambda y e^{-\lambda y} - \left[e^{-\lambda x}\right]_{x=0}^{x=y} \\
&= 1 \;-\; \frac{\lambda^2 y^2}{2} e^{-\lambda y} \;-\; \lambda y e^{-\lambda y} \;-\; e^{-\lambda y}\,.
\end{aligned}
$$

Vorbereitung 2

Sei X eine kontinuierliche Zufallsvariable.

1. Zeigen Sie: Falls $X \sim \mathcal{N}(2, \frac{1}{2})$, dann gilt $2X + 1 \sim \mathcal{N}(5, 2)$.
2. Seien $d_1, d_2, c \in \mathbb{R}$ mit $d_1 < d_2$ und $c > 0$. Berechnen Sie $a, b \in \mathbb{R}$, so dass für $Y = aX + b$ gilt

$$\Pr[d_1 \le X \le d_2] = \Pr[-c \le Y \le c]\,.$$

Lösung

1. Seien μ und σ der Erwartungswert bzw. die Varianz von X, d. h. $\mu = 2$ bzw. $\sigma^2 = \frac{1}{2}$. Nach Satz der Vorlesung ist $2X + 1$ normalverteilt mit Erwartungswert $2\mu + 1 = 5$ bzw. Varianz $2^2\sigma^2 = 2$. W. z. b. w.
2. Sei $a > 0$. Dann gilt

$$d_1 \le X \le d_2 \quad \Longleftrightarrow \quad ad_1 + b \le Y \le ad_2 + b\,.$$

Wir lösen für a, b die Gleichungen

$$ad_1 + b = -c \quad \text{und} \quad ad_2 + b = c\,.$$

Lösung:

$$a = \frac{2c}{d_2 - d_1}\,, \qquad b = \frac{d_1 + d_2}{d_1 - d_2} \cdot c\,.$$

Vorbereitung 3

Wir betrachten unabhängige stetige Zufallsvariablen X und Y, die beide auf dem Intervall $[0, 1] \subseteq \mathbb{R}$ gleichverteilt sind. Sei $Z = \max\{X, Y\}$.

1. Berechnen Sie die Verteilungsfunktion F_Z.
2. Bestimmen Sie eine Funktion $u : [0, 1] \to [0, 1]$, so dass $u(X)$ die gleiche Verteilung wie Z besitzt.

Lösung

1. Die gemeinsame Dichte der Zufallsvariablen X, Y sei $f_{X,Y}(x, y)$. Aufgrund der Unabhängigkeit von X, Y gilt für $(x, y) \notin [0, 1] \times [0, 1]$ die gemeinsame Dichte 0 und für $(x, y) \in [0, 1] \times [0, 1]$

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) = 1 \cdot 1 = 1\,.$$

Wir berechnen die Verteilungsfunktion $F_Z(z)$. Offenbar gilt zunächst $F_Z(z) = 0$ bzw. $F_Z(z) = 1$ für $z \leq 0$ bzw. $1 \leq z$. Für $z \in [0, 1] \subseteq \mathbb{R}$ gilt

$$\begin{aligned} F_Z(z) &= \Pr[\max\{X, Y\} \leq z] \\ &= \Pr[X \leq z, Y \leq z] \\ &= \int_{[0,z]\times[0,z]} f_{X,Y}(x, y)\, \mathrm{d}x\mathrm{d}y \\ &= \int_{[0,z]\times[0,z]} 1\, \mathrm{d}x\mathrm{d}y \\ &= z^2\,. \end{aligned}$$

2. Aus der Vorlesung ist bekannt, dass wir eine Simulation von F_Z aus der Inversen von F_Z erhalten können.

Wir rechnen direkt und setzen die Invertierbarkeit von u voraus. Sei $Y = u(X)$. Dann gilt

$$\begin{aligned} F_Y(y) &= \Pr[Y \leq y] = \Pr[u(X) \leq y] \\ &= \Pr[X \leq u^{-1}(y)] \\ &= F_X(u^{-1}(y)) \\ &= u^{-1}(y)\,. \end{aligned}$$

Aus der Gleichung $F_Z(y) = F_X(u^{-1}(y))$ folgt nun $y^2 = u^{-1}(y)$, mithin

$$u(x) = \sqrt{x}\,.$$

Tutoraufgabe 1

Seien X und Y kontinuierliche Zufallsvariablen mit gemeinsamer Dichtefunktion

$$f_{X,Y}(x,y) = \begin{cases} 6xy^2 & : \quad 0 \le x \le 1,\, 0 \le y \le 1 \\ 0 & : \quad \text{sonst} \end{cases}$$

1. Berechnen Sie die Randdichte $f_X(x)$.
2. Bestimmen Sie den Wert der Verteilungsfunktion $F_{X,Y}(\frac{1}{2}, \frac{1}{2})$.
3. Zeigen Sie die Unabhängigkeit der Variablen X und Y.

Lösung

1. $f_X(x) = 2x$. Berechnung für $0 \le x \le 1$:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y \\ &= \int_0^1 6xy^2\,\mathrm{d}y \\ &= 2x \cdot \left[y^3\right]_0^1 \;=\; 2x\,. \end{aligned}$$

2. $F_{X,Y}(\frac{1}{2}, \frac{1}{2}) = \frac{1}{32}$. Berechnung:

$$\begin{aligned} F_{X,Y}(\tfrac{1}{2}, \tfrac{1}{2}) &= \int_0^{\frac{1}{2}} \left(\int_0^{\frac{1}{2}} 6xy^2\,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_0^{\frac{1}{2}} 2x \cdot \left[y^3\right]_0^{\frac{1}{2}}\,\mathrm{d}x \\ &= \int_0^{\frac{1}{2}} 2x \cdot \frac{1}{8}\,\mathrm{d}x \\ &= \frac{1}{8} \cdot \left[x^2\right]_0^{\frac{1}{2}} \;=\; \frac{1}{32}\,. \end{aligned}$$

3. Mit

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y \\ &= \int_0^1 6xy^2\,\mathrm{d}x \\ &= 3y^2 \cdot \left[x^2\right]_0^1 \;=\; 3y^2\,. \end{aligned}$$

für alle $0 \le y \le 1$ folgt $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$ für alle $0 \le x, y \le 1$.
Ansonsten gilt $f_{X,Y}(x,y) = 0 = f_X(x) \cdot f_Y(y)$.

Tutoraufgabe 2

In einem Unfallkrankenhaus treffen im Schnitt alle 20 Minuten Patienten zur Behandlung ein. Die Zeit zwischen zwei Behandlungsfällen sei exponentialverteilt mit Parameter $\frac{1}{20}$. Wenn 1 Stunde lang kein Patient eingetroffen ist, macht das Personal Ruhepause. Wir wollen wissen, welcher Zeitabstand zwischen zwei Ruhepausen zu erwarten ist.
Seien $T_1, T_2, \ldots$ die Zeitspannen zwischen dem Eintreffen zweier Behandlungsfälle und W die Wartezeit bis zur nächsten Ruhepause.

1. Geben Sie $\mathbb{E}[T_1 \mid T_1 \geq 60]$ an.
2. Geben Sie $\mathbb{E}[W \mid T_1 \geq 60]$ an.
3. Berechnen Sie $\mathbb{E}[W]$.

Lösung

Vorüberlegungen:

a) Es wird offenbar eine Folge von Zeitpunkten z_i betrachtet,
zu denen Patienten im Krankenhaus eintreffen.

b) Die Zeitdifferenzen $T_i = z_i - z_{i-1}$ werden als exponentialverteilt angenommen. Dies bedeutet, dass diese T_i nicht davon abhängen, wie lange noch kein Patient eingetroffen ist (Gedächtnislosigkeit).

Der Parameter λ bedeutet „Anzahl der Patienten pro Zeiteinheit" im Durchschnitt, hier also $\frac{1}{20}$ Patient pro Zeiteinheit als Erwartungswert.

Alle T_i sind unabhängig, d. h. die Menge der T_i ist unabhängig.

c) Falls $T_i > 60$, dann gibt es eine Ruhepause.

d) Die Frage ist, wie lange man warten muss, bis erstmalig $T_N > 60$ festgestellt wird?

Die Wartezeit ist dann

$$W = 60 + \sum_{j=1}^{N-1} T_j$$

oder

$$W = 60 + W'$$

mit

$$W' = \sum_{j=1}^{N-1} T_j \,.$$

1. Sei $\lambda = \frac{1}{20}$. Dann gilt $\mathbb{E}[T_1] = \frac{1}{\lambda} = 20$.

   Da die Exponentialverteilung gedächtnislos ist, gilt

   $$\mathbb{E}[T_1 \mid T_1 \geq 60] = 60 + \mathbb{E}[T_1] = 80\,.$$

   Wir zeigen die Gleichung direkt durch Berechnung wie folgt:

   T_1 ist exponentialverteilt:

   $$F_{T_1}(t) = \begin{cases} 1 - e^{-\lambda t} & : t \geq 0 \\ 0 & : \text{sonst}\,. \end{cases}$$

Verteilung der bedingten Variablen $T_1 \mid T_1 \geq 0$:

$$\begin{aligned} F_{(T_1|T_1 \geq 60)}(t) &= \Pr[(T_1 \mid T_1 \geq 0) \leq t] \\ &= \frac{\Pr[T_1 \leq t, T_1 \geq 60]}{\Pr[T_1 \geq 60]} \end{aligned}$$

Für $t \geq 60$ folgt

$$\begin{aligned} F_{(T_1|T_1 \geq 60)}(t) &= \frac{(1 - e^{-\lambda \cdot t}) - (1 - e^{-\lambda \cdot 60})}{e^{-\lambda \cdot 60}} \\ &= 1 - e^{-\lambda \cdot (t-60)}\,. \end{aligned}$$

Damit ist die Variable $T' = (T_1 \mid T_1 \geq 60) - 60$ gleichverteilt wie T_1, d. h. entsprechend exponentialverteilt.

Es folgt

$$\begin{aligned} \mathbb{E}[(T_1 \mid T_1 \geq 60)] &= \mathbb{E}[T' + 60] \\ &= \mathbb{E}[T_1] + 60 \\ &= 80\,. \end{aligned}$$

2. Offenbar gilt

$$\mathbb{E}[W \mid T_1 \geq 60] = 60\,.$$

3. Es gilt $W = 60 + \sum_{j=1}^{N-1} T_j$ oder $W = 60 + W'$ mit $W' = \sum_{j=1}^{N-1} T_j\,.$

N ist eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit $p = \Pr[T \geq 60]$. Es gilt

$$p = \Pr[T \geq 60] = e^{-3}\,.$$

Berechnung von $\mathbb{E}[W']$:

Wir setzen $T = T_1$.

$$\begin{aligned} \mathbb{E}[W'] &= \sum_{n=1}^{\infty} \mathbb{E}[W'|N = n] \cdot p(1-p)^{n-1} \\ &= \sum_{n=1}^{\infty} \mathbb{E}[T|T \leq 60] \cdot (n-1) \cdot p(1-p)^{n-1} \\ &= \mathbb{E}[T|T \leq 60] \cdot \mathbb{E}[N-1]. \end{aligned}$$

Es gilt

$$\mathbb{E}[N-1] = e^3 - 1\,.$$

$\mathbb{E}[T|T \leq 60]$ erhalten wir aus den Gleichungen

$$\begin{aligned} \mathbb{E}[T] &= \mathbb{E}[T|T \leq 60] \cdot \Pr[T \leq 60] \;+\; \mathbb{E}[T|T \geq 60] \cdot \Pr[T \geq 60] \\ &= \mathbb{E}[T|T \leq 60] \cdot (1 - e^{-3}) + 80 \cdot e^{-3} \\ &= 20\,. \end{aligned}$$

Es folgt

$$\mathbb{E}[T|T \leq 60] \quad = \quad \frac{20 - 80 \cdot e^{-3}}{1 - e^{-3}} \quad = \quad \frac{20 \cdot e^3 - 80}{e^3 - 1}\,.$$

Ergebnis:

$$\begin{aligned} \mathbb{E}[W'] &= 20 \cdot e^3 - 80\,, \\ \mathbb{E}[W] &= 20 \cdot e^3 - 80 + 60 \\ &\approx 382 \text{ (Minuten)}\,. \end{aligned}$$

Tutoraufgabe 3

Wir benutzen die Funktion $h(t) = 0.027 + 0.0025 \cdot (t - 40)^2$ für $t \in \mathbb{R}$, um die „Sterblichkeitsrate“ durch Lungenkrebs von Kettenraucherinnen abzuschätzen, die mindestens $t \geq 40$ Jahre alt sind. Ihre Lebensdauer sei X und es gelte

$$\Pr[X > t \,|\, X > 40] = \exp\left(-\int_{40}^{t} h(s)\mathrm{d}s\right)\,.$$

Wie groß ist die Wahrscheinlichkeit, dass eine 45-jährige Kettenraucherin mindestens 50 Jahre alt wird?

Lösung

Zu berechnen ist offenbar $\Pr[X \geq 50 \mid X \geq 45]$.
Bei stetigen Zufallsvariablen gilt $\Pr[X \geq 50 \mid X \geq 45] = \Pr[X > 50 \mid X > 45]$.
Für $t \geq s \geq 40$ berechnen wir allgemein die folgende bedingte Wahrscheinlichkeit.

$$\begin{aligned} \Pr[X > t \mid X > s] &= \frac{\Pr[(X > t) \cap (X > s)]}{\Pr[X > s]} \\ &= \frac{\Pr[X > t]}{\Pr[X > s]} \\ &= \frac{\Pr[X > t \mid X > 40]}{\Pr[X > s \mid X > 40]} \\ &= \exp\left(-\int_{40}^{t} h(x)\ \mathrm{d}x + \int_{40}^{s} h(x)\ \mathrm{d}x\right) \\ &= \exp\left(-\int_{s}^{t} h(x)\ \mathrm{d}x\right)\,. \end{aligned}$$

Daher gilt

$$\Pr[X > 50 \mid X > 45] \quad = \quad \exp\left(-\int_{45}^{50} 0.027 + 0.0025(t - 40)^2\ \mathrm{d}t\right)\,,$$

woraus wegen

$$\int_{45}^{50} 0.027 + 0.0025(t - 40)^2\ \mathrm{d}t \quad = \quad 0.135 + 2.1875/3 \quad = \quad 0.86417\bar{6}$$

folgt, dass gilt

$$\Pr[X \geq 50 \mid X \geq 45] \quad = \quad \exp(-0.86417\bar{6}) \quad \approx \quad 42.1402575\ldots\,\%\,.$$